NCVS Status and Progress Report
Volume 13/April 1999

The National Center for Voice and Speech is a consortium of institutions--The University of Iowa, The Denver Center for the Performing Arts, The University of Wisconsin-Madison and The University of Utah--whose investigators are dedicated to the rehabilitation, enhancement and protection of voice and speech.

Editorial and Distribution Information

Editor, Ingo Titze
Production Editors, Julie Stark and Julie Ostrem
Technical Editor, Martin Milder

Distribution of this report is not restricted.
However, production was limited to 700 copies.

Correspondence should be addressed as follows:

Editor, NCVS Status and Progress Report
The University of Iowa
330 Wendell Johnson Building
Iowa City, Iowa 52242
(319) 335-6600
FAX (319) 335-8851
e-mail: titze@shc.uiowa.edu
website: http://www.ncvs.org

Primary Sponsorship

The National Institute on Deafness and Other Communication Disorders,
Grant Number P60 DC00976

Other Sponsorship

The University of Iowa
Department of Speech Pathology and Audiology
Department of Otolaryngology - Head and Neck Surgery

The Denver Center for the Performing Arts
Wilbur James Gould Voice Research Center
Department of Public Relations
Department of Public Affairs
Denver Center Media
Department of Development

The University of Wisconsin-Madison
Department of Communicative Disorders
Department of Surgery, Division of Otolaryngology
Waisman Center
Department of Electrical and Computer Engineering

The University of Utah
Department of Otolaryngology - Head and Neck Surgery
LDS Hospital

The University of Illinois
Department of Speech and Hearing Science

NCVS Personnel

Administration

Central Office
Ingo Titze, Director
Julie Ostrem, Program Associate
Julie Stark, Secretary

Area Coordinators

Research - Ingo Titze
Training - Patricia Zebrowski
Continuing Education - Julie Ostrem
Information Dissemination - Thea Carruth

Advisory Board

Katherine Harris, Ph.D.
Minoru Hirano, M.D.
Clarence Sasaki, M.D.
Johan Sundberg, Ph.D.

Investigators, Affiliates and Support Staff

Fariborz Alipour, Ph.D.
Kristin Baker, Ph.D.
David Berry, Ph.D.
Alyson Beylar, B.S.
Florence Blager, Ph.D.
Diane Bless, Ph.D.
Myrna Burt
John Canady, M.D.
Thea Carruth, M.P.H.
Roger Chan, Ph.D.
Geron Coale, M.A.
Christy Dahl, B.A.
Charles Davis, Ph.D.
Hu Ding, M.D., Ph.D.
Mindy Ebert
Anna Edmonds, B.A.
Catherine Emerich, M.S.
Wendy Fick, B.A.
Jeffrey Fields, B.M.
Eileen Finnegan, Ph.D.
John Folkins, Ph.D.
Charles Ford, M.D.
Steven Gray, M.D.
Elizabeth Hammond, M.D.
Henry Hoffman, M.D.
Richard Hurtig, Ph.D.
Michael Karnell, Ph.D.
David Kuehn, Ph.D.
Russel Long, M.S.
Daniel McCabe, D.M.A.
Martin Milder, B.S.
Paul Milenkovic, Ph.D.
Jerald Moon, Ph.D.
Lea Ann Moriarty, B.A.
Lorraine Olson Ramig, Ph.D.
Julie Ostrem, B.S.
Donald Robin, Ph.D.
Robin Samlan, M.S.
Shimon Sapir, Ph.D.
Richard Schmidt, Ph.D.
Nicole Seuer, B.S.
Elaine Smith, Ph.D.
Marshall Smith, M.D.
Martin Spencer, B.A.
Jennifer Spielman, M.A.
Julie Stark, A.A.
Brad Story, Ph.D.
Edie Swift, M.S.
Sue Ann Thompson, Ph.D.
Ingo Titze, Ph.D.
Jane VanDeWiel, B.S.
Katherine Verdolini, Ph.D.
George Woodworth, Ph.D.
Patricia Zebrowski, Ph.D.
Lynn Zimba, Ph.D.
Nicole Zinda, B.S.

Doctoral Students

Christine Bergan, M.M.
Renee Bogschutz, M.S.
Jay Granier, M.A.
Elisa Huff, M.S.
Eric Hunter, M.S.
Melda Kunduk, M.S.
Douglas Montequin, M.A.
Phyllis Palmer, M.A.
Helen Sharp, M.S.
Susan Thibeault, M.S.

Postdoctoral Fellows

Michael Edgerton, D.M.A.
Kirrie Ballard, Ph.D.

Visiting Scholars

Anne-Maria Laukkanen, Finland
Niro Tayama, M.D., Japan

Contents

Part I. Research papers submitted for peer review in archival journals

Part II. Tutorial report

Foreword

The following essay points out the importance of quality of life issues with regard to voice and how our research can help address these issues. This piece originally appeared in the Recurrent Respiratory Papillomatosis Newsletter (Volume 5 No.2) and is reprinted with the permission of its author.

-Ingo R. Titze, NCVS Director

By Andrea Behr

As I began to think about writing my college essays, it soon became clear to me what my topic should be. There is something about me that has shaped my entire life but can only be described through an essay. I have had a rare disease called laryngeal papillomas, since I was four months old, and because of the nature of the disease (small bumps on my throat and vocal cords), I have whispered all my life.

For seventeen years I have found myself forgetting that I am different from other people in one very specific and obvious way - they have a voice and I do not. My small elementary school was primarily responsible for this because it provided me with an environment in which no one ever questioned my ability to do something just as well as others. For instance, in eighth grade I had a speaking part in our class play, something that today seems more unusual than it did back then. Like most people who begin the college admissions process, I was forced to review my failures and accomplishments more closely than ever. As I did this, I found myself wishing that I could have sung in the recent holiday assembly, or acted in the fall play along with many of my friends. I realize, however, that for all of these things I might have done, there are many other aspects about me that can substitute for these losses. I am just as talkative as my friends, I ask just as many questions in class as other students, and I am just as optimistic about my future as anyone else with a voice.

My condition has forced me to undergo approximately sixty operations at The Children's Hospital of Philadelphia. My friends and family have always been supportive during and after these operations, but I am the only one that can understand the awful feeling in the pit of my stomach as I enter the hospital. Although I have always dreaded my own visits to the hospital, which are now less frequent, I am surprisingly interested in the many functions and responsibilities involved with the hospital, including the lives of physicians, nurses, and administration. To satisfy this interest, I decided to volunteer last summer at my hospital. No one understood why I would want to return to a place I had unwillingly been so many times, but I felt a need to be on the other side—a care giver instead of a patient. I worked with babies in the Infant Transitional Unit of the building, holding and feeding them, and cheerfully talking to their parents. Indeed, this experience did not quell my interest in children and hospital, it actually opened the door further.

When I first meet someone, I always hope that he or she would listen to what I am saying and not comment upon my whisper so that I will not have to explain my voice. Of course this never happens, but I have learned that a person I end up liking is the person who immediately accepts me as I am and for whom I am. People are often surprised, first, that I do not have laryngitis and second, that I like to talk, but I am glad that the people who eventually get to know me seem to forget about my voice. Now that I am older, I am often asked what I want to do when I get out of college. I have learned to recognize that my life has been and will be forever affected by my handicap and that because of this my career options may be different from others, but I have also recognized that if I continue to expect as much from myself as I can (and always have), my life may not be all that different.

Part I

Research papers submitted for peer review in archival journals

NCVS Status and Progress Report - 13
April 1999, 1-10

Vocal Fold Extracellular Matrix and Its Biomechanical Influence Part I: The Fibrous Proteins

Steven D. Gray, M.D.
Department of Surgery – Division of Otolaryngology Head and Neck Surgery, The University of Utah School of Medicine
Ingo R. Titze, Ph.D.
Fariborz Alipour, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa
Thomas H. Hammond, B.S.
The University of Utah School of Medicine

Abstract

This state of the art article is the first of two that discusses the molecular composition of the vocal fold and the relation of these molecules to the biomechanical and physiological performance of the tissue. The components of the extracellular matrix may be divided into fibrous proteins and interstitial proteins. The fibrous proteins, consisting of the collagens and elastins, are the focus of this report. Elastin concentration varies by tissue depth in the vocal folds. Variation of elastin by age is reported but some controversy exists. The biomechanical terms of stress and strain (and stress-strain curves of human vocal folds) are related to the fibrous proteins of the vocal folds. The fibrous proteins, their role in stress, and their effect on the dynamic range of vocal pitch is presented.

The lamina propria (**LP**) of the vocal fold has received considerable attention from laryngologists over the last two decades. Although some interest in this tissue was present before that time, it was Hirano's histologic work, coupled with the introduction of the cover-body theory of phonation, which propelled a basic understanding of the **LP** to the laryngologist. His work was unique in that his research group not only identified histologic differences in various regions of the **LP**, but it also related those findings to vocal fold physiology with clinical implications.

The purpose of this present work is two fold:

1) To present a current concept of the composition of the extracellular matrix of the lamina propria as it relates to the vocal folds.

2) To discuss and relate biomechanical concepts of vocal fold physiology to the relevant molecular structure. If this work succeeds in its stated purpose, a third goal will be accomplished: to help students and teachers of laryngology care for the normal and abnormal voice.

Introduction

Hirano found that the lamina propria (**LP**) of the vocal fold are functionally three layers.[3,1] The most superficial layer of the **LP** (**SLLP**) was characterized histologically by loose tissue with few collagen or elastin fibers. The middle or intermediate layer of the **LP** (**MLLP**) was found to have an increased number of elastin fibers and the deep layer of the **LP** (**DLLP**) had an increased number of collagen fibers. (Figure 1; following page) The more superficial tissue was termed the *cover* and the deeper tissue was called the *body*.[2]

The notion of a cover and body fit well with studies of vocal fold vibration showing not only medial to lateral motion of tissue, but also vertical motion of tissue., The physiologic concept of the theory is that the more medial and superficial tissue slides, glides and moves over a more rigid body of tissue. Specifically a wave of *"mucosal upheaval"*, which starts in the immediate infrafold area and then travels superiority, had been identified on high-speed cinematography., This is referred to as the mucosal wave. It was recognized that for the superficial tissue, freedom of

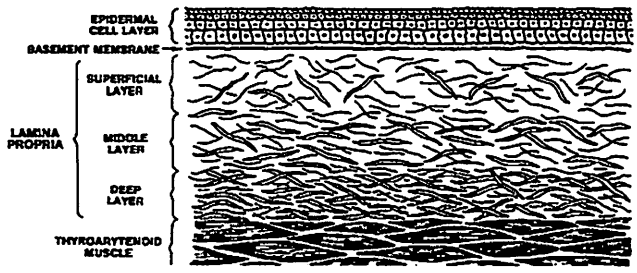


Figure 1. This figure shows the tissue of the vocal fold and its division into layers based on anatomic, functional, or biomechanical differences.(Gray, et al., 1993). [Reprint courtesy: Singular Publishing Group, Inc. (800) 521-8545, 401 West "A" Street, #325 San Diego, CA 92101-7904.]

movement in all directions was critical. The cover-body theory of phonation met that criteria and the studies of the LP provided histologic evidence that the superficial tissues allowed freedom of movement while the deeper tissues were more tightly integrated or bound.

The cover of the vocal fold was described as the epithelium, the SLLP, while the DLLP and vocalis muscle were assigned to the body. The MLLP was appropriately characterized as a transitional layer. Since the amount of tissue in vibration would depend upon intensity and pitch, the cover would appear, at times, to be thicker and would extend to include more of the intermediate layer.[32] Additionally, the normal thickness of the SLLP and MLLP probably varies somewhat according to age and this would affect what layers and how much of the MLLP is involved in the cover.[9]

Titze gave mathematical proof that self-sustained oscillation of the vocal folds was a direct result of a propagating mucosal wave.[10] The energy transfer mechanism from aerodynamics to tissue movement was shown to be in the form of a negative persistence in the nonlinear differential equation for the notion of the body and cover. The mucosa wave velocity in the cover was responsible for the negative resistance. It needs to be remembered that the cover-body concept is one of the simplest physiologic divisions of the anatomic correlates. Another concept, the mucosa-ligament-body involves three functional layers. The vocal ligament was described by Hirano as the MLLP and the DLLP. The elastin and collagen fibers making up the ligament are arranged in a longitudinal fashion so that they are relatively parallel to the thyroarytenoid muscle. In this arrangement they can respond to longitudinal stress when influenced by the intrinsic muscles of the larynx. Although the fine motor control of phonation and the registers of voice (falsetto, chest) are determined by the intrinsic laryngeal musculature, the ligament helps to balance the tensions in adjacent layers, to set the geometrical shape of the tissue to be oscillated, and to be a major bearer of stress in more intense and high pitched phonation.[32]

Extracellular Matrix

Tissue can be divided into cellular and extracellular tissue. This division is particularly useful for the vocal folds since the vast majority of the cover tissue is extracellular, but the majority of body tissue is cellular (muscle). The extracellular tissue is actually a matrix of proteins, carbohydrates, and lipids (Table 1). Although the lipids and carbohydrates are undoubtedly important, they are very difficult to study since we do not have good specific histological ways of identifying lipids or carbohydrates. The proteins are much easier to identify and studies of the extracellular matrix (ECM) have divided the proteins into fibrous and interstitial proteins.

Table 1.
Extracellular Matrix

Fibrous Proteins	Interstitial Proteins	Other Molecules
Collagens	Proteoglycans	Lipids
Elastins	Glycoproteins	Carbohydrates

Fibrous proteins are the collagens and elastins.[11] The interstitial proteins are the proteins located in the space between the fibrous proteins. They are proteoglycans and glycoproteins.[12] Both the interstitial and fibrous proteins have properties that are important to human vocal fold physiology. The fibrous proteins are important for shape, form and are designed to handle stress (a precise biomechanical term discussed later). The interstitial proteins appear to control tissue viscosity, water content, tissue size, and size and density of collagen fibers.

Fibrous Proteins

Elastin

Various types of collagens and elastins are found in the lamina propria of the human vocal fold. Elastin material occurs in three types, dependent on the amount of amorphous to fibrillar elastin. The three types are elaunin, oxytalan, and elastin fibers.[9] Oxytalan is composed of microfibrils 10-12mm in diameter. Elaunin has these microfibrils and a small amorphous component. Elastin fibers have a larger amorphous component as a core with the microfibrils surrounding the core. The elastin fibers are sometimes referred to as mature elastin, since of the three types, the elastin fibers are the most elastic (biomechanically referring to the ability to be stretched roughly two times their length and then return

to the normal length).[13] Elaunin and oxytalan are less stretchable and are found in tissue where stress is higher, such as the superficial layer of tendons and in cartilage.[14] These two types are also common in fetal and perinatal tissue and consequently were initially thought to be immature forms of elastin fibers. That concept is no longer accepted as elaunin and oxytalan have been found in many tissues in mature humans. [13,14,15]

Differentiating elastin types is significant to laryngology in that there is a significant amount of elastin proteins throughout the lamina propria, some of it not in elastin fiber form, but in elaunin and oxytalan form. This is especially true in the **SLLP**. It is important to note that of the three types of elastin material, only the elastin fiber form is stained by most of the histology stains for elastin.[16] Elaunin and oxytalan forms of elastin are not stained well, consequently these are not often visualized and at the present state of technology can only be fully identified with electron microscopy.

The amount of elastin material in all layers of the **LP** can be easily appreciated by the following simple experiment. Elastin material (all forms) does autofluoresce when stained with hematoxylin and eosin, and when the lamina propria is made fluorescent, all layers of the lamina propria show much elastin.

Figure 2 shows such a vocal fold **LP** being made fluorescent. The yellow color is all elastin material, and note that the immediate **SLLP** is brightly shining showing abundant elastin is present in the **SLLP**. This figure shows that the entire **LP** contains a lot of elastin. However, when the Verhoeff's elastin stain is used to identify only the elastin fibers, few elastin fibers are found in the **SLLP** indicating the elastin in the **SLLP** is not in elastin fiber form. It was the elastin stain that was used to originally identify the three layers of the **LP**. Figure 3 shows the **LP** stained for elastin fibers.

A dotted oval is placed around the elastin fibers located in the **MLLP** and **DLLP.**

Elastin Fiber Distribution in the Lamina Propria

Using elastin staining and an image analysis system, the concentration of elastin fibers in the **SLLP** to the vocalis muscle in 20 middle-aged, cadaveric, adults (age 25-50) vocal folds can be assessed. All vocal folds were obtained within 24 hours of death. Vocal folds were removed and cut vertically, perpendicular to the free edge of the vocal cord. A facing midsection was fixed in Carson's Fixative and embedded in paraffin using routine histologic methods.[17]

Four micron-thick sections of normal vocal fold were stained to detect their elastic tissue content using Verhoeff's elastin tissue stain.[18] This stain uses ferric chloride and iodine to detect elastin and acid fuchsin as a counterstain. The slides were analyzed using an image analysis system specially configured for histological and cytological studies. The image analysis system and methodology has been described previous studies.[9,20,18]

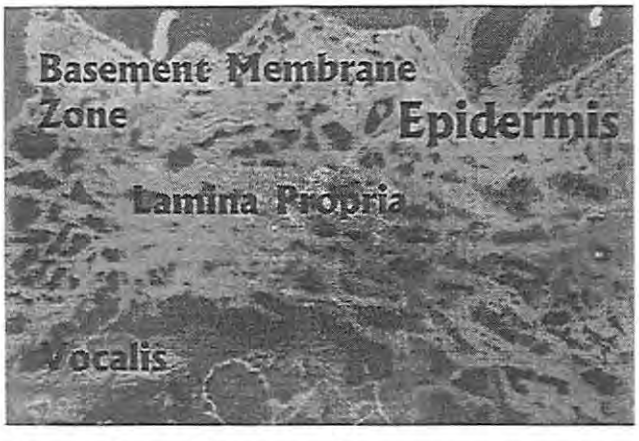


Figure 2. This vocal fold tissue has been stained with hematoxylin and eosin and then made fluorescent. The elastin molecule becomes fluorescent. The picture demonstrates the tissue between the basement membrane zone and the vocalis muscle. This is done on fresh, frozen tissue, which accounts for the gaps or dark spaces in between the fluorescent tissue. The dark spaces are therefore artifactual and due to the technique. This picture, which indicates that elastin protein is present in considerable abundance throughout all layers of the lamina propria, is in contrast to Figure 3, showing that the elastin fiber component is more predominant in the middle and deep layers.

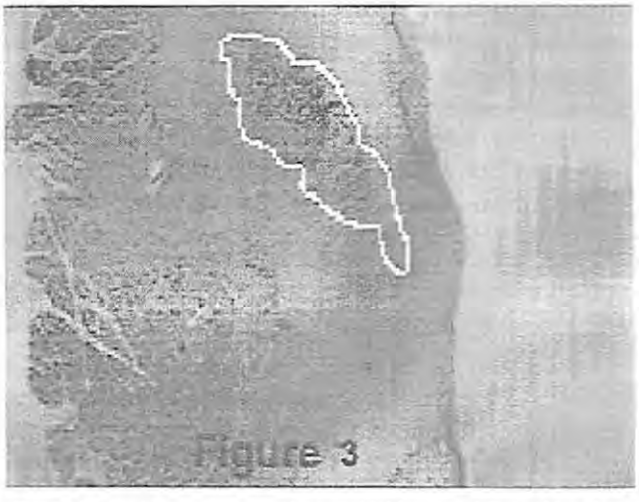


Figure 3. This figure demonstrates vocal fold material stained for elastin fibers. There is a preponderance of elastin fibers in the intermediate and deep layer. This picture uses an elastin Verhoeff's stain. Based on figure 2 and 3, one can see why the entire lamina propria has elastin qualities whereas the vocal ligament portion has biomechanical properties for stress and strain as described later in the text.

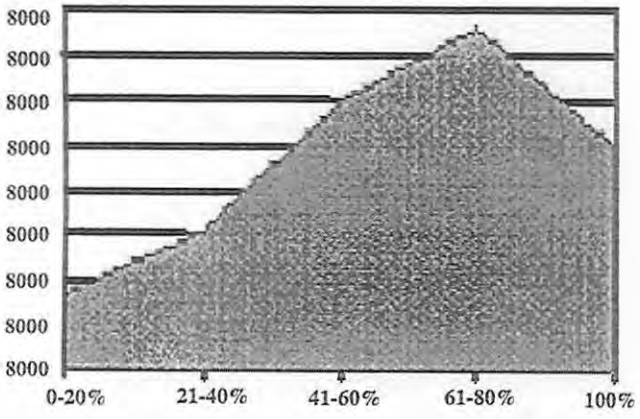


Figure 4. This figure demonstrates the elastin amount in relative units plotted by depth from the basement membrane zone to the vocalis muscle. 0% percent is the basement membrane zone whereas 100% represents the start of the vocalis muscle. Elastin fiber concentration peaks at around 60-80% depth in the lamina propria. This slide represents a composite of 20 middle-aged adults (Hammond, et al., in press). [Reprint courtesy: Mosby Year Book, Inc. Otolaryngology Head and Neck Surgery Journal.]

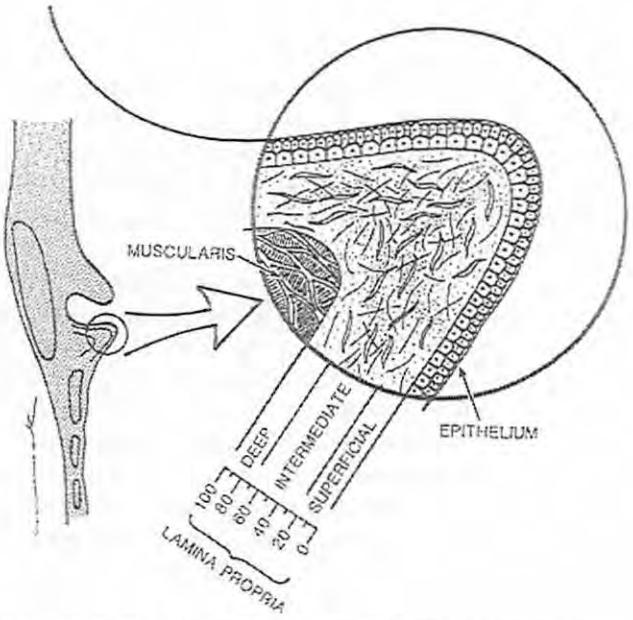


Figure 5. This is a schematic of the estimated width of the lamina propria layers based upon elastin fiber concentration.

A summary graph of the elastin fiber concentration by depth of the **LP** is presented (Figure 4). This study demonstrates the magnitude of the difference in concentration of elastin fibers between the **SLLP** and the vocal ligament. It also shows the depth the **SLLP** extends into the lamina propria. This may be important in guiding us towards issues of replacement therapy of the **LP**. It is felt by some that an elastin mixture would be the best biomechanical **LP** substitute. As you can see from the graph, the slope of elastin fiber concentration starts increasing at about 25% depth and begins to level off at about 45% depth into the **LP**. It continues to stay high until the last 15-20%. From these data it appears that the **SLLP** consists of about 25-35% of the initial depth of the **LP**, with the **MLLP** making up the next 45-55%. (Figure 5) Although our initial tendency was to call the last 20% the **DLLP**, that is not entirely accurate since the study did not assess collagen fiber density. No gender difference in the amount of elastin fibers across the **LP** was identified.[16]

Elastin Fiber Across Age Groups

Using similar methodology, elastin fiber concentration in infants, middle age, and geriatric specimens has been studied.[9] (Figure 6) The main objective of this earlier study was to identify the concentration of elastin fibers between ages using quantitative methodology. The numbers were small, but no gender differences were found. Elastin concentration was much higher in the geriatric samples and lowest in the infants. It was noted that the

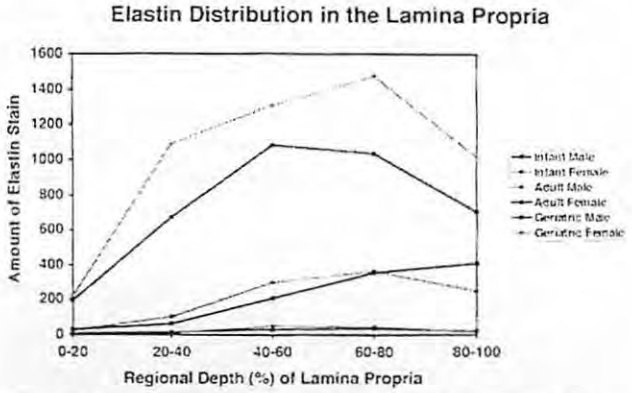


Figure 6. This graph shows increasing elastin fiber concentration by age. Note that even in the superficial layer, the elastin content is quite a bit higher in geriatric folds than what younger adults experience.

increase in the elastin fiber concentration occurs more superficially in the **SLLP** as we get older, and the thickness of the **SLLP**, measured by this histologic method, becomes considerably thinner in the geriatric patient. Three representative photos are shown to illustrate these two points: the thinning of the **SLLP** and the increasing elastin fiber concentration. These figures received a Verhoeff elastin stain that stains elastin fibers black. The elastin fibers appear as tiny black dots in the **LP**. Figure 7A is from an infant larynx; note that minimal elastin is present. Figure 7B is a middle-aged male. Note that the **MLLP** shows moderate staining of black elastin fibers. Figure 7C is a geriatric male larynx and the abundant black staining is apparent almost right up to the epithelium.

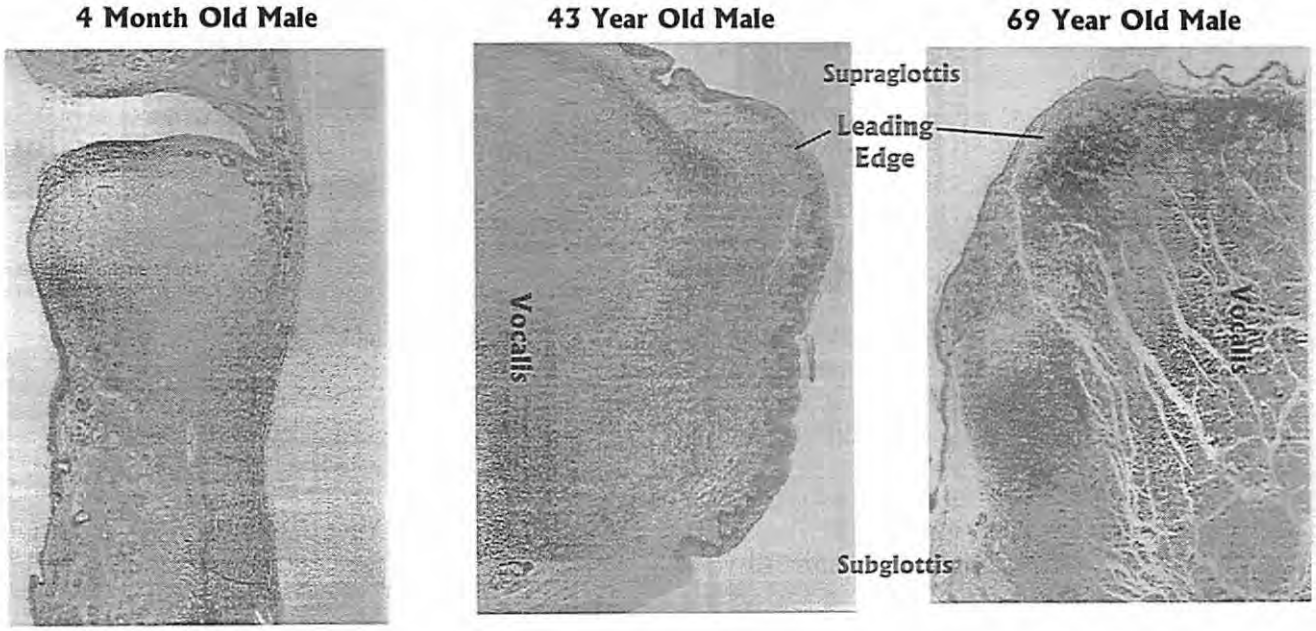


Figure 7. Using a Verhoeff's elastin stain that stains elastin fibers black, 3 ages are examined. 7A is a four month old male, 7B is a 43 year old male, and 7C is a 69 year old male. Note the continued shrinkage of the superficial layer and the more predominant the black color becomes in the intermediate layer. Also note that the fullness or total width of the intermediate layer is not only large at the leading edge, but remains of substantial width slightly below the leading edge in the immediate infrafold portion.

If one believes that limiting surgery to the **SLLP** avoids scarring, then this study would indicate that the patients at risk for scarring are in the geriatric population. Clinically, this concept may not bear out, as it seems from clinical experience that geriatric patients are not more predisposed to iatrogenic laryngeal injury. It may be that it is not the collagens and elastins of the **ECM** that determine limits of surgical safety, but rather the presence or amount of decorin or hyaluronic acid in the **SLLP** that leads to the ability to operate on the **SLLP** without inducing scar. Decorin and hyaluronic acid are discussed in part two of this report. If concentration in decorin or hyaluronic acid in the **SLLP** is low then even minor surgical intervention could potentially lead to scar formation in the **SLLP**, despite great surgical skill. However, this thinning of the SLLP may have relevance to formation of a sulcus or contribute to some cases of bowed vocal folds.

Collagen

Collagen Type I, II, and III have all been identified in the **LP**.[19] Collagen Type IV and VII are located in the basement membrane zone.[20,21,22,23] Collagen studies are somewhat more difficult to do because within 24 hours of death, collagens lose their antigenicity, making collagen specific studies unreliable even when done within a few hours postmortem. We don't have quantitative studies of how the collagens are distributed across the **LP**. The biologic function of collagens is fairly singular: to provide strength. Collagens don't stretch well. They do hold things together well. Their general spatial orientation is directed longitudinally, from anterior to posterior, which is similar to the elastin fibers. This orientation allows them to bear the stress of intrinsic laryngeal muscles.

Fibrous Proteins and Senescence

The finding of increased elastin in the **LP** with aging is consistent with other studies of connective tissue. As we age, we tend to accumulate both collagen and elastin fibers in our extracellular matrix because the body breaks down and turns over less of these fibrous proteins. Kohn repeatedly pointed out that age-related changes in most organ systems are more expressed in the extracellular matrix than in the actual cells.[24] This is demonstrated by tendons becoming more resistant to heat denaturation and less soluble and digestible by collagenase as we age.[25] The elastic modulus of the lung becomes stiffer and decreases linearly with age.[26] Myocardium becomes stiffer and arteries become less elastic with age.[27,28] These are functions of the components of the extracellular matrix. More elastin or collagen does not necessarily reflect better elasticity or function of the tissue, since most of the elastin and collagen in the aged tissue is less functional due to cross-linking.[29]

Vocal Fold Lamina Propria and Senescence

These findings are not in complete agreement with other **LP** studies of elastin. Sato showed cross-linking of

elastin fibers in the geriatric specimens, but reported that the elastin fibers seem to become less numerous than in younger adults.[30] Hirano's earlier studies also reflected a decrease in elastin fibers between normal and geriatric vocal folds.[31] We do not know how to explain these differences with Hammond's findings.[9] Similar confusion has been found in other literature where elastin content has been reported to increase, decrease or remain the same through senescence. Although both groups employed elastin stains, the methodology used was quite different. Hammond et al. used a quantitative image analysis system and fewer specimens, while the other two reports used more specimens and qualitative visual assessment. Both methods have some advantages and disadvantages. Perhaps racial differences could be present.

These studies help define how the microarchitecture of the vocal fold change as we age. Clinically, the knowledge of increased or decreased elastin fibers in the **LP** is not very useful without knowledge of how that finding may effect vocal performance, which at this time is lacking. The finding of the shift to the left for the steep portion of the elastin curve in geriatric specimens has several possible implications. Is the **SLLP** shrinking (atrophy) or is it simply being invaded (conversion) by elastin? Whatever the mechanism, the result is a proportional increase in the total width of the **MLLP** and decrease in the width of the **SLLP**. This decreases the amount of cover in relation to the body of the vocal fold as we age. In the infant, this scenario is reversed, with a very large cover to body ratio.

Although we do not have a knowledge of what these vocal folds looked like in vivo, occasional specimens have been found that appeared to have a sulcus vergeture, and in these specimens it seems as though the **SLLP** is lacking, with the **MLLP** and **DLLP** being relatively normal. The **MLLP**'s superficial border seems to be nearly the basement membrane zone in severe sulcus cases. This could explain the profound effect a sulcus may have on the voice. By loss of the **SLLP**, they have lost most of their functional cover of tissue. Instead, their epithelium really is attached directly to a transitional layer and body of the vocal fold. Anecdotally, we have not seen any cases where atrophy of the **MLLP** seems to occur. From our observations, it also appears that some conversion of **SLLP** to **MLLP** is occurring. This concept would be supported by general theories of aging in which it has been reported that as we age the elastin content of the **ECM** in some tissues slowly increases.[25,35] It is likely that aging may produce both atrophy of the **SLLP** and some conversion (infiltration) of the **SLLP** by **MLLP** proteins. We need to reemphasize that we do not know how these folds appeared and performed when in vivo, and these are strictly histologic observations.

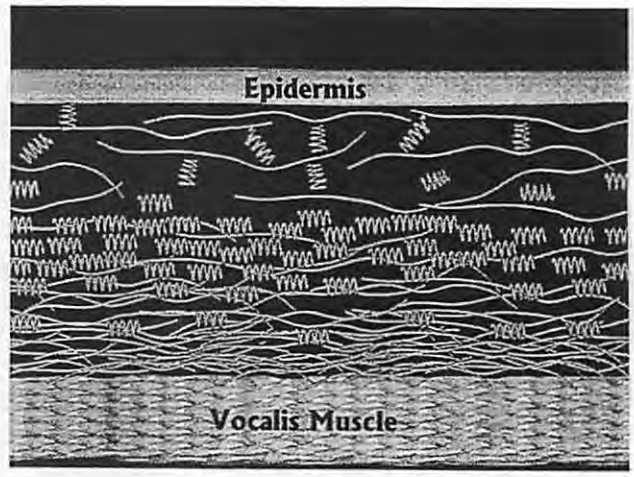


Figure 8. This is a computer rendering performed on Animation Master, a software package of 3D modeling and rendering produced by HASH, Inc. The rope like structures are to represent collagen fibers and the spring like structures are to represent elastin fibers. The picture demonstrates the fibrous proteins only of the vocal fold. Please compare this with other computer renderings done similarly in part 2 of this report showing other extracellular matrix molecules.

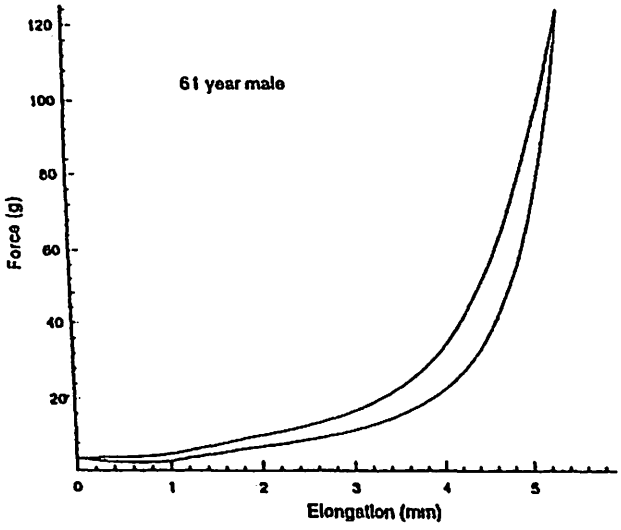
Building the Lamina Propria (Fibrous Proteins)

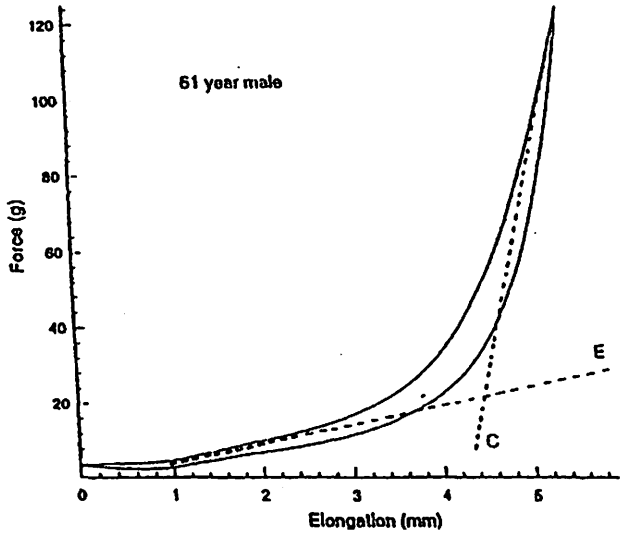
The extracellular matrix is conceptually built as shown in Figure 8. Thus far we have discussed only the fibrous proteins of the **ECM**, the collagens and the elastins. We have reviewed the distribution of elastin fibers through the lamina propria and as it relates to age. Note that there are more collagen fibers deeper in the fold. Note also that the overall arrangement of the fibrous proteins is parallel to the edge, although individual fiber molecules may not be so directed. Recall that elastin material is present throughout all layers although the elastin fiber form is most concentrated in the **MLLP**.

Biomechanics of the Fibrous Proteins

Biomechanically, stress-strain curves are frequently used to describe properties of fibrous materials.[32] **Stress** can be defined as the amount of force per unit area applied to a substance to induce a physical, morphologic change. **Strain** is defined as corresponding deformation (the change in length - original length). Studies on the vocal ligament have measured how much stress leads to a certain level of strain.[33,34] Figure 9 (stress-strain) illustrates this relationship. These stress-strain curves were obtained according to standard methodology.[34]

Figure 9A is a stress-strain curve of a 60 year old man. Figure 9B is the same curve but with lines drawn in to show linear approximations to the collagen and elastin fiber responses. The last stress-strain curve is from a younger 31 year old man for comparison. (See Figure 9B) Note that the two curves are not exactly similar. The stress-strain curve of





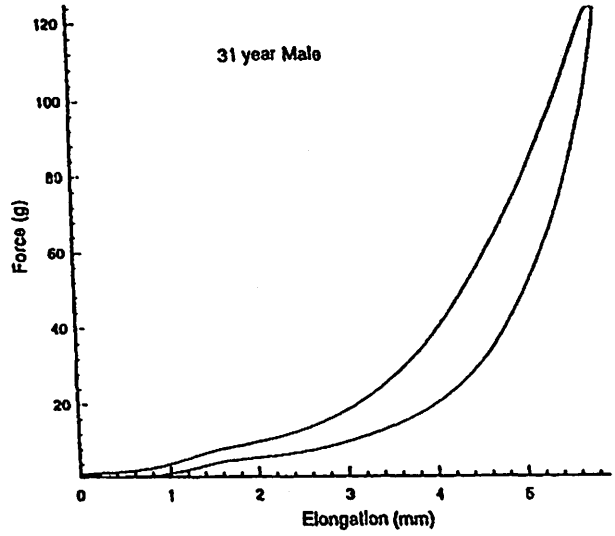


Figure 9. These are stress-strain curves of human vocal ligaments. 9A (upper left) is a 61 year old male. 9B (left) is the same 61 year old male with functions of elastin and collagen drawn in. 9C (above) is a 31 year old male for comparison with the 61 year old male.

the 61 year old man shows a steeper slope in the collage portion (increased stiffness) and less area between the lines indicating less tissue creep (increased viscosity) as compared to the 31 year old male.

By applying stress longitudinally to the fibers of the vocal ligament, the ligament stretches to a longer length. Note from the stress-strain curves that the vocal ligament stretches easily at low level of stress, but upon reaching a longer length the stress required for further lengthening increases remarkably. This type of nonlinear response is attributed to the mix of collagen and elastin fibers.[35] The resistance to stretch due to elastin alone, making the first part of the stress-strain curve fairly linear (spring-like) characterizes the initial part of the stress-strain curve in the human vocal ligament. The early part of the stress-strain curve mirrors elastin fiber performance. After a certain amount of strain has occurred (about 30%), the stress-strain curve takes a noticeable upward slope. In stress-strain relationships of ligaments and tendons, the point at which the upward slope occurs is often referred to as the *"break-point"*, because it represents the place of ligament strain where collagen fibers are being recruited to resist further strain.[36] Once the break-point area is reached, further stress does not result in much more length change within the vocal ligament. The late part of the stress-strain curve is similar to collagen fiber performance. Note that in the 61 year old male, the part of the curve corresponding to collagen fiber performance is more steep than in the 31 year old. This likely indicates increased stiffness in the collagen fibers of the older man, a common finding as we age. Also note the area within the 31 year old curve is greater than that of the 61 year old. The area within the curve is an indirect indication of viscosity. The 31 year old vocal ligament likely has lower viscosity and more tissue creep than that of the 61 year old male.

The steeper part of the stress-strain curve is designed to resist strain. The particular composition of the vocal ligament varies individually and with age. With respect to vocal performance, it is interesting to reflect on how the response of elastin is meshed with the response of collagen to make the stress-strain curve of the vocal ligament. The interaction between elastin fiber performance and the collagen fiber performance may be influenced by proteoglycans.

Dynamic Range Properties/Fibrous Proteins

For speaking pitches, the more functional part of the stress-strain curve, the part of the curve that most of us vocally use daily is dominated by elastin fiber performance.

For singing at high pitches, however, the contribution of collagen to the stress-strain curve may be significant. To understand this, we appeal to the formula for the frequency of a vibrating string:

$$F = \frac{1}{2L}\frac{\sqrt{\sigma}}{\rho}$$

where σ is stress, ρ is tissue density and L is the length of the string.[32] The authors recognize that the vocal folds cannot be adequately described as vibrating strings. But the mathematical relationships described by this equation do have some merit in helping us understand biomechanical relationships in phonation. As can be seen from the equation, length has an inverse relationship to frequency. That is, although the vocal folds get longer as our pitch goes up, the longer vocal folds should vibrate at a lower frequency. To raise the frequency, we have to overcome the increase in vocal fold length by increased stress. Tissue density is nearly a constant (1.0-1.05 g/cm^3). By increasing tissue stress by a factor of four, we double the frequency (note the square root relation). But this rise is diminished by the length increase, giving us less than a doubling of frequency for a four-fold increase in stress. Now it is apparent why the stress-strain curve is so relevant to vocal production. It is critical that we easily generate high levels of tissue stress without much increase in length. This ability to easily generate high levels of tissue stress is a product of the collagenous part of the vocal ligament and is a main mechanism of providing humans with a large dynamic pitch range.

Summary

In summary, the proteins of the extracellular matrix of the vocal folds are classified into fibrous and interstitial vocal folds. The collagens and elastins comprise the fibrous proteins and influence stress-strain properties of the vocal folds. Collagens are proteins that bear stress and contribute strength to the tissue, whereas elastins are proteins that allow the tissue to deform and return to its original shape. These are key properties for proper function of the vocal folds. Tissue stress, the ability to carefully control it and precisely generate it, is important for dynamic pitch range in humans.

Acknowledgments

This work was supported by grant P60 DC00976-09 to the National Center for Voice and Speech from the National Institute on Deafness and Other Communicative Disorders (NIDCD).

References

1. Hirano M: Structure of the vocal fold in normal and disease states. Anatomical and Physical Study. ASHA Report. 1981; 11: 11-30.

2. Hirano M, Kakita Y: Cover-body theory of vocal fold vibration. In: Daniloff RG, ed. Speech Science. San Diego, CA: College-Hill Press; 1985:1-46.

3. Hirano M: Phonosurgery: basic and clinical investigations. Otologia Fukuoka. 1975;21:239-442.

4. Saito S: Phonosurgery: basic study on the mechanism of phonation and endolaryngeal microsurgery. Otologia Fukuoka. 1977;23:171-384.

5. Hirano M, Yoshida T, Tanaka S: Vibratory behavior of human vocal folds viewed from below. In: Gauffin J, Hammarberg B, ed. Vocal fold physiology. San Diego, CA: Singular Publishing Group; 1991:1-6.

6. Matsushita H: Vocal fold vibration of excised larynges. A study with ultra-high speech cinematography. Otologia Fukuoka. 1969;15:127-142.

7. Matsushita H: The vibratory mode of the vocal folds in the excised larynx. Folia Phoniatr. 1975;27:7-18.

8. Baer T: Investigation of phonation using excised larynx.. Cambridge, MA: Institute of Technology; 1975. Ph.D. Thesis.

9. Hammond TH, Gray SD, Butler J, Zhou R, Hammond E: A study of age and gender related elastin distribution changes in human vocal folds. Otolaryngology HNS J. In press.

10. Titze IR: The physics of small-amplitude oscillation of the vocal folds. J Acoust Soc Am. 1988;83:1536-1552.

11. Hirano M: Structure of the vocal fold in normal and disease states: anatomical and physical study. ASHA Report. 1981; 11:11-30.

12. Pawlak A, Hammond T, Hammond E, Gray S: Immunocytochemical study of proteoglycans in vocal folds. Ann Otol Rhinol Laryngol. 1996; 105:6-11.

13. Porto LC, Chevalier M, Peyrol S, Guerret S, Grimaud JA: Elastin in human, baboon, and mouse liver: an immunohistochemical and immunoelectron microscopic study. Anat Rec. 1990;228:392-404.

14. Ferreira JMC, Caldini EG, Montes GS: Distribution of elastic system fibers in the peripheral nerves of mammals. Acta Anat (Basal). 1987; 130:168-73.

15. Caldini EG, Caldini N, De-Pasquale V, Strocchi R, Guizzardi S, Ruggeri A, Montes GS: Distribution of elastic system fibres in the rat tail tendon and its associated sheaths. Acta Anat (Basel). 1990;139:341-8.

16. Hammond TH, Zhou R, Hammond EH, Pawlak A, Gray SD: The intermediate layer: A morphologic study of the elastin and hyaluronic acid constituents of normal human vocal folds. J of Voice. 1997;11:59-66.

17. Carson FL, Martin JH, Lynn JA: Formalin fixation for electron microscopy: a reevaluation. Am J Clin Pathol. 1973;59:365-7.

18. Carson FL: Histotechnology: a self-instructional text. ASCP. 1990; 140-1.

19. Gray SD, Hirano M, Sato K: Mollecular and cellular structure of vocal fold tissue. Vocal Fold Physiology: Frontiers of Basic Science, edn 1. Titze IR, ed. San Diego: Singular Publishing; 1993:1-34.

20. Mossallam I, Kotby M, Ghaly A, Nassar A, Barakah M: Histopathological aspects of benign vocal fold lesions associated with dysphonia. Vocal Fold Histopathology: A Symposium, edn 1. Edited by Kierchner J. San Diego: College Hill Press; 1986:65-80.

21. Courety M, Shohet J, Scott M, Ossoff R: Immunohistochemical characterization of benign laryngeal lesions. Ann Otol Rhinol Laryng. 1996; 105:525-53 1.

22. Gray SD:, Pignatari SSN, Harding P: Morphologic Ultrastructure of anchoring fibers in normal vocal fold basement membrane zone. J Voice. 1994;8(l):48-52.

23. Gray SD: Basement membrane zone injury in vocal nodules. In: Gauffin J, Hammarberg B, eds. Vocal Fold Physiology Conference 21-28. San Diego, CA: Singular Publishing Group Inc.

24. Kohn R: Evidence against cellular aging theories. In: Testing the Theories of Aging. Adelman RC,Roth GS, eds. Boca Raton, FL: CRC Press 1982, 221.

25. Schneider SL and Kohn RR: Effects of age and diabetes mellitus on the solubility of collagen from human skin, tracheal cartilage and dura mater. Exp. Gerontol. 1982;17:185.

26. Niewoehner DE, Kleinerman J, Liotta L: Elastic behavior of postmortem human lungs: effects of aging and mild emphysema. J Appl Physiol. 1975;39:943.

27. Borg TK, Ranson WSF, Moslehy FA, Caulfield JB: Structural basis of ventricular stiffness. Lab Invest. 1981; 44:49.

28. Pillsbury HC, Hung W, Kyle MC, Freis ED: Arterial pulse waves velocity and systolic time intervals in diabetic children. Am Heart J. 1974;87:783.

29. Monniere VM, Sell DR: Collagen as a biomarker of aging. In: Practical Handbook of Human Biological Age Determination. Balin AK (eds.), CRC pub. 1994.

30. Sato K, Hirano M: Age related changes of elastic fibers in the superficial layer of the lamina propria of vocal folds. Ann Otol Rhinol Laryngol. 1997;106:44-48.

31. Hirano M, Kurita S, Nakashima T: Growth, development and aging of human vocal fold. In: Bless DM, Abbs JH, eds. Vocal Fold Physiology 22-43, San Diego, Ca: College Hill Press.

32. Titze IR: Principles of Voice Production. Englewood Cliffs, NJ: Prentice Hall; 1994:31-44.

33. Hirano M: Structure and vibratory behavior of the vocal folds. In: Sawashima M. Franklin SC, eds. Dynamic aspects of speech production. Tokyo: University of Tokyo Press, 1977:13-30.

34. Perlman AL, Titze IR, Cooper DS: Elasticity of canine vocal fold tissue. J Speech and Hear Res. 1984;27:212-219.

35. Cox RH: Passive mechanics and connective tissue composition of canine arteries. Am J Physiol. 1978;234:H533-H531.

36. Armentano RL, Levenson J, Barra JG, Fisher EIC, Breitbart GJ, Pichel RH, Simon A: Assessment of elastin and collagen contribution to aortic elasticity in conscious dogs. Am J Physio. 1991;260(6pt2):1870-77.

NCVS Status and Progress Report - 13
April 1999, 11-20

Vocal Fold Extracellular Matrix and Its Biomechanical Influence Part II: The Interstitial Proteins

Steven D. Gray, M.D.
Department of Surgery – Division of Otolaryngology Head and Neck Surgery, The University of Utah School of Medicine
Ingo R. Titze, Ph.D.
Fariborz Alipour, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa
Thomas H. Hammond, B.S.
The University of Utah School of Medicine

Abstract

This is the second part of a two part series to describe the extracellular matrix of the human vocal fold and associated biomechanical concepts. The key tissue element of the vibratory portion of the vocal folds is the extracellular matrix. Although the fibrous components of that matrix have been described for two decades, the remaining portions of the extracellular matrix are not well understood. This article discusses the interstitial proteins of the lamina propria, their location in the vocal fold, and their biologic and biomechanical effects. Just as part 1 discussed stress-strain relationships and dynamic vocal range with respect to the fibrous proteins, this paper discusses viscosity, and its effect on "phonatory threshold pressure" and "energy expended due to phonation" as it relates to the interstitial proteins. The role of hyaluronic acid in determining tissue viscosity is presented. The small chain proteoglycans, decorin and fibromodulin are introduced. Computer generated schematics of the extracellular matrix are presented.

Introduction

This is the second part of a discussion on the extracellular matrix proteins of the lamina propria of the vocal folds and the biomechanical properties. The first part of the discussion dealt with the fibrous proteins: the collagens and elastins. The fibrous proteins have received a lot of attention. It is interesting, and not unexpected, that some of the interstitial proteins also have a layered orientation in the vocal folds. Their mechanical properties determine many of the oscillatory characteristics of the vocal folds. They likely affect *viscosity, fluid content, thickness of the lamina propria layers, and even collagen fiber population density and size* (See Table of Proteoglycans following page).The interstitial proteins include the proteoglycans and the glycoproteins.

Proteoglycans

The proteoglycans consist of a protein core that uses a linkage protein to attach to the core other proteins, carbohydrates and lipids.[10] This allows the proteoglycans to carry a variety of structural molecules with various biologic activities, including water, or simple linear sugars, to the most highly sulfated polysaccharide found in nature: heparin.[11,10,12,13] In this way the extracellular matrix (**ECM**) can regulate concentrations of some carbohydrate, protein and lipid molecules. These attached chains of molecules to the protein core are termed glycosaminoglycan chains and determine a wide range of biologic activity.[10] For instance, a component of the chain in hyaluronic acid creates a

Table 1.
Proteoglycans Found in Human Vocal Folds[17]

Proteoglycans	Function	Localization in Vocal Folds
Hyaluronic Acid	Creates and is the means for control of tissue viscosity, effects tissue flow resistance and tissue osmosis, provides space filling and space occupying molecules, probably help determines lamina propria layer thickness[41,1,2]	Found throughout the ECM of the lamina propria, slightly more intense in the intermediate layer of the LP, found in macrophages and fibroblasts of the LP, evidence suggests gender specificity. Males> females
Decorin	Binds to collagen fibers resulting in delayed fibril formation and thinner fibril formation[3,4], may help reduce fibrosis and scar following injury	Found in the ECM of the lamina propria, may be more concentrated in the SLLP
Fibromodulin	Binds to collagen fibers resulting in delayed and thinner collagen fibers [41,42], both decorin and fibromodulin may effect ligament performance	Found in the ECM of the lamina propria, found mainly in the intermediate and deep layers of the LP, seems to be concentrated around the vocal ligament[5,6]
Versican	Has ability like hyaluronic acid to fill space, bind and organize water molecules[7]	Found being manufactured by the fibroblasts and macrophages in the LP
Heparan Sulfate Proteoglycan	Binds to fibronectin, collagen IV, and laminin[8,9], may play a role in tissue morphogenesis [46,47]	Found in the basement membrane zone of the vocal folds, found in the fibroblasts and macrophages of the LP

Aggregan was not found in vocal folds
Biglycan has not been searched for in the vocal folds

ECM-extracellular matrix, LP-lamina propria, SLLP-superficial layer of the lamina propria

polarity which seems to be particularly suited for binding of water molecules, thereby increasing water and ion content of the **ECM** as well as increasing the total volume or space the molecule occupies.[10,41] Another example is the glycosaminoglycan chain in decorin and fibromodulin which has a strong affinity for binding to collagen fibers, thus helping to regulate collagen fiber thickness and morphology.[41,42]

Proteoglycans have reached remarkable levels of sophistication and multifunctional roles. They provide growth supportive or suppressive function, modulate wound repair, bind and deliver growth factors, act as major biologic filters, surround and likely influence **ECM** communication with the fibroblast, and have major adhesive properties.[41,10,3,13] Importantly, proteoglycans seem to have the tools to help or regulate the biologic performance of other **ECM** proteins.

Proteoglycans do not occur often in isolation. They are usually in a mixture of other **ECM** molecules and their concentrations effect **ECM** molecules in different ways.[14,13] Proteoglycans are divided into groups based on similarities between the protein cores and their glycosaminoglycan chains and consequently their biologic activities.[10,13] One group consists of large proteoglycans, which tend to aggregate: these are aggregan and versican. Hyaluronic acid differs from these proteoglycans in that it is not covalently attached to proteins and therefore is usually considered in a unique, separate group of proteoglycans.[13] But because it is large, aggregating, and a major determinant in tissue viscosity *(like aggregan and versican)* we will include hyaluronic acid in this group of large chain proteoglycans. Aggregan is a major constituent of cartilage.[15,16]

The large proteoglycans affect tissue viscosity and are space occupying. These large proteoglycan molecules are self-repeating and can be huge, thereby filling up lots of space in a gel-like arrangement. As mentioned, they have an affinity toward water molecules and when possible create a strongly hydrated gel. Because of these properties, their concentration levels can have profound effects on tissue flow resistance and tissue osmosis.[17,11,10,13,41,42] Large chain proteoglycans are found wherever the body is concerned about functions such as shock absorption *(cartilage-aggregan)*, viscosity and fluid movement *(synovial fluid-hyaluronic acid)*, and space filling *(vitreous humor of the eye- hyaluronic acid)*. Since the same biologic needs are relevant in the vocal folds, large chain proteoglycans likely play an important role.

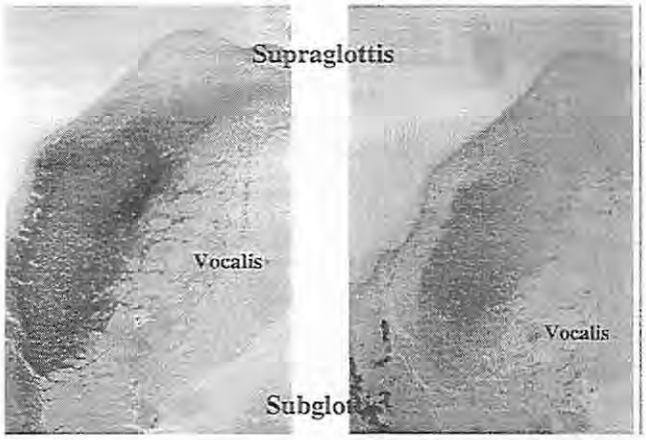
Small chain proteoglycans are another group of proteoglycans and are characterized by a leucine-rich protein core and a small, compact molecular size.[11,10] These small proteoglycans are many but we are mostly concerned with decorin, fibromodulin and to a less extent, biglycan. All these small proteoglycans bind to collagen Type I, although not necessarily at the same site. Studies have shown that collagen synthesis in the presence of decorin and fibromodulin yields smaller and thinner fibrils.[41,42] Concentration levels of these proteoglycans have been found to be higher out in the **ECM** than closer to the regulating cells for the matrix and it has been suggested that these regulating cells *(chondrocytes for cartilage and fibroblasts for the vocal folds)* use these small proteoglycans as factors to regulate collagen assembly. An important corollary exists that a decrease in tissue concentration of fibromodulin or decorin results in more collagen synthesis with bigger fibers. Biglycan will only be mentioned briefly here, because its role in normal vocal folds is likely minimal. Biglycan also regulates collagen synthesis but unlike decorin seems to promote ongoing fibrosis.

Hyaluronic Acid

Hyaluronic acid (**HA**) can have marked effects on tissue viscosity, tissue flow, tissue osmosis, tissue dampening *(shock absorption)* and space filling.[41] Part of the mechanism that allows this to happen is the configuration or shape of the **HA** molecule. It is estimated that an **HA** molecule takes a 1,000 times in space or volume beyond what would be expected based on molecular composition and weight.[41] Unlike the fibrous proteins of the vocal folds (collagen and elastin), which fibers can be arranged linearly and stacked compactly into tendinous or ligamentous structures, **HA** likes to be porous, loosely folded and globular.[41,18] Remembering that these molecules have the propensity and ability to aggregate, it is apparent why these molecules can influence tissue size by essentially filling up a tissue space. Similarly, this ability to fill up space consisting of loosely arranged molecules makes this molecule appropriate for shock absorption. Their large, voluminous size, but porous quality combined with a high water content makes this proteoglycan biomechanically ideal for enduring the pounding and vibratory collision to which vocal folds are subject.

Identification of Hyaluronic Acid

Hyaluronic acid (**HA**) in vocal folds can be studied through two common methods. The most common method is using an acid mucopolysaccharide stain (**AMP**) with and without enzymatic treatment by testicular hyaluronidase.[19] To perform this stain, two consecutive four-micron-thick sections of human vocal folds are taken *(in our work this is usually done with a vertical orientation through the mid-portion of the membranous vocal fold)*. One section is treated first with the enzyme testicular hyaluronidase, which removes the **HA** from the tissue specimen. Controls can be any tissue that contains **HA** such as cartilage, skin, and liver. Following removal of the **HA** with hyaluronidase, that tissue section is stained at the same time and solution with



(A) With Hyaluronic Acid

(B) Without Hyaluronic Acid (Removed by hyaluronic enzymetrically)

Figure 1. This figure shows consecutive slices through the mid-membranous portion of a human vocal fold stained with acid mucopolysaccharide stain. Part A shows the stain with the hyaluronic acid present and Part B shows the tissue with the hyaluronic acid removed enzymatically. The difference between the blue intensity in A and B is the hyaluronic acid present in Part A but removed in Part B. Note that in Part B that there was still a greenish-blue color, which represents other proteoglycans or carbohydrates present in the lamina propria of the vocal folds.

the other section (the one not enzymatically exposed) for acid mucopolysaccharides (proteoglycans and mucopolysaccharides). Acid mucopolysaccharides are usually stained a dark blue, sometimes green tinged color. The difference between the blue intensity of the two stains should be the **HA** present in the tissue section not enzymatically removed as compared to the lack of the blue color present in the enzymatically treated one.

The remaining faint blue-green color in the tissue section which had the **HA** removed reminds us that other acid mucopolysaccharides *(some of which are probably other proteoglycans but can include carbohydrates)* are present in normal human vocal folds. An example of this stain and the amount of **HA** present in human vocal folds is shown in Figure 1. Notice that there is a fair amount of intense blue staining in the **middle layer** of the **LP (MLLP)**; the **MLLP** is the layer that is probably the thickest of the three **LP** layers. Of interest, note that the **HA** seems to be slightly more intense in the immediate infrafold area as opposed to the concentration at the exact leading edge of the vocal fold. Although not quantified yet, it has been our observation of over 100 histologically stained larynges that many of them (though not all) have this propensity for a slight increase in **HA** amount in the immediate infrafold area. This area corresponds to the region where the mucosal wave begins in its vertical travel upward.

A second method to study **HA** is performed by using immunocytochemical techniques which utilized antibodies to certain **HA** receptors. Pawlak used this technique

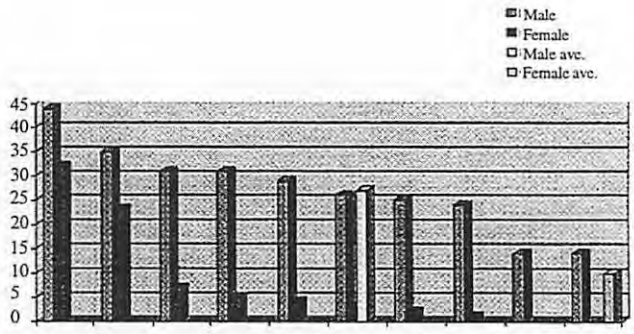


Figure 2. This shows a graph of hyaluronic acid by gender from 10 males and 8 females. Concentration of hyaluronic acid is in relative value units. Note that even though the average male has about three times of hyaluronic acid than the average female, there are some female vocal folds with greater hyaluronic acid concentration than some male vocal folds.

in her identification of **HA** in human vocal folds.[17] The advantage of this latter technique is that it may be much more specific. This approach does not directly measure **HA**. The disadvantage is that some immunocytochemical techniques are more difficult to perform and sometimes much more difficult to quantitate with image analysis systems. On the other hand, the **AMP** stain described before is much easier to use quantitatively for image analysis systems. Unfortunately the **AMP** stain approach does not give information about specific types of **HA**.

Hyaluronic Acid and Gender

By using the **AMP** staining method to identify **HA**, Hammond quantified the amount of **HA** in males and females.[20] He analyzed the mid-section of 18 (10 males, 8 females, 20 to 60 year-old) membranous folds and found that despite moderate individual variation in the **HA** content per area, males have about 3:1 more **HA** than females. Using his data we have generated the following graph showing the gender difference in **HA**. (See Figure 2)

Since a quantitative gender difference has not been found with the fibrous proteins, this gender difference in **HA** is provocative.[21] This may offer insights as to male: female vocal differences in normal voice production and pathologies when one considers the biomechanical functions of **HA** and the consequences of having more **HA** in the vocal fold. Notice from the graph that even though males tend to have more **HA,** there is overlap among the gender such that some females have considerable **HA** and in some cases more than males. In Part I: The Fibrous Proteins, of this **ECM** discussion we presented a schematic of the fibrous proteins, now we add **HA**. Because of molecular configuration and biologic effect we have chosen a large voluminous shape for **HA** molecules. Additionally, please note that the addition of the molecules has expanded the **ECM** and filled up more space. (see Figure 3A and 3B; following page)

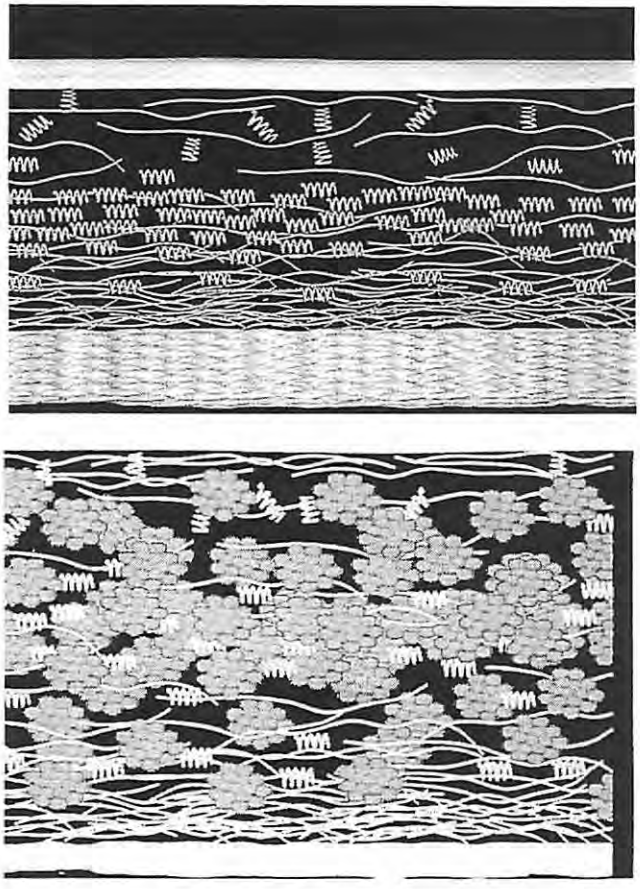


Figure 3: This figure represents a computer generated schematic using Animation Master by HASH, Inc. which is useful in generating 3-D animation. Figure 3A (top) represents the extracellular matrix with fibrous proteins only. As in Part 1, the rope like structures represent collagen fibers and the spring like structures represent elastin fibers. The large blue molecules represent hyaluronic acid. Note that between figure 3A and figure 3B (bottom), the extracellular matrix is a larger and more expanded space due to the addition of the proteoglycan-hyaluronic acid.

Small-Chain Proteoglycans

Decorin

The small proteoglycans have been described to be predominantly layer specific, with decorin being in the **SLLP** and fibromodulin in the **MLLP** and **DLLP**[17]. As shown in the chart on proteoglycans, both fibromodulin and decorin bind to and have influence on the fibrous proteins of the **ECM** (See Table of Proteoglycans) Both have been shown to regulate collagen fibril rate of formation and thickness. Decorin appears to be found more in the **SLLP** based on Pawlak's[17] studies using immunocytochemical techniques. Although the function is speculative it is certainly possible that its presence may be the reason why the **SLLP** is sparsely populated with collagen fibers as opposed to the deeper layers of the **LP**.

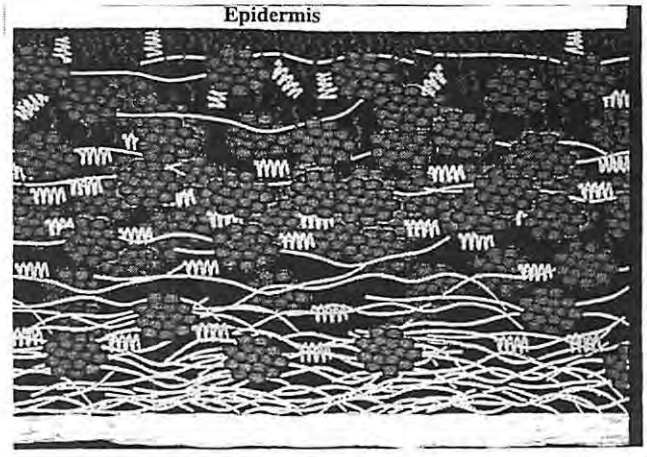


Figure 4. This represents the addition of decorin to the schematic model. Decorin is represented using green squiggles since that is how they actually appear using immunofluorescent techniques for identification. It is found predominantly in the superficial layer of the lamina protein and affects collagen fiber formation and density.

In our schematic model of the lamina propria (Figure 4), we will add in decorin, which is predominantly in the **SLLP** and functions to decrease collagen fiber size and density as well as having an ameliorating effect on tissue injury and healing.

Decorin and Scarring

Decorin may have substantial effects upon fibroblastic response to tissue injury. A study regarding pulmonary fibrosis was performed using decorin as an independent variable and measures of fibrosis as the outcome.[22] In this model bleomycin was instilled into rat lungs leading to severe interstitial fibrosis; the experimental group was infused with decorin in the post injury period. The group with decorin infusion had scar measurements (such as fibronectin deposition) that were similar to normal lungs rather than the non-infused fibrosis group. Again it is speculative, but the presence of decorin may be one of the reasons surgeons can operate on the **SLLP** without incurring significant fibrosis, whereas surgery to the deeper layers of the **LP** often results in fibrosis and scarring.

Biglycan

Biglycan is a small chain proteoglycan that will be only briefly discussed here. Studies thus far have not looked for biglycan in vocal folds. It is a proteoglycan more commonly found in scar tissue. It is mentioned only to point out a relationship to decorin. The presence of biglycan usually is associated with a decrease in decorin concentration. Frequently these two proteoglycan's presence and distributions are inversely related[23] (negatively correlated). Biglycan concentrations are high *(decorin concentration*

Fibromodulin

Figure 5. This figure shows a histologic slide of the distribution of fibromodulin used by Pawlak. Fibromodulin identification requires fresh frozen tissue to be used because antigenicity is lost with other techniques. The fibromodulin molecule is detected by the golden brown color. The remaining tissue looks shredded because of the necessity of performing the stain on frozen (otherwise unfixed) tissue. The fibromodulin is located in the same area as the vocal ligament. Note that the majority of the protein is again in the immediate infrafold region.

low) in hypertrophic scars and immature scars, with decorin concentrations being high in normal skin and normal mature scars. A similar study of burns demonstrated high levels of biglycan in immature scars, with more normal levels of decorin and low levels of biglycan found in the more mature scars.[24] Again normal to high levels of decorin in the **ECM** resulted in less scarring.

Fibromodulin

Although fibromodulin is structurally similar to decorin, it has not been associated with the same biologic properties on wound repair as decorin. Fibromodulin appears to be found more in association with tendons and ligaments. It is believed that fibromodulin influences aspects of ligament and tendon performance, although conclusive studies are lacking. In the vocal folds, fibromodulin has a unique distribution that surrounds the area of the vocal ligament (See Figures 5&6). Figure 5 shows a histologic slide of the distribution of fibromodulin used by Pawlak. Fibromodulin identification requires fresh frozen tissue to be used because antigenicity is lost with other techniques.

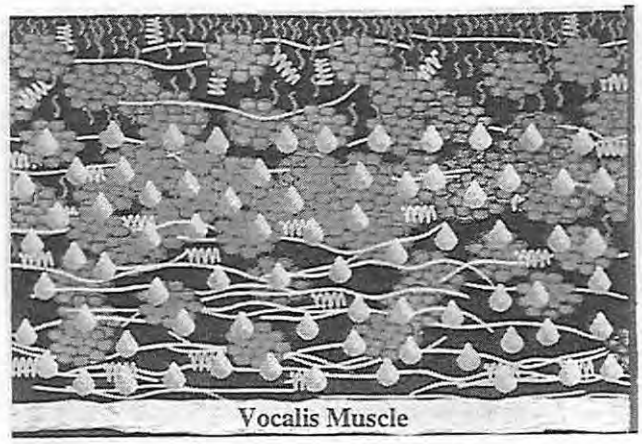


Figure 6. This represents a computer schematic of the fibromodulin as small blue oil droplets since it is likely that the function is to grease the collagen and elastin fibers of the vocal ligament to enhance tissue ligament performance for specific vocal needs.

The technique in this stain uses an antibody directed against fibromodulin and an avidin-biotin immunoperoxidase method with diaminobenzidine as a substrate. Therefore the fibromodulin molecule is detected by the golden brown color. The remaining tissue looks shredded because of the necessity of performing the stain on frozen (otherwise unfixed) tissue. Note the location of the fibromodulin with respect to the remaining, although hard to identify vocal fold **LP**. The fibromodulin is located in the same area as the vocal ligament. Note that the majority of the protein is again in the immediate infrafold region. Fibromodulin in figure 6 is represented by small blue oil droplets since it is likely that the function is to *"grease"* the collagen and elastin fibers of the vocal ligament to enhance tissue ligament performance for specific vocal needs.

Glycoproteins (Fibronectin), Carbohydrates & Lipids

These are all molecular components of the ECM. At this time few of these have received any investigation because of a variety of reasons, not the least is that carbohydrates and lipids *(not being proteins and therefore not antigenic)* are very difficult to study and require unique research training to do so. This does not mean they are not important. Chan and Titze's finding[37] that fat is similar to the viscosity of normal vocal folds underscores the concept that in our quest for understanding the biomechanical effects of the ECM we may yet be coming back to lipids and carbohydrates.

Fibronectin, the most common of the glycoproteins in the **ECM**, is somewhat ubiquitous in vocal fold **LP**.[25] Its functions are varied and still not fully described. It serves to help bind proteins together and provide molecular strength and adhesion between molecules. We do know that fibronectin plays an important role in wound healing and

these have been described with respect to laryngology.[26,27] Fibronectin is present in normal vocal folds and it seems to be in greater concentration in some disease states such as nodules and probably scarring.[26,28]

Viscosity

Viscosity is a measure of how poorly a substance flows. In the cgs system of measurements, the unit of viscosity is the poise (dyn-s/cm2). In the mks system, the unit is the pascal-second (Pa-s). Viscosity of vocal fold tissue is critical in certain aspects of vocal fold oscillation. Among those aspects are **a)** phonation threshold pressure and **b)** power or energy used by keeping the vocal folds oscillating.[29,30,31,32]

Phonation Threshold Pressure

Phonation threshold pressure (**PTP**) is defined as the minimum subglottic pressure required to produce vocal fold oscillation.[33,34] It is a measurable quantity that is an objective indicator of how easy the vocal folds can begin to oscillate. Titze described an equation that predicts PTP for a rectangular glottis:

$$PT = VDW / T$$

where **V** is the mucosal wave velocity (which is proportional to shear elasticity), **D** is the tissue-damping coefficient (which is proportional to tissue viscosity), **W** is the prephonatory glottal width and **T** is the thickness of the vocal folds.[33] Notice that these are four factors that influence and define how easy it is for us to get phonation started. When disease effects the vocal folds and a lot of effort is required to get their voice started, it is likely that one of the above factors has been changed enough that the **PTP** has increased. In severe cases **PTP** can increase enough so that vocal fold oscillation becomes virtually unobtainable. To lower **PTP**, and consequently make it easier for us to produce sound we have essential four options available: **a)** decrease mucosal wave velocity, **b)** decrease the prephonatory glottal width, **c)** increase the thickness of the vibrating portion of vocal fold, and d) decrease the tissue viscosity.

In view of this, it becomes apparent that pathologies that lower **PTP** usually either decrease tissue viscosity or increase the thickness of the vocal fold. An example is Rheinke's edema, in which often both viscosity is lowered and thickness of the vibrating portion of the vocal folds is increased. Of course tissue viscosity in Rheinke edema can vary dramatically, in some cases being less than that of normal tissue while in others being greater than that of normal tissue. Rheinke's edema may represent a disorder in which relative proportions of the proteoglycans of the vocal folds are altered. The problem with this disorder is not getting the vocal folds to oscillate, but to maintain a stability of oscillation with the extra fluids and proteins. As can be seen from the equation for **PTP**, tissue viscosity has a linear relationship with **PTP**.

Although not a subject of this paper, it is worth noting that **PTP** is proportional to the prephonatory glottal width. This may also have clinical importance.[35] A more in depth discussion of **PTP** and its factors with more accurate and less simplified equations is provided by Titze.[33]

Energy Required To Sustain Phonation

Sustaining phonation requires a constant supply of energy. The energy is provided by pressure and flow delivered from the lung. Energy is lost during phonation because frictional forces need to be overcome when viscous tissue is set into oscillation. The energy is lost in the form of heat. Theoretically it is possible to measure how much energy each individual uses to sustain vocal fold oscillation by measuring the heat produced in the tissue, but this is difficult in practice. A quantitative relationship has been derived from basic principles to estimate the energy dissipated by vocal fold oscillation[36]:

$$E = (LT / D)\eta\omega^2\xi^2$$

where **E** is the energy dissipated, **L, T, D** are the length, thickness and depth of the vocal folds respectively, η is the tissue viscosity, ω is the angular frequency of oscillation ($2\pi F_0$) and ξ is the vibrational amplitude. To decrease this energy dissipation in sustaining oscillation (or in other words, the amount of effort it requires for us to keep talking) we have the following options:

A) Decrease length and thickness of the vocal folds, increase depth of the vocal folds or decrease frequency of oscillation. These items have been grouped together because they all describe what happens to vocal fold configuration when pitch changes, meaning that elements of vocal fold configuration associated with higher frequency oscillation lead to more energy lost. As described by Titze, *"...this is simply because higher frequencies are associated with more rapid movement of the tissue, which in turn produces more friction."* [36]

B) Decrease vibrational amplitude by decreasing loudness of our voice. This equation shows that increasing vocal loudness, which is the result of increasing vibrational amplitude, leads to more energy loss since greater tissue movement results in greater frictional loss.

C) Lower tissue viscosity.

Fundamentally this equation shows that to expend less energy for continued vocal fold oscillation, you should keep your voice soft, at lower pitch or lower tissue viscosity. This equation helps us understand why voice therapy is useful treating certain voice disorders, and why surgeons often asked patients to reduce pitch and loudness while

vocal folds are healing since doing so requires the least amount of energy *(and consequently less tissue friction)* for vocal production. For purposes of this paper, this equation describes the critical importance of tissue viscosity for vocal production. An increase in tissue viscosity results in greater energy loss for sound production. This has ramifications in people with symptoms of vocal fatigue, a common symptom of voice disorders. In many of those voice disorders, beside speech therapy options dealing with the above items of loudness and pitch, the only real option that will be curative is changing the tissue viscosity.

Hyaluronic Acid And Vocal Fold Viscosity

Viscosity and Various Tissues

As we have seen, hyaluronic acid (**HA**) is a major component of the **ECM** in the vocal folds, and **HA** is probably the major determinant of tissue viscosity. To look at other substances and the viscous properties associated with them, Chan and Titze investigated Gelatin, GAX collagen, non-cross-linked collagen, abdominal fat and vocal fold mucosa in two specimens.[37] Figure 7 shows their results. The graph is arranged to show dynamic viscosity, which describes the viscous property as the substance is oscillated at different frequencies. As the graph shows, the viscosity goes down as frequency increases. Note that both axies are logarithmic displays, not linear. Chan and Titze found that collagen substances were much more viscous than normal vocal folds. It can be seen that the viscosities of the vocal fold mucosa at 10 Hz measured around 1-3 Pa-s *(Pascal second, which is 10-30 poise)* while abdominal fat was similar. Collagen substances measured around 8 to 12 Pa-s. Generally the viscosity measurements of **HA** have been measured at O.4 Pa-s at 10Hz[38,39]. Although not measured in their experiment, **HA** viscosity can be varied to almost any degree by changing the side chains attached to the protein core or by increasing the concentration of the **HA**.[40] An increasing concentration of **HA** leads to increasing tissue viscosity in nearly an exponential function until a maximum concentration is reached *(about 1 gram/dl)*.[41,42] At that level, the **HA** molecules are essentially packed together resulting in a very viscous gel.

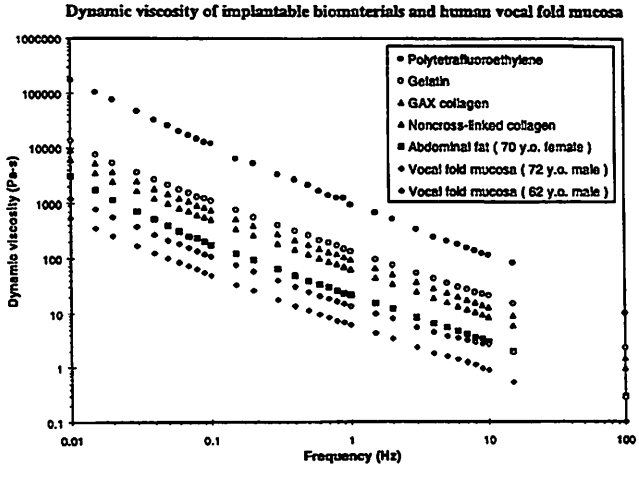


Figure 7. This graph shows the dynamic relationships of various implantable biomaterials. The graph is logarithmic and not linear. Reprint Permission of the above graph granted by Lippincott-Raven Publishers, 1997.

Effect of Hyaluronic Acid on Vocal Fold Viscosity

In order to determine the effect of **HA** on vocal fold viscosity, the following experiment was performed: samples of the cover of the vocal fold from two human specimens were removed microscopically using ear dissection instruments and an operating microscope. The epithelium and superficial layer of the lamina propria were included. The vocal ligament area was not included. The samples were then placed in a Bohlin Controlled Stress *(CS-50)* Rheometer using the same methodology previously described by Chan and Titze.[37]

Following determination of the dynamic viscosity for the two specimens, both specimens were bathed in testicular hyaluronidase solution, mixed according to the manufacturer. Testicular hyaluronidase is specific for enzymatically digesting **HA**. The specimens were bathed for one hour at a temperature controlled at 37 degrees C. The specimens were then removed from the bath and rinsed twice with a physiologic saline solution. Noticeable change in the cover was apparent with a much more white color *(similar to the fibrous protein collagen color)*. The two specimens were then placed back in the Rheometer and dynamic viscosities were again measured.

Results

The results from the two specimens are shown in Figure 8. Since two samples are obviously too few in number, individual variation cannot be interpreted reliably. But we know from the above discussion and Hammond's data that the amount of **HA** in each person's vocal fold is individual. Also, most of the **HA** is found in the intermediate layer, which is not part of the dissected cover in this experiment. Therefore, this experiment on the cover may not show the magnitude of change possibly expected if the intermediate layer were to be included. Nevertheless, the purpose of the experiment was to determine the dynamic viscosity change of the cover associated with the removal of **HA** in order to test the hypothesis that **HA** does influence vocal fold viscosity. From the data it appears that dynamic viscosity is affected by **HA** and that the removal of **HA** leads to increased tissue viscosity. Note that in specimen 2611 the viscosity increased nearly 3 to 4 fold, whereas in the other specimen the increase was a factor of two to three. On a logarithmic scale this may not seem much, but **PTP** and

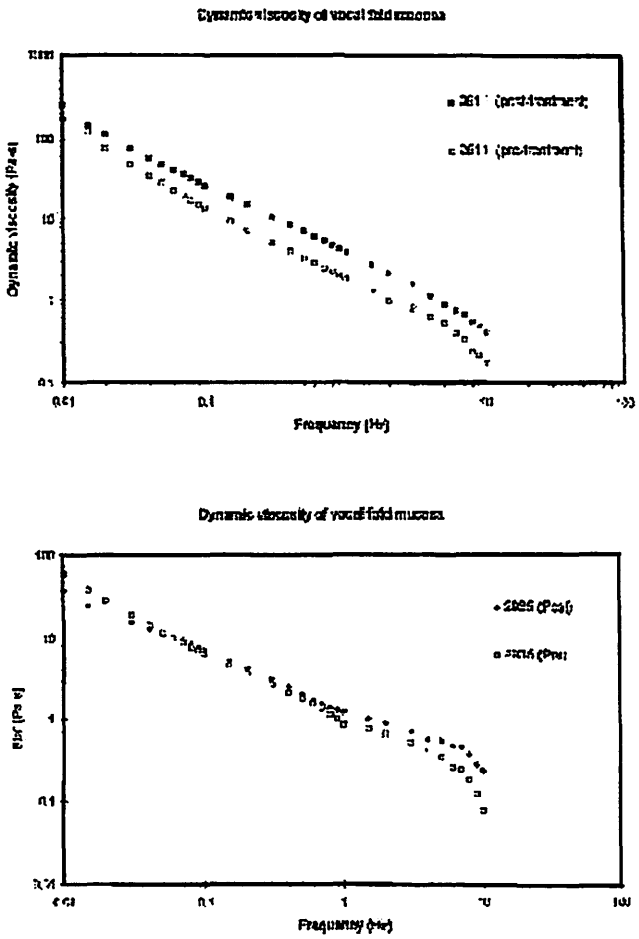


Figure 8. With dynamic viscosity plotted on a logarithmic scale to the y-axis and frequency being placed on the x-axis, this figure demonstrates two human vocal fold samples before and after the removal of hyaluronic acid. Figure 8A (top) shows that the viscosity increased following the removal of hyaluronic acid. Figure 8B (bottom) demonstrates similar results. The differences become more apparent at higher frequencies. The current instrumentation only allows us to go up to 10 hertz whereas vocal folds vibrate at much higher frequencies. At 10 hertz, an increase in viscosity of 3-4 fold is already apparent following the removal of hyaluronic acid.

energy required for sustained phonation have linear relationships with vocal fold viscosity, in which case a factor of 2 or 3 is substantial.

Future

Dynamic vocal range was discussed in the Fibrous Proteins: Part I. The key to a large dynamic vocal range is being able to generate tissue stress easily, but the present discussion underscores that low energy dissipation is also important. For easy phonation we would ideally have a very loose, very pliable, low viscosity tissue in the mucosa. But to have dynamic range, we need highly stressed tissue in the ligament of the muscle. In some ways, these are opposing concepts that have been dealt with wonderfully by the human body and the vocal folds. Otolaryngologists are challenged to recreate a similar structure surgically when it is lost due to disease. Simply creating a nice low viscous mucosal layer would improve ease of phonation, but may result in significant loss of dynamic range if too much of the natural properties of the ligament are compromised.

One of the intents of this paper was to provide a discussion of the **ECM** of the lamina propria and its related biomechanical properties. In the future, we hope to learn how to manipulate tissue viscosity, since it is the one biomechanical item that potentially could lead to better restoration of vocal function. Proteoglycans appear to be key molecules in this regard, particularly **HA**. Fortunately, laryngology is not alone in this research arena and **HA** research is a big enterprise. **HA** research has significant interest in synovial fluid diseases, such as arthritis or the area of synthetic cartilage substitutes. Researchers have learned that viscosity of hyaluronic acid is determined by the size of the protein, carbohydrate or lipid chain attached to the protein core. Larger chains *(molecules)* create a more viscous environment whereas small chains lead to lower tissue viscosity. On the other hand, increased concentration of HA also raises tissue viscosity. These molecular chains attached to hyaluronic acid may be hydrophobic or hydrophilic, which may affect matrix interactions and water content of the **ECM**. In the future both chain size and **HA** concentration will be used to manipulate tissue viscosity. Matrix interaction with the **HA** will likely be modulated by attaching different pharmacologic side chains. Various drugs, including cortisone, and lipids can be attached to these side chains, giving us further methods to influence matrix interactions (like scarring or inflammation) and effect tissue viscosity.

Conclusion

The vocal fold lamina propria is composed of a number of proteins, which make up a matrix in the lamina propria. This matrix determines the biomechanical properties of the lamina propria. Most of the proteins show some layer association, meaning the deposition of that protein varies according to the lamina propria layer. Collagens and elastins are important fibrous proteins of the extracellular matrix and help determine certain aspects of vocal performance. Particular stress-strain properties of the vocal ligament are likely reflective of collagen and elastin fiber interaction.

Interstitial proteins such as the proteoglycans fill in between the fibrous proteins and therefore probably have a strong effect on biomechanical performance. Tissue viscosity is an important determinant of properties for tissue oscillation. Hyaluronic acid has been shown to be a key molecule influencing tissue viscosity. Hyaluronic acid deposition shows gender specificity, with men having more HA than women. Manipulation of some of these **ECM** components, particularly **HA**, may be valuable therapeutic tools in the future.

Acknowledgments

This work was supported by grant P60 DC00976-09 to the National Center for Voice and Speech from the National Institute on Deafness and Other Communicative Disorders (NIDCD).

Bibliography

1. Ruoslahti E: **Structure and biology of proteoglycans.** *Annu Rev Cell BioL* 1988;4:229-55.

2. Couchman JR, Woods A: **Structure and biology of pericellular proteoglycans.** In: Roberts DD, Mecham RP, eds. *Cell surface and extracellular glycoconjugates: structure and function.* San Diego, CA: Academic Press, 1993:33-82.

3. McCarthy KJ, Accavifti MA, Coughman JR: **Immunological characterization of a basement membrane specific chondroitin sulfate proteoglycan.** *JCellBioL* 1989;109:3187-98.

4. Lander Ad: **Proteoglycans. Guidebook to the extracellular matrix and adhesion proteins.**InKreisT, ValeR,eds. NewYork,NY: Oxford University Press,1993:12-16.

5. Labat-Robert J,Bihari-Varga M,Robert L: **Extracellular matrix.** FEBSLett.1990;268:386-93.

6. Heinegard D, Franze A, Hedbom E, Sommarin Y: **Common structures of the core proteins of interstitial proteoglycans.** *Ciba Found Symp.* 1986:124:69-88.

7. Kimura JH, Shinomura T, Thonar EJ: **Biosynthesis of cartilage proteoglycan and link protein.** Methods *EnzymoL* 1987;144:372-93.

8. Pawlak AS, Hammond TH, Hammond E, Gray SD: **Immunocytochemical study of proteoglycans in vocal folds.** *Ann Otol Rhino & Laryng.* 1996;105:6-1.

9.Scott JE, Heatley F, Hull WE: **Secondary structure of hyaluronate in solution.** *Biochem J* 1984;220:197-205.

10. Carson FL: **Histotechnology: a self-instructional text.** Chicago: ASCP, 1990:127-9.

11. Hammond TH, Zhou R, Hammond EH, Pawlak A, Gray SD: **The intermediate layer: a morphologic study of the elastin and hyaluronic acid constituents of normal human vocal folds.** *J Voice.* 1997;1 1:5966.

12. Hammond TH, Gray SD, Butler J, Zhou R, Hammond E: **A study of age and gender related elastin distribution changes in human vocal folds.** *Oto HNS J.* Accepted for publication.

13. Westergren-Thorsson G, Hernnas J. Sarnstrand B. Oldberg A. Heinegard D, Malmstrom A: **Altered expression of small proteoglycans, collagen, and transforming growth factor-beta 1 in developing bleomycin-induced pulmonary fibrosis in rats.** *J of Cli Invest.* 1993;92:632-7.

14. Garg HG, Siebert JW, Garg A, Neame PJ: **Inseparable iduronic acid-containing proteoglycan PG(ldoA) preparations of human skin and post-burn scar tissues: evidence for elevated levels of PG(ldoA)-i in hypertrophic scar by N-terminal sequencing.**

15. Scott PG, Dodd CM, Tredget EE, Ghahary A, Rahemtulla F: **Immunohistochemical localization of the proteoglycans decorin, biglycan and versican and transforming growth factor-beta in human post burn hypertrophic mature scars.** *Histopath.* 1995;26:423-31.

16. D'Ardenne AJ, Burns J, Sykes BC, Kirkpatrick P: **Comparative distribution of fibronectin and type III collagen in normal human tissues.** *J Patho L* 1983;141:55-69.

17. Gray SD, Hammond E, Hanson DF: **Benign pathological responses of the larynx.** *Ann Otol, Rhino & Laryng.* 1995;104:13-18.

18. Grinnell F: **Fibronectin and wound healing.** *J Cell Biochem.* 1984; **26:107-16.**

19. Courey MS, Shohet JA, Scott MA:, Ossoff RH: **Immunohistochemical characterization of benign laryngeal lesions.** *Ann Otol Rhinol Laryng,* 1996; 105(7):525-31.

20. Finkelhor BK, Titze IR, Durham PL: **The effect of viscosity changes in the vocal folds on the range of oscillation.** *J Voice.* 1987;1:320-325.

21. Titze IR, Schmidt SS, Titze MR: **Phonation threshold pressure in a physical model of the vocal fold mucosa.** *J Acoust Soc Am.* 1995;97:3080-3084.

22. Verdolini - Marston K, Titze IR, Druker DG: **Changes in phonation threshold pressure with induced conditions of hydration.** *J Voice.* 1990;4:142-151.

23. Verdolini K, Titze IR, Fennell A: **Dependence of phonatory effort on hydration level.** *J Speech Hear Res.* 1994; **37:1001-1007.**

24. Titze IR: **The physics of small-amplitude oscillation of the vocal folds.** *J Acoust Soc Am.* 1988; 83:1536-1552.

25. Titze IR: **Phonation threshold pressure: A missing link in glottal aerodynamics.** *J Acoust Soc Am.* 1992; 91:2926-2935.

26. Gray, SD, Bielamowicz SA, Titze IR, et al: **Experimental approaches to vocal fold alteration: Introduction to the minthyrotomy.** *Ann Otol, Rhinol, Latyng,* accepted for publication 1998.

27. Titze IR: **Heat generation of the vocal fold and possible effect on vocal endurance.** In: Lawrence V, *ed. Transcripts of the 100 symposium. care of the professional voice.* Philadelphia, PA: Voice Foundation: 1982:52-59.

28. Chan RW, Titze IR: **Viscosities of injectable biomaterials in vocal fold augmentation surgery.** *Laryngoscope.* Submitted for publication.

29. Balazs EA, Gibbs Da: **The Theological properties and biological function of hyaluronic acid.** In Balazs EA, ed. *Chemistry and molecular biology of the intercellular matrix* (vol. 3). New York: Academic Press, 1970:1241-1253.

30. Zhu WB, Mow VC: **Viscometric properties of proteoglycan solutions at physiological concentrations.** In Mow VC, Ratcliffe A, Woo S, eds. *Biomechanics of diarthrodial joints.* New York: Springer-Verlag, 1990:313-344.

31. Bothner H, Wik 0: **Rheology of hyaluronate.** *Acta Otolaryngol (Stockh).* 1987; 442:25-30.

32. Laurent TC, Laurent UBG, Fraser JR: **Functions of hyaluronan.** *Annals of Rheu Dis.* 1995;54:429-432.

33. Morris ER, Rees DA, Welsh EJ: **Conformation and dynamic interactions in hyaluronate solutions.** J *Mol BioL* 1980; 138:383-400.

NCVS Status and Progress Report - 13
April 1999, 21-36

Viscoelastic Shear Properties of Human Vocal Fold Mucosa: Measurement Methodology and Empirical Results

Roger W. Chan, Ph.D.
Ingo R. Titze, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa

Abstract

A standard method for the empirical rheological characterization of viscoelastic materials was adopted to measure the viscoelastic shear properties of human vocal fold mucosal tissues (the superficial layer of lamina propria). A parallel-plate rotational rheometer was employed to measure shear deformation of viscoelastic tissue samples, which were deformed between two rigid circular plates rotating in small-amplitude sinusoidal oscillations. Elastic and viscous shear moduli of the samples were then quantified as a function of oscillation frequency (0.01 Hz to 15 Hz) based on shear stresses and strains recorded by the rheometer. Data were obtained from 15 excised human larynges (10 male and 5 female). Results showed that the elastic shear modulus μ and the damping ratio ζ of human vocal fold mucosa were relatively constant across the range of frequencies observed, while the dynamic viscosity η decreased monotonically with frequency (i.e., shear-thinning). Intersubject differences in μ and η as large as an order of magnitude were observed, part of which may reflect age-related and gender-related differences. Some molecular interpretations of the findings are discussed.

Introduction

Mechanical properties of vocal fold tissues are important in the study of the acoustics and biomechanics of voice production. This has been demonstrated by computer simulation of speech and voice, where knowledge of tissue mechanical properties is required to predict tissue forces and deformation (through constitutive equations). For example, Titze (1976) and Titze and Talkin (1979) assumed transverse isotropy in a linear, elastic continuum to model vocal fold tissues, which in principle requires five independent tissue mechanical constants for the constitutive equations. These elastic constants include the longitudinal Young's modulus E' (an indication of tensile elasticity or stiffness along the direction of tissue fibers), the transverse and the longitudinal Poisson's ratios (ν and ν', respectively, as indications of tissue compressibility or the amount of tissue volume change upon deformation), and the transverse and the longitudinal shear moduli (μ and μ', respectively, as indications of tangential elasticity or stiffness). Berry and Titze (1996) later showed that the two shear moduli were sufficient if tissue is assumed to be incompressible (such that the Poisson's ratios are known *a priori*) and there is no vibration in the direction of tissue fibers (such that E' does not enter the equations of motion).

Among the different layers of vocal fold tissues, the vocal ligament and the thyroarytenoid muscle may be assumed to be transversely isotropic, where tissue mechanical properties are independent of the orientation of deformation in the plane transverse to the essentially parallel tissue fibers, but different in the (longitudinal) direction of the fibers. On the other hand, the superficial layer of lamina propria (mucosa) can be assumed to be isotropic, where tissue properties are independent of the orientation of deformation in all planes, because of the basically random orientation of macromolecules in the lamina propria extracellular matrix (Hirano, 1981; Matsuo *et al.*, 1984; Pawlak *et al.*, 1996; Hammond *et al.*, 1997). Hence only *one* shear modulus is needed for the superficial layer, as the transverse and the longitudinal shear moduli become identical. A fair amount of empirical data on tissue fiber tension (for ligament and muscle) has been gathered (e.g., Alipour-Haghighi and Titze, 1991; Min *et al.*, 1995). This tension can be related to μ'. However, little is known about μ, neither for the isotropic mucosa nor the transversely isotropic ligament

and muscle. The present study is directed toward this deficiency in the knowledge of vocal fold tissue properties. In fact, shear properties of other orofacial tissues necessary for solving the mechanics of other speech production models are also lacking (e.g., finite element models of the tongue and the velum, Wilhelms-Tricarico, 1995; Berry *et al.*, 1998).

Knowledge of the mechanical properties of vocal fold tissues is also important clinically for biomechanical implants. The effect of tissue properties on vocal fold oscillation should be considered in the surgical management of vocal fold disorders (phonosurgery), where various types of synthetic or biological substitute materials have been used as surgical implants to replace or augment natural tissues in pathological vocal folds. In the management of vocal fold mucosal defects, e.g., scarring and atrophy, biomaterials such as collagen and fat are implanted close to or directly into the vocal fold mucosa as a filler substance, so as to repair focal defects, to smooth the vocal fold margin, and to soften any stiffened scar tissue (Benninger *et al.*, 1996; Ford *et al.*, 1992; Sataloff *et al.*, 1997; Shaw *et al.*, 1997). It is obvious that the introduction of these implantable biomaterials into the vocal fold can change its mechanical properties and alter the mechanics of vocal fold oscillation, which could make phonation either easier or more difficult. This is particularly true when the vocal fold mucosa is directly involved in repair, because the mucosa is the major vibratory portion of the vocal fold, especially in small-amplitude oscillations like phonation onset and offset (Hirano, 1975; Saito *et al.*, 1985; Fukuda *et al.*, 1991). *Shear* properties of vocal fold tissues and implantable biomaterials are the most relevant in this context, because oscillation of the mucosa involves the propagation of a surface mucosal wave, which is a shear wave (Baer, 1975; Hirano, 1975). Therefore, shear properties become important factors in the choice of optimal biomaterials for mucosal repair (Chan, 1998; Chan and Titze, 1998a). In other situations, like the management of unilateral vocal fold paralysis, where the main objective of surgery is to restore glottal competence by medializing and augmenting the paralyzed vocal fold, the properties of biomaterials like Teflon and gelatin are less important because they are usually implanted lateral to the vibratory portion of the vocal fold (Arnold, 1963; Schramm *et al.*, 1978).

Given the lack of information on the shear properties of vocal fold tissues in the literature, this paper attempts to provide some empirical data on the shear properties of human vocal fold mucosal tissues. A standard method commonly used in rheology (see definition below) was adopted to characterize the shear properties of human vocal fold mucosa, namely *sinusoidal oscillatory shear deformation*. Details of this *in vitro* method for the empirical measurement of shear properties are described below, and the linear viscoelastic theory behind the method is reviewed. Empirical data obtained with oscillatory shear experiments are interpreted in terms of some molecular findings on vocal fold tissues and other extracellular connective tissues in the biomedical literature. Theoretical modeling of viscoelastic shear properties, helpful for data extrapolation and parametric descriptions of vocal fold tissues, will be addressed in a forthcoming paper (Chan and Titze, 1998b).

Measurement of Viscoelastic Shear Properties

In material science, ideal solids are described as "Hookean elastic" and ideal fluids as "Newtonian viscous". Many materials, however, are *viscoelastic*, exhibiting both elastic and viscous properties. The behaviors of viscoelastic materials depend on the rate of deformation, in addition to thermodynamic (temperature and pressure) influences (Ferry, 1970; Fung, 1993). *Rheology*, defined by the Society of Rheology as the scientific study of the deformation and flow of matter, is related to the study of viscoelasticity, particularly viscoelastic shear properties. Standard rheological techniques can be employed to empirically determine the shear properties of viscoelastic materials (Ferry, 1970; Bird *et al.*, 1977; Whorlow, 1980; Barnes *et al.*, 1989). Two such standard methods are discussed below: *steady shear* deformation and *oscillatory shear* deformation.

Steady Shear Deformation

Under steady shear (i.e., continuous, uniform and non-oscillatory shear) deformation, the viscosity (*steady-shear viscosity*) of a purely viscous (Newtonian) fluid is defined as follows:

$$\tau = \eta_S \dot{\gamma} \qquad (1)$$

where τ is shear stress (in Pa), η_S is steady-shear viscosity (in Pascal-second, Pa-s), γ is shear strain and $\dot{\gamma}$ is strain rate (in s^{-1}). The "dot over" notation represents a time-derivative. The SI unit Pascal-second is ten times larger than the somewhat more commonly used unit Poise (P). Water has a steady-shear viscosity of 0.001Pa-s, or 0.01P. For a Newtonian fluid like water, the steady-shear viscosity is a constant independent of strain rate (but dependent on temperature). For non-Newtonian fluids (including most polymeric materials), however, viscosity is a function of strain rate, and there can also be an elastic component not described by Eq. (1). Hence steady shear deformation alone is not the condition sufficient to characterize the rheological or shear properties of viscoelastic materials (Ferry, 1970; Bird *et al.*, 1977; Fung, 1993).

Biological tissues and materials are all viscoelastic (Fung, 1993). Their elastic and viscous shear properties are revealed when they are subject to sinusoidal oscillatory shear deformation, which is summarized next.

Oscillatory Shear Deformation

Viscoelastic shear properties are quantified by the *complex shear modulus* (or dynamic shear modulus), which includes an elastic component, the *elastic shear modulus*, and a viscous component, the *dynamic viscosity*. These components are additive for viscoelastic materials, according to the theory of *linear* viscoelasticity (Ferry, 1970; Whorlow, 1980; Barnes *et al.*, 1989; Fung, 1993). Therefore, for a viscoelastic material, the general linear constitutive equation relating shear stress with shear strain and strain rate is

$$\tau = \mu\gamma + \eta\dot{\gamma} \tag{2}$$

where μ is elastic shear modulus or storage shear modulus (in Pa) and η is dynamic viscosity (in Pa-s) of the viscoelastic material. Consider the application of a sinusoidal oscillatory shear strain (with a given amplitude and frequency) to a sample of the material. According to the constitutive equation, it can be shown that steady-state conditions of oscillation will be reached for *small-amplitude* oscillations, for which the sinusoidal shear strain (input) will result in a sinusoidal shear stress (output) of the material at the *same* frequency of oscillation. The resulting shear stress will have an amplitude proportional to the strain amplitude, and a phase lag independent of stress or strain amplitude (Ferry, 1970; Whorlow, 1980; Fung, 1993). The same linear theory also applies when a sinusoidal shear *stress* is being applied, resulting in a sinusoidal shear strain. Using complex number notations, the shear strain and shear stress can be represented as:

$$\gamma^* = \gamma_0 e^{i\omega t} \tag{3}$$

$$\tau^* = \tau_0 e^{i(\omega t+\delta)} \tag{4}$$

where γ^* is complex shear strain (the asterisk superscript represents complex number notation), γ_0 is shear strain amplitude, τ^* is complex shear stress, τ_0 is shear stress amplitude, i is the imaginary number $\sqrt{-1}$, ω is angular frequency of oscillation, t is time, and δ is the phase shift (lead) of shear stress relative to shear strain. The ratio of the shear amplitudes (τ_0/γ_0) and the phase shift (δ) are related to the elastic shear modulus μ and the dynamic viscosity η, which are functions of the frequency of oscillation and thermodynamic variables like temperature and pressure.

By time differentiation of the complex shear strain, the complex strain rate is

$$\dot{\gamma}^* = i\omega\gamma_0 e^{i\omega t} \tag{5}$$

According to the constitutive equation [Eq. (2)], application of the sinusoidal shear strain [Eq. (3)] on the viscoelastic material sample yields a shear stress of

$$\begin{aligned}\tau^* &= \mu\gamma_0 e^{i\omega t} + i\omega\eta\gamma_0 e^{i\omega t} \\ &= (\mu + i\omega\eta)\gamma_0 e^{i\omega t} \\ &= \mu^*\gamma^*\end{aligned} \tag{6}$$

where μ^* is the complex shear modulus (also called the dynamic shear modulus or the modulus of rigidity; Fung, 1993):

$$\mu^* = \mu + i\omega\eta \tag{7}$$

It is clear from Eq. (6) that the complex shear modulus can also be defined as the ratio of complex shear stress to complex shear strain. Using a polar coordinate representation, μ^* can be expressed in terms of its *magnitude* $|\mu^*|$ and phase angle δ, which is just the phase shift between stress and strain (Ferry, 1970; Fung, 1993):

$$\mu^* = |\mu^*| e^{i\delta} \tag{8}$$

where

$$|\mu^*| = \sqrt{\mu^2 + (\omega\eta)^2} \tag{9}$$

and

$$\delta = \tan^{-1}\frac{\omega\eta}{\mu} \tag{10}$$

Expressing the complex shear stress in terms of $|\mu^*|$ and δ according to Eqs. (6) and (8),

$$\tau^* = |\mu^*|\gamma_0 e^{i(\omega t+\delta)} \tag{11}$$

Thus the application of a sinusoidal shear strain on a viscoelastic material described by the linear constitutive equation does result in a sinusoidal shear stress at the same frequency of oscillation (ω). The resulting shear stress is expressed exactly in the form shown in Eq. (4), with

$$\tau_0 = |\mu^*|\gamma_0 \tag{12}$$

Note that the shear stress amplitude τ_0 is proportional to the shear strain amplitude γ_0, with the proportionality constant being the magnitude of the complex shear modulus $|\mu^*|$. In other words, the magnitude of the complex shear modulus turns out to be the ratio of the shear amplitudes (τ_0/γ_0).

The elastic shear modulus μ and the dynamic viscosity η can also be expressed in terms of the shear amplitudes and the phase shift measurable by oscillatory shear deformation. According to Eq. (12) and Euler's relation, the polar form of the complex shear modulus [Eq. (8)] can be expanded as

$$\mu^* = \frac{\tau_0 \cos\delta}{\gamma_0} + i\frac{\tau_0 \sin\delta}{\gamma_0} \quad (13)$$

which yields the following expressions for its real part and imaginary part [c.f. Eq. (7)]:

$$\mu = \frac{\tau_0 \cos\delta}{\gamma_0} \quad (14)$$

$$\eta = \frac{\tau_0 \sin\delta}{\omega\gamma_0} \quad (15)$$

The elastic shear modulus μ is proportional to the elastically stored strain energy (or internal energy) in the viscoelastic material over one cycle of oscillation. The dynamic viscosity η, on the other hand, is proportional to the energy dissipated or lost in the viscoelastic material, typically as heat. It characterizes the mechanical opposition to (shear) flow in the material and is often a monotonically decreasing function of frequency for viscoelastic and polymeric materials.

In the rheology and viscoelasticity literature (Ferry, 1970; Bird *et al.*, 1977; Whorlow, 1980; Barnes *et al.*, 1989; Fung, 1993), these viscoelastic shear properties have been commonly represented using the symbol G:

$$G^* = G' + iG'' \quad (16)$$

where $G^* = \mu^*$, $G' = \mu$ and $G'' = \omega\eta$. G'' is called the *viscous shear modulus* or loss shear modulus, defined as the product of angular frequency and dynamic viscosity. The dynamic viscosity η is commonly represented by η' because it can be defined as the real part of the complex dynamic viscosity η^*. Using the symbol G in polar form,

$$G^* = |G^*| e^{i\delta} \quad (17)$$

The magnitude of the complex shear modulus has been used as an indication of the overall shear elasticity, stiffness, or rigidity of a viscoelastic material, while the phase shift is an indication of the amount of energy loss relative to the amount of energy stored in the material. The *damping ratio* ζ (also called *loss tangent* or *loss factor*) can be defined as the tangent of the phase shift:

$$\zeta = \tan\delta = \frac{\omega\eta}{\mu} \quad (18)$$

A material is purely elastic if $\zeta = 0$ (when $\eta = 0$), purely viscous if $\zeta = \infty$ (when $\mu = 0$), and viscoelastic if ζ lies somewhere in between (when $0° < \delta < 90°$, i.e., when there are both viscous and elastic components). In the context of oscillation, a system is underdamped if $\zeta < 1.0$, critically damped if $\zeta = 1.0$, and overdamped if $\zeta > 1.0$. The typical values of damping ratio for human vocal fold tissues have been reported to be on the order of 0.1-0.2 at 30-40 Hz (Kaneko *et al.*, 1972) or 0.2-0.4 at 130 Hz (Isshiki, 1977). Vocal folds of excised human larynges were perturbed with mechanical impulses and their damped oscillations were measured in those studies. Accordingly, in the two-mass model of vocal fold oscillation, Ishizaka and Flanagan (1972) used a damping ratio of 0.1 for the lower mass and 0.6 for the upper mass (an equivalent value of 0.16 for the bulk of the vocal fold) to simulate self-sustained oscillation at a reasonable range of model parameters (e.g., subglottal pressure, vocal fold length, mass, and spring constants). However, these data on damping ratio were based on damped oscillations of whole vocal folds, not individual tissue layers. No data is available on the damping ratios of different vocal fold tissue layers, particularly the vocal fold mucosa which is the major vibratory portion in small-amplitude oscillations. Also, it is known that damping ratio is a function of frequency (albeit a weak function for most biological soft tissues; c.f. Fung, 1993), yet data are only available at a few frequencies. Therefore the present study attempts to provide a more comprehensive set of data on the damping ratio of vocal fold mucosal tissues, as well as their elastic shear modulus and dynamic viscosity.

Method

Vocal Fold Tissue Samples

A rheometer was used to measure the complex shear modulus of human vocal fold mucosal tissues. Tissue samples were obtained from two sources, 1) fresh excised larynges from the University of Iowa Hospitals and Clinics (UIHC) autopsy unit within 24 hours post-mortem; and 2) quick-frozen excised larynges shipped from the University of Utah (courtesy Dr. Steven Gray). Most of the larynges from UIHC autopsy were obtained within 18 to 20 hours post-mortem. All of the larynges from Utah were quick-frozen with liquid nitrogen within 18 hours post-mortem and stored at a temperature below -20°C. Larynges were harvested from ten male and five female cadaver subjects. None of them had a history of laryngeal pathology, laryngeal intubation or smoking. Table 1 summarizes the subject information.

Since it was recognized that post-mortem changes and/or freezing might have introduced significant differences in the mechanical properties of the vocal fold mucosal tissues, an experiment was done to compare the shear properties of fresh canine vocal fold tissues with those of frozen tissues (Chan, 1998). Canine tissues were used because they could be obtained immediately post-mortem

Table 1
Characteristics of Human Cadaver Subjects

Subject Number	Sex	Race	Age	Source
1	M	Polynesian	28	Utah
2	M	white	30	Utah
3	M	white	31	Utah
4	M	n.a.	34	UIHC autopsy
5	M	white	36	Utah
6	M	white	50	UIHC autopsy
7	M	white	59	UIHC autopsy
8	M	white	60	UIHC autopsy
9	M	n.a.	62	UIHC autopsy
10	M	n.a.	72	UIHC autopsy
11	F	white/Hispanic	25	UIHC autopsy
12	F	n.a.	55	UIHC autopsy
13	F	n.a.	55	UIHC autopsy
14	F	white	71	UIHC autopsy
15	F	n.a.	86	UIHC autopsy

Note: n.a. = race information not available

from cardiovascular laboratories in the UIHC, while it was not feasible to obtain fresh enough human larynges from UIHC autopsy. Results showed that up to 24 hours of post-mortem changes at room temperature and up to one month of freezing storage following quick freezing of mucosal tissues (using liquid nitrogen) did not affect tissue shear properties significantly.

For each excised larynx, the vocal fold epithelium and the superficial layer of lamina propria were carefully dissected. A small incision on the superior surface of the vocal fold epithelium, just above the boundary between the superficial layer of lamina propria and the vocal ligament, was made by a surgical blade (No. 15). A fine tissue grasping forcep was used to slightly lift up the mucosa. The superficial layer of the lamina propria was separated from the intermediate layer (the vocal ligament) by a Bouchayer spatula for blunt dissection. The dissection process was similar to the lateral microflap technique in phonosurgery (Courey and Ossoff, 1995). Under this procedure, the epithelium always remained attached to the superficial layer during dissection, as well as during the rheometric experiments. Its purpose was to serve as a natural boundary of attachment between the rheometer and the superficial layer of lamina propria.

Except for some frozen hemilarynges from Utah, two samples of mucosal tissues were obtained from each larynx, one from the left and the other from the right vocal fold. Symmetry in shear properties between two vocal folds of the same larynx was assessed in a few subjects. The volume of each tissue sample was approximately 0.1 cc, and

Figure 1. Photo of the Bohlin CS-50 Rheometer.

the gap size between the two plates of the rheometer was adjusted to fit this sample volume (see below).

Instrumentation and Rheometric Measurements

The Rheometer

A computer controlled rheometer (Bohlin CS-50, Bohlin Instruments Inc., Cranbury, NJ) was used for data acquisition and analysis (Fig. 1). A parallel-plate (plate-on-plate) testing geometry was used, consisting of a stationary lower plate and a rotating upper plate (diameter = 20 mm). With this geometry, the sample volume required to fill the gap between the two plates could be as small as 0.1 cc when a gap size of 0.3 mm was used. Since the sample volume of vocal fold mucosa obtained from one single vocal fold was typically on the order of 0.1 cc, the gap size of 0.3 mm was appropriate (Fig. 2; following page).

In the rheometer, the sample was subject to a precisely controlled sinusoidal torque from the upper plate, driven by a drag cup motor with a torque range of 0.001 to 10 mNm (milli Newton-meter) and a resolution of 0.0002 mNm. The resulting angular displacement and angular velocity of the upper plate were monitored by a sensitive

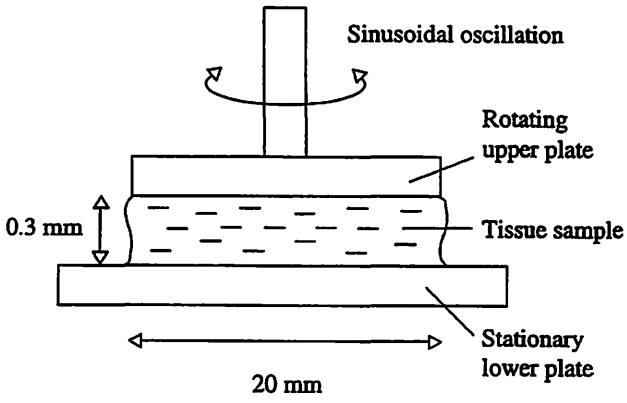


Figure 2. Schematic showing a tissue sample between the two plates of the rheometer (cross-sectional view not to scale).

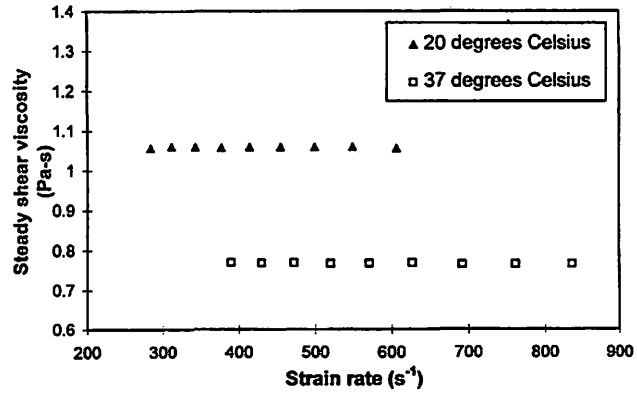


Figure 3. Calibration of rheometer with silicone (standard viscosity = 1 Pa-s at 20°C).

transducer (with a fully linear angular deflection range, and a resolution of 2×10^{-5} rad) as functions of time. Sinusoidally varying shear stress, shear strain and strain rate were calculated from the prescribed torque and the measured angular velocity by the computer, based on conversion constants of the testing system geometry in use. For the parallel-plate geometry with a diameter of 20 mm and a gap size of 0.3 mm, the shear stress-torque conversion constant was 477,465 Pa/Nm. The strain rate-angular velocity conversion constant was 25 s^{-1}/rad s^{-1}. The rotational moment of inertia of the upper plate about the long axis was 2.2×10^{7} kgm^2. From the shear stress and strain functions, the shear amplitudes (τ_0 and γ_0) and phase shift (δ) were measured by the controlling software. From these, the elastic shear modulus, the dynamic viscosity and the damping ratio were computed according to the analytical expressions discussed in the previous section. The software used a window of approximately 20 cycles of oscillation for the calculations.

Oscillatory Shear Experiments

Measurements of shear properties were made at a frequency range of 0.01Hz to 15Hz, covering 32 frequencies over three decades. Testing at higher frequencies was not meaningful, because linearity of oscillations could not be ensured when rotor (upper plate) and sample inertial effects became significant at high frequency (see below). Both upward and downward frequency sweeps were performed for a few samples in order to check for possible differences between the two. No significant differences were observed between them (Chan, 1998). Hence, in order to be efficient, only upward frequency sweeps were performed for all measurements.

Calibration of the rheometer was done by the manufacturer and was double checked by measuring the steady-shear viscosities of (1) standard Newtonian silicone (polydimethylsiloxane) solutions (Dow Corning 360 Medical Fluids, Dow Corning Company, Midland, MI) with known viscosities of 0.1, 0.35, 1, and 12.5 Pa-s at 20°C; (2) standard fluid asphalt (also Newtonian) supplied by the manufacturer (Bohlin Instruments Inc.) with a known viscosity of 35 Pa-s at 20°C. The measured viscosities showed only a small deviation from the stated viscosities (< 6%) in all cases. Fig. 3 shows the typical results for silicone (with a standard viscosity of 1 Pa-s).

Throughout the experiments, samples in the rheometer were maintained at a temperature of 37°C ± 0.1°C by a Bohlin Temperature Control Unit which circulated distilled water into the mounting of the lower plate. Hydration levels of vocal fold tissue samples were also kept relatively constant by bathing them in physiological (0.9%) saline solution during the experiments.

Linearity of Oscillation

In order to ensure stress-strain linearity, small-amplitude oscillation was maintained by setting the shear strain amplitude (γ_0) at a value small enough for the linear constitutive equation [Eq. (2)] to hold. Recall from Section II that linearity is characterized by a proportionality between the shear amplitudes, and a phase shift being independent of the shear amplitudes (Ferry, 1970; Whorlow, 1980; Fung, 1993). That is, viscoelastic shear properties are independent of the strain amplitude in the linear region, when shear strain is small. The target shear strain amplitude was located by strain sweep tests done prior to frequency sweep experiments, where the amplitude of shear strain was slowly varied and the viscoelastic responses (μ and η) were obtained as a function of strain amplitude at selected frequencies.

Fig. 4 shows typical results of a strain sweep test on a sample of human vocal fold mucosa (from Subject 8). It can be seen that the elastic shear modulus μ was fairly independent of the shear strain amplitude γ_0 at small strain, up to an amplitude of around 0.01-0.05 rad. Therefore a

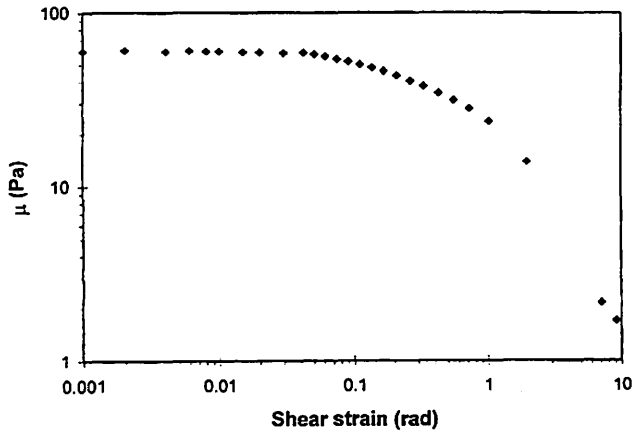


Figure 4. Elastic shear modulus μ of human vocal fold mucosa as a function of shear strain amplitude γ_0 (60 y.o. male); frequency = 1 Hz.

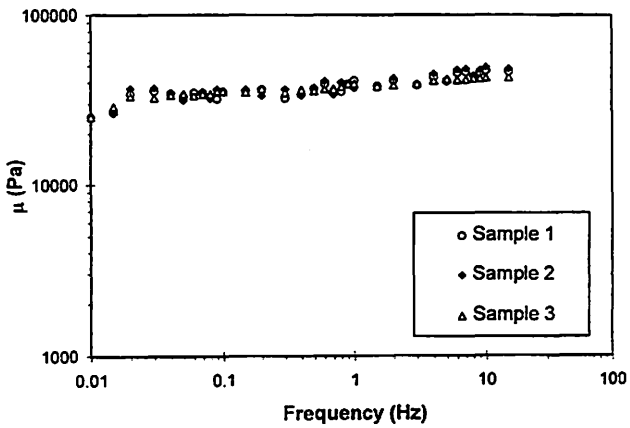


Figure 5. Elastic shear modulus μ of three similar Teflon samples as a function of frequency. Discrepancies between the samples represent an estimation of the magnitude of experimental error.

target strain amplitude of 0.01 rad was set for all small-amplitude frequency sweep experiments. It was maintained by the controlling software through automatic adjustments of the shear stress amplitude at each oscillation frequency.

Another feature of the software that helped to ensure linearity was its automatic introduction of delay time periods up to 100 seconds prior to the calculations of shear stress and strain rate functions at each frequency of oscillation. This delay period guaranteed that measurements were made from steady-state sinusoidal oscillations, rather than from onset transients.

Sample Inertia

An assumption underlying the linear viscoelastic theory is that the inertia of the sample is negligible in oscillatory shear deformation. It requires that the sample thickness be small compared with the wavelength of shear waves propagated through the medium at the frequency of oscillation (Ferry, 1970). This is called the "gap loading" condition, as the driving upper plate would "feel" like loading an empty gap with no sample inertia. This condition is achieved when there are negligible inertial forces in the sample during oscillation. According to Ferry (1970), inertial forces are small when

$$h << \sqrt{\frac{2\pi\eta}{\rho f}} \qquad (19)$$

where h is sample thickness (in m), ρ is density of the sample (in kg/m^3), and f is frequency of oscillation (in Hz). Thus a thick sample (a large gap size), a low sample viscosity, or a high frequency oscillation could all invalidate the assumption of negligible sample inertial effects. Assuming a sample dynamic viscosity of 0.5 Pa-s, which was approximately the smallest viscosity value observed

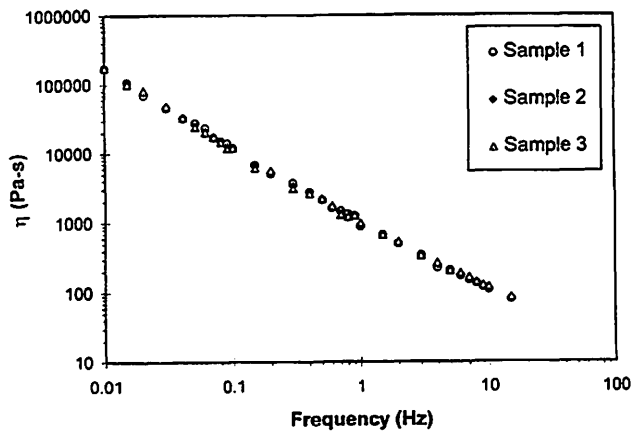


Figure 6. Dynamic viscosity η of three similar Teflon samples as a function of frequency. Discrepancies between the samples represent an estimation of the magnitude of experimental error.

for the average vocal fold mucosa, a sample density of 1100 kg/m^3 (vocal fold tissue density reported by Perlman, 1985), and an oscillation frequency of 15 Hz (close to the highest possible frequency allowed by the rheometer), the sample thickness (h) should be much smaller than 13.8 mm in order for inertial effects to become negligible. For the 0.3-mm gap size of the rheometer, it was about 46 times smaller than 13.8 mm. In order not to significantly violate the assumption of negligible inertial effects, it was arbitrarily decided that the cut-off ratio would be 50 times. With this cut-off, meaningful measurements were obtainable at frequencies up to 10 Hz, with measurements at 15 Hz considered as marginally acceptable.

Estimation of Experimental Error

Experimental error associated with the rheological experimental procedure was estimated by mounting

and testing similar material samples for more than one trial. Figs. 5-6 show the elastic shear modulus and the dynamic viscosity, respectively, of three similar samples of Teflon (polytetrafluoroethylene) as a function of oscillation frequency, plotted on a log-log scale. Because the Teflon samples came from the same source (the same tube ready for vocal fold injection, from Mentor Inc., Hingham, MA), discrepancies in results between the samples reflected an estimation of the repeatability or experimental error of the oscillatory shear experiments. Possible sources of errors included differences in the process of sample mounting (e.g., incomplete sample filling between the plates, excessive removal of excess sample material) and errors of the rheometer (e.g., inertia of the rotating plate and the sample, errors in rotor torque output, errors in measurement of angular velocity). According to the largest magnitude of differences observed in Figs. 5-6, experimental error associated with the measurement procedure was estimated to be within 10% of the measured viscoelastic data values across all frequencies.

Repeated trials of experiments were also performed on samples of human vocal fold mucosa. Any differences in shear properties between the trials would partly reflect the above experimental error, and partly reflect the effect of the rheological testing procedure *itself* on shear properties of tissue samples. Following exactly the same experimental procedure as described above, two trials of measurements were made on each of several samples of human vocal fold mucosa, separated by a duration of 30 minutes during which the tissue samples were kept in saline solution at room temperature. Figs. 7-8 show the effect of this repeated testing on the elastic shear modulus and the dynamic viscosity, respectively, of a mucosal tissue sample from Subject 15. It is clear that shear properties of the sample remained

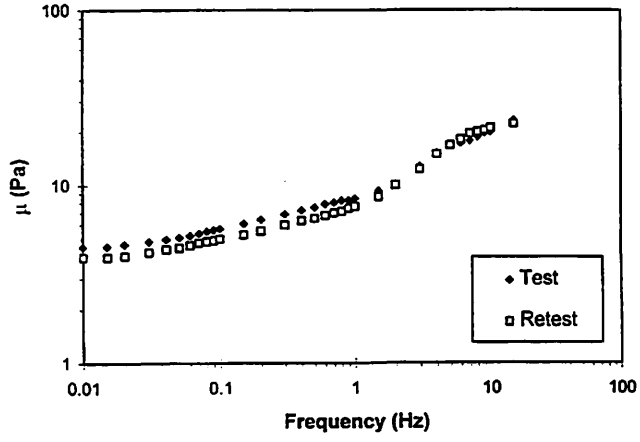


Figure 7. Effect of rheological testing on elastic shear modulus μ of human vocal fold mucosa (86 y.o. female).

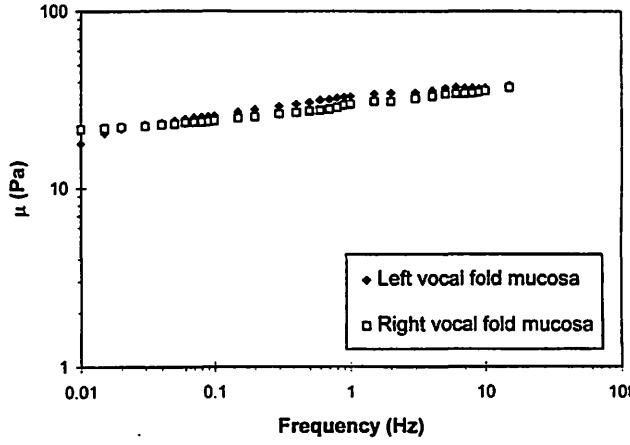


Figure 9. Left-right symmetry in elastic shear modulus μ of human vocal fold mucosa (71 y.o. female).

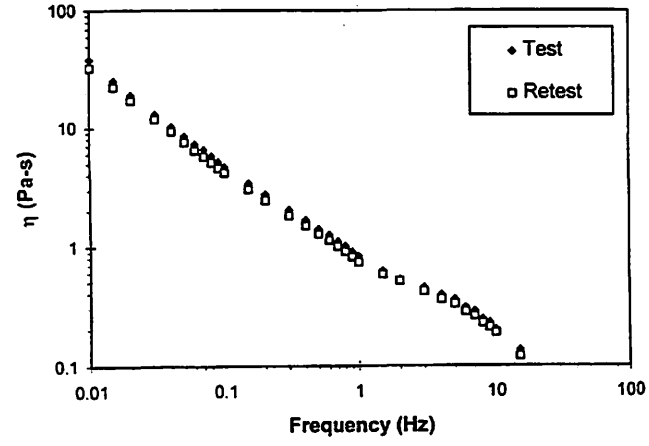


Figure 8. Effect of rheological testing on dynamic viscosity η of human vocal fold mucosa (86 y.o. female).

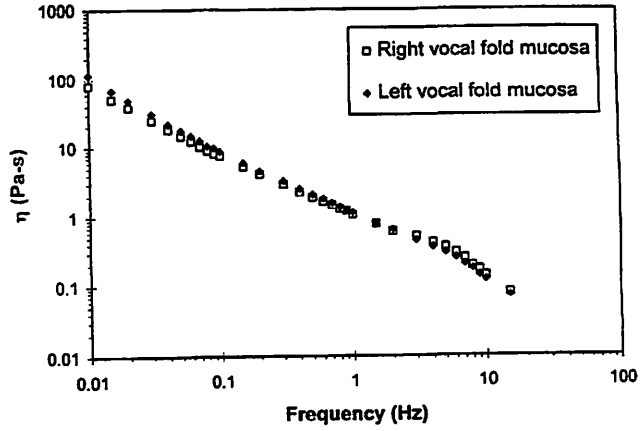


Figure 10. Left-right symmetry in dynamic viscosity η of human vocal fold mucosa (71 y.o. female).

basically the same in the two trials, with differences between the trials falling within the range of experimental error at most frequencies. This finding was also consistently observed with other test samples. Hence the experimental procedure itself did not significantly affect the shear properties of the tissue samples.

Results and Discussion

Symmetry in Shear Properties

Figures 9-10 show the elastic shear modulus μ and the dynamic viscosity η, respectively, of human vocal fold mucosal tissue samples obtained from the left and the right vocal folds of the same larynx (from Subject 14), as a function of oscillation frequency on a log-log scale. Logarithmic scales are used so that decades of changes in shear properties and in frequency could be clearly shown. Results showed that very similar values of μ and η were obtained for the samples from the two vocal folds. The same finding was also observed in other subjects, where measurements were made on both vocal folds. Hence it was concluded that there were no indications of mechanical asymmetry in the shear properties between mucosal tissues of contralateral vocal folds. This finding was not surprising, given that symmetric or near-symmetric vocal fold vibration is commonly observed in normal human subjects, and that not much asymmetry in vocal fold histology and molecular constituents has been observed in the literature.

Shear Properties of Human Vocal Fold Mucosa

Figures 11 to 16 show the viscoelastic shear properties of human vocal fold mucosal tissues from all subjects, including the elastic shear modulus μ (Figs. 11-12), the dynamic viscosity η (Figs. 13-14), and the damping ratio ζ (Figs. 15-16). Data for male and female subjects are shown in separate figures, and they are again plotted as a function of oscillation frequency on a log-log scale.

The elastic shear modulus μ was a relatively flat function of frequency and showed a very similar pattern of slope changes across the frequency range observed (Figs. 11-12). For male subjects, it remained basically constant at low frequency but increased slowly with frequency at higher frequencies (> 1 Hz) (Fig. 11). An exception was the 72-year-old male, whose μ was increasing with frequency across the entire frequency range and thus had a slightly greater overall slope. Another exception was the 30-year-old male, who did not show the slow increase in μ at frequencies above 1 Hz. For females, μ was relatively flat at low frequency (< 1 Hz) (although it was less flat than the male data), but again it was increasing slowly with frequency at higher frequencies (Fig. 12). Despite these similarities across different subjects, there were huge intersubject differences in the magnitude or numerical value of the elastic shear modulus, where differences were sometimes as large as one or two orders of magnitude. For instance, the male data ranged from approximately 10 to 100 Pa at low frequency and from approximately 40 to 1000 Pa at relatively high frequency (15 Hz) (Fig. 11), while for female subjects they ranged from approximately 3 to 20 Pa at low frequency and from approximately 10 to 40 Pa at high frequency (15 Hz) (Fig. 12). A closer look at these data showed that *in general* the male vocal fold mucosa was more elastic (stiffer) than the female vocal fold mucosa, and that older subjects seemed to have a more elastic vocal fold mucosa than younger subjects (see next section for further discussion).

The dynamic viscosity η was a monotonically decreasing function of frequency for all subjects (Figs. 13-

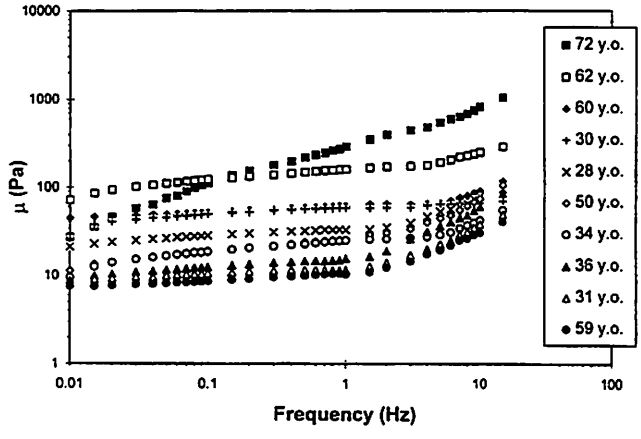


Figure 11. Elastic shear modulus μ of human vocal fold mucosa as a function of frequency (male subjects).

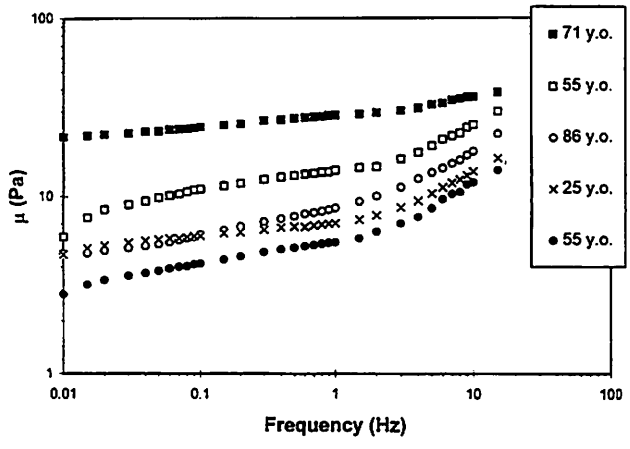


Figure 12. Elastic shear modulus μ of human vocal fold mucosa as a function of frequency (female subjects).

14), a *shear thinning* effect commonly observed in biological soft tissues and polymeric materials (Ferry, 1970; Fung, 1993). The decrease with frequency was approximately linear on the log-log scale, which indicated that the relationship between dynamic viscosity and frequency could be modeled by a power law:

$$\eta = k f^n \tag{20}$$

or

$$\log \eta = \log k + n \log f \tag{21}$$

where f is frequency in Hz, k and n are constants. Data were averaged across male and female subjects and then fitted to Eq. (21) by simple linear least-squares regression. Results are summarized in Table 2, where values of the coefficient k and the exponential index n of the power law are given. The square of the correlation coefficient r (the coefficient of determination) is also given as an estimate of the goodness-of-fit. It can be seen that data are very well matched by the regression equation, yielding correlation coefficients above 0.99. Dynamic viscosity decreased with frequency at approximately the same rate for different subjects (with similar values of the slope n), but there was considerable difference between the subjects in vertical separation between the curves (Figs. 13-14). Similar to μ, huge intersubject differences in η as large as an order of magnitude were again evident for most subjects, with older and/or male subjects generally showing a more viscous vocal fold mucosa than younger and/or female subjects.

Table 2.
Results of Linear Least-Squares Regression Curve-Fitting for the Dynamic Viscosity (η) of Human Vocal Fold Mucosa According to Eq. (21).

Tissue	k	n	r^2
Male vocal fold mucosa (N=10)	3.1440	-0.8755	0.999
Female vocal fold mucosa (N=5)	0.7038	-0.8345	0.996

The damping ratio ζ was basically a relatively flat function of frequency for most subjects (Figs. 15-16), which has been commonly observed in other biological soft tissues as well (Fung, 1993). Most subjects showed a consistent pattern of changes in damping ratio across the frequency range, with ζ decreasing very slightly with frequency at low frequencies (< 0.1 Hz), remaining constant or increasing very slightly with frequency at intermediate frequencies (0.1 to 3-5 Hz), and decreasing more rapidly at relatively high frequency (> 3-5 Hz). Two exceptions to this pattern were the 72-year-old male, whose ζ was basically a monotonically decreasing function with a slightly more negative overall slope (Fig. 15), and the 86-year-old female, whose ζ was slightly increasing with frequency across most of the frequency range (Fig. 16). Regarding intersubject differences, it was interesting to note that there was an apparent "convergence" of the data or a relatively small range of ζ for different subjects, despite huge (orders of magnitude) differences observed in the other shear properties (μ and η). Comparing to the data of μ and η, ζ was within a very small range of values for most subjects, approximately 0.2-0.5 at low frequency and 0.1-0.3 at higher frequencies (10-15 Hz) (except the 72-year-old male and the 86-year-old female). These damping ratio values were close to those reported previously by Kaneko *et al.* (1972) and Isshiki (1977), whose measurements were based on whole vocal folds

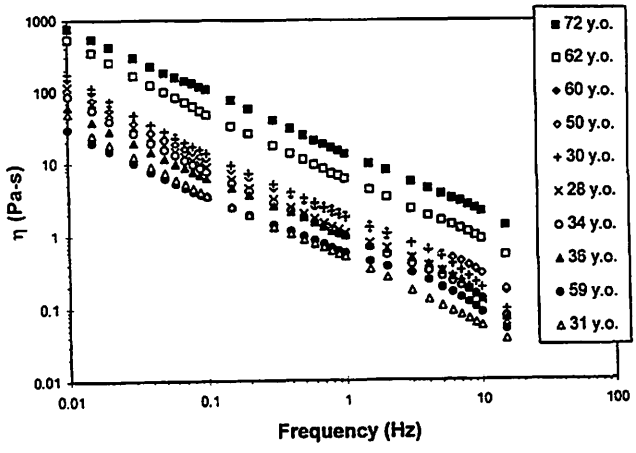


Figure 13. Dynamic viscosity η of human vocal fold mucosa as a function of frequency (male subjects).

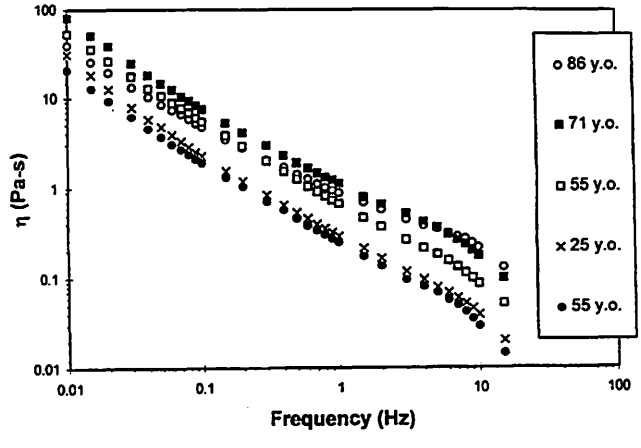


Figure 14. Dynamic viscosity η of human vocal fold mucosa as a function of frequency (female subjects).

excited quite differently (by mechanical impulses) at quite different frequencies (30-40 Hz and 130 Hz). These results suggested that vocal fold oscillation would remain underdamped ($\zeta < 1.0$) for most subjects despite the fact that their vocal fold mucosal tissues were often very different in elasticity and viscosity.

The Effects of Age and Sex

Some of the large intersubject differences in the elastic shear modulus and the dynamic viscosity of human vocal fold mucosa may be attributed to the effects of subject age and sex. On any given figure, subjects at the upper end of the range often showed a shear property three or four times, or even an order of magnitude larger than those at the lower end of the range. For example, at higher frequencies (10-15 Hz) where extrapolations to audio frequencies might be possible, the 60-year-old male and the 50-year-old male had an elastic shear modulus μ approximately three times larger than that of the 31-year-old and the 59-year-old (Fig. 11), while μ was an order of magnitude above for the 72-year-old and the 62-year-old. Comparably large differences were also observed for η and in female subjects.

Part of these intersubject differences may simply be normal intersubject variability, but part of them may also reflect an apparent age effect. For males, older subjects generally showed larger μ and η than younger subjects at *higher* frequencies (10-15 Hz), especially when ages at or above 50 were compared with ages below 50 (Figs. 11 and 13). An obvious exception to this trend was the 59-year-old male, who often showed the smallest μ and η among all male subjects. For females, there was also a rather clear age effect for η (except one of the 55-year-old female), but the trend was less clear for μ (Figs. 12 and 14).

In summary, it appeared that the viscoelastic shear properties of human vocal fold mucosa generally increased with age. Older subjects seemed to show a more elastic (stiffer) and more viscous vocal fold mucosa, especially for male subjects. The few exceptions to this trend could reflect the dependence of tissue shear properties on *physiological age* rather than chronological age of the subjects. This is only speculative, however, as data on physiological fitness and general health of the subjects were not collected (e.g., cardiovascular and pulmonary data).

There was also an apparent sex difference. The values for μ and η at higher frequencies (10-15 Hz), as well as their *average* data across frequencies for males were often three to five times larger than those for females at *comparable ages* (Figs. 11-14). Sometimes the difference between the average female and subjects at the upper end of the male range became as large as an order of magnitude. For example, at relatively high frequency (> 1 Hz), μ of the 72-year-old male and the 62-year-old male was an order of magnitude above those of the female subjects (Figs. 11-12). For η there was a comparably large difference across the entire frequency range.

Rheologists generally believe that viscoelastic shear properties are macroscopic reflections and realizations of microscopic and molecular events (Ferry, 1970; Bird *et al.*, 1977; Whorlow, 1980; Barnes *et al.*, 1989). Therefore, an interpretation of the above empirical findings with a molecular approach is given next.

Molecular Interpretations of Tissue Shear Properties

Recall that for a viscoelastic material, the elastic shear modulus μ is a quantification of the energy storage in the material. A highly elastic material is capable of near-

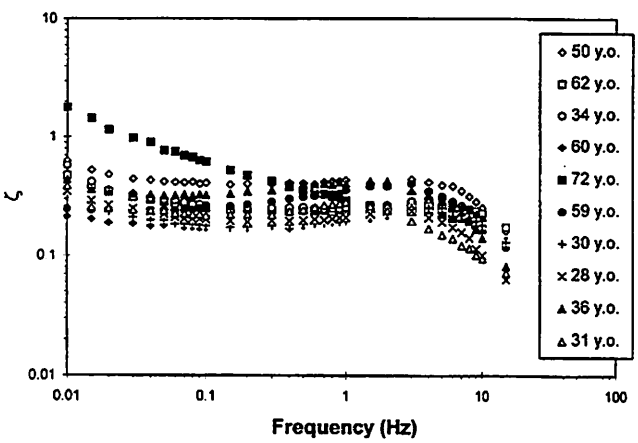


Figure 15. Damping ratio ζ of human vocal fold mucosa as a function of frequency (male subjects).

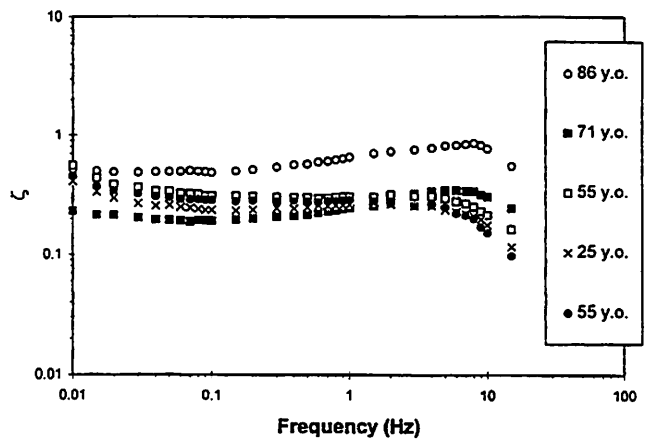


Figure 16. Damping ratio ζ of human vocal fold mucosa as a function of frequency (female subjects).

complete storage and recovery of internal energy during deformation. The dynamic viscosity η, on the other hand, is associated with dissipation of internal energy, typically as heat. A highly viscous material flows slowly and dissipates more energy in the flow process.

At the microscopic level, shear elasticity and viscosity of a material depend on the ease of relative displacement and slippage between (or within) molecules, which is determined by the number and strength of different kinds of *intramolecular* and *intermolecular* interactions. Some of the most significant molecular interactions include physical entanglement, electrostatic forces (e.g., H bond), hydrophilic and hydrophobic interactions (of polar and nonpolar chemical functional groups), and other physio-chemical interactions (e.g., covalent cross-link formation) (Bird *et al.*, 1977; Barnes *et al.*, 1989; Lehninger *et al.*, 1993). The more frequent and the stronger the molecular interactions, the more resistant the molecules are to relative displacement and slippage, and the higher the elasticity and viscosity.

These interactions are especially important in polymeric materials, where long chains of macromolecules occupy a large space relative to their atomic dimensions. For example, more densely packed polymeric chains have a higher probability of physical overlap and entanglements, chemical cross-link formation, and hence stronger electrostatic, hydrophilic and hydrophobic interactions. Longer and more highly charged polymeric chains are also more likely to have physical entanglements, chemical cross-links, and more significant electrostatic interactions. Therefore *concentration, molecular weight* and *molecular structure* of the macromolecules are all important in dictating molecular interactions and subsequently material elasticity and viscosity (Ferry, 1970; Barnes *et al.*, 1989). As a variety of biomacromolecules are found in the vocal fold lamina propria, possible differences at these molecular levels could account for the apparent age and sex effects in tissue shear properties found in the present study.

Vocal fold molecular composition has been investigated to some depth. Cellular structures found in the lamina propria include mainly fibroblasts and macrophages, and some myofibroblasts (for tissue repair and wound healing), while the extracellular matrix of lamina propria consists of biomacromolecules like fibrous proteins (collagen and elastin), glycosaminoglycans (e.g., hyaluronate, keratan sulfate, heparan sulfate, chondroitin sulfate and dermatan sulfate), proteoglycans (glycosaminoglycans covalently attached to a protein core, e.g., versican, decorin and fibromodulin), and structural glycoproteins (e.g., fibronectin) (Hirano, 1981; Matsuo *et al.*, 1984; Gray *et al.*, 1993; Pawlak *et al.*, 1996; Hammond *et al.*, 1997, 1998; Catten *et al.*, 1998). As cellular structures are relatively sparse in the lamina propria, including the superficial layer (vocal fold mucosa), the extracellular matrix probably plays an important role in metabolism, as well as in dictating tissue shear properties. Among the many constitutents found in the extracellular matrix, the fibrous proteins collagen and elastin are important mechanically because they form a scaffolding meshwork structure which supports interactions with other macromolecules in the lamina propria. For the superficial layer, non-fibrous or interstitial proteins like hyaluronate (physiological form of hyaluronic acid at a *p*H of 7.0-7.4), proteoglycans and structural glycoproteins are probably more important as fibrous proteins are sparsely found (Hirano, 1981; Gray *et al.*, 1993; Hammond *et al.*, 1997, 1998). They are important mechanically because of their huge molecular weights (often on the order of 10^5-10^6), high (negative) charge densities and capabilities of extensive molecular interactions (Lehninger *et al.*, 1993; Pawlak *et al.*, 1996).

Collagen

In the aging literature, it is well known that the triple helical polypeptide chains of collagen become increasingly cross-linked (covalently) as well as increasingly resistant to enzymatic (proteolytic) activity of collagenase (which tends to decrease) in old age. These findings have been repeatedly observed in different connective tissues across species (Sell and Monnier, 1995). An increase in covalent cross-linking would lead to an increase in the number and strength of molecular interactions among collagen fibers, which would in turn lead to an increase in elasticity and viscosity of aging tissues. Also, increased resistance to proteolysis and decreased collagenase activity imply a slower collagen turnover and a gradual accumulation of collagen and fragments of polypeptide chains, which could lead to higher collagen concentration and an increase in tissue elasticity and viscosity in old age. One other age-related change of collagen that has been observed is an increase of collagen concentration in the human vitreous humor after 50 years of age (Sebag, 1987). It may have been due to the decrease in collagen turnover discussed above. If such age-related changes were also true for collagen in the vocal fold mucosa, they may contribute to an age-related increase in tissue shear properties, especially for male subjects where there was a rather clear difference between ages above and below 50.

Elastin

Fragmentation, degeneration and loss of elastin fibers have been common findings in aging of human aorta and skin, which are probably the results of a decrease in resistance to proteolysis as well as increased enzymatic (elastase) activity (Sell and Monnier, 1995). A reduction in the biosynthesis of tropoelastin molecules has also been reported in old aged human skin fibroblasts, which would lead to decreased elastin content (Sephel and Davidson,

1986). In addition, there is also evidence for an age-related increase of covalent cross-links among elastin fibers in aorta, as well as cross-links between elastin and other proteins like collagen (Fujimoto, 1982; Powell *et al.*, 1992). Coupled with a simultaneous increase in relative collagen content (concentration) in old age, these findings suggest that the mechanical properties of these tissues tend to change towards more collagen-like and less elastin-like, including typically a decrease in extensibility (ultimate deformation) and an increase in elastic modulus or stiffness.

Some of these age-related changes of elastin have also been reported for human vocal fold tissues. Similar abnormalities, atrophy and loss of elastin fibers have been found in the lamina propria, especially for male subjects (Hirano *et al.*, 1989). Hence the relative changes in collagen content versus elastin content described above are also possible in the old vocal fold mucosa. Such age-related molecular changes could be responsible for an increase in stiffness and viscosity of the vocal fold mucosa with age, as they tend to transform the lamina propria mechanically into more collagen-like and less elastin-like.

Using elastin-van Gieson staining (EVG or Verhoeff's elastic tissue stain) and electron microscopy, Hammond *et al.* (1997) quantified the relative amount of elastin across different layers of the lamina propria in male and female subjects. Ultrastructural observations using electron microscopy revealed substantial quantities of relatively immature forms of elastin fibers in the superficial layer, including oxytalan (a fibrillar form) and elaunin (an amorphous form mixed with fibrillar components). They found that the relative amount of elastin varied between male and female. Higher levels of overall elastin content in older subjects have also been observed recently (Hammond *et al.*, 1998). These findings suggest that sex- and age-related differences in elastin concentrations could contribute to the differences in tissue shear properties reported in the present study.

Hyaluronic Acid

Age-related changes of hyaluronate reported in the aging literature have been commonly obtained from connective tissues like human articular cartilages and intervertebral disks. One consistent change of hyaluronate as part of an aggregating collagen-proteoglycan network structure has been the selective proteolytic breakdown of hyaluronate chains with higher molecular weights (on the order of 10^6). The consequences of such degradations include a gradual decrease of the average molecular weight of hyaluronate chains (from > 10^6 at age 20 to approximately 5×10^5 at age 80), a slow accumulation of partially degraded hyaluronate (with lower molecular weights), and a progressive increase of hyaluronate content or concentration (approximately a two-fold change from age 40 to age 80), the last of which may also be a consequence of increased biosynthesis (Holmes *et al.*, 1988; Sell and Monnier, 1995).

Possible effects of these hyaluronate changes on tissue shear properties depend on the balance between the effects of concentration and the effects of molecular size. An increase in hyaluronate content or concentration would facilitate molecular interactions and lead to an increase in elasticity and viscosity, while a decrease in the molecular weight of hyaluronate would tend to reduce molecular interactions and hence a decrease in shear properties. If the facilitative effects on molecular interactions due to higher concentrations do outweigh the effects of decreased molecular weights, these age-related changes of hyaluronic acid likely contribute significantly to an increase in tissue elasticity and viscosity with age, because the hyaluronate chains have relatively huge molecular weights, high charge densities and thus much potential for molecular interactions.

Hammond *et al.* (1997) quantified the relative amount of hyaluronate in male versus female subjects by image analysis of histological sections of vocal fold lamina propria stained with acid mucopolysaccharide (AMP), with and without hyaluronidase treatment. They found that males had a relative hyaluronate content three times higher than that of females, which suggested a higher probability of molecular interactions and an increase in tissue elasticity and viscosity. Recall that μ and η for males were found to be three to five times higher than those for females at comparable ages. This molecular finding of sex-related difference in hyaluronate content is therefore very well correlated with the apparent sex difference in tissue shear properties found in the present study.

Proteoglycans

The role of proteoglycans in aging is not well understood (Sell and Monnier, 1995). Nonetheless, one of the often reported age-related changes has been the progressive decline in tissue hydration with age (e.g., in human articular cartilages and intervertebral disks), which suggests a reduction of water-imbibition, water-binding and swelling capabilities of the collagen-proteoglycan structural network (Theocharis, 1985). This reduction is likely caused by (1) structural changes in proteoglycans (e.g., decreased fixed charge density, decreased molecular size due to chain fragmentation) (Adams and Muir, 1976; Urban and McMullin, 1985) and (2) increased cross-linking of collagen which restricts water binding and swelling of proteoglycans in the molecular network (Thonar *et al.*, 1986). An increase in cross-linking between collagen and proteoglycan has also been observed, thereby reducing water affinity of the network structure (Sell and Monnier, 1995).

Most of these age-related changes of proteoglycans and the collagen-proteoglycan network structure tend to facilitate molecular interactions among the macromolecules,

except the decrease in molecular size of proteoglycans which tends to reduce molecular interactions. If similar changes were also to happen for proteoglycans in the superficial layer, particularly decorin which has been found abundantly (Pawlak *et al.*, 1996), they could contribute to an age-related increase in tissue shear properties.

In summary, the effects of age and sex on shear properties of vocal fold mucosal tissues observed in the present study seem to be correlated with many of the age- and sex-related molecular differences in the vocal fold lamina propria and other extracellular connective tissues reported in the literature. However, more data on the effects of age and sex on major molecular constituents of the superficial layer (e.g., proteoglycans) and their molecular interactions are certainly needed in order to establish clearer relationships between molecular and viscoelastic data.

Summary and Conclusion

Using a parallel-plate rotational rheometer, small-amplitude oscillatory shear experiments were performed to measure the viscoelastic shear properties of human vocal fold mucosal tissues (the superficial layer of lamina propria). Elastic shear modulus, dynamic viscosity and damping ratio of tissue samples from 15 excised human larynges (ten male and five female) were quantified as a function of oscillation frequency (0.01 Hz to 15 Hz). Intersubject differences in elastic shear modulus and dynamic viscosity as large as an order of magnitude were found. Part of this large intersubject variability may be attributed to age differences. Tissue samples for older subjects were generally stiffer and more viscous than those of younger subjects. There was also an apparent sex effect, stiffer and more viscous mucosal tissues being observed for male subjects. These findings may be interpreted in terms of the age- and gender-related molecular differences in the vocal fold lamina propria and other extracellular connective tissues reported in the literature.

The major limitation of the present investigation was that rheological data were obtained at or below 15 Hz. Because of limitations of the rheometer's plate and sample inertia, the highest frequency at which linear small-amplitude oscillations were possible and meaningful measurements could be made was 15 Hz. As this was approximately an order of magnitude below audio frequencies or typical frequencies of vocal fold oscillation (usually >100 Hz), extrapolations of the data to audio frequencies are needed in order that the shear properties become relevant to the context of vocal fold oscillation. Theoretical modeling of tissue shear properties helpful to such data extrapolations will be addressed in a follow-up paper (Chan and Titze, 1998b).

Another limitation was that measurements of tissue shear properties were based on the tissue samples of a relatively limited number of subjects (ten males and five females). Although age- and gender-related differences in shear properties were observed, further studies with an improved experimental design are needed to specifically assess age and sex effects. For example, age effect could be assessed by comparing the data of different age groups (e.g., with ages <40, 40-60, >60), each of which has a reasonable number of subjects of the same sex. On the other hand, sex effect may be assessed by comparing the data of two sex groups with age-matched subjects.

Acknowledgments

This study was supported by a grant (No. P60 DC00976) from the National Institute on Deafness and Other Communication Disorders, a division of the National Institutes of Health. The authors gratefully acknowledge the contributions of Dr. Steven D. Gray of the University of Utah, who kindly supplied us some of the larynges and useful dissection instruments. He also made valuable suggestions and comments on the medical and biological aspects of the project. We also thank Dr. Brad H. Story for his contributions in rheological measurements.

References

Adams, P., and Muir, H. (1976). Qualitative changes with age of proteoglycans of human lumbar discs. *Ann Rheumatoid Disord, 35,* 289-296.

Alipour-Haghighi, F., and Titze, I.R. (1991). Elastic models of vocal fold tissues. *J Acoust Soc Am, 90,* 1326-1331.

Arnold, G.E. (1963). Alleviation of aphonia or dysphonia through intracordal injection of Teflon paste. *Ann Otol Rhinol Laryngol, 72,* 384-395.

Baer, T. (1975). *Investigation of phonation using excised larynges.* Unpublished Ph.D. dissertation, MIT, Cambridge, MA.

Barnes, H.A., Hutton, J.F., and Walters, K. (1989). *An introduction to rheology.* Amsterdam, The Netherlands: Elsevier, pp. 1-10; 37-54; 97-114.

Benninger, M.S., Alessi, D., Archer, S., Bastian, R., Ford, C., Koufman, J., Sataloff, R.T., and Spiegel, J.R. (1996). Vocal fold scarring: Current concepts and management. *Otolaryngol - Head Neck Surg, 115,* 474-482.

Berry, D.A., and Titze, I.R. (1996). Normal modes in a continuum model of vocal fold tissues. *J Acoust Soc Am, 100,* 3345-3354.

Berry, D.A., Moon, J.B., and Kuehn, D.P. (1998). A histologically-based finite element model of the soft palate. *Cleft Palate-Craniofacial J* (in press).

Bird, R.B., Armstrong, R.C., and Hassager, O. (1977). *Dynamics of polymeric liquids. Vol.1: Fluid mechanics.* New York: John Wiley & Sons, pp. 129-204; 275-303.

Catten, M., Gray, S.D., Hammond, T.H., Zhou, R., and Hammond, E.H. (1998). An analysis of cellular location and concentration in vocal fold lamina propria. *Otolaryngol - Head Neck Surg* (in press).

Chan, R.W. (1998). *Shear properties of vocal fold mucosal tissues and their effect on vocal fold oscillation.* Unpublished Ph.D. dissertation, The University of Iowa, Iowa City, IA.

Chan, R.W. and Titze, I.R. (1998a). Viscosities of implantable biomaterials in vocal fold augmentation surgery. *Laryngoscope, 108,* 725-731.

Chan, R.W. and Titze, I.R. (1998b). Theoretical modeling of viscoelastic shear properties of human vocal fold mucosa. *J Acoust Soc Am,* manuscript in preparation.

Courey, M.S., and Ossoff, R.H. (1995). Surgical management of benign voice disorders. In J.S. Rubin, R.T. Sataloff, G.S. Korovin, W.J. Gould (eds.)*Diagnosis and treatment of voice disorders.* New York: Igaku-Shoin, pp. 366-382.

Ferry, J.D. (1970). *Viscoelastic properties of polymers.* 2nd edn. New York: Wiley, pp.1-33.

Ford, C.N., Bless, D.M., and Loftus, J.M. (1992). Role of injectable collagen in the treatment of glottic insufficiency: a study of 119 patients. *Ann Otol Rhinol Laryngol, 101,* 237-247.

Fujimoto, D. (1982). Aging and crosslinking in human aorta. *Biochemical and Biophysical Research Communication, 109,* 1264-1269.

Fukuda, H., Kawasaki, Y., Kawaida, M., Shiotani, A., Oki, K., Tsuzuki, T., Fujioka, T., and Takayama, E. (1991). Physiological properties and wave motion of the vocal fold membrane viewed from different directions. In J. Gauffin and B. Hammarberg (eds.) *Vocal fold physiology: Acoustic, perceptual, and physiological aspects of voice mechanisms.* San Diego: Singular Publishing Group, pp. 7-14.

Fung, Y.C. (1993). *Biomechanics. Mechanical properties of living tissues.* (2nd edition). New York: Springer-Verlag, pp. 23-65; 242-320.

Gray, S.D., Hirano, M., and Sato, K. (1993). Molecular and cellular structure of vocal fold tissue. In I.R. Titze (ed.) *Vocal fold physiology: Frontiers in basic science.* San Diego: Singular Publishing Group, pp. 1-35.

Hammond, T.H., Gray, S.D., Butler, J., Zhou, R., Hammond, E. (1998). Age- and gender-related elastin distribution changes in human vocal folds. *Otolaryngol - Head Neck Surg, 119,* 314-322.

Hammond, T.H., Zhou, R., Hammond, E.H., Pawlak, A., and Gray, S.D. (1997). The intermediate layer: A morphologic study of the elastin and hyaluronic acid constituents of normal human vocal folds. *J Voice, 11,* 59-66.

Hirano, M. (1975). Phonosurgery: Basic and clinical investigations.*Otologia (Fukuoka), 21,* 239-440.

Hirano, M. (1981). Structure of the vocal fold in normal and disease states anatomical and physical studies. In C.L. Ludlow, and M.O. Hard (eds.) *Proceedings of the conference on the assessment of vocal pathology.* Rockville, MD: American-Speech-Language-Hearing Association, pp. 11-30.

Hirano, M., Kurita, S., Sakaguchi, S. (1989). Aging of the vibratory tissue of human vocal folds. *Acta Otolaryngologica, 107,* 428-433.

Holmes, M.W.A., Bayliss, M.T., and Muir, H. (1988). Hyaluronic acid in human cartilage: Age-related changes in content and size. *Biochemical J, 250,* 435-441.

Ishizaka, K., and Flanagan, J.L. (1972). Synthesis of voiced sounds from a two-mass model of the vocal cords. *Bell System Tech J, 51,* 1233-1268.

Isshiki, N. (1977). *Functional Surgery of the Larynx.* Kyoto, Japan: Kyoto University, pp.62-67.

Kaneko, T., Asano, H., Naito, J., Kobayashi, N., Hayashi, K., and Kitamura, T. (1972). Biomechanics of the vocal cords - On damping ratio. *J Japan Soc Bronchoesophagol, 25,* 133-138.

Lehninger, A.L., Nelson, D.L., and Cox, M.M. (1993). *Principles of biochemistry.* (2nd edition). New York: Worth Publishers, pp. 298-321.

Matsuo, K., Watanabe, S., Hirano, M., Kamimura, M., Tanaka, Y., and Takazono, I. (1984). Acid mucopolysaccharide and glycoprotein in the vocal fold: Alterations with aging. *Practica Otolaryngologica Kyoto, 77,* 817-822.

Min, Y.B., Titze, I.R., and Alipour-Haghighi, F. (1995). Stress-strain response of the human vocal ligament. *Ann Otol Rhinol Laryngol, 104,* 563-569.

Pawlak, A.S., Hammond, T.H., Hammond, E.H., and Gray, S.D. (1996). Immunocytochemical study of proteoglycans in vocal folds. *Ann Otol Rhinol Laryngol, 105,* 6-11.

Perlman, A.L. (1985). A technique for measuring the elastic properties of vocal fold tissue. Unpublished Ph.D. dissertation, The University of Iowa, Iowa City, IA.

Powell, J.T., Vine, N., and Crossman, M. (1992). On the accumulation of D-aspartate in elastin and other proteins of the ageing aorta. *Atherosclerosis, 97,* 201-208.

Saito, S., Fukuda, H., Kitahara, S., Isogai, Y., Tsuzuki, T., Muta, H., Takayama, E., Fujioka, T., Kokawa, M., and Makino, K. (1985). Pellet tracking in the vocal fold while phonating - Experimental study using canine larynges with muscle activity. In I.R. Titze and R.C. Scherer (eds.) *Vocal Fold Physiology.* Denver, CO: Denver Center for the Performing Arts, pp. 169-182.

Sataloff, R.T., Spiegel, J.R., Hawkshaw, M.J., Rosen, D.C., and Heuer, R.J. (1997). Autologous fat implantation for vocal fold scar: A preliminary report. *J Voice, 11,* 238-246.

Schramm, V. L., May, M., and Lavorato, A. S. (1978). Gelfoam paste injection for vocal fold paralysis: Temporary rehabilitation of glottic incompetence. *Laryngoscope, 88,* 1268-1273.

Sebag, J. (1987). Ageing of the vitreous. *Eye, 1,* 254-262.

Sell, D.R., and Monnier, V.M. (1995). Aging of long-lived proteins: extracellular matrix (collagens, elastins, proteoglycans) and lens crystallins. In E.J. Masoro (ed.) *Handbook of physiology. Section 11: Aging.* New York: Oxford University Press, pp. 235-305.

Sephel, G.C., and Davidson, J.M. (1986). Elastin production in human skin fibroblast cultures and its decline with age. *J Investigative Dermatol, 86,* 279-285.

Shaw, G.Y., Szewczyk, M.A., Searle, J., and Woodroof, J. (1997). Autologous fat injection into the vocal folds: Technical considerations and long-term follow-up. *Laryngoscope, 107,* 177-186.

Theocharis, D.A. (1985). Comparisons between extracted and residual proteoglycans on the glycosaminoglycan level and changes with ageing. *International J Biochem, 17,* 155-160.

Thonar, E.J.-M.A., Bjornsson, S., and Kuettner, K.E. (1986). Age-related changes in cartilage proteoglycans. In K. Kuettner, R. Schleyerbach and V.C. Hascall (eds.) *Articular cartilage biochemistry.* New York: Raven Press, pp. 273-288.

Titze, I.R. (1976). On the mechanics of vocal-fold vibration. *J Acoust Soc Am, 60,* 1366-1380.

Titze, I.R., and Talkin, D.T. (1979). A theoretical study of the effects of various laryngeal configurations on the acoustics of phonation. *J Acoust Soc Am, 66,* 60-74.

Urban, J.P.G., and McMullin, J.F. (1985). Swelling pressure of the intervertebral disc: Influence of proteoglycan and collagen contents. *Biorheology, 22,* 145-157.

Whorlow, R.W. (1980). *Rheological techniques.* Chichester, West Sussex, England: Ellis Horwood, pp. 243-307.

Wilhelms-Tricarico, R. (1995). Physiological modeling of speech production: Methods for modeling soft-tissue articulators. *J Acoust Soc Am, 97,* 3085-3098.

NCVS Status and Progress Report - 13
April 1999, 37-52

Theoretical Modeling of Viscoelastic Shear Properties of Human Vocal Fold Mucosa

Roger Chan, Ph.D.
Ingo Titze, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa

Abstract

The viscoelastic shear properties of human vocal fold mucosa were previously measured as a function of frequency [Chan and Titze, J. Acoust. Soc. Am., in review], but data were obtained only in a frequency range of 0.01-15 Hz, an order of magnitude below typical frequencies of vocal fold oscillation (on the order of 100 Hz). This study represents an attempt to extrapolate the data to higher frequencies based on two viscoelastic theories, (1) a quasi-linear viscoelastic theory widely used for the modeling of viscoelastic properties of biological tissues [Fung, *Biomechanics* (Springer-Verlag, New York, 1993), pp. 277-292], and (2) a molecular (statistical network) theory commonly used for the viscoelastic modeling of polymeric materials [Zhu *et al.*, *J. Biomechanics* **24**, 1007-1018 (1991)]. Analytical expressions of elastic and viscous shear moduli, dynamic viscosity, and damping ratio based on the two theories with specific model parameters were applied to curve-fit the empirical data. Results showed that the theoretical predictions matched the empirical data reasonably well, allowing for parametric descriptions and extrapolations of the data to audio frequencies.

Introduction

Previous studies have shown that the viscoelastic shear properties of the vocal fold mucosa are not only important for the theoretical study of voice production, but also for the surgical management of vocal fold mucosal disorders (Chan, 1998; Chan and Titze, 1998, 1999a, 1999b). Specifically, data on the elastic properties (elastic shear moduli) and viscous properties (viscous shear modulus and dynamic viscosity) of the vocal fold mucosa and other tissue layers are needed for the computer simulation of vocal fold vibration (Titze and Talkin, 1979; Berry and Titze, 1996). Viscous properties of the mucosa are also important for determining phonation threshold pressure, an objective measure of the "ease" of phonation (Titze, 1988). Clinically, a low phonation threshold pressure following a surgical procedure is an indication of surgical success in terms of restoring normal vocal fold oscillation (Chan and Titze, 1998, 1999a).

In our previous study, a parallel-plate rotational rheometer was employed to deform tissue samples of the vocal fold mucosa in oscillatory (dynamic) shear (Chan and Titze, 1999b). Data on elastic shear modulus, dynamic viscosity and damping ratio were quantified as a function of oscillation frequency. However, dynamic shear data were only obtained at relatively low frequencies (0.01-15 Hz) because of limitations of the rheometer's plate inertia and sample inertia. As this was approximately an order of magnitude below typical frequencies of vocal fold oscillation (usually >100 Hz), extrapolations of the data to audio frequencies are needed in order that the data become relevant for vocal fold oscillation.

In addition to data extrapolation, parametric descriptions of tissue shear properties are possible with theoretical modeling of the viscoelastic data. If dynamic shear data of the vocal fold mucosa agree with theoretical predictions based on a model with a specific set of parameters, viscoelastic behaviors of the vocal fold mucosa can be efficiently represented by the model's constitutive equations and the particular set of model parameters. Dynamic shear data can be conveniently reproduced, while other viscoelastic data (e.g., stress relaxation and strain creep data) may also be predicted from the constitutive equations and the model parameters.

This paper attempts to describe the viscoelastic shear properties of human vocal fold mucosa with two commonly used modeling approaches, namely a quasi-linear viscoelastic model and a molecular model (statistical network model). The quasi-linear viscoelastic theory proposed by Fung (1972, 1981, 1993) has been successfully applied to describe the viscoelastic properties of many biological tissues, including the shear properties of soft tissues. The statistical network theory (the revised version of Zhu *et al.*, 1988, 1991) has been extensively applied to describe the rheological properties of polymeric materials, as well as soft tissues like extracellular connective tissues. Prior to a discussion of these theories, however, it is helpful to first review the traditional approach to viscoelastic modeling.

Traditional Linear Viscoelastic Models

Traditional models for viscoelasticity are one-dimensional lumped-element models made of basic elastic and viscous elements, namely linear springs and dashpots (or dampers), respectively. Different combinations of springs and dashpots have been used to model different time-dependent and strain history-dependent viscoelastic behaviors, e.g., hysteresis, stress relaxation, and strain creep. The simplest combinations are the Maxwell model (a spring and a dashpot connected in series), the Voigt model (or Kelvin-Voigt model; a spring in parallel with a dashpot), and the standard linear model (or Kelvin model; a Maxwell model in parallel with a spring; see Fig. 1) (Ferry, 1980; Barnes *et al.*, 1989; Fung, 1993). Many other variations have also been developed by adding more springs and dashpots in series and in parallel.

Consider the standard linear model (Kelvin model) shown in Fig. 1. An external force F is applied to the model to produce an elongation ψ, or alternatively, an elongation ψ is applied to generate a force F. The series spring (k_s) and the parallel spring (k_p) are the elastic elements, while the dashpot (η_d) is the viscous element. Assuming that inertial forces are negligible (in slowly varying deformations), two force equilibrium equations may be written using the free body concept at node A and at node B, respectively:

$$F = k_p \psi + k_s (\psi - \psi_i) \tag{1}$$

$$\eta_d \dot{\psi}_i = k_s (\psi - \psi_i) \tag{2}$$

where F, ψ and ψ_i are functions of time t, ψ_i is the internal elongation at node B, and the "dot over" notation represents time differentiation. Solving for ψ_i in Eq. (1) and then substituting it into Eq. (2), the governing differential equation for the Kelvin model can be written as

$$\frac{\eta_d}{k_s}\dot{F} + F = \eta_d \frac{k_p + k_s}{k_s}\dot{\psi} + k_p \psi \tag{3}$$

This governing equation can also be expressed in terms of a series time constant τ_s and a parallel time constant τ_p:

$$\tau_s \dot{F} + F = k_p (\tau_p \dot{\psi} + \psi) \tag{4}$$

where τ_s and τ_p are defined in terms of the spring constants k_s and k_p, and the dashpot viscosity η_d:

$$\tau_s = \frac{\eta_d}{k_s}, \qquad \tau_p = \eta_d \left(\frac{k_p + k_s}{k_p k_s}\right) \tag{5}$$

Consider the viscoelastic behavior of *stress relaxation*, where a step increase of elongation (or deformation in general) is applied to a viscoelastic sample described by the Kelvin model. As a result, there is an instantaneous increase of force (and hence stress) due to the elastic component of the model, followed by a time-dependent reduction of the initial force as dictated by a relaxation time constant. This gradual reduction of force (or stress) as a function of time can be quantified by solving for F in Eq. (4), with a step input of elongation:

$$\begin{aligned} \psi(t) &= 0 \qquad \text{when } t < 0 \\ &= \psi_0 \qquad \text{when } t \geq 0, \end{aligned} \tag{6}$$

where ψ_0 is a constant. With this step elongation, the derivative of the elongation $\dot{\psi} = 0$ and the governing differential equation [Eq. (4)] becomes

$$\dot{F} + \frac{1}{\tau_s} F = \frac{k_p}{\tau_s} \psi_0 \tag{7}$$

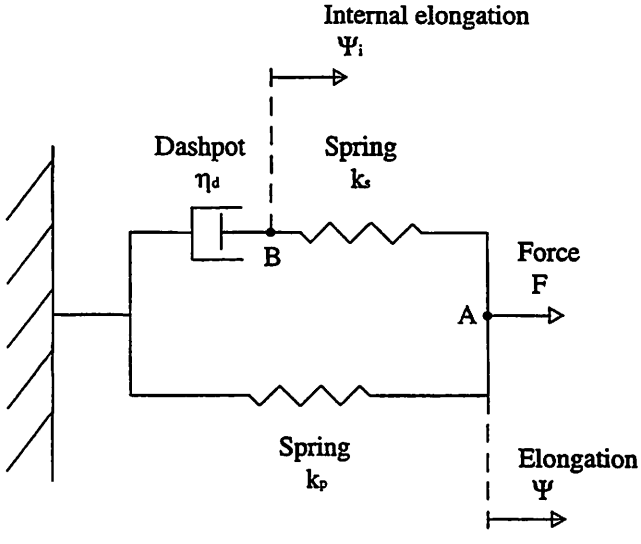


Figure 1. The standard linear model (Kelvin model) of viscoelasticity.

This is a first-order linear differential equation in $F(t)$, with the forcing function being a constant. The solution for this type of equation may be obtained by the method of Laplace transform, or by solving for the homogeneous solution (the general solution of the corresponding homogeneous equation) and a particular solution satisfying the equation. The homogeneous solution for Eq. (7) is $C_1 e^{-\frac{t}{\tau_s}}$, where C_1 is a constant, while a particular solution is $k_p\psi_0$. Thus the general solution is

$$F = k_p \psi_0 + C_1 e^{-\frac{t}{\tau_s}} \tag{8}$$

C_1 can be solved with the knowledge of the initial condition for Eq. (7), which corresponds to the instantaneous force generated by the step increase of elongation at $t = 0$:

$$F(0) = (k_p + k_s)\,\psi_0 \tag{9}$$

This initial force arises purely from the elastic component (springs) of the model as the viscous component (dashpot) cannot respond instantaneously. Substituting Eq. (9) into Eq. (8) at $t = 0$, we obtain $C_1 = k_s\psi_0$. Hence the solution for Eq. (7) is

$$F(t) = \left[k_p + k_s\, e^{-\frac{t}{\tau_s}}\right]\psi_0 \tag{10}$$

or, in terms of the two time constants [c.f. Eq. (5)],

$$F(t) = k_p\left[1 + (\frac{\tau_p}{\tau_s} - 1)\, e^{-\frac{t}{\tau_s}}\right]\psi_0 \tag{11}$$

Thus the Kelvin model relaxes from a force of $F = (k_p + k_s)\psi_0$ at $t = 0$ exponentially to $F = k_p\psi_0$ at $t = \infty$, as dictated by the series time constant τ_s which is the *relaxation time constant*. Eqs. (10) and (11) describe the force-elongation relationship of the model in relaxation, which can be transformed to a stress-strain relationship in order to formally describe stress relaxation. It is done by defining a stress σ_k from the force F, a strain ε_k from the elongation ψ, and an elastic modulus E from the parallel spring constant k_p:

$$\sigma_k = \frac{F}{A}\,,$$

$$\varepsilon_k = \frac{\psi}{L_0} \qquad (\text{hence } \varepsilon_0 = \frac{\psi_0}{L_0}\,), \tag{12}$$

$$\text{and} \qquad E = \frac{k_p L_0}{A}$$

where A is the cross-sectional area and L_0 is the reference length of the sample described by the model. With these definitions, the governing differential equation for the Kelvin model [Eq. (4)] becomes the constitutive equation:

$$\tau_s \dot{\sigma}_k + \sigma_k = E\,(\tau_p\, \dot{\varepsilon}_k + \varepsilon_k) \tag{13}$$

and Eq. (11) becomes the stress-strain relationship characterizing stress relaxation:

$$\sigma_k(t) = E\left[1 + (\frac{\tau_p}{\tau_s} - 1)\, e^{-\frac{t}{\tau_s}}\right]\varepsilon_0 \tag{14}$$

A *relaxation function* $k(t)$ can be defined accordingly:

$$k(t) = \frac{\sigma_k(t)}{\varepsilon_0} = E\left[1 + (\frac{\tau_p}{\tau_s} - 1)\, e^{-\frac{t}{\tau_s}}\right] \tag{15}$$

It is clear that the relaxation function of the Kelvin model is an exponential function with a relaxation time constant τ_s. In fact, the relaxation functions of linear viscoelastic models are always composed of exponential terms, each with a time constant τ_i. By adding more springs and dashpots to a model, more derivative terms are introduced into its constitutive equation, yielding more exponential terms in the relaxation function. In general, the relaxation function can be expressed as a sum of N exponential terms with N characteristic time constants, each associated with a characteristic coefficient (Fung, 1993):

$$k(t) = \sum_{i=1}^{N} a_i\, e^{-\frac{t}{\tau_i}} \tag{16}$$

where τ_i are the characteristic time constants and a_i are the characteristic coefficients associated with τ_i. A series of

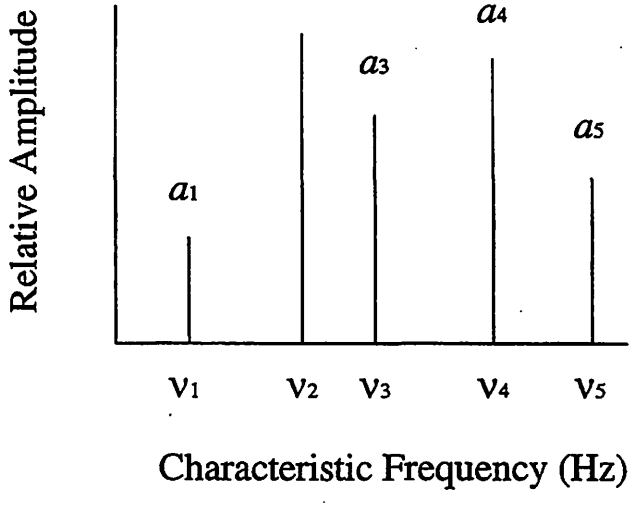


Figure 2. A discrete spectrum of the relaxation function (after Fung, 1993).

characteristic frequencies υ_i can be defined as the inverse of the characteristic time constants τ_i. The discrete function $a_i(\upsilon_i)$ is called the spectrum of the relaxation function, or *relaxation spectrum*, which is a density or weighting function of τ_i. This spectrum is *discrete* for models with a finite number of exponential terms, or a finite number of mechanical elements (see Fig. 2 for an example).

A discrete relaxation spectrum is not adequate for the description of viscoelastic properties of many biological soft tissues (Fung, 1993). In particular, traditional linear models with discrete relaxation spectra are associated with a finite number of peaks (at the characteristic frequencies) in their frequency-dependent damping characteristics. However, relatively flat or frequency-insensitive damping characteristics have been commonly observed for biological soft tissues, which can only be accurately described with a *continuous* relaxation spectrum (Fung, 1993). Continuous relaxation spectra have been implemented in the contexts of the quasi-linear viscoelastic theory and the statistical network theory, which are summarized next.

The Quasi-Linear Viscoelastic Theory

Fung (1972, 1981, 1993) proposed a quasi-linear viscoelastic theory (QLV) to describe commonly observed nonlinear viscoelastic properties of biological tissues, including nonlinear stress-strain characteristics, differences in stress response between loading and unloading (hysteresis phenomena), and the general dependence of stress on time and strain history. Many biomechanics researchers have successfully applied the theory to model the time-dependent and strain history-dependent viscoelastic properties of various animal and human soft tissues in the passive state (i.e., relaxed state for contractile tissues), including tendons (Woo *et al.*, 1982, 1993; Chun and Hubbard, 1986), ligaments (Woo *et al.*, 1981, 1982, 1993), mesentery (Fung, 1972), aorta (Tanaka and Fung, 1974), cardiac muscles (Pinto and Fung, 1973; Pinto and Patitucci, 1980), smooth muscles (Price *et al.*, 1979), spinal cords (Bilston and Thibault, 1996), and articular cartilages (Woo *et al.*, 1980; Simon *et al.*, 1984; Zhu *et al.*, 1986; Spirt *et al.*, 1989). Fung first developed the theory in the context of stress relaxation, but formulations of the theory can be generalized for other modes of deformation, including oscillatory shear deformation. For instance, the theory has been used for the modeling of viscoelastic shear properties of articular cartilage (e.g., Mak, 1986; Zhu *et al.*, 1986; Spirt *et al.*, 1989). Let us first review Fung's basic formulations of the theory in the context of stress relaxation, with the use of a continuous relaxation spectrum. Applications of the theory to oscillatory shear deformation is then discussed.

Formulation of the Theory

Consider a viscoelastic tissue sample subject to an ideal step increase of strain, with an infinite strain rate at time $t = 0$:

$$\begin{aligned} \varepsilon(t) &= 0 \quad \text{when } t < 0 \\ &= \varepsilon_0 \quad \text{when } t \geq 0 \end{aligned} \tag{17}$$

where $\varepsilon(t)$ is strain and ε_0 is a constant. The fundamental assumption of the QLV is that the stress relaxation function K is both time-dependent and strain history-dependent and can be expressed as the product of a function of time and a function of strain:

$$K \equiv K[\varepsilon(t),t] = G(t)\sigma^e[\varepsilon(t)] \tag{18}$$

where $K[\varepsilon(t),t]$ is the relaxation function, $G(t)$ is the *reduced relaxation function*, which is a function of time only, and $\sigma^e[\varepsilon(t)]$ is the *elastic response* (the superscript "e" stands for "elastic"), which is defined as the *instantaneous* stress developed in the sample in response to the ideal step increase of strain. The reduced relaxation function $G(t)$ is defined as the time-dependent relaxation function normalized by the relaxation function at $t = 0^+$:

$$G(t) = \frac{K(t)}{K(0^+)}, \quad G(0^+) = 1 \tag{19}$$

This function is dimensionless and by definition dependent on time only. With this definition, the relaxation function K is equal to the elastic response $\sigma^e[\varepsilon]$ at $t = 0^+$, which is an instantaneous stress response due to the sample's elastic component [c.f. Eq. (18)].

Consider the strain history as a series of infinitesimal step increases in strain. According to the Boltzmann superposition principle of linearity, which states that the effects of sequential (step) changes in strain are additive, the stress contributions from all individual step strains at past times can be summed up to yield the stress at time t (Ferry, 1980; Bird *et al.*, 1977a; Fung, 1993). Therefore the stress σ at time t can be expressed in the form of a hereditary or convolution integral between the reduced relaxation function $G(t)$ and the rate of change of the elastic response $\sigma^e[\varepsilon]$:

$$\sigma(t) = \int_0^t G(t-\tau)\frac{\partial \sigma^e[\varepsilon]}{\partial \varepsilon}\frac{\partial \varepsilon}{\partial \tau}d\tau \tag{20}$$

where $\sigma(t)$ is the overall stress response (Lagrangian stress) at time t due to a history of strain $\varepsilon(t)$ (Green's strain), for a time history beginning at $t=0$. It is clear that this constitutive equation is mathematically similar to that of the generalized linear viscoelastic model, i.e., $G(t)$ and $\sigma^e[\varepsilon]$ are playing the roles of relaxation function and strain in linear viscoelasticity, respectively (Ferry, 1980; Bird *et al.*, 1977a):

$$\sigma(t) = \int_0^t k(t-\tau)\frac{d\varepsilon}{d\tau}d\tau \qquad (21)$$

These analogies explain why the theory is "quasi-linear", which allows the simple yet powerful mathematical tools of analysis for linear systems to be applicable, e.g., in the context of oscillatory shear deformation. Prior to such linear analysis, however, a continuous relaxation spectrum has to be implemented so that biological tissues may be adequately described.

Continuous Relaxation Spectrum

The damping characteristics of many biological soft tissues have been found to remain relatively insensitive to the rate of deformation (frequency or strain rate), often across several decades of variation (Fung, 1993). As discussed in Section II, this frequency insensitivity cannot readily be accounted for by the finite number of exponential terms in the relaxation functions of traditional linear viscoelastic models. Rather, a generalized model with an infinite number of springs and dashpots must be considered. Equivalently speaking, a *continuous* relaxation spectrum corresponding to an infinite number of exponential terms in the relaxation function should be used. In order to accommodate such a continuous relaxation spectrum in the context of the QLV, Fung (1972, 1981) derived the following *generalized* reduced relaxation function:

$$G(t) = \frac{1+\int_0^\infty S(\tau)e^{-\frac{t}{\tau}}d\tau}{1+\int_0^\infty S(\tau)d\tau} \qquad (22)$$

where $S(\tau)$ is the continuous relaxation spectrum, τ is a continuous variable playing the role of "time constants" in the exponential terms, while in fact there are no time constants in the continuous integrals. Note that $G(0^+) = 1$, fulfilling the definition of the reduced relaxation function in Eq. (19). A specific continuous relaxation spectrum $S(\tau)$ was proposed by Fung (1972, 1981) to account for frequency-insensitive damping:

$$\begin{aligned} S(\tau) &= \frac{c}{\tau} && \text{for } \tau_1 \le \tau \le \tau_2 \\ &= 0 && \text{elsewhere} \end{aligned} \qquad (23)$$

where c is a dimensionless constant. The constants c, τ_1, and τ_2 are the model parameters of the QLV, to be adjusted for the theory to match the empirical data. Fig. 3 shows a generic example of the continuous spectrum with $c = 0.1$, $\tau_1 = 0.01$ seconds, and $\tau_2 = 100$ seconds, which results in a rather flat damping for $\tau_1 \le 1/\omega \le \tau_2$ (Fung, 1993). It has been shown repeatedly that the viscoelastic properties of many biological soft tissues are closely described by the QLV using this spectrum (for references, see the paragraph at the beginning of this section). A similar spectrum had also been used for the descriptions of hysteresis and damping in other fields of scientific studies, e.g., in the descriptions of dielectricity (Wagner, 1913), metal electromagnetism (Becker and Doring, 1939; Neubert, 1963), and airplane flutter (Theodorsen and Garrick, 1940).

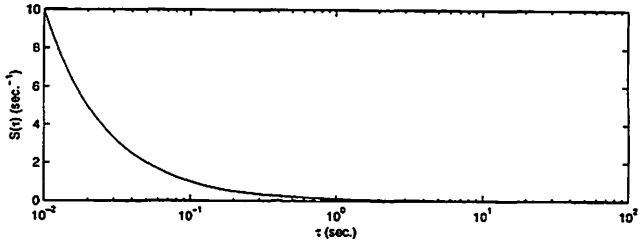


Figure 3. A continuous spectrum of the relaxation function [c.f. Eq. (23)].

Based on this specific continuous relaxation spectrum, the QLV can be applied to the experimental condition of oscillatory shear deformation, which is illustrated below.

Application to Oscillatory Shear Deformation

Consider the condition of oscillatory shear deformation, where the overall stress response $\sigma(t)$ and the elastic response $\sigma^e(t)$ can be expressed as complex harmonic functions of time:

$$\sigma^*(t) = \sigma_0 e^{i(\omega t+\delta)} \qquad (24)$$

$$\sigma^{e*}(t) = \sigma^e{}_0 e^{i\omega t} \qquad (25)$$

where σ_0 and $\sigma^e{}_0$ are amplitudes, i is the imaginary number $\sqrt{-1}$, ω is angular frequency of oscillation, and δ is the phase shift between the two complex quantities. Since the elastic response is playing the role of strain [Eqs. (20) and (21)], $\sigma^{e*}(t)$ can be regarded as the sinusoidal input to a linear viscoelastic system, yielding a sinusoidal stress output $\sigma^*(t)$ at the same oscillation frequency but shifted by a phase of δ. Therefore, the definition of the complex shear modulus as the ratio of complex shear stress to complex shear strain in the context of linear viscoelasticity can be applied here (Ferry, 1980; Fung, 1993; Chan and Titze, 1999b):

$$\mu_q^*(\omega) = \frac{\sigma^*}{\sigma^{e*}} \qquad (26)$$

where $\mu^*{}_q(\omega)$ is the complex shear modulus of elasticity (the subscript "q" stands for the QLV). On substituting the complex overall stress and complex elastic response [Eqs. (24) and (25)] into the constitutive equation [Eq. (20)],

$\mu^*_q(\omega)$ can be expressed in terms of the reduced relaxation function $G(t)$:

$$\mu^*_q(\omega) = \frac{\sigma_0}{\sigma^e_0} e^{i\delta} = i\omega \int_0^t G(t) e^{i\omega t}\, dt \qquad (27)$$

Using the generalized reduced relaxation function $G(t)$ and the continuous relaxation spectrum $S(\tau)$ described above [Eqs. (22) and (23)], and integrating from τ_1 to τ_2 with respect to τ, the final expression for the complex shear modulus is (Fung, 1981, 1993):

$$\mu^*_q(\omega) = \frac{1+\frac{c}{2}[\ln(1+\omega^2\tau_2^2)-\ln(1+\omega^2\tau_1^2)]+ic[\tan^{-1}(\omega\tau_2)-\tan^{-1}(\omega\tau_1)]}{1+c\ln\frac{\tau_2}{\tau_1}} \qquad (28)$$

By definition, the real part of $\mu^*_q(\omega)$ is the elastic shear modulus μ_q, while the imaginary part is the viscous shear modulus $\omega\eta_q$:

$$\mu_q = \frac{1+\frac{c}{2}[\ln(1+\omega^2\tau_2^2)-\ln(1+\omega^2\tau_1^2)]}{1+c\ln\frac{\tau_2}{\tau_1}} \qquad (29)$$

$$\omega\eta_q = \frac{c[\tan^{-1}(\omega\tau_2)-\tan^{-1}(\omega\tau_1)]}{1+c\ln\frac{\tau_2}{\tau_1}} \qquad (30)$$

Once μ_q and $\omega\eta_q$ are computed, the dynamic viscosity η_q and the damping ratio ζ_q can be readily derived ($\eta_q = \frac{\omega\eta_q}{\omega}$ and $\zeta_q = \frac{\omega\eta_q}{\mu_q}$). These analytical expressions represent theoretical predictions of the viscoelastic shear properties of biological soft tissues based on the QLV. The three model parameters c, τ_1, and τ_2 can be adjusted to curve-fit these predictions to the empirical data of different tissues, including those of the vocal fold mucosa.

According to a sensitivity analysis of the QLV, changes of the model parameters c, τ_1 and τ_2 are critically related to the stress relaxation function and the frequency-dependent damping ratio predicted by the theory (Sauren and Rousseau, 1983). Their results showed that the parameter c indicates the amount of viscous effects present with respect to elastic effects, such that larger values of c correspond to a higher damping ratio across a wide range of frequency (10^{-3}-10^3 Hz), a more "sharply tuned" damping ratio curve (i.e., more frequency-sensitive), and an increase in stress relaxation magnitude and rate. The other model parameters τ_1 and τ_2 are "time constants" which define the range of the continuous relaxation spectrum and hence the frequency-insensitive range of the damping curve. The smaller τ_1 governs the "fast viscous phenomenon", i.e., the high-frequency limit of the relatively flat range of the damping curve, analogous to the upper cut-off frequency of the frequency response of a band-pass filter. The larger τ_2 governs the "slow viscous phenomenon", i.e., the low-frequency limit of the relatively flat range of damping, analogous to a lower cut-off frequency (Sauren and Rousseau, 1983). A decrease in τ_1 or an increase in τ_2 tends to expand the "bandwidth" of the damping curve, making it more "broadly tuned" or frequency-insensitive.

Molecular Theories

The molecular structure of the vocal fold lamina propria has been studied considerably (see summary in Chan and Titze, 1998, 1999a). As most of the molecular constituents found in the lamina propria extracellular matrix are biomacromolecules (or biopolymers), it seems logical to apply rheological molecular theories to describe the viscoelasticity of the vocal fold mucosa.

Molecular theories attempt to model the rheological properties of macromolecular or polymeric materials based on their molecular dynamics and kinetics. The viscoelastic properties of polymeric materials and polymers in solutions arise from both intramolecular and intermolecular interactions among the macromolecules, as well as intermolecular interactions between the macromolecules and the solvent (usually water) molecules (Ferry, 1980; Bird *et al.*, 1977b). Intramolecularly, there is a resting *equilibrium* state associated with each macromolecule in which the orientations of chemical bond vectors are distributed randomly, resulting in a state of minimum internal energy (Bird *et al.*, 1977b; Barnes *et al.*, 1989). When the macromolecule is deformed by an external force (or stress), the orientations of chemical bond vectors are changed and elastic energy is stored in the new orientations. Elastic recovery represents reorientation of the bond vectors back to the equilibrium state. On the other hand, resistance to flow and dissipation of internal energy (viscous behaviors) are caused by a molecule's interactions with other molecules and solvent molecules, which introduce frictional drag on the molecules and delay the process of bond vector reorientation (Bird *et al.*, 1977b; Barnes *et al.* 1989). There are many types of molecular models, one of which is the transient network model or *statistical network model.*

Statistical Network Theory

The statistical network theory for the rheological modeling of polymeric fluids and materials has been developed extensively by Lodge (1968), Bird *et al.* (1977a, 1977b), De Kee and Carreau (1979), and Zhu and his coworkers (Zhu and Mow, 1990; Zhu *et al.*, 1988, 1991). Based on earlier analytical work by Lodge (1968) and De Kee and Carreau (1979), Zhu *et al.* (1988, 1991) described a network theory capable of relating the rheological properties of macromolecular materials to structural parameters defining a molecular network model of the material. In order to describe the transient non-Newtonian flow proper-

ties of concentrated macromolecular solutions not predicted by previous network theories (e.g., thixotropic responses[1] of proteoglycan solutions), Zhu *et al.* (1988, 1991) developed a non-linear second-order constitutive equation for the network theory. This updated and extended version of the statistical network theory has been shown to closely describe the viscoelastic properties of some biopolymeric materials, including proteoglycan solutions (Zhu and Mow, 1990; Zhu *et al.*, 1988, 1991) and collagen-proteoglycan mixtures (Zhu *et al.*, 1996).

The theory assumes that polymeric materials form a macromolecular *network* through intermolecular interactions of chains of macromolecules, particularly physical entanglement. The network is composed of *chain segments* (or network segments) and *junction sites* (or interaction sites). A chain segment is defined as the molecular chain between two adjacent junction sites. A junction site is defined as a common point where molecular chains are constrained to coincide and where forces can be transferred between molecules. It is assumed that these junction sites are confined to local interactions only and that they are constrained to move as part of the continuum of macromolecules. They are *transient* junctions but not permanent bonds. During deformation, both chain segments and junction sites are created and lost continuously. At any given time, the network consists of a distribution of chain segments created at different past times. This distribution is a function of the history of strain, or *memory* of the network. A memory function is incorporated into the model's constitutive equation and is defined in such a way that it depends on the rate of creation and the rate of loss of junction sites (or chain segments), as illustrated in the next section.

Constitutive Equation and Complex Shear Modulus

Fig. 4 shows the schematic of the molecular network model. Part (a) illustrates the connections between several chain segments and junction sites, while part (b) shows the details of one chain segment. The chain segment is modeled as a molecular chain which consists of N beads connected together with (N - 1) rigid rods of length L. Each chain segment can be described by a configuration vector $\mathbf{R}$, defined in terms of the junction position vectors $\mathbf{r}_i$ and $\mathbf{r}_j$, at the ith and jth junctions respectively (Fig. 4). According to statistical mechanics, a time-average force is developed in the chain segment because of random molecular (Brownian) motions and the constrains of the chain at the junction sites. The time-average force makes the chain segment behave like a Hookean spring with a spring constant H (Bird *et al.*, 1977b):

[1] Thixotropic materials show a gradual decrease of shear stress (and viscosity) *with time* under a constant strain rate, due to a breakdown of the molecular network during shearing (e.g., tomato ketchup).

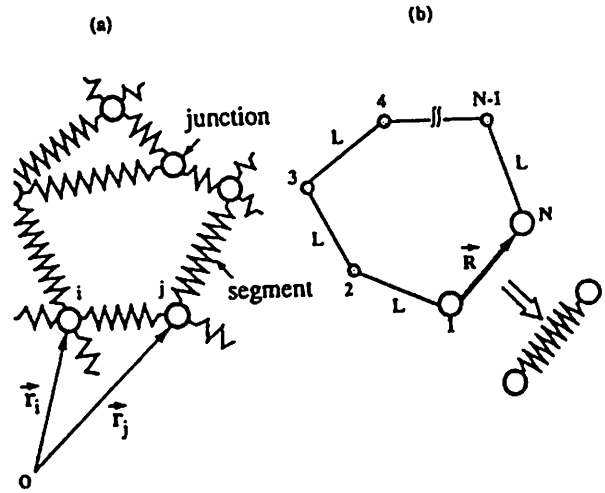


Figure 4. Schematic of the statistical network model (after Zhu et al., *1991).*

$$H = \frac{3kT}{(N-1)L^2} \tag{31}$$

where k is the Boltzmann constant ($= 1.3807 \times 10^{-23}$ JK^{-1}) and T is the absolute temperature (in Kelvin). Elastic energy can be stored in these network springs during deformation. At any given point in time, the network consists of a continuous distribution of chain segments with various configuration vectors $\mathbf{R}$, which depends on the creation and loss of individual chain segments at different past times, i.e., the strain history which is described by a *memory function*. The overall stress developed in the network is assumed to be the sum of individual stress contributions from all chain segments existing at the present time t and created at all past times t'. Using tensor notations, the second-order constitutive equation for the network developed by Zhu *et al.* (1988, 1991) is

$$\tau = \int_{-\infty}^{t} m(t, t')\Gamma(t, t')dt' \tag{32}$$

where τ is the extra stress tensor, $m(t, t')$ is the memory function, and $\Gamma(t, t')$ is the Finger strain tensor. Note that τ includes the stress contributions from all chain segments but not the stress contributions from the interstitial fluid. In fact, the total stress tensor is given by $-p\mathbf{I} + \tau$, where p is the interstitial fluid pressure associated with the fluid's incompressibility, and $\mathbf{I}$ is the identity tensor. The Finger strain tensor Γ is defined by

$$\Gamma = \mathbf{F}^{-1}(\mathbf{F}^{-1})^{T} - \mathbf{I} \tag{33}$$

where $\mathbf{F}$ is the deformation gradient relating a past junction position $\mathbf{r}_i'$ to the current junction position $\mathbf{r}_i$ ($\mathbf{r}_i = \mathbf{F}^{-1}\mathbf{r}_i'$). The memory function $m(t, t')$ in the constitutive equation is

expressed as a function of the creation rate and the loss rate of the chain segments, which are assumed to be continuous functions of the second invariant[2] of the first Rivlin-Ericksen tensor[3] (Π_1) as well as the second invariant of the second Rivlin-Ericksen tensor[3] (Π_2):

$$m(t, t') = \int_0^{\infty} C(\Pi_1, \Pi_2, s) e^{-\int_{t'}^{t} \frac{dt}{L(\Pi_1, \Pi_2, s)}} ds \qquad (34)$$

where C is the rate of creation and L is the rate of loss of chain segments (and junction sites), both continuous functions of Π_1 and Π_2. The second invariant of a tensor is an indication of the magnitude of the tensor, thus Π_1 quantifies the rate of deformation of the network, while Π_2 quantifies the accelaration of deformation. It becomes clear that the constitutive equation is indeed second-order, because of the dependence of $m(t, t')$ on both Π_1 and Π_2. s is called the "segment complexity", a spectral variable showing the diversity of chain segment types on a continuous scale, i.e., the distribution of the chain segment configuration vectors **R**. With the above constitutive equation and memory function, Zhu *et al.* (1988, 1991) obtained the following expressions for the real part and the imaginary part of the complex shear modulus μ^* under small-amplitude oscillatory shear conditions:

$$\mu(\omega) = \int_0^{\infty} \frac{\eta(s)\lambda(s)\omega^2 ds}{1+[\lambda(s)\omega]^2} \qquad (35)$$

$$\omega\eta(\omega) = \int_0^{\infty} \frac{\eta(s)\omega ds}{1+[\lambda(s)\omega]^2} \qquad (36)$$

where μ is the elastic shear modulus, $\omega\eta$ is the viscous shear modulus, $\lambda(s)$ is the network relaxation spectrum (with the dimension of time) and $\eta(s)$ is a function of $\lambda(s)$ (with the dimension of viscosity). They are given by

$$\lambda(s) = \frac{\lambda_0}{(1+s)^{\alpha}} \qquad (37)$$

$$\eta(s) = \frac{\eta_0 \lambda(s)}{\int_0^{\infty} \lambda(s) ds} \qquad (38)$$

where λ_0 is maximum relaxation time of the network, α is a spectral parameter greater than 1.0 (see below), and η_0 is zero shear-rate viscosity. Note that the network relaxation spectrum $\lambda(s)$ is *continuous*, as s is a continuous variable. This implementation of a continuous relaxation spectrum is similar to that in Fung's quasi-linear viscoelastic theory, as discussed in Section III. This similarity suggests that the statistical network theory is capable of the description of a frequency-insensitive damping in biological soft tissues, like the QLV. The spectral parameter α can be expressed as

$$\alpha = 1 + \frac{\ln(\eta_0/\eta_{\infty})}{\ln(1+s_0)} \qquad (39)$$

where η_{∞} is infinite shear-rate viscosity and s_0 is the maximum number of "segment complexity" which defines the range of chain segments under shear deformation, such that $0 < s \le s_0$.

In summary, four independent parameters (η_0, η_{∞}, λ_0, s_0) and one dependent parameter (α) are needed for the description of linear viscoelastic shear properties using the present version of the statistical network theory. Among the shear properties, the elastic shear modulus μ and the viscous shear modulus $\omega\eta$ are computed from Eqs. (35) and (36), while the dynamic viscosity η and the damping ratio ζ can be derived from μ and $\omega\eta$. The model parameters can be adjusted to curve-fit these theoretical predictions to the empirical data of specific polymeric materials and tissues, including vocal fold mucosal tissues. Knowledge of the parameters provides insights on how the network model of molecules might interact and contribute to deformation and shear properties. Specifically, the total number of chain segments per unit volume (number density of chain segments) in the network, $N(t)$, may be expressed as a function of time and strain history. When the molecular network is at rest (momentarily), the chain segments are randomly oriented and a maximum of $N(t)$ is maintained. The maximum number density of chain segments (N_0) corresponds to this peak of $N(t)$ and is related to the model parameters by

$$N_0 = \frac{\eta_0 s_0 (\alpha - 1)}{kT\lambda_0} \qquad (40)$$

This number denotes an estimation of the effective amount (or extent) of molecular interactions in a polymeric material (Zhu and Mow, 1990; Zhu *et al.*, 1991). In the next section, the viscoelastic data of human vocal fold mucosa reported previously (Chan and Titze, 1999b) are interpreted in terms of this index of molecular interactions, as well as the statistical network theory and the quasi-linear viscoelastic theory in general.

Results and Discussion

Quasi-Linear Viscoelastic Theory

Figs. 5-6 show the viscoelastic shear properties of human male vocal fold mucosa as a function of oscillation frequency, including the elastic shear modulus μ, the viscous shear modulus $\omega\eta$, the dynamic viscosity η and the damping ratio ζ. The data points (in circles) represent the average empirical data of ten male subjects, while the solid lines are the "best-fit" theoretical curves based on the QLV. Agreement between the empirical data and the theoretical predictions for *each* of the shear properties was assessed by

[2] The second invariant (Π) of a tensor A is defined as $\Pi = [(tr(A))^2 - tr(A^2)]/2$.

[3] The first Rivlin-Ericksen tensor $A_1 = L + L^T = 2D$, where L is the velocity gradient and D is the rate of strain or "stretching" tensor. The second Rivlin-Ericksen tensor $A_2 = \dot{L} + L^T + 2L^T L$.

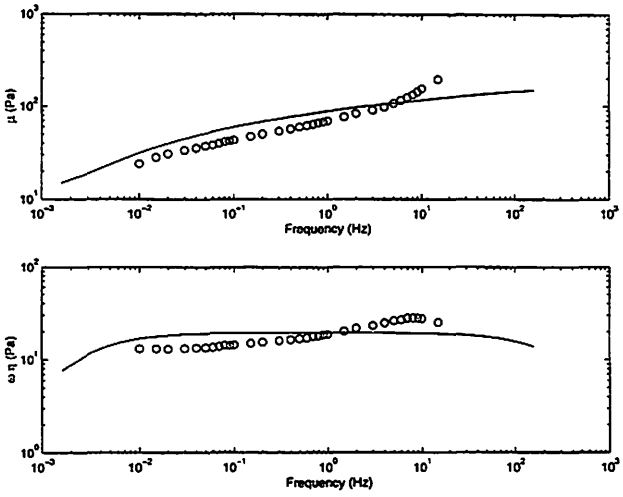


Figure 5. Curve-fitting of elastic shear modulus μ (upper graph) and viscous shear modulus ωη (lower graph) of human male vocal fold mucosa by the QLV.

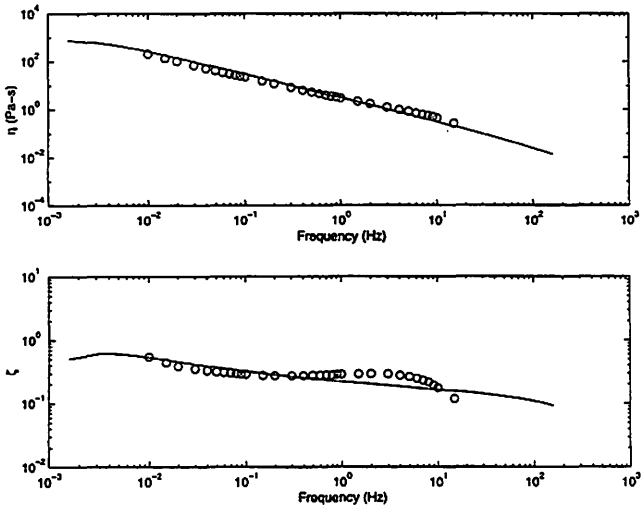


Figure 6. Curve-fitting of dynamic viscosity η (upper graph) and damping ratio ζ (lower graph) of human male vocal fold mucosa by the QLV.

computing a correlation coefficient r as the goodness-of-fit statistic:

$$r = \sqrt{1 - \frac{\sum_i (x_i - y_i)^2}{\sum_i (x_i - \bar{x})^2}} \qquad (41)$$

where x_i are the empirical data of one of the shear properties, y_i are the corresponding model predictions, and $\bar{x}$ is the average of x_i. Two steps were involved in the curve-fitting procedure: (1) Beginning with an initial set of parameters that correspond to the generic continuous relaxation spectrum shown in Fig. 3 ($c = 0.1$, $\tau_1 = 0.01$ seconds, $\tau_2 = 100$ seconds), an approximate graphical fit was found empirically by varying the model parameters (c, τ_1 and τ_2) individually; (2) Next, the "best fit" was found by fine-tuning the model parameters until the *sum* of the correlation coefficients (for different shear properties) was maximized. For the average data of ten human male subjects, the best fit was achieved with $c = 1.0$, $\tau_1 = 0.0005$ seconds, and $\tau_2 = 70$ seconds. Table I summarizes the correlations between theory and data using this set of model parameters.

Table I.
Correlation Coefficients of Curve-Fitting with QLV Modeling of Human Male Vocal Fold Mucosa (c.f. Figures 5-6)

Shear properties	r
μ	0.8584
$\omega\eta$	0.2453
η	0.9069
ζ	0.6596

Figs. 5-6 show that in general the theoretical curves of the QLV fit the empirical data quite well, with reasonably high correlation coefficients in three of the four shear properties, except the viscous shear modulus ωη (Table I). There were certain discrepancies between theory and data, however, especially at higher frequencies (around 10-15 Hz). For instance, the empirical elastic shear modulus data (μ) appeared to be increasing with frequency at a slightly larger slope when higher frequencies were approached, yet the QLV predicted a smooth curve increasing only slowly with frequency, without any significant change of slopes (Fig. 5). The result was an under-prediction of the data at high frequency. Predictions of the QLV did not match very well for ωη and ζ, again particularly at higher frequencies (Figs. 5-6), yielding correlation coefficients of only 0.2453 and 0.6596, respectively. The viscous shear modulus ωη was actually slightly increasing with frequency across the whole frequency range, as opposed to the very flat (frequency-independent) curve predicted by the QLV. This discrepancy was clearly shown by the low correlation coefficient (0.2453). Meanwhile, there was an excellent match between theory and data for η (Fig. 6), whose correlation coefficient was higher than 0.9.

Figs. 7-8 (following page) show the average empirical data for ten female subjects (in circles) and the best-fit theoretical predictions of μ, ωη, η and ζ based on the QLV. A similar curve-fitting procedure as described above was performed. The best fit was achieved with $c = 10$, $\tau_1 =$

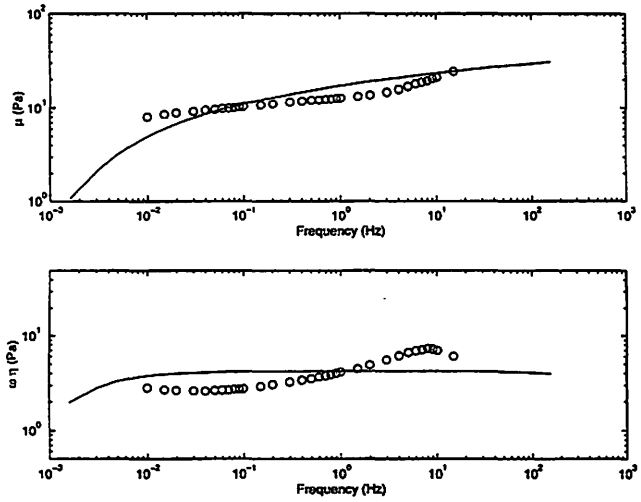


Figure 7. Curve-fitting of elastic shear modulus μ (upper graph) and viscous shear modulus ωη (lower graph) of human female vocal fold mucosa by the QLV.

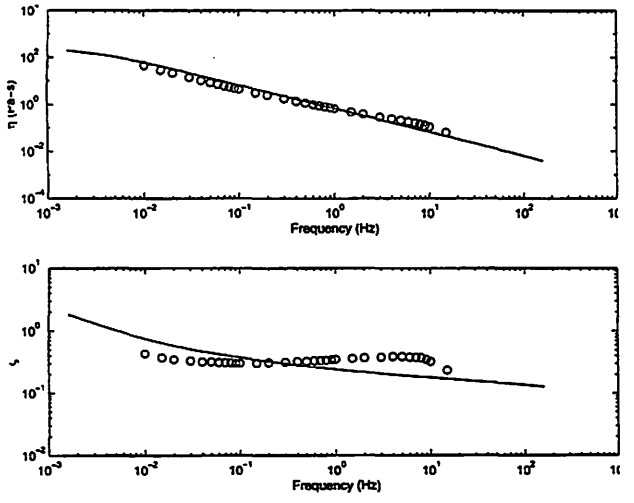


Figure 8. Curve-fitting of dynamic viscosity η (upper graph) and damping ratio ζ (lower graph) of human female vocal fold mucosa by the QLV.

0.0001 seconds, and τ_2 = 90 seconds. Table II shows the correlations between theory and data associated with this set of model parameters.

Comparing with the results for the male data, the theoretical curves of the QLV generally fit the female data *less* well, with lower correlation coefficients across all shear properties (Table II). Discrepancies between theory and data were quite significant for ωη and ζ, where only a general trend with an approximate slope was predicted in each case. The theory failed to closely predict the course of slope changes across the frequency range, yielding rather low correlation coefficients (0.2223 and 0.5951). Both ωη and ζ actually showed a "peak" at around 10 Hz, which was more

Table II.
Correlation Coefficients of Curve Fitting with QLV Modeling of Human Female Vocal Fold Mucosa (c.f. Figures 7-8)

Shear property	*r*
μ	0.6328
$\omega\eta$	0.2223
η	0.8587
ζ	0.5951

prominent than the "peak" observed in the male data and hence a larger deviation from the flat curves predicted by the QLV. On the other hand, there was a much better match between theory and data for η (Fig. 8) where the correlation coefficient was again the highest.

Differences in the model parameters for achieving best-fits for the male and the female data may be interpreted in terms of the findings of the sensitivity analysis of the QLV (Sauren and Rousseau, 1983). Recall that the parameter *c* indicates the relative amount of viscous effects, where a larger *c* corresponds to a higher damping ratio and a more sharply tuned damping curve (i.e., more frequency-sensitive). Also recall that the other parameters τ_1 and τ_2 define the range of the continuous relaxation spectrum and the relatively frequency-insensitive range of the damping curve. A decrease in τ_1 or an increase in τ_2 is associated with a broader or more flat damping curve. Our findings showed that a larger value of *c* was obtained for modeling the female data (c = 10) than for the male data (c=1.0), suggesting a higher and a more frequency-sensitive damping ratio for female than for male. On the other hand, the female data were modeled with a smaller τ_1 and a larger τ_2 (τ_1 = 0.0001 sec., τ_2 = 90 sec.) than the male (τ_1 = 0.0005 sec., τ_2 = 70 sec.), suggesting a less frequency-sensitive damping ratio curve. Thus it seemed that the model parameters of the QLV were apparently capable of offsetting the effects of one another on the frequency sensitivity of damping ratio, yielding a damping ratio with similar frequency dependence for both male and female. It also seemed that the QLV was capable of predicting a larger damping ratio for female based on the magnitude of the continuous relaxation spectrum, i.e., a larger parameter *c*. These theoretical predictions on the damping ratio of male and female vocal fold mucosal tissues indeed agreed with their empirical data, where (1) ζ was found to be larger for female at higher frequencies, and (2) there was a comparably flat range of ζ for both male and female data (c.f. Chan and Titze, 1999b).

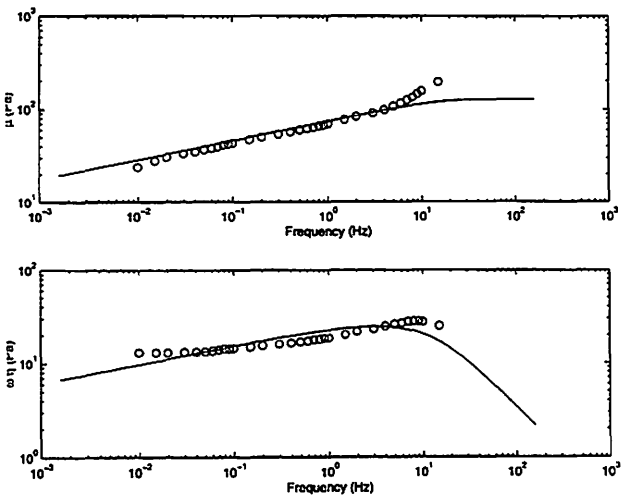


Figure 9. Curve-fitting of elastic shear modulus μ (upper graph) and viscous shear modulus ωη (lower graph) of male vocal fold mucosa by the statistical network model.

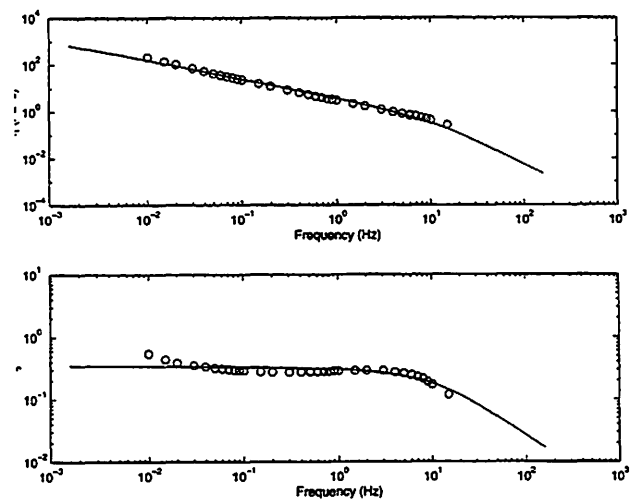


Figure 10. Curve-fitting of dynamic viscosity η (upper graph) and damping ratio ζ (lower graph) of male vocal fold mucosa by the statistical network model.

Table III.
Correlation Coefficients of Curve Fitting With Statistical Network Modeling of Human Vocal Fold Mucosa (c.f. Figures 9-10)

Shear property	r
μ	0.9042
$\omega\eta$	0.7547
η	0.9757
ζ	0.7265

Statistical Network Theory

Figs. 9-10 show the average empirical data for ten male subjects (in circles) and the theoretical predictions of viscoelastic shear properties (μ, ωη, η and ζ) based on the statistical network theory. The "best fit" between the theoretical curves and the empirical data was again determined empirically in a way similar to that described for the QLV, i.e., by adjusting the model parameters individually until the correlation coefficients were maximized. For the average data of ten human male subjects, the best fit was achieved with the four independent model parameters η_0=3.2×10^7 Pa-s, η_∞ = 0.32 Pa-s, λ_0 = 10^8 seconds, s_0 = 100, and with α = 4.99138. Table III summarizes the correlations between theory and data with a list of correlation coefficients for the shear properties.

It is observed that theoretical curves of the statistical network theory generally fit the empirical data quite well, with all correlation coefficients higher than 0.7 (Table III). Similar to the case of the QLV, there were some discrepancies between theory and data at higher frequencies (around 10-15 Hz), especially for μ and ωη. The empirical μ data appeared to be increasing with frequency at a slightly larger slope at high frequency than that at lower frequencies, yet a smooth curve was predicted with a slightly smaller slope at high frequency than that at lower frequencies (Fig. 9). The result was an under-prediction of the empirical data above 10 Hz. This relative disagreement between theory and data at higher frequencies was similar to the deviation observed for the QLV. Also similar to what was observed for the QLV, predictions of the statistical network theory did not match too well with the empirical ωη and ζ data, again particularly at higher frequencies (Figs. 9-10). However, curve-fitting based on the statistical network theory seemed to be better than the fit based on the QLV (c.f. Figs. 5-6). In fact, correlation coefficients associated with the best fit of the statistical network theory were 0.7547 for ωη and 0.7265 for ζ, compared to only 0.2453 and 0.6596, respectively, for the QLV.

Fig. 9 also shows that the statistical network theory was more capable of predicting the slightly increasing trend of ωη with frequency, compared to the QLV which predicted a very flat or frequency-independent ωη curve. Predictions of ζ based on the statistical network theory (Fig. 10) also showed the decreasing trend of the data at high frequency (>5 Hz) better than those of the QLV. Meanwhile, the theory matched the empirical dynamic viscosity data (η) as well as the QLV, with an almost excellent match and a correlation coefficient higher than 0.97.

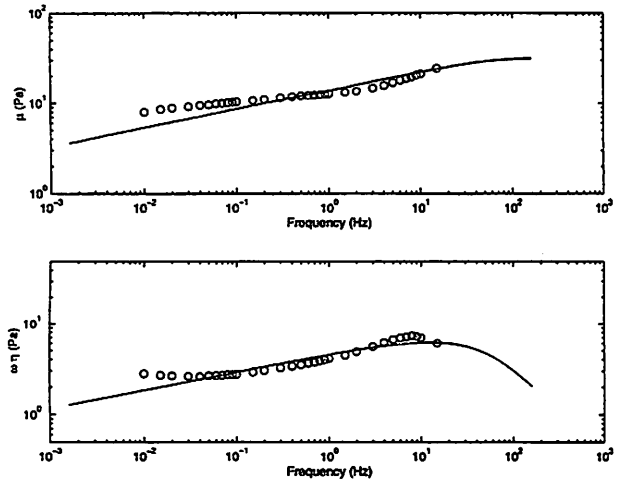


Figure 11. Curve-fitting of elastic shear modulus μ (upper graph) and viscous shear modulus ωη (lower graph) of female vocal fold mucosa by the statistical network model.

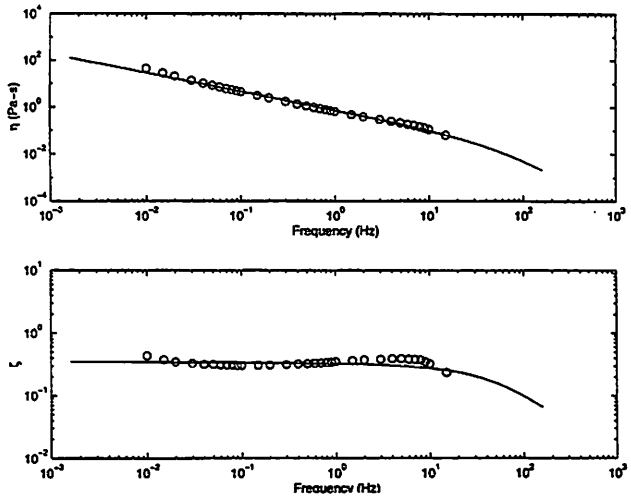


Figure 12. Curve-fitting of dynamic viscosity η (upper graph) and damping ratio ζ (lower graph) of female vocal fold mucosa by the statistical network model.

Figs. 11-12 show the average empirical data for ten female subjects (in circles) and the best-fit theoretical curves of μ, ωη, η and ζ based on the statistical network theory. The best fit was achieved with the model parameters η_0=2.0×10^6 Pa-s, η_∞=0.02 Pa-s, λ_0=2.5×10^7 seconds, s_0 = 100, and with α= 4.99138. Table IV lists the correlation coefficients for the female data using this set of model parameters.

Comparing with the results of curve-fitting for the male data, Figs. 11-12 show that predictions of the statistical network theory fit the female data even better than the male data, with all correlation coefficients higher than 0.8 (Table IV). In particular, the theoretical curves matched the female μ and ωη data rather closely at high frequency, compared to the male data (c.f. Figs. 9 and 11). This finding was in direct

Table IV.
Correlation Coefficients of Curve Fitting with Statistical Network Modeling of Human Female Vocal Fold Mucosa (c.f. Figures 11-12)

Shear property	r
μ	0.8989
$\omega\eta$	0.9370
η	0.9515
ζ	0.8395

contrast with that of the QLV, where the agreement between theory and data was better for the male data than for the female data. Also contrary to what was observed for the QLV, predictions of the statistical network theory matched quite well with the female ωη and ζ data (Figs. 11-12). Compared to the QLV, the theory was again more capable of predicting the slightly increasing trend of ωη with frequency and the decreasing trend of ζ at high frequency, similar to what was observed for the male data. Indeed, there seemed to be actually a better match between theory and data for ωη and ζ for the female than the male data, both in terms of visual comparisons and correlation coefficients. Meanwhile, the statistical network theory matched the female η data as well as the male data, with a good agreement and a correlation coefficient higher than 0.95.

With the two sets of model parameters known for the average male and the average female vocal fold mucosa, the maximum number density of chain segments (N_0) in the molecular network model can be computed according to Eq. (40). N_0 is an indication of the extent of molecular interactions of the network. Zhu and his colleagues (1991, 1996) found that N_0 increases with the concentration of proteoglycan solutions, as well as the relative content of collagen in collagen-proteoglycan mixtures. They also found that N_0 is positively correlated with the shear elasticity and viscosity of the proteoglycan solutions and the collagen-proteoglycan mixtures, supporting the hypothesis that the viscoelastic shear properties of polymeric materials increase with the extent of molecular interactions.

For the male vocal fold mucosa, N_0 = 2.9841×10^{22} m^{-3} while for the female vocal fold mucosa, N_0 = 7.4602×10^{21} m^{-3}, which was approximately four times smaller than that of the male. These findings suggested that the male vocal fold mucosa has effectively four times more molecular interactions sites than the female mucosa, in the language of the statistical network theory. This may be related to the mo-

lecular finding that the male vocal fold lamina propria has higher contents (concentrations) of certain molecular constituents than the female, particularly hyaluronic acid, and possibly collagen, elastin and proteoglycans (see the discussion in Chan and Titze, 1999b). Such molecular differences and the larger number of interaction sites in the male vocal fold mucosa could be responsible for the gender-related difference in viscoelastic shear properties reported previously (Chan and Titze, 1999b).

Zhu *et al.* (1991, 1996) also found that the parameters η_0, η_∞ and λ_0 increase with increasing concentration of proteoglycan solutions, while η_0 and λ_0 increase with the relative content of collagen in collagen-proteoglycan mixtures. A comparison between the model parameters for the male and the female vocal fold mucosa showed that η_0, η_∞ and λ_0 were all larger for male than for female. This finding further supported the hypothesis that the male vocal fold mucosa can be modeled as a molecular network with a larger number of interaction sites than the female mucosa, as a result of higher concentrations of key molecular constituents like hyaluronic acid and proteoglycans. It also suggested that the higher elasticity and viscosity of the male vocal fold mucosa may be partly due to a higher relative content of collagen in the lamina propria extracellular matrix. Higher concentration of hyaluronic acid in the male vocal fold lamina propria has been reported (Hammond *et al.*, 1997), but further studies on the relative content of other molecular constituents are needed to elucidate their roles in the determination of viscoelastic shear properties in the male versus female vocal fold mucosa.

Theoretical Extrapolations of Tissue Shear Properties

Based on the two theories, extrapolations of the theoretical predictions of the viscoelastic shear properties to frequencies of phonation (> 100 Hz) are possible, which allow for the predictions of shear properties at frequencies relevant to vocal fold oscillation. It has been shown analytically that the effective amount of damping in small-amplitude vocal fold oscillation depends critically on the viscous shear modulus $\omega\eta$ or the dynamic viscosity η of the vocal fold mucosa (Titze, 1988; Chan, 1998). Specifically, it was shown that phonation threshold pressure is directly related to $\omega\eta$ (or η). On the other hand, the damping ratio ζ of the vocal fold mucosa is critical in determining whether vocal fold oscillation is at all possible, whether oscillation would remain underdamped ($\zeta < 1.0$) across different frequencies.

For both male and female vocal fold mucosa, extrapolation of the theoretical curves of $\omega\eta$ to a frequency above 100 Hz using Fung's (1993) QLV showed that $\omega\eta$ remains basically constant as frequency is increased to the range of phonation (Figs. 5 and 7). Extrapolation of the theoretical curves of ζ showed that ζ continues to decrease slowly with frequency, always remaining to be < 1.0 up to phonational frequencies, where it has a numerical value of approximately 0.1 (Figs. 6 and 8).

On the other hand, extrapolation of the theoretical curves of $\omega\eta$ and ζ to a frequency above 100 Hz using Zhu's (1991) statistical network theory showed that both $\omega\eta$ and ζ begin to decrease with frequency quite dramatically, at a much faster rate than that predicted by the QLV (Figs. 9-12). It can be seen that for both male and female vocal fold mucosa ζ tends to decrease with frequency monotonically and approaches a numerical value on the order of 0.01 at phonational frequencies (Figs. 10 and 12).

These findings suggested that vocal fold oscillation would remain underdamped and be able to sustain at a frequency above 100 Hz, since ζ remains to be < 1.0. As frequency is increased, phonation threshold pressure, which is directly related to the viscous shear modulus $\omega\eta$, would remain basically constant according to the QLV, or it would decrease with frequency according to the statistical network theory, *if* other factors affecting phonation threshold pressure did not change with frequency. That is, vocal fold oscillation should remain as easy to initiate and sustain as frequency is increased. If such trends of extrapolations persist into even higher frequencies (200-1000 Hz), this may be the basis why voice production is possible over a relatively wide range of fundamental frequencies in singing situations (i.e., 2-3 octaves). As long as the viscous and damping properties of the vocal fold mucosa (as quantified by $\omega\eta$ and ζ) do not increase with frequency, the effective amount of damping on vocal fold oscillation would remain surmountable, resulting in realistic and achievable phonation threshold pressures even for phonation at high pitches. Moreover, individual differences in the ability to sing at high pitches may reflect differences in tissue shear properties of the vocal fold mucosa among different subjects, which may in turn reflect molecular and possibly genetic differences of the vocal fold lamina propria. This speculation is quite possible as large intersubject differences in tissue shear properties of the vocal fold mucosa were indeed observed in our previous study (Chan and Titze, 1999b). Certainly, further empirical measurements at higher frequencies need to be done in order to validate these arguments as well as the extrapolation of the theoretical curves.

Solely based on the goodness of fit between theory and data in the frequency range where empirical measurements were made (0.01-15 Hz), it seemed that the statistical network theory was somewhat more accurate than the QLV in describing the shear properties of vocal fold mucosal tissues. However, the statistical network theory predicted that $\omega\eta$ and ζ to continue to decrease with frequency quite dramatically, which suggested that the effective amount of

damping and phonation threshold pressure may become unrealistically low at high frequencies (i.e., phonation being too easy with too high oscillation amplitudes). In contrast, the QLV predicted a relatively frequency-insensitive damping curve, which suggested a more appropriate amount of damping at high frequencies and more realistic phonation threshold pressures. The relatively flat damping curves were also more consistent with the findings for other biological soft tissues reported in the literature. It is not completely clear from these findings which of the two theories is more appropriate for the description of tissue shear properties of the vocal fold mucosa. Again, further empirical measurements should be made at higher frequencies so that the theoretical extrapolations can be validated and that this issue may be resolved.

Summary and Conclusion

Empirical data on the viscoelastic shear properties of human vocal fold mucosa were previously reported at a frequency range of 0.01-15 Hz (Chan and Titze, 1999b). The present investigation attempted to model these data with a quasi-linear viscoelastic theory (Fung, 1972, 1981, 1993) and a statistical network theory (Zhu *et al.*, 1988, 1991). The quasi-linear viscoelastic theory has been successfully applied to describe the viscoelastic properties of many biological tissues, including the shear properties of soft tissues. The statistical network theory, on the other hand, has been extensively applied to describe the rheological properties of polymeric materials, as well as soft tissues like extracellular connective tissues. Both of these theories provide a unique mathematical framework in which a continuous relaxation spectrum may be used to model the generally frequency-insensitive damping characteristics of soft tissues. Based on these theories, extrapolations of the dynamic shear data to frequencies of phonation (> 100 Hz) are possible, which allow for the predictions of shear properties at frequencies relevant to vocal fold oscillation.

Results of viscoelastic modeling showed that there was generally reasonable agreement between the theoretical predictions and the empirical data within the frequency range of 0.01-15 Hz for both theories, although there was some discrepancies between theory and data at higher frequencies. These results suggested that the shear properties of the human vocal fold mucosa may be described parametrically by both theories under small-amplitude oscillation conditions. Theoretical extrapolations of the data to frequencies of phonation suggested that the damping properties of vocal fold mucosa may remain relatively constant and insensitive to frequency, or may even decrease with frequency. If similar patterns persist into higher frequencies (up to 1000 Hz), this may be the basis why vocal fold oscillation may remain underdamped over a relatively wide range of fundamental frequencies, i.e., 2-3 octaves in singing situations. Further studies should involve the empirical measurements of viscoelastic shear properties at higher frequencies, so as to validate the theoretical extrapolations.

Acknowledgment

This study was supported by a grant (No. P60 DC00976) from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.

References

Barnes, H.A., Hutton, J.F., and Walters, K. (1989). *An introduction to rheology.* Amsterdam, The Netherlands: Elsevier, pp. 1-10; 37-54; 97-114.

Becker, R., and Doring, W. (1939). *Ferronmagnetismus.* Springer, Berlin.

Berry, D.A., and Titze, I.R. (1996). Normal modes in a continuum model of vocal fold tissues. *J Acoust Soc Am, 100,* 3345-3354.

Bilston, L.E., and Thibault, L.E. (1996). The mechanical properties of the human cervical spinal cord *in vitro. Ann Biomed Engineering, 24,* 67-74.

Bird, R.B., Armstrong, R.C., and Hassager, O. (1977a). *Dynamics of polymeric liquids. Vol.1: Fluid mechanics.* New York: John Wiley & Sons, pp. 129-204; 275-303.

Bird, R.B., Hassager, O., Armstrong, R.C., and Curtiss, C.F. (1977b). *Dynamics of polymeric liquids. Vol.2: Kinetic theory.* New York: John Wiley & Sons, pp. 471-580; 701-721.

Chan, R.W. (1998). *Shear properties of vocal fold mucosal tissues and their effect on vocal fold oscillation.* Unpublished Ph.D. dissertation, The University of Iowa, Iowa City, IA.

Chan, R.W. and Titze, I.R. (1998). Viscosities of implantable biomaterials in vocal fold augmentation surgery. *Laryngoscope, 108,* 725-731.

Chan, R.W. and Titze, I.R. (1999a). Hyaluronic acid (with fibronectin) as a bioimplant for the vocal fold mucosa. *Laryngoscope* (in press).

Chan, R.W. and Titze, I.R. (1999b). Viscoelastic shear properties of human vocal fold mucosa: Measurement methodology and empirical results. *J Acoust Soc Am* (in review).

Chun, K.J. and Hubbard, R.P. (1986). Development of reduced relaxation function and stress relaxation with paired tendon. *Adv Bioengineering (ASME, New York), 13,* 162.

De Kee, D., and Carreau, P.J. (1979). A constitutive equation derived from Lodge's network theory. *J Non-Newton Fluid Mechan, 6,* 127-143.

Ferry, J.D. (1980). *Viscoelastic properties of polymers.* (3rd edn.). New York: Wiley.

Fung, Y.C. (1972). Stress-strain history relations of soft tissues in simple elongation. In Y.C. Fung, N. Perrone, and M. Anliker (eds.) *Biomechanics: Its foundations and objectives.* Englewood Cliffs, NJ: Prentice-Hall, pp. 181-208.

Fung, Y.C. (1981). *Biomechanics. Mechanical properties of living tissues.* New York: Springer-Verlag, pp. 22-61.

Fung, Y.C. (1993). *Biomechanics. Mechanical properties of living tissues.* (2nd edition). New York: Springer-Verlag, pp. 23-65; 242-320.

Hammond, T.H., Zhou, R., Hammond, E.H., Pawlak, A., and Gray, S.D. (1997). The intermediate layer: A morphologic study of the elastin and hyaluronic acid constituents of normal human vocal folds. *J Voice, 11,* 59-66.

Lodge, A.S. (1968). Constitutive equations from molecular network theories for polymer solutions. *Rheologica Acta, 7,* 379-392.

Mak, A.F. (1986). Apparent viscoelastic behavior of articular cartilage - Contributions from intrinsic matrix viscoelasticity and interstitial fluid flow. *J Biomechan Engineering, 108,* 123-130.

Neubert, H.K.P. (1963). A simple model representing internal damping in solid materials. *Aeronautical Quarterly, 14,* 187-197.

Pinto, J.G. and Fung, Y.C. (1973). Mechanical properties of the heart muscle in the passive state. *J Biomechanics, 6,* 597-616.

Pinto, J.G. and Patitucci, P.J. (1980). Viscoelasticity of passive cardiac muscle. *J Biomechan Engineering, 102,* 57-61.

Price, J.M., Patitucci, P.J., and Fung, Y.C. (1979). Mechanical properties of resting taenia coli smooth muscle. *Am J Physiol (Cell Physiol), 236,* C211-C220.

Sauren, A.A.H.J. and Rousseau, E.P.M. (1983). A concise sensitivity analysis of the quasi-linear viscoelastic model proposed by Fung. *J Biomechan Engineering, 105,* 92-95.

Simon, B.R., Coats, R.S., and Woo, S.L-Y. (1984). Relaxation and creep quasilinear viscoelastic models for normal articular cartilage. *J Biomechan Engineering, 106,* 159-164.

Spirt, A.A., Mak, A.F., and Wassell, R.P. (1989). Nonlinear viscoelastic properties of articular cartilage in shear. *J Orthopaedic Res, 7,* 43-49.

Tanaka, T.T., and Fung, Y.C. (1974). Elastic and inelastic properties of the canine aorta and their variation along the aortic tree. *J Biomechanics, 7,* 357-370.

Theodorsen, T., and Garrick, E. (1940). *Mechanism of flutter.* U.S. National Advisory Committee on Aeronautics, Report B685.

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

Titze, I.R., and Talkin, D. (1979). A theoretical study of the effects of various laryngeal configurations in the acoustics of phonation. *J Acoust Soc Am, 66,* 60-74.

Wagner, K.W. (1913). Zur theorie der unvoll Kommener dielektrika. *Ann Physics, 40,* 817-855.

Woo, S.L-Y., Gomez, M.A., and Akeson, W.H. (1981). The time and history-dependent viscoelastic properties of the canine medial collateral ligament. *J Biomechan Engineering, 103,* 293-298.

Woo, S.L-Y., Gomez, M.A., Woo, Y-K., and Akeson, W.H. (1982). Mechanical properties of tendons and ligaments. I. Quasi-static and nonlinear viscoelastic properties. *Biorheology, 19,* 385-396.

Woo, S.L-Y., Johnson, G.A., and Smith, B.A. (1993). Mathematical modeling of ligaments and tendons. *J Biomechan Engineering, 115,* 468-473.

Woo, S.L-Y., Simon, B.R., Kuei, S.C., and Akeson, W.H. (1980). Quasi-linear viscoelastic properties of normal articular cartilage. *J Biomechan Engineering, 102,* 85-90.

Zhu, W., Iatridis, J.C., Hlibczuk, V., Ratcliffe, A., and Mow, V.C. (1996). Determination of collagen-proteoglycan interactions in vitro. *J Biomechanics, 29,* 773-783.

Zhu, W.B., Lai, W.M., and Mow, V.C. (1986). Intrinsic quasi-linear viscoelastic behavior of the extracellular matrix of cartilage. *Trans Orthopaedic Res Soc, 11,* 407.

Zhu, W.B., Lai, W.M., and Mow, V.C. (1988). A second order rheological model to predict the time-dependent behavior of proteoglycan solutions. *Adv Bioengineering (ASME, New York), 8,* 187-190.

Zhu, W.B., Lai, W.M., and Mow, V.C. (1991). The density and strength of proteoglycan-proteoglycan interaction sites in concentrated solutions. *J Biomechanics, 24,* 1007-1018.

Zhu, W.B., and Mow, V.C. (1990). Viscometric properties of proteoglycan solutions at physiological concentrations. In V.C. Mow, A. Ratcliffe, and S. Woo (eds.). *Biomechanics of diarthrodial joints.* New York: Springer-Verlag, pp. 313-344.

NCVS Status and Progress Report - 13
April 1999, 53-62

Hyaluronic Acid (with Fibronectin) as a Bioimplant for the Vocal Fold Mucosa

Roger W. Chan, Ph.D.
Ingo R. Titze, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa

Abstract

Objectives: To measure the viscoelastic shear properties of hyaluronic acid, with and without fibronectin, and to compare them with those of the human vocal fold mucosa and other phonosurgical biomaterials.

Methods: Viscoelastic shear properties of various implantable biomaterials (Teflon, gelatin, collagen, fat, hyaluronic acid, and hyaluronic acid with fibronectin) were measured with a parallel-plate rotational rheometer. Elastic and viscous shear properties were quantified as a function of oscillation frequency (0.01 Hz to 15 Hz) at 37°C.

Results: The shear properties of hyaluronic acid were relatively close to those of human vocal fold mucosal tissues reported previously. Hyaluronic acid at specific concentrations (0.5-1.0%), with or without fibronectin, was found to exhibit viscous shear properties (*viscous shear modulus* and *dynamic viscosity*) similar to those of the average male and female vocal fold mucosa.

Conclusions: According to a theory that establishes the effects of tissue shear properties on vocal fold oscillation, phonation threshold pressure (a measure of the ease of phonation) is directly related to the viscous shear modulus of the vibrating vocal fold mucosa. Therefore, our findings suggest that hyaluronic acid, either by itself or mixed with fibronectin, may be a potentially optimal bioimplant for the surgical management of vocal fold mucosal defects and lamina propria deficiencies (e.g., scarring) from a biomechanical standpoint.

Introduction

In phonosurgery, various types of synthetic or biological substitute materials (implantable biomaterials) have been used as bioimplants to replace and augment natural tissues in pathological vocal folds. Basically, implantable biomaterials may be divided into two groups according to the site of implantation (often injection): the "non-mucosal group" and the "mucosal group". Biomaterials in the "non-mucosal group" are those which are usually implanted outside the vocal fold mucosa (defined as the epithelium and the superficial layer of the lamina propria), mostly lateral to the thyroarytenoid muscle, but sometimes close to the lamina propria. Medialization of a paralyzed vocal fold has been the frequent objective for phonosurgery involving this group of biomaterials, including silicone (in Japan), polytetrafluoroethylene (Teflon or Polytef; Mentor Inc., Hingham, MA), gelatin (Gelfoam; Upjohn Co., Kalamazoo, MI), collagen (Zyderm, Phonagel, or Vocalogen; Collagenesis Inc., Beverly, MA), and adipose tissue (human abdominal subcutaneous fat). On the other hand, biomaterials in the "mucosal group" are those which can be implanted directly into the mucosa, frequently to soften a scar tissue, to smooth the vocal fold margin, to augment an atrophy, or to treat other types of mucosal disorders[1,2]. Adipose tissue or fat has been the most commonly used biomaterial in this group[3,4], while collagen has also been used for such purposes to some extent[2]. Unfortunately, the use of fat and collagen for the management of vocal fold mucosal defects like scarring has not always yielded satisfactory clinical results, e.g., the restoration of an adequate mucosal wave[1,3,4].

Knowledge of the mechanical properties of vocal fold tissues and implantable biomaterials has important implications for phonosurgery. Of particular importance are their *viscoelastic* (viscous and elastic) *shear* properties. It is obvious that the surgical introduction of implantable biomaterials into the vocal folds can change their viscoelastic shear properties and hence alter the mechanics of phonation, at least during the initial postoperative stages (before

there is extensive biological assimilation of the implant). This is especially true for surgical procedures involving the correction of vocal fold mucosal defects (e.g., scarring), where mucosal group biomaterials are implanted directly into the mucosa. Since the vocal fold mucosa is the major vibratory portion of the vocal fold[5-7], biomaterials implanted directly into the mucosa affect the mechanics of phonation more directly and extensively. The viscoelastic shear properties of vocal fold mucosal tissues and implantable biomaterials, as well as their effect on vocal fold oscillation, should therefore be a special factor of concern for the surgical management of vocal fold mucosal defects. Specifically, *phonation threshold pressure* is an important parameter of phonation that describes how vocal fold oscillation may be affected by the implant. Phonation threshold pressure has been defined as the minimum subglottal pressure required to initiate and/or sustain vocal fold oscillation[8,9]. It is a measure of the "ease" of phonation, an objective indication of vocal function, and can be useful as a diagnostic parameter of vocal health[10,11]. The effect of tissue shear properties on phonation threshold pressure has been studied theoretically[8,9,12] as well as empirically, using excised larynges[13], a physical model of the vocal fold mucosa[12,14,15], and human subjects[10,11]. Theoretical results showed that phonation threshold pressure is directly related to the *viscous shear modulus* of the vocal fold mucosa[12], while empirical results showed that it is directly related to the *dynamic viscosity* of the vocal fold mucosa[10-15]. These findings agree with each other because the viscous shear modulus and the dynamic viscosity are derivable from each other (see definitions below).

With none of the existing common biomaterials being ideal for the management of vocal fold mucosal defects like scarring, it is necessary to search for new and more optimal biomaterials with considerations of their *biomechanical* effects on vocal fold oscillation, in addition to biological effects like immunological concerns. Logically, materials made up of biomacromolecules found in the human vocal fold mucosa should be potential candidates for this search. Two such molecular constituents of the lamina propria are targeted in the present investigation, namely the glycosaminoglycan *hyaluronic acid* (or hyaluronate, the physiological form in the extracellular matrix at a *p*H of 7.0-7.4) and the structural glycoprotein *fibronectin*.

Significant amounts of hyaluronic acid and fibronectin have been found in various layers of the normal human vocal fold lamina propria[16,17], including the superficial layer (Steven D. Gray, unpublished data). These macromolecules have very high molecular weights (on the order of 10^5-10^6) and are capable of various kinds of intra- and intermolecular interactions, e.g., physical entanglement among coils of hyaluronate chains[18,19], and adhesive interactions between fibronectin and other extracellular matrix molecules like collagen, hyaluronic acid, and proteoglycans[20,21]. These molecular interactions are likely important in contributing to the viscoelastic shear properties of the normal vocal fold mucosa[12], which suggests that hyaluronic acid mixed with fibronectin may be a potentially optimal bioimplant for vocal fold mucosal defects. Indeed, hyaluronic acid (together with the polysaccharide dextran) has already been used for implantation into the rabbit vocal fold, in order to assess its clinical and histological suitability for vocal fold augmentation[22]. It was found that significant resorption of the implanted hyaluronic acid occurred within a week, as would be expected from its approximately 5-day half-life. However, the authors also found that the implant promoted endogenous soft tissue growth and development in the vocal fold, presumably through the recruitment of fibroblasts and the trigger of collagen synthesis[22].

In our previous study addressing the mechanical properties of implantable biomaterials[23], the *dynamic viscosity* of Teflon, gelatin, collagen, and fat was measured. However, hyaluronic acid and fibronectin were not among the biomaterials studied, and data on elastic and viscous shear moduli were not reported. This study attempts to provide a more comprehensive set of data on the viscoelastic shear properties of these currently and potentially useful biomaterials.

Materials and Methods

Viscoelastic Shear Properties

Linear viscoelastic shear properties of various biomaterial samples were measured using the rheological technique of *sinusoidal oscillatory shear deformation* (see below), including (1) *elastic shear modulus*, (2) *dynamic viscosity*, and (3) *viscous shear modulus*. Detailed definitions and descriptions of these shear properties can be found elsewhere[12,24]. Briefly, for a viscoelastic material, the elastic shear modulus μ is a quantification of the energy storage component, or elasticity (stiffness) of the material in shear. The dynamic viscosity η is a quantification of the energy loss component of the material. As a derived quantity, the viscous shear modulus is defined as the product of the angular frequency of oscillation and the dynamic viscosity ($\omega\eta$). It is also a quantification of the energy loss of the material in oscillatory shear deformation. For viscoelastic materials, including polymeric and biological materials, these shear properties vary with the rate of deformation, i.e., they are functions of oscillation frequency in oscillatory shear deformation[25-27].

Biomaterial Samples

Details of the Teflon, gelatin, collagen, and adipose tissue samples have been described elsewhere[12,23]. Briefly, Teflon (Mentor Incorp., Hingham, MA) samples were obtained in form of an injectable paste suspended in

glycerin (740 mg/ml). Gelatin samples were prepared by mixing Gelfoam (Upjohn Co., Kalamazoo, MI) absorbable gelatin powder with saline solution (250 mg/ml). Collagen samples included both 0.0075% glutaraldehyde cross-linked (GAX) bovine dermal collagen suspension (Phonagel or Zyplast; Collagen Corp., Palo Alto, CA) and the noncross-linked Zyderm (Collagen Corp.) (35 mg/ml). Human abdominal subcutaneous adipose tissue samples were harvested from five cadavers within 24 hours post-mortem. Considerable intersubject differences in viscoelastic shear properties were found for the fat samples[12]. Besides, it was found that shear properties were affected by (1) the source of the fat samples (whether they were taken from the superficial or the deep fat tissue layer)[28], and (2) the method of sample preparation (whether they were mechanically agitated or not)[12]. For the present report, data for adipose tissue were averaged across subjects and conditions.

Hyaluronic acid samples of four different concentrations were prepared by mixing 0.1 mg, 1.0 mg, 5.0 mg or 10.0 mg sodium hyaluronate powder (Mr $\approx 3\times$ to 6×10^6; from human umbilical cord; Sigma Chemical Co., St. Louis, MO) with 1.0 cc phosphate-buffered solution (pH = 7.0), resulting in hyaluronate solutions of 0.1 mg/ml (0.01%), 1.0 mg/ml (0.1%), 5.0 mg/ml (0.5%), and 10 mg/ml (1.0%). These concentrations were chosen because they correspond to different extents of molecular entanglement and network formation. Hyaluronic acid molecules with a molecular weight of 3-4 millions have a length of about 10 m and a radius of gyration of about 200 nm[18,19]. Therefore, in solution, chains of these huge and highly charged molecules assume a highly extended random coil conformation which is approximately spherical in shape. The coil structure is stabilized by hydrogen bonding, with water molecules mechanically immobilized inside[18,21]. There is little or no overlap or entanglement among the hyaluronate chains and coils at a concentration of 0.01%, but some entanglements begin at a concentration of 0.06-0.1%[18,19]. As the concentration increases, there are increasingly more entanglements and the molecules gradually form a dynamic molecular network. At a concentration of approximately 1.0% the molecules are almost completely (96%-100%) entangled[18,29]. Hence hyaluronate samples were prepared at the concentrations of 0.01%, 0.1%, 0.5% and 1.0% in order to assess the biomechanical effects of varying molecular entanglements and network formation on viscoelastic shear properties.

There are well-known adhesive molecular interactions between fibronectin and hyaluronic acid[20,21]. As these molecular interactions likely contribute more to the viscoelastic shear properties of vocal fold mucosa than fibronectin molecules alone[12], samples of hyaluronate solutions mixed with fibronectin were prepared. Samples of three different concentrations were prepared by mixing 0.33 mg fibronectin powder (Mr $\approx 2\times10^5$; from human foreskin fibroblasts; Sigma Chemical Co.) with 1.0 cc hyaluronate solutions of three concentrations (0.01%, 0.1%, and 1.0%). With these preparations, possible effects of fibronectin on the shear properties of hyaluronic acid may be assessed, as well as possible changes of such effects under different extents of hyaluronate molecular entanglements. The concentration of fibronectin was chosen to be 0.33 mg/ml (0.033%) because it was found to be the approximate normal level of fibronectin in human synovial fluid[30]. Because fibronectin level in the vocal fold mucosa was not known, it seemed justified to use a concentration which is physiologic in another extracellular connective tissue in the body.

Rheometer and Oscillatory Shear Experiments

Viscoelastic shear properties of biomaterials were empirically measured using a standard experimental method in *rheology* (the science of the deformation and flow of matter), which has been described in details previously[12,23,24]. Briefly, small-amplitude sinusoidal oscillatory shear deformation of the sample was performed with a rheometer (Bohlin CS-50 Rheometer; Bohlin Instruments Inc., Cranbury, NJ). A parallel-plate (plate-on-plate) testing geometry was used, with a stationary lower plate and a rotating upper plate (diameter = 20 mm), and a gap size of 0.3 mm. A sample was placed in the gap between the two plates, and was subject to a precisely controlled sinusoidal torque from the upper plate, at a specific oscillation frequency. The resulting angular displacement and angular velocity of the upper plate were monitored by a sensitive transducer as functions of time. Shear stress, shear strain and strain rate associated with the oscillatory shear deformation were computed from the prescribed torque and the measured angular velocity by a computer, and the viscoelastic shear properties were obtained based on these functions.

Measurements were made at a frequency range of 0.01 to 15 Hz, covering 32 frequencies over three decades. Testing at a higher frequency was not meaningful because of limitations of the rheometer. These limitations were stress-strain linearity and sample inertial effects[12,24]. Calibration of the rheometer had been done by the manufacturer and was double checked by measuring the steady-shear viscosities of standard Newtonian silicone (polydimethylsiloxane) solutions (Dow Corning 360 Medical Fluids, Dow Corning Co., Midland, MI) with known viscosities. The measured viscosities showed only a small deviation from the stated viscosities (< 6%) in all cases[12,24]. Throughout the experiments, material samples were maintained at a temperature of $37^{\circ}C \pm 0.1^{\circ}C$ in the rheometer by a temperature control unit which circulated distilled water into the mounting of the lower plate.

Estimation of Experimental Error

Repeatability or experimental error associated with the rheological experimental procedure has been reported previously[12,23,24]. It was estimated by testing similar material

samples for multiple trials, including similar Teflon samples and similar gelatin (Gelfoam) samples. Because the samples came from the same sources of preparation, discrepancies in results between the samples reflected an estimation of errors. Possible sources of errors included differences in the process of sample mounting (e.g., incomplete sample filling between the plates, excessive removal of excess sample material) and machine errors of the rheometer (e.g., errors in rotor torque output, measurement of angular displacement). It was estimated that experimental error was within 10% of the measured viscoelastic data values across all frequencies[12,23,24].

Results

Figures 1-3 show the viscoelastic shear properties (the elastic shear modulus μ, the dynamic viscosity η, and the derived viscous shear modulus $\omega\eta$, respectively) of hyaluronic acid (hyaluronate solutions) at four concentrations (0.01%, 0.1%, 0.5%, and 1.0%), plotted as functions of frequency on a log-log scale. Logarithmic scales are used in order to accomodate decades of changes in shear properties and in frequency. It can be seen that the elastic shear modulus (Fig. 1) and the viscous shear modulus (Fig. 3) of hyaluronic acid at all concentrations were increasing monotonically with frequency, especially at lower frequencies (< 0.1 Hz) where the rate of increase was the highest. There was a lower rate of increase, or even a plateau, at higher frequencies (> 1 Hz), especially for the viscous shear modulus. On the other hand, the dynamic viscosity (Fig. 2) of hyaluronic acid at higher concentrations (0.5-1.0%) was relatively constant at low frequency (< 0.1 Hz), but it became a monotonically decreasing function at higher frequencies (> 0.1 Hz) (a phenomenon known as *shear-thinning*). The dynamic viscosity of hyaluronic acid at lower concentrations (< 0.5%) followed the same trend, except that it was increasing with frequency at low frequency (< 0.1 Hz) and did not begin to decrease until a relatively high frequency (above 1 Hz).

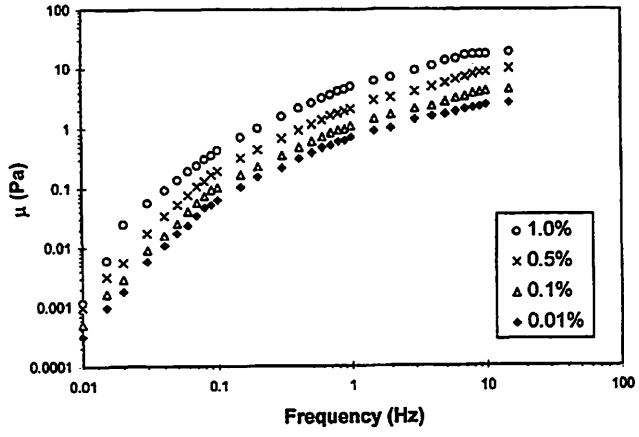


Figure 1. Elastic shear modulus of hyaluronic acid (at four concentrations) as a function of frequency.

Figs. 1-3 also show that the concentration of hyaluronic acid did not essentially affect the slopes of the curves, but only the vertical positions of the curves (the elasticity and viscosity at specific frequencies). It is clearly seen that the viscoelastic shear properties always increased with increasing hyaluronate concentration. In particular, the elastic shear modulus showed a small but gradual difference (approximately evenly distributed) among the four concentrations (Fig. 1). The dynamic viscosity and the viscous shear modulus showed only a small difference between the two lower concentrations (0.01% and 0.1%) and between the two higher concentrations (0.5% and 1.0%), but a large difference between the lower and the higher concentrations (approximately two orders of magnitude) (Figs. 2-3). These

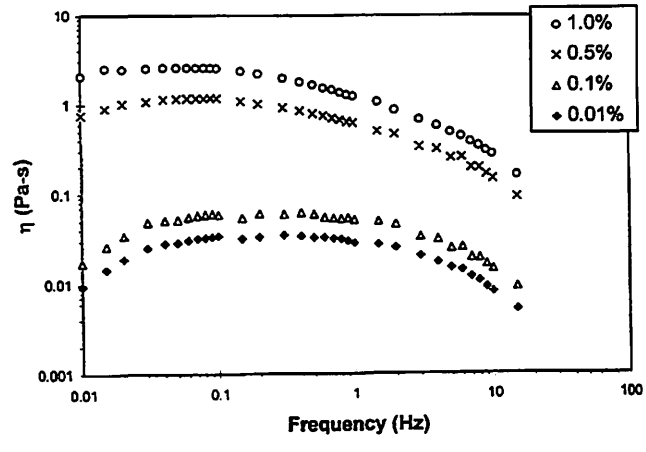


Figure 2. Dynamic viscosity of hyaluronic acid (at four concentrations) as a function of frequency.

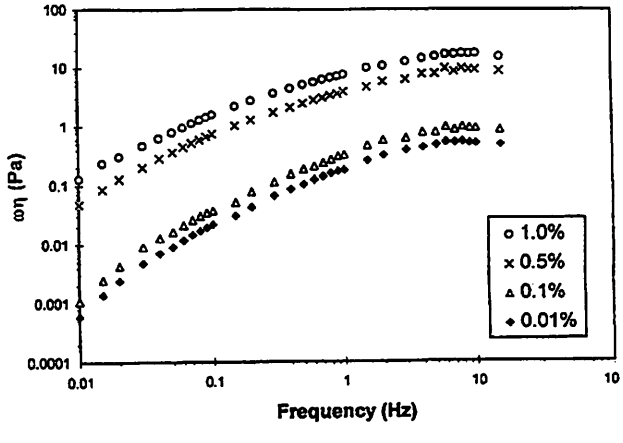


Figure 3. Viscous shear modulus of hyaluronic acid (at four concentrations) as a function of frequency.

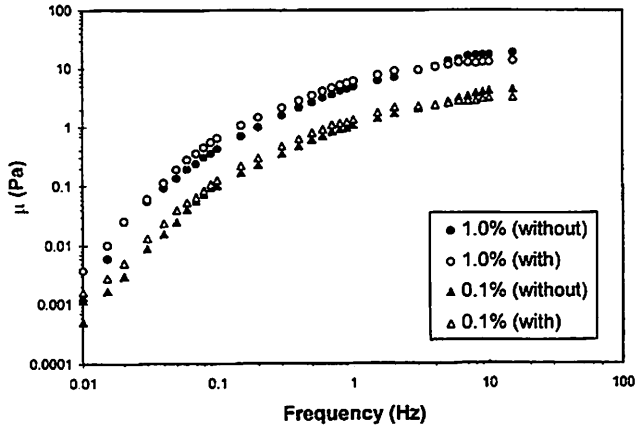


Figure 4. Elastic shear modulus of hyaluronic acid (at two concentrations) with and without fibronectin as a function of frequency.

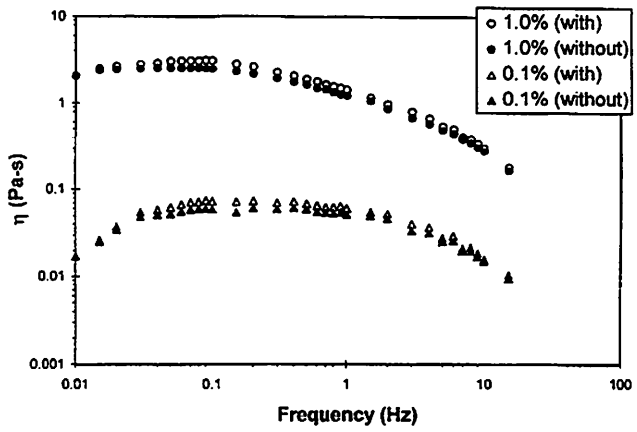


Figure 5. Dynamic viscosity of hyaluronic acid (at two concentrations) with and without fibronectin as a function of frequency.

findings were also consistently observed for hyaluronate samples mixed with fibronectin, to be discussed below.

Figs. 4-6 summarize the effect of fibronectin on the viscoelastic shear properties of hyaluronic acid at two concentrations (0.1% and 1.0%). The results of only two concentrations of hyaluronic acid were included for visual clarity (data for hyaluronic acid at a concentration of 0.01% were very similar to those at 0.1%). Surprisingly, the data showed that the elastic shear modulus, the dynamic viscosity, and the dervied viscous shear modulus of hyaluronic acid were not affected much by the presence of fibronectin. The differences observed between samples with fibronectin and samples without fibronectin were on the order of the magnitude of experimental error (< 10%). Therefore, the introduction of fibronectin into hyaluronic acid did not seem to change its viscoelastic shear properties, at least across the three molecular concentrations measured (0.01%, 0.1%, 1.0%).

To facilitate visual comparison, data on viscoelastic shear properties of hyaluronic acid at four different concentrations were plotted on the same graphs with those of other biomaterials (Teflon, gelatin, collagen, fat), as well as those of the average human vocal fold mucosa reported elsewhere[12,23,24]. Hyaluronic acid with fibronectin was not included, as it was just shown that fibronectin did not have a significant effect on shear properties. Figs. 7-9 (following pages) show the elastic shear modulus, the dynamic viscosity, and the viscous shear modulus of all the materials respectively.

It is clear that the elastic and viscous shear moduli of Teflon, gelatin, collagen, fat, as well as human vocal fold mucosa were relatively weak functions of frequency, as they remained relatively constant across the entire frequency range of measurements. This finding was quite different from that for hyaluronic acid, which showed a rather strong

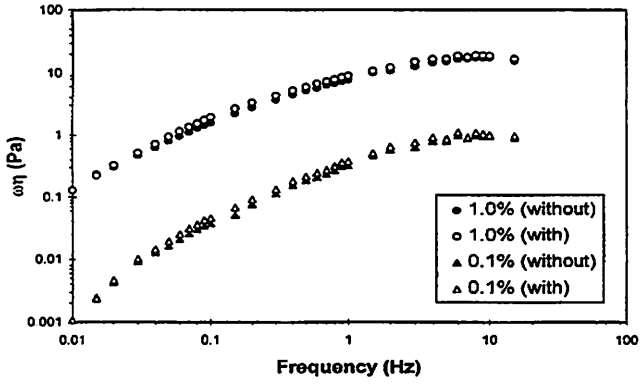


Figure 6. Viscous shear modulus of hyaluronic acid (at two concentrations) with and without fibronectin as a function of frequency.

dependence on frequency (steeper slopes of the curves). The corresponding dynamic viscosity was therefore a monotonically decreasing function of frequency for all the biomaterials, except for hyaluronic acid, which was relatively flat, especially at intermediate frequencies (0.1-1 Hz, Fig. 8). In other words, most of the biomaterials and the human vocal fold mucosa showed the phenomenon of shear-thinning, similar to most polymeric and biologic materials (e.g., proteoglycans solutions)[25,26,31]. It is believed that shear-thinning may be the consequence of a decrease in the number and strength of transient interaction sites formed in a molecular network as frequency or shear rate is increased[25,26]. On the other hand, hyaluronic acid was the most *Newtonian-like* among all the biomaterials, as the viscosity of a Newtonian fluid (like water) is independent of frequency or shear rate (i.e., there is no shear-thinning, nor "shear-thickening"). Note that the viscosity of pure water is 0.001 Pa-s, which is the baseline in Fig. 8. Thus, the x-axis becomes the reference viscosity in Fig. 8, against which comparisons with the various

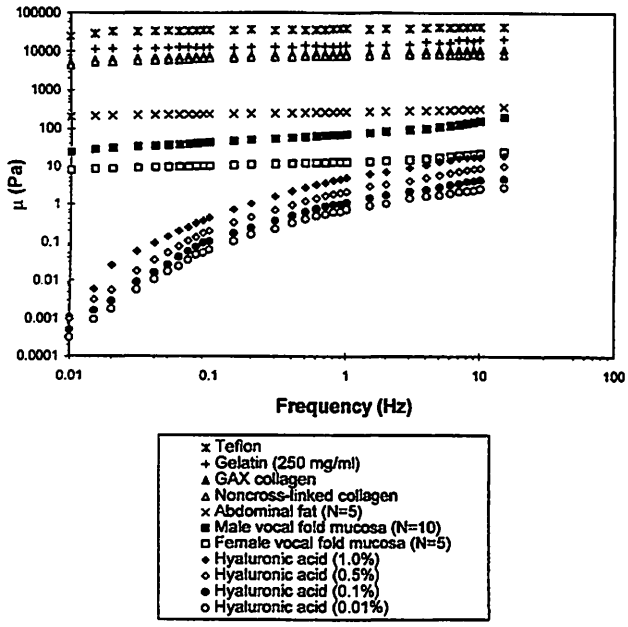


Figure 7. Elastic shear modulus of implantable biomaterials and human vocal fold mucosa as a function of frequency.

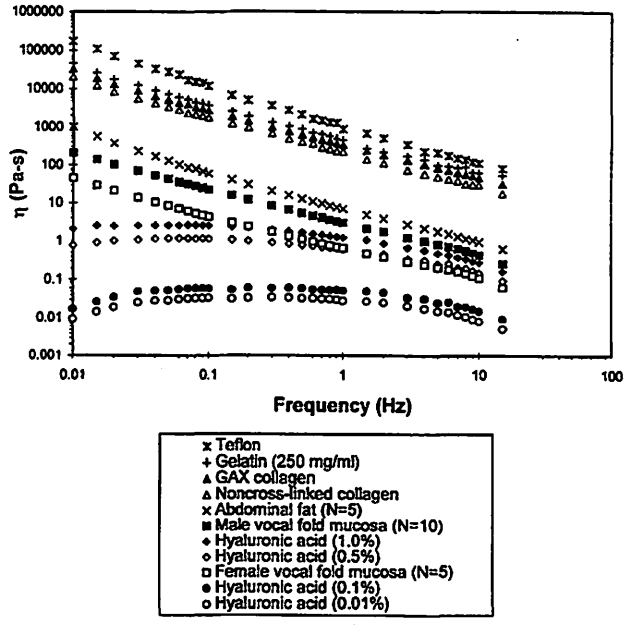


Figure 8. Dynamic viscosity of implantable biomaterials and human vocal fold mucosa as a function of frequency.

biomaterials can be made. For instance, at a frequency of 10 Hz, hyaluronic acid at higher concentrations (0.5-1.0%) was over 100 times more viscous than water, while Teflon was over 100,000 times more viscous than water.

Most of the biomaterials and the human vocal fold mucosa showed very similar shapes and slopes in their viscoelastic shear properties, with vertical separations between the curves being the major difference. Hyaluronic acid was the exception, again. Teflon had the largest shear elasticity (or stiffness) and viscosity, followed by gelatin, GAX collagen, noncross-linked collagen, abdominal fat, and the human vocal fold mucosa (note that there was a considerable sex difference; for details see previous reports[12,24]). It is also clear that the viscoelastic shear properties of the non-mucosal group biomaterials (Teflon, gelatin, collagen) were often orders-of-magnitude higher than those of the mucosal group biomaterials (fat and hyaluronic acid), which were relatively close to the shear properties of the human vocal fold mucosa. Finally, and probably most importantly, the viscoelastic shear properties of hyaluronic acid at higher concentrations (0.5-1.0%) seemed to approach those of the human vocal fold mucosa as frequency was increased, despite the unique shapes and slopes of the curves of hyaluronic acid. This was particularly true for the dynamic viscosity and for the derived viscous shear modulus, for which there was much similarity at frequencies > 1 Hz. Hyaluronic acid at a concentration of 1.0% behaved like the average male vocal fold mucosa, and hyaluronic acid at a concentration of 0.5% behaved like the average female vocal fold mucosa.

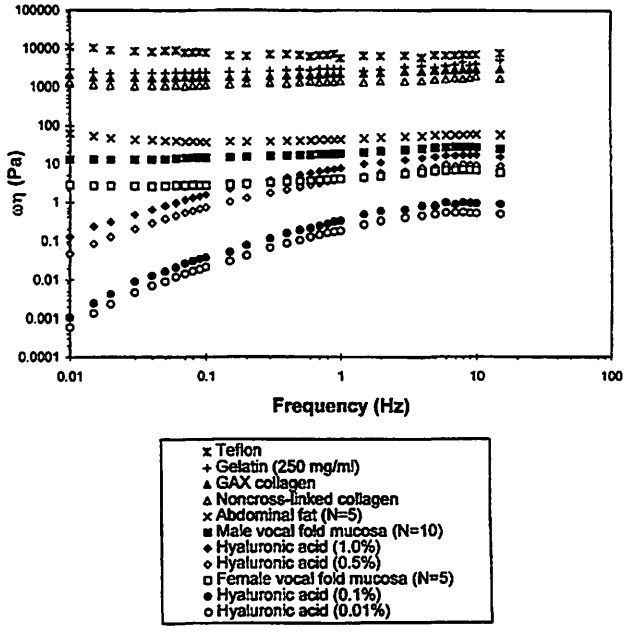


Figure 9. Viscous shear modulus of implantable biomaterials and human vocal fold mucosa as a function of frequency.

Discussion

Hyaluronic acid (or the physiological form hyaluronate) is a heteropolysaccharide with a disaccharide monomer, a very long linear polymer with a very high molecular weight (on the order of 10^6). The hyaluronate molecule assumes an extended conformation in solution, forming a

random coil structure[18,21]. The coils are capable of various molecular interactions, with the extent of interactions depending on their molecular weight and concentration in solution[19]. In particular, they physically entangle with one another to form a dynamic molecular network that contributes to shear elasticity and viscosity[19]. Hyaluronic acid has been widely used as an ophthalmic surgical aid, e.g., for cataract surgery where it is injected as a replacement for vitreous humor, to maintain the intraocular space and to cushion the corneal endothelium[32,33]. It has also been widely used as an implant in orthopaedic and osteoarthritic surgery, to lubricate arthritic joints[34,35].

It was found in the present study that the viscous shear properties of hyaluronic acid are highly dependent on concentration. This finding could be interpreted molecularly in terms of the formation of a dynamic molecular network, where hyaluronic acid at a low concentration (0.01-0.1%) produced little entanglement of hyaluronate chains and coils (< 5% intra- and inter-molecular overlap), while hyaluronic acid at a higher concentration (0.5-1.0%) produced much higher levels of entanglement (>80% overlap)[18,19]. It is likely that the extent of molecular interactions in the network is directly proportional to the level of molecular entanglement and overlap[19,25]. Therefore, hyaluronic acid at a higher concentration showed a much higher viscosity (and elasticity) because of the much higher level of molecular entanglement.

Fibronectin is a mosaic or modular protein covalently attached to oligosaccharide chains, with a fairly high molecular weight (on the order of 10^5)[36]. One of the main biological functions of fibronectin is to promote adhesive interactions and information interchange between cells and the extracellular matrix. It has binding sites interacting with receptor proteins on the cell membrane of fibroblasts (e.g., integrin), as well as macromolecules in the extracellular matrix like collagen, hyaluronic acid, and proteoglycans[20,21]. It also plays an important role in promoting wound healing, as it is intensely deposited in sites of tissue injury[37]. Because of its adhesive interactions with other macromolecules, which often cause changes in their molecular configurations[20,21], it was reasoned that fibronectin probably contributes to the shear properties of the hyaluronate molecular network. However, it was found that fibronectin did *not* have much effect on the shear properties of hyaluronic acid (Figs. 4-6), despite the well known presence of hyaluronate-binding sites on fibronectin molecules[20]. Our findings were comparable to those of a previous study by Hsu *et al.*[38], who studied the viscoelastic shear properties of collagen gels mixed with extracellular macromolecules, including fibronectin. They also found that fibronectin did *not* affect the shear properties of collagen, despite the presence of collagen-binding sites on fibronectin[39].

Two possible molecular interpretations may account for these findings. First, molecular size of the fibronectin used (Mr ≈ 2×10^5) might have been too small for the hyaluronate-binding sites to develop effective *biomechanical* interactions with the high-molecular-weight hyaluronate molecules (Mr ≈ 3 to 6×10^6), although it was large enough for biochemical adhesive interactions[20]. Second, concentrations of the fibronectin used (0.033% in this study, 0.002% in the study by Hsu *et al.*[38]) might have been too low to produce any appreciable changes in the molecular configurations of hyaluronic acid and its shear properties. Although the concentration used in this study (0.033%) was a physiological level for human synovial fluid[30], the role of fibronectin in synovial fluid might have been more biochemical than biomechanical. Further studies on the effect of fibronectin on viscoelastic shear properties need to be done when more information on the concentration and molecular weight of fibronectin in the human vocal fold mucosa are available.

One limitation of the present study was that the viscoelastic shear properties of implantable biomaterials were measured *in vitro*. Although these *in vitro* shear properties do affect vocal fold oscillation when the biomaterials are implanted into the vocal fold mucosa[12], their *in vivo* shear properties will likely change over time because of biological assimilation of the implant. In order to assess the short-term and long-term *in vivo* effects of the implant on the shear properties of vocal fold mucosa, the *change* of tissue shear properties as a function of time should be studied in future. This could be done in animal experiments, where the shear properties of vocal fold mucosa undergoing implantation may be measured at different preoperative and postoperative stages.

Another limitation was that viscoelastic data were only obtained at relatively low frequencies, despite a frequency range of three decades (0.01 Hz to 15 Hz). Because of limitations of the rheometer, the highest frequency at which meaningful measurements could be made was 15 Hz. As this was approximately an order of magnitude below typical frequencies of vocal fold oscillation (usually >100 Hz), empirical and theoretical extrapolations of the data to audio frequencies need to be done[12,24], and these extrapolations have to be validated by further empirical measurements at higher frequencies.

Assuming that the results of the present study could be extrapolated to frequencies of phonation, our data suggest that hyaluronic acid is potentially useful as a bioimplant for the surgical management of vocal fold mucosal defects and lamina propria deficiencies. Hyaluronic acid at a concentration of 0.5-1.0% seems to show the optimal shear properties (particularly viscous shear properties) for vocal fold oscillation and "easy onset" of phonation. The inclusion

of fibronectin in hyaluronic acid may make the implant more biologically optimal for the vocal fold mucosa because of the positive role fibronectin plays in wound healing[37], which may be helpful for the metabolic repair of tissue damage in the lamina propria. Further studies should be conducted to investigate whether the potentially facilitative effects of hyaluronic acid could last beyond the initial postoperative periods, after it is being assimilated by body tissue reactions.

In order to use hyaluronic acid as a bioimplant for phonosurgery, immunological and other safety concerns (like airway compromise and foreign body reactions) should be carefully addressed in future animal and clinical studies. However, no serious risk is expected because of the fact that hyaluronic acid has been extensively used as a bioimplant in other surgical specialties for many years, including ophthalmic, orthopaedic and osteoarthritic surgery[32-35].

Conclusion

Using a parallel-plate rotational rheometer, sinusoidal oscillatory shear experiments were performed to quantify the viscoelastic shear properties of hyaluronic acid, including the elastic shear modulus, the dynamic viscosity, and the derived viscous shear modulus. These shear properties were compared to those of biomaterials commonly used in treating vocal fold paralysis or vocal fold mucosal defects (Teflon, gelatin, collagen, and fat). It was found that hyaluronic acid at specific concentrations (0.5-1.0%) had viscous shear properties similar to those of the average human vocal fold mucosa, and that fibronectin did not significantly affect the shear properties of hyaluronic acid. These findings suggest that, biomechanically speaking, hyaluronic acid (with or without fibronectin) may be a potentially optimal bioimplant for the surgical management of vocal fold mucosal defects and lamina propria deficiencies. It likely facilitates vocal fold oscillation, particularly the "ease of vibration onset", at least during the initial postoperative periods. Further research should be directed toward understanding the short-term and long-term *in vivo* effects of hyaluronic acid and fibronectin on the vocal fold mucosa.

Acknowledgment

This study was supported by a grant (No. P60 DC00976) from the National Institute on Deafness and Other Communication Disorders, a division of the National Institutes of Health. We would like to thank Dr. Steven D. Gray of the University of Utah, for his generosity in sharing with us many wonderful ideas, suggestions, and comments.

Bibliography

1. Benninger MS, Alessi D, Archer S, et al. Vocal fold scarring: Current concepts and management. *Otolaryngol Head Neck Surg* 1996; 115:474-482.

2. Ford CN, Bless DM. Collagen injection in the scarred vocal fold. *J Voice* 1987; 1:116-118.

3. Sataloff RT, Spiegel JR, Hawkshaw MJ, Rosen DC, Heuer RJ. Autologous fat implantation for vocal fold scar: A preliminary report. *J Voice* 1997; 11:238-246.

4. Shaw GY, Szewczyk MA, Searle J, Woodroof J. Autologous fat injection into the vocal folds: Technical considerations and long-term follow-up. *Laryngoscope* 1997; 107:177-186.

5. Hirano M. Phonosurgery: Basic and clinical investigations. *Otologia (Fukuoka)* 1975; 21: 239-440.

6. Saito S, Fukuda H, Kitahara S, et al. Pellet tracking in the vocal fold while phonating - Experimental study using canine larynges with muscle activity. In: Titze IR, Scherer RC, eds. *Vocal Fold Physiology.* Denver, CO: Denver Center for the Performing Arts; 1985:169-182.

7. Fukuda H, Kawasaki Y, Kawaida M, et al. Physiological properties and wave motion of the vocal fold membrane viewed from different directions. In: Gauffin J, Hammarberg B, ed . *Vocal fold physiology: Acoustic, perceptual, and physiological aspects of voice mechanisms.* San Diego: Singular Publishing Group; 1991:7-14.

8. Titze IR. The physics of small-amplitude oscillation of the vocal folds. *J Acoust Soc Am* 1988; 83:1536-1552.

9. Lucero JC. The minimum lung pressure to sustain vocal fold oscillation. *J Acoust Soc Am* 1995; 98:779-784.

10. Verdolini K, Titze IR, Fennell A. Dependence of phonatory effort on hydration level. *J Speech Hear Res* 1994; 37:1001-1007.

11. Verdolini-Marston K, Titze IR, Druker DG. Changes in phonation threshold pressure with induced conditions of hydration. *J Voice* 1990; 4:142-151.

12. Chan RW. *Shear properties of vocal fold mucosal tissues and their effect on vocal fold oscillation.* Unpublished Ph.D. dissertation. Iowa City, IA: The University of Iowa; 1998.

13. Finkelhor BK, Titze IR, Durham PL. The effect of viscosity changes in the vocal folds on the range of oscillation. *J Voice* 1987; 1:320-325.

14. Titze IR, Schmidt SS, Titze MR. Phonation threshold pressure in a physical model of the vocal fold mucosa. *J Acoust Soc Am* 1995; 97:3080-3084.

15. Chan RW, Titze IR, Titze MR. Further studies of phonation threshold pressure in a physical model of the vocal fold mucosa. *J Acoust Soc Am* 1997; 101:3722-3727.

16. Pawlak AS, Hammond TH, Hammond EH, Gray SD. Immunocytochemical study of proteoglycans in vocal folds. *Ann Otol Rhinol Laryngol* 1996; 105:6-11.

17. Hammond TH, Zhou R, Hammond E, Pawlak A, Gray S. The intermediate layer: A morphologic study of the elastin and hyaluronic acid constituents of normal human vocal folds. *J Voice* 1997; 11:59-66.

18. Laurent TC. Biochemistry of hyaluronan. *Acta Otolaryngol (Stockh)* 1987; 442:7-24.

19. Bothner H, Wik O. Rheology of hyaluronate. *Acta Otolaryngol (Stockh)* 1987; 442:25-30.

20. Ruoslahti E. Fibronectin and its receptors. *Annual Rev Biochem* 1988; 57:375-413.

21. Lehninger AL, Nelson DL, Cox MM. *Principles of biochemistry.* New York: Worth Publishers; 1993:298-321.

22. Hallen L, Dahlqvist A, Laurent C. Dextranomeres in hyaluronan (DiHA): A promising substance in treating vocal cord insufficiency. *Laryngoscope* 1998; 108:393-397.

23. Chan RW, Titze IR. Viscosities of implantable biomaterials in vocal fold augmentation surgery. *Laryngoscope* 1998; 108:725-731.

24. Chan RW, Titze IR. Viscoelastic shear properties of human vocal fold mucosa: Measurement methodology and empirical results. *J Acoust Soc Am*, submitted for publication.

25. Ferry JD. *Viscoelastic properties of polymers.* New York: Wiley; 1970:1-33.

26. Bird RB, Armstrong RC, Hassager O. *Dynamics of polymeric liquids. Vol.1: Fluid mechanics.* New York: John Wiley & Sons; 1977:129-204.

27. Fung YC. *Biomechanics. Mechanical properties of living tissues.* 2nd ed. New York: Springer-Verlag; 1993:23-65.

28.Johnson D, Cormack GC, Abrahams PH, Dixon AK. Computed tomographic observations on subcutaneous fat: Implications for liposuction. *Plast Reconstr Surg* 1996; 97:387-396.

29. Morris ER, Rees DA, Welsh EJ. Conformation and dynamic interactions in hyaluronate solutions. *J Mol Biol* 1980; 138:383-400.

30. Abdullin AR, Litvinov RI, Zinkevich OD, Salikov IG. Quantitative and qualitative changes in the fibronectin level in rheumatoid arthritis. [Russian]. *Terapevticheskii Arkhiv* 1988; 60:89-94.

31. Mow VC, Mak AF, Lai WM, Rosenberg LC, Tang LH. Viscoelastic properties of proteoglycan subunits and aggregates in varying solution concentrations. *J Biomechanics* 1984; 17:325-338.

32. Balazs EA. Sodium hyaluronate and viscosurgery. In Miller D, Stegmann R, eds. *Healon (sodium hyaluronate): A guide to its use in opthalmic surgery.* New York: John Wiley & Sons; 1983:5-28.

33. Pollack F. Healon (Na hyaluronate): A review of the literature. *Cornea* 1986; 5:81-93.

34. Balazs EA, Denlinger JL. Viscosupplementation: A new concept in the treatment of osteoarthritis. *J Rheumatol* 1993; 20 (supplement 39):3-9.

35. Peyron J. Intraarticular injections in the treatment of osteoarthritis: State of art. *J Rheumatol* 1993; 20 (supplement 39):10-15.

36. Potts JR, Campbell ID. Structure and function of fibronectin modules. *Matrix Biol* 1996; 15:313-320.

37. Grinnell F. Fibronectin and wound healing. *J Cellular Biochem* 1984; 26:107-116.

38. Hsu S, Jamieson AM, Blackwell J. Viscoelastic studies of extracellular matrix interactions in a model native collagen gel system. *Biorheology* 1994; 31:21-36.

39. Hayashi M, Yamada KM. Domain structure of the carboxyl-terminal half of human plasma fibronectin. *J Biol Chem* 1983; 258:3332-3340.

NCVS Status and Progress Report - 13
April 1999, 63-70

Geometric Structure of the Human and Canine Cricothyroid and Thyroarytenoid Muscles for Biomechanical Applications

Karin A. Cox, M.A.
Fariborz Alipour, Ph.D.
Ingo R. Titze, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa

Abstract

The geometric structure of the cricothyroid muscle (CT) and thyroarytenoid muscle (TA) was quantified in six human and three canine larynges. Each muscle was divided into a series of fiber bundles. Using a three-dimensional micrometer probe, the coordinates of the origin and insertion of each bundle was measured before dissection. It was found that the mass of the CT muscle in the canine was 1.463 ± 0.280 grams, which was significantly greater than the 0.9423 ± 0.123 grams in human. This was a result of the cross-sectional area of the canine CT being 105.3 ± 11.6 mm^2 instead of 73.8 ± 7.4 mm^2 for human. However, the ratios of CT/TA mass and cross-sectional area between two groups were not significantly different, suggesting that the two muscles grow proportionally.

Introduction

The intrinsic muscles of the larynx are essential in the control of loudness, pitch, and voice quality in speaking and singing. These muscles are part of a cooperative system used for the purpose of widening the glottis (abduction), closing the glottis (adduction) and lengthening the vocal folds. The mechanism of lengthening the vocal folds is of particular interest in this study; therefore two intrinsic muscles, the cricothyroid (CT) and thyroarytenoid (TA), were chosen for a comparative analysis of fiber orientation and size. These two muscles are responsible for most mechanical actions within the larynx that affect pitch control.

Because the thyroarytenoid and cricothyroid are in direct opposition to each other, they are used differentially as regulators of fundamental frequency[1]. Many studies have clarified the function of the cricothyroid muscle in phonation using electromyography[2-5](EMG), morphological techniques[6] and experiments on canines[7-9]. These studies indicate that increased cricothyroid activity raises fundamental frequency, all else being constant. The cricothyroid muscle usually acts in a "non isometric" mode and decreases its length during contraction. This decrease of length could reach up to 50% at high pitches[10, 11]. Furthermore, in the singer's chest voice, the activity of the cricothyroid muscle is increased with fundamental frequency and only in falsetto register does it reach its maximum[3, 10].

The function of the thyroarytenoid muscle, however, is less clear in fundamental frequency (Fo) control. Increased activity of the cricothyroid muscle raises Fo whereas increased activity in the thyroarytenoid muscles may either raise or lower Fo.[4,5,9,12] Tension of the vocal fold is controlled by the vocal fold length. Greater contraction of the muscle will lower Fo because the active tension in the thyroarytenoid is outweighed by reduced tension in the cover because of a decrease in length.

The cricothyroid and thyroarytenoid work together as agonist/antagonist pairs[10,13] as they cause rotation and gliding around the cricothyroid joint. In order to solve the dynamics of this rotation and gliding, it is necessary to know the net clockwise and counterclockwise torques that are produced by the contractions of both the cricothyroid and thyroarytenoid muscles. Their individual torque depends on the magnitude of force, the line of action and the moment arm. If it is assumed that muscle bundles are activated equally then the force can be proportional to the cross-sectional area of the muscle, the effective torque can be calculated from the individual bundle's geometry (coordinates of origin, insertion, and cross-sectional area). Thus

the knowledge of muscular fiber length and orientation is an essential in any estimate of torque and moment arm of the cricothyroid rotation.

A comparative analysis between the anatomy of the human and canine larynx is important when studying these muscles because much of the knowledge we gain about the vocal mechanism is obtained from animal experimentation. The canine larynx has been a popular substitute for the human larynx in a variety of studies because of its similar size and gross structure. It is critical, however, to identify differences that occur between the two species before application can be made to human function.

The purpose of this study was to examine the differences between orientation, length, mass and cross-sectional area of the canine and human larynx and discuss their implications for pitch control. This will be done mainly in terms of ratios between cricothyroid and thyroarytenoid measures, which are most helpful for identifying differences in overall function.

Methodology

Human and Canine Tissue Collections

Human larynges from autopsy were obtained from the Department of Pathology at the University of Iowa or from the University of Utah. The six male larynges were collected and slowly frozen 5 to 19 hours after death and after the stages of rigor mortis in the neck region were completed. Their ages were 36, 65, 74, 78, 41, 71 years respectively. All larynges were obtained from individuals with no history of head or neck trauma and no laryngeal disease. The canine larynges were obtained from the University of Iowa Animal Care Unit after having been sacrificed for experimentation in a cardiovascular research laboratory. The three male specimens were obtained immediately after death and submerged in 0.9% saline solution. The ages of the specimens were estimated by a University Veterinarian on the basis of dental development and were found to be no more than one to two years old. All specimens were mixed Labrador and German Shepherd breeds ranging in size from 25 to 26.1 Kg. All larynges were obtained from canines with no evidence of head or neck trauma or disease of the larynx.

Preparation of Tissue

The human larynges were frozen slowly following autopsy and stored for several days in a freezer (-4°C). They were also thawed slowly at room temperature before measurement. In order to minimize the effects of dehydration, the fluid balance of the dead tissue was maintained by keeping the larynx submerged in a 0.9% saline solution at room temperature before each measurement. Throughout the dissection procedure and after each fiber was removed, the saline solution was applied to maintain hydration of each muscle and all dissected fibers. To prevent excessive hydration the fiber was placed on absorbable paper for a brief moment to soak up surface saline before being weighed. The remaining tissue was submerged in .9% saline solution and stored in the refrigerator (+4° C) after each muscle was dissected.

The canine larynges were also submerged in 0.9% saline solution following removal and stored in the refrigerator (+4°C) for 20 to 48 hours before measurement. During dissection, the canine larynges were protected against dehydration using the same procedure used for the human larynges. Thus, the processes immediately after harvesting were different for canine and human samples (freezing in human), but storage during dissection was similar.

Fiber Dissection

Universal safety precautions were used for all dissection procedures. Micro dissection scissors, lightweight sand papers, scalpels, and tweezers were used to remove unwanted fat, fascia, and nerve tissue. The excised larynges were prepared such that only the major cartilages (thyroid, cricoid, arytenoid, and a piece of the trachea) and intrinsic muscles remained as the framework. Excess tissue was removed to expose the cricothyroid muscle (CT), in particular the pars recta (PR) and the pars obliqua (PO). Superiorly, all the tissue was removed to the level of the true vocal folds. The vocal fold mucosa and ligament were also removed so that all the thyroarytenoid muscle (TA) fibers were visible. To minimize tissue alterations, the same person performed all fiber dissections.

The second step in this procedure was to mount the larynx on a lab bench such that the cricoid cartilage was fixed. Rotation of the other cartilage allowed the axis of rotation (through the cricothyroid joint) to be determined. A pin was inserted through both cricothyroid joints to establish the origin of the coordinate system and to keep the joints from rotating and slipping. The cricothyroid joint constituted one of the primary reference points. The midline of the lower border of the anterior arch of the cricoid cartilage constituted the second primary reference point and the anterior commissure constituted the third primary reference point. A pin was also inserted through the cricoarytenoid joint to keep the arytenoid from sliding or rocking in any direction. The three-dimensional geometry of the larynx was quantified with a three-dimensional mechanical positioning system composed of three orthogonally-mounted micrometers. Vernier markings allowed the position of each slide to be read to an accuracy of 0.1 mm. The range of measurement was determined so that the larynx could remain in a fixed position until all the data for that muscle was collected. Several additional reference points on the major cartilages (thyroid, cricoid and arytenoid) were recorded with the device to indicate where the origins and insertions

were in reference to the whole laryngeal framework. It should be noted that the left CT muscle and the right TA muscle were dissected for each specimen. Ease of dissection and apparatus limitations were the primary reasons that muscles of the same side were not used for each specimen. Symmetry was thus assumed in each case for analysis, but a test for symmetry in terms of bulk mass was performed (see below).

The third step in this procedure was to isolate fiber bundles on two separate portions of the CT muscle (PR and PO) and on the TA muscle (vocalis and muscularis). The vocalis and muscularis portions were not easily defined; therefore, there was no distinction made between the two portions in the TA. Muscle fiber bundles were defined by removing the thin layer of fascia and gently rubbing the muscle across the grain to separate the various fascicles. The fiber bundles were removed according to muscle definition and ease of separation. Superficial layers were separated and removed first to expose deeper layers of bundles. Three-dimensional coordinates of origin ($X1, Y1, Z1$) and insertion ($X2, Y2, Z2$) were recorded for each fiber bundle.

Mass Measurement

Each dissected muscle fiber bundle was weighed on an electronic balance (Mettler model AE 100, 0.1-mg accuracy) after removal from the muscle. The mass of each muscle fiber bundle was recorded and accumulated to represent the total muscle mass of the CT muscle, PR and PO bellies of the CT muscle and the TA muscle. The contralateral muscles were also weighed on the balance and recorded for five of the six human larynges and three of the three canine larynges.

Length and Angle Calculations

The length of each dissected muscle fiber bundle (L) was determined from the three- dimensional coordinates of origin and insertion points as

$$L = \sqrt{(X2 - X1)^2 + (Y2 - Y1)^2 + (Z2 - Z1)^2}$$

After the lengths were recorded for each muscle fiber, the average length was determined for the complete muscle. The angle θ of each muscle bundle in XY plane relative to was calculated from

$$\theta = \tan^{-1}\left(\frac{Y2 - Y1}{X2 - X1}\right)$$

In other planes such as YZ plane, it was calculated similarly (see Figures 1 and 2). Then the range of angles for the complete muscle and its average was calculated for CT, TA, PR and PO muscles.

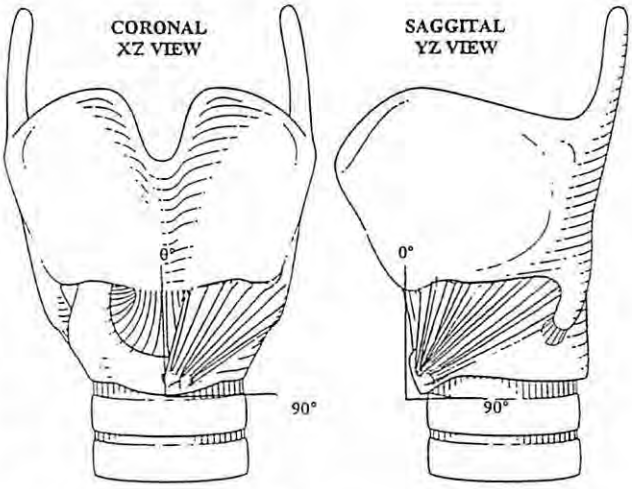


Figure 1. XZ and YZ views of the thyroarytenoid muscle.

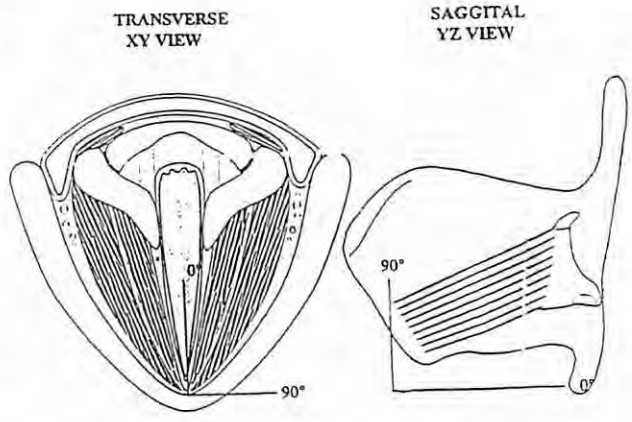


Figure 2. XZ and YZ views of the cricothyroid muscle.

Area Calculation

The cross-sectional area of each bundle, A was calculated from

$$A = \frac{m}{\rho L},$$

where m is the bundle mass, ρ is the density of muscle tissue, and L is the bundle length. The area of each muscle fiber bundle was recorded and added together to represent the total muscle area of the complete muscle.

Statistical Analysis

The CT muscle, PR, PO and TA muscle were measured with respect to average length, total muscle mass and total cross-sectional area of the muscle. In order to account for the discrepancies in sample's geometric structure and make appropriate between-group comparisons, a nonparametric statistical test with the level of significance

value of p=0.05 was used (Mann-Whitney U Test[14]). It compared these muscles in the human and canine on their characteristics: 1) Muscle mass, 2) CT/TA mass ratio, 3) Muscle length, 4) Muscle cross-sectional area, 5) CT/TA area ratio, and 6) Orientation angle.

Results

Comparative Mass Results

The total mass of CT, its PR and PO bellies and TA muscles were measured for six human larynges and three canine larynges. The total mass of the contralateral muscle was also measured for five of the human larynges and three of the canine larynges. Figure 3 shows that average total mass of the human CT muscle (denoted HCT on the bar graph) is 0.9423 g (SD=0.1228 g). The average total mass of the canine CT muscle (denoted CCT) is 1.4632 g (SD=0.2802 g). Results of the Mann Whitney U Test revealed significant between-group differences (p=0.0155) in total mass for the CT muscle. Therefore the canine CT muscles used in this study (large mongrel dogs) had about 0.5 g more mass than the human CT muscles. Similarly, the canine TA muscles had about 0.56 g more mass than human TA (p=0.001).

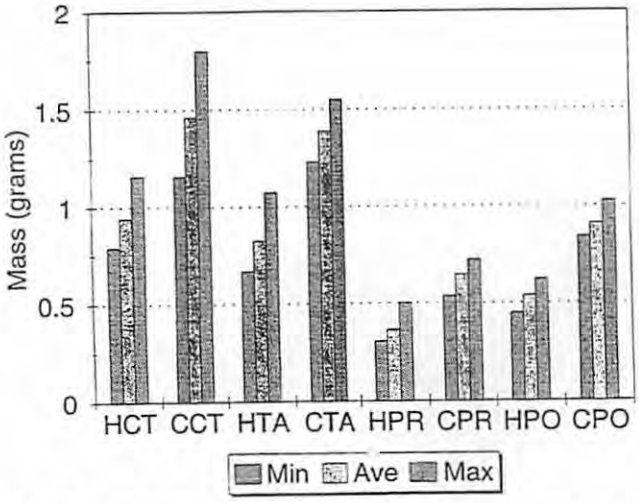


Figure 3. Comparative mass of CT and TA muscles. On the X-axis, the first character H is for human, C for canine. The second and third characters identify the muscle (TA for the thyroarytenoid, CT for the cricothyroid, PR for the pars recta, and PO for the pars obliqua). The legend Min stands for minimum, Ave for average, and Max for maximum.

Comparative Length Results

The average lengths of the CT muscle, its PR and PO bellies, and TA muscles were calculated for six human larynges and three canine larynges. Figure 4 shows that the average length of the human CT muscle is 13.8 mm (SD=1.0 mm). The average length of the canine CT muscle is 15.2 mm (SD=0.4 mm). There is a significant between-group difference (p=0.0478) in average length for the CT muscle; the canine CT muscles used in this study were longer than the human CT muscles by about 1-2 mm. Similarly, the canine TA muscles were about 3.6 mm longer than human TA muscles (p=0.0476).

Comparative Cross-sectional Area Results

Figure 5 shows that the average cross-sectional area of the human CT muscle is 73.8 mm^2 (SD=7.4 mm^2) and the average cross-sectional area of the canine CT muscle is 105.3 mm^2 (SD=11.6 mm^2). There is a significant between-group differences (p=0.0238) in total muscle cross-sectional area for the cricothyroid muscle, with the canine CT cross-sectional area being about 30 mm^2 greater than the human. Similarly, the canine TA muscle with average cross-sectional area of 63.8 mm^2 is about 50 % thicker than human TA with 40.9 mm^2 cross-sectional area (p=0.0238).

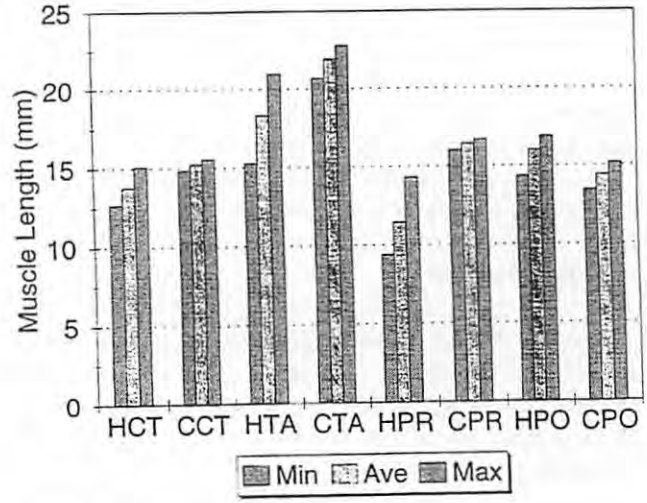


Figure 4. Comparative length of the TA and CT muscles.

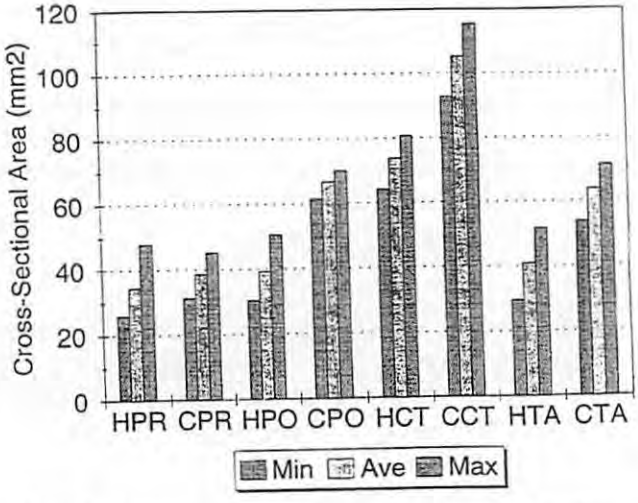


Figure 5. Comparative cross-sectional area of TA and CT muscles.

Comparative Mass and Cross-sectional Area Ratio

The ratios of total CT muscle mass to total TA muscle mass were calculated to eliminate the effects of muscle bulk and atrophy due to the age of each specimen. Ratios were also useful in comparing the relationship between the total mass of the two agonist/antagonist muscle pairs within each larynx. Mass ratios were calculated for six human larynges and the three canine larynges. The ratios of the contralateral muscles were also calculated for five of the human larynges and three of the canine larynges. Figure 6 shows that the average CT/TA mass ratio of the human muscles is 1.10506 (SD=0.0999). The ratio for the canine is 1.0466 (SD=0.1078) which does not show any significant between-group difference (p=0.119). Cross-sectional area ratios were also calculated for six human larynges and the three canine larynges and shown in the same Figure. The CT/TA cross-sectional area ratio of the human muscles is 1.8 (SD=0.3). This ratio for the canine is 1.7 (SD=0.1), showing no significant between-group differences (p=0.381).

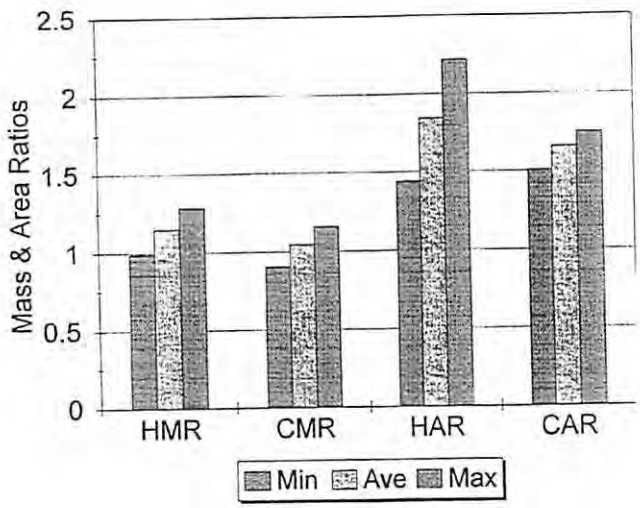


Figure 6. Comparative mass and area ratio in TA and CT muscles. MR and AR on the X-axis stand for mass and area ratio of CT muscle to TA muscle.

Comparative Angle Results

Figure 7 shows that the average orientation angle (from vertical position) of the pars recta (PR) belly of the human cricothyroid is 31.2° (SD=2.5°) in the XZ plane (coronal). The same angle for the canine is 25.5° (SD=2.8°) with a significant between-group differences (p=0.0238) of about 5-6 degrees. This means that the canine CT muscle lies more vertical. The pars oblique bellies of human and canine have average values of 47.2° (SD=9.3°) and 38.3° (SD=6.6°) in the same plane which indicate human PO muscle is more horizontal than canine.

These muscles have different angles in YZ plane (midsagital). Figure 8 shows that the pars recta belly has an average angle of 24.1° for human and 8.2° for canine samples. This is a significant difference (p=0.0476) that indicate the canine PR muscle is more vertical in mid-saggital plane than human PR muscle. These angles for human and canine pars oblique muscles are 63.9° and 40° showing the human pars oblique is more horizontal in the YZ plane.

Figure 9 (following page) shows the average angles of TA muscles of the human and canine samples. These angles are based on the measurement in the XY plane (transverse view) and saggital YZ view as shown in the Figure 2. The average angle for the human TA is 11.0° (SD=3.3) and for the canine TA is 17.4° (SD=6.3) which indicates canine vocal folds at equilibrium position is more abducted than human vocal folds. The average TA angles in the saggital plane are 13.9° and –0.8° indicating that canine

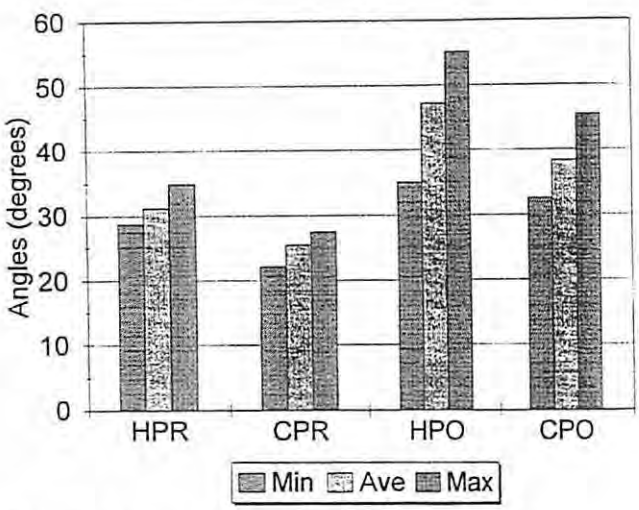


Figure 7. Orientation angles of CT muscle in XZ plane.

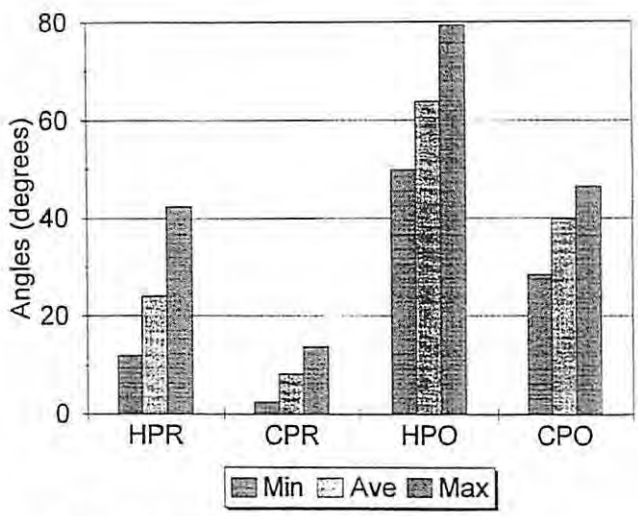


Figure 8. Orientation angles of CT muscle in YZ plane.

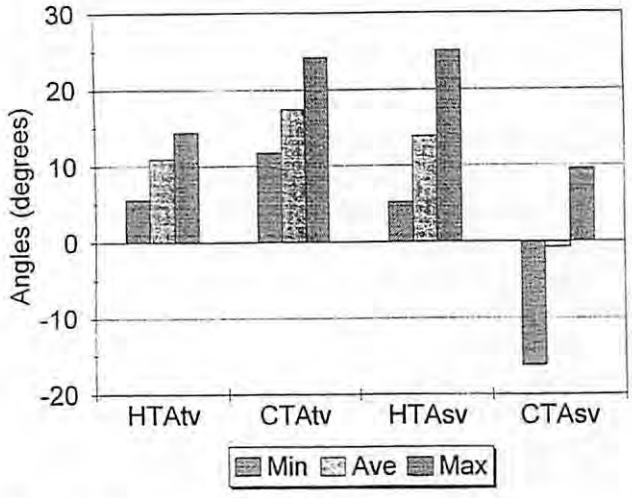


Figure 9. Orientation angles of TA muscle in XY and YZ planes.

TA is almost horizontal but human TA is tilted upward from anterior to posterior. However these TA angle differences are not statistically significant.

Discussion

Mass and cross-sectional area were significantly larger in our canine larynges than in our human larynges. However, the mass and cross-sectional area ratios show that the relationship between these two muscles is similar in the two species. If we assume that all the muscle bundles are equally activated, the force produced by each bundle is proportional to its cross-sectional area. Studying the mass and cross-sectional area ratios allows us to predict that the cricothyroid is capable of producing greater forces in the elongation and rotation processes[11]. Cross-sectional area calculations, however, assume that the fiber bundles are of uniform diameter along their length. This was not necessarily the case. The fiber bundles are thinner at their origin and insertion than in the center. This should be considered for modeling procedures in the future.

When comparing the mass measurements in this study to those of Faaborg-Andersen[2], the total mass of the cricothyroid was remarkably similar (942 mg compared to 954 mg for human CT). Cross-sectional area calculations were also similar (74 mm^2 compared to 60 mm^2) for the cricothyroid muscle. The mass of the thyroarytenoid was not similar because Faaborg-Anderson[2] used the thyrovocalis portion only, while our study used the whole thyroarytenoid (823 mg compared to 318 mg).

Length

The length measurements in this study are within the range of length measurements reported by Hirano et al[15]. Their length measure included both the membranous and cartilaginous portions. Their data suggested that the vocal fold length range from 17 to 21 mm in males, whereas in this study the average length of the human thyroarytenoid muscle was 18.3 mm. The canine thyroarytenoid muscle in this study was 21.9 mm, which is longer than the range of 15-19 mm reported by Alipour & Titze[16]. The exterior portion of the cricothyroid muscle has longer muscle fiber bundles than the deeper portion of the muscle. Therefore, the difference in length measurements with those reported by Alipour & Titze[16] is likely due to the fact that their samples were prepared from the medial portion of canine TA muscle. Overall, the lengths of the cricothyroid and thyroarytenoid muscles were significantly longer for our canines than our humans. The exception was the pars obliqua belly of cricothyroid. Although not statistically significant, it was longer in the human than in the canine.

Orientation and Pitch Control Issues

Differences between human and canine orientation angles of these muscles may be due to phylogenetic effects within the human race[17,18]. A more horizontal orientation of both of the cricothyroid bellies (PR & PO) may contribute to more backward displacement or retraction of the cricoid arch[10,11]. Many investigators[17-21] believe that this ventrodorsal sliding significantly affects changes in vocal fold length. It can be hypothesized then that the human mechanism is more capable than the canine mechanism to produce dramatic changes in length. Perhaps that is why the human instrument is more versatile and can produce octave jumps in pitch within a second during singing. It may be noteworthy that the orientation of the thyroarytenoid was not significantly different between the human and canine specimens.

Important Similarities and Differences

Pressmen & Kelemen[22] stated that identifying certain details in the comparative anatomy of the larynx is of great importance to the physiologist, especially if laboratory experimentation is concerned. The greatest part of our scientific knowledge of the physiology of the larynx comes from observations of laboratory animals. Considering this fact, it is important to identify these differences in structure from one species to the next, even though they may appear very similar. Differences in innervation, quantitative muscle relationships, patterns of cartilaginous structure, and position and size are important. These differences contribute to the difficulty of applying knowledge gained by animal experimentation to human conditions.

The larynx of large canines is similar to the human larynx in terms of size and vocal fold histology; however, there are a few additional differences related to the two muscles that may be noteworthy to mention after dissection. An obvious difference was the absence of the vocal ligament in the canine. Titze et al.[4] comment that the canine larynx is

not an ideal model of the human larynx when discussing pitch control and vocal fold tissue morphology because the vocal ligament is absent. After examining both the human and canine cricoid cartilages and their relationship to the cricothyroid muscle, another obvious difference is noted. The human cricothyroid muscle appears to be set into a groove within the cricoid cartilage and the canine cricothyroid muscle appeared to be more exposed and bulky. Perhaps this could be related to body proportion differences within the two species. The neck of a canine is large and muscular in proportion to the rest of its body, while the neck of a human is proportionally much thinner. This could be related to each species' normal anatomical position and gravitational effects on that position. Perhaps it could also be related to the uniqueness of the human muscular framework that results in a more versatile and specialized instrument.

In spite of these differences, much of the information on the mechanics of vocal fold vibration has emerged from the study of excised canine larynges. A number of studies suggest that excised larynges do not accurately reproduce physiologic conditions of vocal fold tension and mass during vibration[23]. Moore & Berke[8] believe that in vivo canine models appear to be better than excised larynges because they maintain blood flow and intrinsic laryngeal muscular tension while preventing postmortem deterioration of the tissue. Choi et al.[9] also suggests that "despite differences in structure between human and canine larynges, in vivo canine studies are well suited for examining questions about fundamental frequency in phonation, because the basic gross function of major laryngeal muscles are quantitatively the same for the two species." Hast[24] was a strong supporter of the use of the canine larynx and commented that the physiological similarity of the canine supports the canine's value for research on the mechanisms of the larynx.

Moore & Berke[8] recognized that "the use of animals to study laryngeal function provides a setting in which new concepts can be tested, while at the same time allowing manipulation of variables not easily controlled in humans." Traditionally, investigators have used the canine as their principal animal model on which laryngeal studies have been based. That is why it is so important to know if the canine model truly represents a human when discussing pitch control.

Conclusions

Mass and cross-sectional area ratios of the human and canine specimens in this study suggest that the basic gross function of the two muscles is the same for the two species. However, the differences in fiber orientation could alter the biomechanical function between the two muscles. Data collected in this study from the cricothyroid muscle will be useful for calculating the torque acting on the cricothyroid joint to change vocal fold length. These calculations will help establish the functional relationship between the muscles in pitch control and provide a three-dimensional picture of the dynamics of rotation. The data will also be used for a three-dimensional reconstruction of the muscular attachments in the laryngeal framework. As always, it would be beneficial to have more subjects of different age and sex to interpret differences in laryngeal control in the developmental and aging periods.

Acknowledgment

The authors would like to thank Roger Chan for assistance in the sample preparation. This work was supported by the National Institute on Deafness and Other Communication Disorders, National Institute of Health (grant # P60 DC00976), and performed in accordance with the PHS Policy on Human care and Use of Laboratory Animals, the NIH Guide for the Care and Use of Laboratory Animals, and the Animal welfare Act (7U.S.C. et seq.); the animal use protocol was approved by the Institutional Animal Care and Use Committee (IACUC) on 10-22-1993, The University of Iowa.

References

1. Titze IR. Phonation threshold pressure: a missing link in glottal aerodynamics. J Acoust Soc Am 1992;91(5):2926-2935.

2. Faaborg-Andersen, K. Electromyographic investigation of intrinsic laryngeal muscles in humans. Acta Physiol Scand 1957;41(Suppl 140).

3. Hirano M, Vennard W, Ohala J. The function of laryngeal muscles in regulating fundamental frequency and intensity of phonation. J Speech Hear Res 1969;12:616-28.

4. Titze IR, Luschei ES, Hirano M. Role of the thyroarytenoid muscle in regulation of fundamental frequency. J Voice 1989;3:213-24.

5. Larson CR, Kempster GB. Voice fundamental frequency changes following discharge of laryngeal motor units. In: Titze IR & Scherer RC, Eds. Vocal Fold Physiology: biomechanics, acoustics, and phonatory control. The Denver Center for the Performing Arts, Denver, CO, 1983:91-104.

6. Hirano M. Morphological structure of the vocal cord as a vibrator and its variations. Folia Phoniat 1976;26:89-94.

7. Rubin H. Experimental studies on vocal pitch and intensity in phonation. Laryngoscope 1963;73:973-1015.

8. Moore DM, Berke GS. The effect of laryngeal nerve stimulation on phonation: a glottographic study using an in vivo canine model. J Acoust Soc Am 1988;83:705-15.

9. Choi HS, Berke GS, Ye M, Kreiman J. Function of the thyroarytenoid muscle in a canine laryngeal model. Ann Otol Rhinol Laryngol 1993;102:769-76.

10. Arnold GE. Physiology and pathology of the cricothyroid muscle. Laryngoscope 1961;71:687-753.

11. Alipour-Haghighi F, Perlman AL, Titze IR. Tetanic response of the cricothyroid muscle. Ann Otol Rhinol Laryngol 1991;100(8):626-631.

12. Kempster G, Larson C, Kistler M. Effects of electrical stimulation of cricoarytenoid and thyroid muscles on fundamental frequency. J Voice 1988;2:221-9.

13. Negus BE. The Comparative Anatomy and Physiology of the Larynx. Grune & Stratton , New York, NY 1949.

14. Siegel S. Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill, New York, NY 1956.

15. Hirano M, Kurita S, Nagata K. Growth, development and aging of human vocal folds. In: Bless DM & Abbs JH, Eds. Vocal Fold Physiology. College-Hill Press, San Diego, CA, 1983:22-43.

16. Alipour-Haghighi F, Titze IR. Elastic models of vocal fold tissues. J Acoust Soc Am 1991;90(3):1326-31.

17. Sonninen A. The role of the external laryngeal muscles in length-adjustment of the vocal cords in singing. Acta Otolaryngol (Stockh) 1956;(Suppl 130):1-102.

18. Stone RE, Nuttall AL. Relative movements of the thyroid and cricoid cartilages assessed by neural stimulation in dogs. Acta Otolaryngol (Stockh) 1974;78:135-40.

19. Hollien H. In search of vocal frequency control mechanisms. In: Bless DM & Abbs JH, Eds. Vocal Fold Physiology. College-Hill Press, San Diego, CA, 1983:361-7.

20. Vilkman E. An apparatus for studying the role of the cricothyroid articulation in the voice production of excised human larynges. Folia phoniat 1987;39:167-77.

21. Zaretsky LS, Sanders I. The three bellies of the canine cricothyroid muscle. Ann Otol Rhinol Laryngol 1992;101:3-16.

22. Pressman J, Kelemen G. Physiology of the Larynx. Arch Otolaryngol 1955;35:506-54.

23. Perlman AL, Titze IR, Cooper DS. Elasticity of canine vocal fold tissue. J Speech Hear Res 1984;27:212-219.

24. Hast MH. The primate larynx: a comparative physiological study of intrinsic muscles. Acta Otolaryngol (Stockh) 1969;67:84-92.

NCVS Status and Progress Report - 13
April 1999, 71-82

Contributions of Individual Muscles to the Submental Surface Electromyogram During Swallowing

Phyllis M. Palmer, Ph.D.
Erich S. Luschei, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa
Debra Jaffe, M.D.
Timothy M. McCulloch, M.D.
Department of Otolaryngology, The University of Iowa Hospitals and Clinics

Abstract

Submental surface electromyographic recordings are commonly used in the investigation of swallowing disorders. The measured electromyography is thought to reflect the actions of floor-of-mouth muscles. While this is a reasonable assumption, to date there have been no investigations to delineate which muscles contribute to this surface recording. The primary goal of this experiment was to determine which muscles contribute most to the submental surface. Electromyography was recorded simultaneously from the submental surface as well as from five individual muscles: mylohyoid, anterior belly of the digastric, geniohyoid, genioglossus and platysma. Three analysis methods were performed to estimate individual muscle contributions: correlation, numeric and analytic. For the numeric and analytic analyses, a linear model was defined and used to represent the relationship between the surface and intramuscular recordings. Muscles that received a high correlation, numeric and/or analytic value were considered to be primary contributors to the submental recording. Regardless of analysis approach, the primary contributions to the submental surface recording were the mylohyoid, anterior belly of the digastric, and the geniohyoid muscles. Contributions from the genioglossus and the platysma muscles were minimal. Contributions as a function of bolus volume and viscosity are also discussed.

Introduction

Electromyography (EMG) has been utilized to estimate muscle activity in people with normal and disordered swallowing. Electromyographic signals are detected by using electrodes that are placed either intramuscularly, to provide information about an individual muscle, or on the surface of the skin near the muscle or muscles of interest. Using the latter detection method, activity may be recorded from a number of muscles in the region of the electrode (Koole, de Jongh, & Boering, 1991; Perry, Easterday & Antonelli, 1981). In the investigation of swallowing and its disorders, typically surface electrodes are used and placed on the submental skin's surface to measure muscle activity (see for example, Barofsky, 1995, Bryant, 1991; deLarminat, Montravers, Dureuil, & Desmonts, 1995; Huffman, 1978; Shaker et al., 1995; Sukthankar, Reddy, Canilang, Stephenson & Thomas, 1994). These surface recordings serve as an *estimate* of the combined activity of muscles in that region. Due to the proximity of floor-of-mouth musculature to the submental surface electrodes, a burst in EMG activity recorded from the submental region is considered to be the result of the combined activity from individual floor-of-mouth muscles, specifically the mylohyoid, geniohyoid, and anterior belly of the digastric muscles. While there are data available regarding the individual actions of these various muscles during the swallow (Doty & Bosma, 1956; Gay, Rendell, & Spiro, 1994, Hrycyshyn & Basmajian, 1972), the contribution of these individual muscles to the submental surface recording has not been investigated. In the absence of

experimental evidence, investigators that use submental surface EMG must consider that the recorded activity may not solely reflect the behavior of the aforementioned muscles. Other muscles in the submental region may also provide input to the recordings obtained from the submental surface. For example, due to their close proximity to the submental surface recording site, the platysma or genioglossus muscles may contribute to signals recorded from the submental region.

Both bolus volume and viscosity have been shown to alter the submental recording. Cook and colleagues (1989) reported that the duration and peak of the submental recording increased with bolus volume. Ertekin and colleagues (1997) confirmed these findings. Contrary to the above two investigations are the data of Dantas and Dodds (1990). They evaluated bolus volumes from 2 ml to 20 ml and did not find any volume-specific increase in the duration of the submental EMG.

Although less often investigated, the results from studies that have evaluated the effects of bolus viscosity are more consistent across studies. With few exceptions, regardless of the instrumentation used or variables evaluated, previous investigators have reported statistically significant alterations in the swallow behavior as a function of bolus viscosity (Dantas & Dodds, 1990; Ertekin et al., 1997; Pouderoux & Kahrilas, 1995; Reimers-Neil, Logemann & Larson, 1994; Ren et al., 1993). With regard to submental EMG, increased viscosity resulted in increased magnitude and duration of the recorded signal (Dantas & Dodds, 1990; Ertekin et al., 1997; Reimers-Neil et al., 1994).

The purpose of this experiment was to examine the relative contributions of five individual muscles in the submental region to the electromyographic recordings obtained from the submental surface during various swallowing tasks. The contributions of various muscles to the submental surface recording were explored by performing EMG recordings from individual muscles while simultaneously recording EMG from the submental surface. EMG contributions were evaluated as a function of bolus volume and viscosity. A secondary goal of this study was to test a linear model that had been developed to represent the relationship between the individual muscles and the surface recording.

Methods

Subjects

Seven healthy adult volunteers without any history of speech or swallowing disorders between the ages of 21 and 37 participated in this investigation. Male subjects were instructed to shave their chin and neck region prior to their participation. All procedures were performed in accordance with the policies of the Human Subjects Committee of the University of Iowa.

Procedure

Upon completion of human subject consent forms, subjects received a brief head and neck exam performed by an otolaryngologist. No head and neck abnormalities were identified. Subjects were then seated in a semi-reclined position. After scrubbing the skin on the forehead, a reference electrode was adhered to the forehead using prefabricated monitoring electrodes. The wire lead from the monitoring electrode was connected to the reference input of an electrically isolated preamplifier.

Electrode Placement

Bipolar hooked-wire electrodes were prepared using 0.002 inch diameter insulated stainless steel wires that were passed through the shaft of a hypodermic needle (0.5 inch, 26 gauge needle or 1.5 inch, 25 gauge needle). Insulation was removed from 1.5 mm of the ends of the wires protruding from the tip of the needle; these ends were then bent back to form hooks. Electrodes were sterilized with an autoclave prior to use.

Bipolar hooked wire electrodes were used in a near-field configuration, and inserted into the following muscles using the following insertion criteria.

Mylohyoid Muscle. Using a technique similar to Lehr et al. (1971), insertion for the mylohyoid muscle was made approximately 3.3 cm posterior to the genium, slightly off midline. The needle was inserted in a shallow and lateral direction with an average depth insertion of 1.2 cm

Geniohyoid Muscle. In a site next to the insertion for mylohyoid, bipolar hooked-wire electrodes were placed into the geniohyoid muscle using a transcutaneous insertion similar to Hirose (1971). The needle was advanced slightly lateral with an average depth of 2.7 cm.

Anterior Belly of the Digastric Muscle. While asking the subject to open his/her mouth against resistance, the anterior belly of the digastric muscle was palpated and the electrode was placed (Hirose, 1971). Electrode insertion occurred approximately 3.4 cm posterior to the genium and approximately 1.6 cm from the midline. The average depth of insertion was 1.3 cm.

Genioglossus Muscle. Using a technique similar to Sauerland et al. (1981) a needle was inserted at midline into the under-surface of the tongue and advanced in a lateral direction. The electrode wires were routed through the buccal cavity and diverted out the corner of the mouth.

Platysma Muscle. Bipolar hooked wire electrodes were implanted above the surface of the hyoid bone, approximately 3-4 cm lateral of the midline. The electrodes were placed just beneath the surface of the skin into this thin muscle layer with an average insertional depth of 0.5 cm.

Table 1.
Muscle-Specific Tasks Used to Characterize Electrode Placement

- Open mouth
- Tongue to alveolar ridge
- Protrude tongue
- Tongue to nose
- Retract lips up (big smile)
- Say /pip/
- Open mouth against resistance
- Tongue to chin
- Retract tongue
- Retract lips down (frown)
- Say /kik/
- Water swallow

Placement of intramuscular electrodes was characterized by having the subject perform a series of muscle-specific tasks. Performance on these tasks was graded and compared to previous investigations to note comparisons. If data obtained from a particular electrode were considered uncharacteristic of that muscle, then another insertion attempt was performed. Table 1 includes a list of the tasks used to characterize the above muscles.

Bipolar surface electrodes (Beckman surface electrodes with a diameter of 11 mm) were placed just lateral to each side of the midline of the submental region, and posterior to the mental protuberance approximately 40% of the way between the midpoint of the mentis and the hyoid bone. The center-to-center inter-electrode distance was 1 cm.

Tasks Performed

Five trials of five swallow tasks (3 ml water bolus, 10 ml water bolus, 20 ml water bolus, 3 ml bolus of applesauce, and a 3 ml bolus of peanut butter) were performed by each subject in random order. For all tasks except peanut butter, quantities were measured and delivered by the experimenter using a syringe. For peanut butter swallows, quantities were measured and placed on a spoon. The subject was handed the spoon and asked to place the bolus in the oral cavity and hold it there until asked to swallow.

Data Acquisition

Electromyographic signals were high pass filtered (30 Hz) and amplified using a variable gain amplifier (B466C Speech Physiology System, University of Iowa Bioengineering). After amplification, the signals were low pass filtered (2500 Hz, 8 pole Butterworth filter) to prevent aliasing. Data were digitized at a sampling frequency of 5000 Hz using WINDAQ (DATAQ Instruments, Inc., Akron, Ohio) which was operating on a personal computer. Auditory monitoring of individual channels of EMG signals was possible throughout the acquisition period.

Data Analysis

Acquired data files were reviewed using WINDAQ software (DATAQ Instruments, Inc., Akron, Ohio). Individual tokens were identified and saved into separate data files. Processing of EMG signals was performed using Matlab (Mathworks, Inc., Englewood Cliffs, N.J.). For each experimental token, amplification gains that were applied during data collection were removed. The signals were rectified by taking the absolute value of each data point. Rectified signals were smoothed using a 20 Hz low pass sixth-order butterworth filter function.

Three methods were used to assess the contribution of the intramuscular EMG signals to the signal recorded from the submental surface: correlation, analytic and numeric.

Method 1: Correlation Analysis

For the correlation analysis, the filtered and rectified signals of a given subject were normalized to the maximum value recorded across all tasks. A zero-lag cross correlation was performed between the surface recording and each of the individual muscle recordings. Swallow tasks differed by both volume and viscosity. Analysis of variance was performed to identify any significant task differences across the correlation values. For tasks that were shown to be statistically different, post-hoc testing was performed. In this investigation, only bolus viscosity was shown to have a significant effect on the outcome of the correlation values.

To further evaluate the effect of bolus viscosity on the contribution of individual muscles to the submental surface recording, peak values for each muscle were extracted from each token. Also, integration was performed on each token to serve as a gross measure of overall muscle activity. Integration and peak values were averaged across the seven subjects and compared across the three viscosities.

Methods 2 and 3: Analytic and Numeric Analyses

The Linear Model. It was hypothesized that the electromyographic activity recorded from each individual muscle represented a portion of the EMG signal recorded at the submental surface. Bouisset and Maton (1972) compared integrated waveforms of surface EMG and intramuscular EMG from the biceps brachii muscle. They noted a linear relationship between the two recording methods. In support of the work by Bouisset and Maton, as well as the desire to use a parsimonious approach, a linear model of submental activity was defined as follows:

$$\hat{E}_{sm}(t) = \sum_{i=1}^{5} w_i e_i(t) \qquad (1)$$

where i represents the five muscles, $e_i(t)$ are the time-dependent EMG signals from the intramuscular electrodes, and w_i represents the weighting coefficient for the activity of a specific muscle.

How well the modeled submental ($\hat{E}_{sm}(t)$) matches the recorded submental signal ($E_{sm}(t)$) can be assessed by summing the difference between the two across the time of the swallow token, yielding an error measure as follows.

$$error = \sum_{t=1}^{n} [E_{sm}(t) - \hat{E}_{sm}(t)]^2 \quad (2)$$

A minimization of this error measure can be used to search for the best combination of weights for the individual muscles.

Data Normalization. In EMG studies, the filtered and rectified signals of a given subject are typically normalized to the maximum activity recorded across all tasks (e.g., Palmer, Perlman, VanDaele & McCulloch,1993; Perlman, Luschei & Dumond, 1989). In fact this was the method used in the correlation analysis. However, for the linear model used in this experiment (see Equation 1) the assumption was that the time-dependent shape of the five EMG signals for a given task can be scaled by the weights and summed to give a close approximation of the surface recording. Thus, it is desired that *all* of the scaling be contained within the weighting values. The typical method of normalization by maximum activity across tasks would impose an undesired scaling factor prior to the weight determination. Instead, each experimental token, after being filtered and rectified, was normalized to its own largest amplitude value giving each signal in a token a final maximum value of 1. This same normalization was also applied to the surface EMG signal.

Analytic Analysis. The model defined in Equation 1 contains five unknown variables: the weight values of each of the five input muscles. Thus five equations are required to perform an analytical solution. By combining Equation 1 and Equation 2, error can be redefined as follows:

$$error = \sum_{t=1}^{N} [E_{SM}(t) - w_1 e_1(t) - w_2 e_2(t) - w_3 e_3(t) - w_4 e_4(t) - w_5 e_5(t)]^2 \quad (3)$$

where $e_i(t)$ represents the processed EMG from the i^{th} muscle during the t^{th} time sample, and w_i represents the weight of that muscle.

If error is differentiated with respect to the five weight values, the derivative of error with respect to w_1 could be expressed mathematically as follows:

$$\frac{\partial error}{\partial w_1} = \sum_{t=1}^{N} [E_{SM}(t) - w_1 e_1(t) - w_2 e_2(t) - w_3 e_3(t) - w_4 e_4(t) - w_5 e_5(t)] e_1(t) \quad (4)$$

Similarly, error is differentiated with respect to w_2 through w_5 resulting in five equations.

As the goal is to identify the best fitting weights for the assumed solution (Equation 1), Equation 4 is solved to produce a minimum error. Minimum error with respect to w_i is achieved when the differential of error with respect to w_i is 0. Thus, to calculate the minimum error, Equation 4 is set to equal 0. Once simplified, the equation can be represented as follows:

$$\sum_{t=1}^{N} e_y(t) E(t) = w_1 \sum_{t=1}^{N} e_y(t) e_1(t) + w_2 \sum_{t=1}^{N} e_y(t) e_2(t) + w_3 \sum_{t=1}^{N} e_y(t) e_3(t) + w_4 \sum_{t=1}^{N} e_y(t) e_4(t) + w_5 \sum_{t=1}^{N} e_y(t) e_5(t) \quad (5)$$

By substituting y with the numbers 1 through 5, five equations can be defined.

When one considers the formula for raw correlation values,

$$C_{xy} = \sum_{t=1}^{N} x(t) y(t) \quad (6)$$

then a relationship between correlation and the linear model becomes evident. Weight values noted in the linear model are a summation of correlation values. A matrix representing the five simultaneous equations can be used to calculate the five weights and can be expressed as follows:

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} \\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} \\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} \\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} \end{bmatrix}^{-1} \begin{bmatrix} C_{1E} \\ C_{2E} \\ C_{3E} \\ C_{4E} \\ C_{5E} \end{bmatrix} \quad (7)$$

where C_{ij} represents the correlation between two signals, the numbers 1 through 5 represent the five intramuscular recordings, and E represents the signal recorded from the submental surface. Using this set of equations it becomes clear that the weights can be calculated using a multiple regression approach.

A Numeric Analysis Utilizing Optimization. Optimization provides an automated method of searching for the best-fitting values or weights for undefined variables in an equation or model. For this experiment, the linear model used during this searching procedure is displayed in Equation 1. Weight values for the intramuscular recording were optimized to produce the smallest error between $E_{sm}(t)$ and $\hat{E}_{sm}(t)$ as noted in Equation 2.

In the optimization of weights for this model, weight values were constrained to range from 0 to 1.0. By defining a lower boundary of zero, negative coefficients could not be obtained. Negative coefficients were not permitted based on the assumption that an individual muscle

can contribute to the submental recording, but it can not take away from the recorded surface activity. The definition of an upper boundary, which was defined to be 1.0, was based on the assumption that during a strong contraction, the activity measured with the surface electrode would not exceed the recording obtained from the intramuscular electrode (Person, 1963). Thus it was assumed that these constraints would make the model more physiologically relevant.

Comparison of Analytic and Numeric Analyses. When a good optimization algorithm is devised, this procedure can be used to verify that the analytic solution has indeed identified the best weights. In this scenario, these two methods would produce similar results. However, when the accuracy of the optimizer is reduced either by defining a poor optimizer or using poor start weights, then the results generated from optimization will differ from the analytic solution. The most common error in optimization is the identification of a local minimum, which is assumed by the optimizer to be a global minimum.

Testing the Model

Across all experimental tokens, two subgroups of tokens were defined. The first subgroup of tokens was comprised of tokens 1, 3 and 5 from each of the five tasks across all seven subjects. Data obtained from this subgroup were averaged and used as weights for testing the model. The second subgroup of tokens, which was comprised of tokens 2 and 4 from each of the tasks across all subjects, was used to test the weights in the model. Thus, average values from the first subgroup of tokens were used as weights in the second subgroup of tokens. Using Equation 1, this procedure yielded a modeled submental waveform. For a given token, the modeled waveform was correlated with the actual submental recording. In addition to correlation values, an error value was calculated to quantify the difference between the actual submental recording and the submental waveform generated by the model. Correlation and error values for the two subgroups of tokens were compared.

Development of Weights. Three sets of weights were developed using data obtained from the three analysis methods. As noted above, for each analysis method weights were generated by averaging values from tokens 1, 3, and 5 across all five tasks and all seven subjects. This resulted in a single set of weights (one for each muscle) for each analysis method across all subjects and tasks. These weights were applied to token subgroup 2.

Using data from the correlation method, average correlation values for tokens 1, 3, and 5, could range from −1.0 to 1.0. Negative values were adjusted to zero based on the model's assumption that while an individual muscle can contribute to the submental recording, it can not take away from the recorded surface activity. All positive correlation values were retained. These adjusted average correlation values served as weights to test the model.

Although it was possible to obtain values outside the desired range of 0 to 1 using the analytic analysis, average values did not exceed these physiologic limits, and so no adjustments were made. The average analytically-derived values served as weights to test the model. For the optimization method, values were constrained such that no values could go outside the predetermined physiologic range of 0 to 1. Therefore, average optimization values were used as weights without any adjustment.

Calculation of the Error Value. The error value was defined as the sum of the squared difference between the modeled and actual submental activity for each data point in a token. Error was calculated using Equation 2. To compare the error values obtained for each token across tasks and subjects, an error sum per data point was calculated by dividing the error sum for an individual token by the number of data points in that token.

Results

Correlation Analysis

Average correlation values obtained between the submental recording and each of the individual muscles are displayed in Table 2 and ranged from 0.77 for the mylohyoid muscle to −0.06 for the genioglossus muscle. Based on these average correlations the primary contributors to the submental surface recording were the mylohyoid, geniohyoid, and anterior belly of the digastric muscles. Analysis of variance revealed that correlation values obtained between the submental recording and each of the individual muscles differed significantly across subjects ($p<0.0001$) and tasks ($p<0.02$).

Table 2.
Average, Minimum and Maximum Correlation Values from the Correlation Analysis, and Average, Minimum and Maximum Weight Values From the Analytic and Numeric Analyses

	MH	AD	GH	GG	PT
Correlation Group Mean	0.77	0.75	0.75	-0.06	0.35
Correlation Minimum Value	0.47	0.34	0.07	-0.81	-0.24
Correlation Maximum Value	0.95	0.94	0.98	0.62	0.83
Numeric Group Mean	0.31	0.27	0.37	0.07	0.10
Numeric Minimum Value	0.00	0.00	0.00	0.00	0.00
Numeric Maximum Value	1.00	1.00	0.97	0.55	0.49
Analytic Group Mean	0.31	0.29	0.36	0.06	0.10
Analytic Minimum Value	-0.44	-0.39	-0.46	-0.28	-0.22
Analytic Maximum Value	1.25	1.14	0.96	0.57	0.50

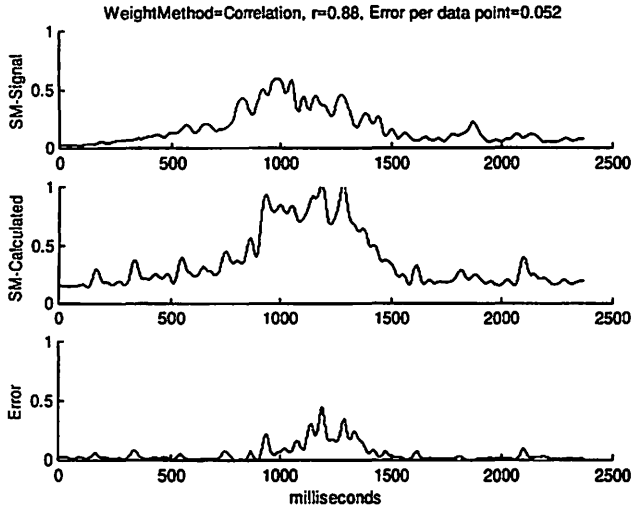


Figure 1. Average correlation values for the seven subjects. The dashed line represents the average correlation values across all subjects.

Correlation Values by Subject

For five of the seven subjects (TM, MB, CF, JB, CH) the highest average positive correlation occurred between the submental recording and the mylohyoid muscle (range = 0.88 for subject CF to 0.76 for subject TM) (Figure 1). For the remaining two subjects (SS, DJ), the highest average positive correlation occurred between the submental recording and the geniohyoid muscle (0.90 for subject DJ, 0.85 for subject SS). For six of the seven subjects (TM, MB, DJ, CF, JB, CH) the anterior belly of the digastric muscle showed a high average positive correlation with the submental surface recording (range = 0.84 for subject DJ to 0.72 for subject MB). For the remaining subject (SS), the anterior belly of the digastric muscle resulted in an average correlation of 0.59 with the submental surface. For one subject (SS), a high positive correlation was observed between the submental recording and the platysma muscle (0.70). For the remaining subjects, a positive correlation was maintained between the submental recording and the platysma muscle, however, the correlation values were weaker (range = 0.40 for subject JB to 0.15 for subject MB).

Overall, there were no strong negative correlation values between the submental recording and any of the five muscles investigated across all the subjects. However, for one subject (JB), a moderate negative correlation was noted between the submental recording and the genioglossus muscle (-0.44). For the remaining six subjects, correlation values between the submental recording and the genioglossus muscle were weaker and ranged from -0.12 for subjects CF and CH to 0.13 for subject MB.

Table 3.
Average Correlation Values for the Five Swallow Tasks

	SM-MH	SM-AD	SM-GH	SM-GG	SM-PT
3 ml Water	0.79	0.78	0.77	-0.15	0.32
10 ml Water	0.83	0.76	0.82	-0.11	0.37
20 ml Water	0.82	0.78	0.85	-0.08	0.41
3 ml Applesauce	0.71	0.71	0.66	0.06	0.30
3 ml Peanut Butter	0.70	0.72	0.64	0.00	0.34

Correlation Values as a Function of Volume and Viscosity

As displayed in Table 3, the mylohyoid, anterior belly of the digastric, and geniohyoid muscles showed similar high correlation values with the submental surface for all water swallows regardless of volume. For applesauce and peanut butter swallows, the correlation values for these three muscles fell slightly. For the genioglossus muscle, water swallows resulted in low negative correlation values, which ranged from –0.08 for 20 ml water swallows to –0.15 for 3 ml water swallows. For more viscous swallows, correlation values between the genioglossus muscle and the submental recording remained low, however correlation values were positive. For the platysma muscle, the highest average correlation between it and the submental recording occurred during 20 ml swallows (0.41), while the lowest correlation occurred during the applesauce swallows (0.30).

Analysis of variance revealed a significant task effect for correlation values between three of the individual muscles and the submental surface recording: $p<0.01$ for the mylohyoid and the geniohyoid muscles, and $p<0.03$ for the genioglossus muscle. Pair-wise comparisons revealed that for the geniohyoid muscle, peanut butter swallows differed significantly from all water swallows ($p < 0.04$). Correlation values obtained for applesauce swallows differed significantly from 10 ml and 20 ml water swallows ($p < 0.01$). For the mylohyoid muscle, significant differences were noted for applesauce versus 10 ml water swallows, peanut butter versus 10 ml water swallows, and peanut butter versus 20 ml water swallows ($p<0.04$). For the genioglossus muscle, the primary difference occurred between applesauce and 3 ml water swallows ($p<0.03$).

Correlation values that were significantly affected by changes in viscosity were analyzed further by calculating the peak values as well as the integration of the processed submental EMG. The submental activity increased in both the area under the curve (Table 4) and the peak value of the curve (Table 5) as viscosity increased from water to applesauce to peanut butter. This increased activity in the surface recording was complimented by increased activity in all muscles with regard to area under the curve of the EMG. Peak values increased across the three substances for all muscles except the geniohyoid muscle. For the geniohyoid muscle an increase was noted from the least

Table 4.
Average Integration Values (mv * ms) for 3 ml Tokens of Different Viscosities for the Submental Recording(SM), and the Five Individual Muscles

	SM	MH	AD	GH	GG	PT
Water	8.1	25.2	10.2	20.1	5.3	0.6
Applesauce	11.0	28.7	11.4	21.8	12.5	1.0
Peanut Butter	19.4	46.5	20.0	22.7	26.1	3.1

Table 5.
Average Peak Values (mv) for 3 ml Tokens of Different Viscosities for the Submental Recording (SM) and the Five Individual Muscles

	SM	MH	AD	GH	GG	PT
Water	0.035	0.130	0.052	0.085	0.033	0.003
Applesauce	0.038	0.140	0.056	0.083	0.047	0.004
Peanut Butter	0.051	0.173	0.076	0.090	0.051	0.007

viscous to the most viscous, however peak values obtained for water and applesauce were similar.

Analytic Analysis

Using a matrix of equations (Equation 7) developed by combining Equations 1 and 2, weights were calculated for each of the five muscles. Individual muscles that received a high weight were considered to be primary contributors to the submental surface recording. Average weight values for each of the individual muscles are displayed in Table 2 and ranged from 0.36 for the geniohyoid muscle to 0.06 for the genioglossus muscle. Based on these weight values, the mylohyoid, geniohyoid and anterior belly of the digastric muscles can be considered primary contributors to the submental surface recording.

Optimization Method Used for the Numeric Analysis

Individual muscles that received a high optimization weight were considered to be primary contributors to the submental surface recording. The optimization values for the mylohyoid and the geniohyoid muscles spanned the allowable range (0-1). Average, minimum and maximum values are listed in Table 2 and ranged from 0.37 for the geniohyoid muscle to 0.07 for the genioglossus muscle. Based on the average optimization values, the mylohyoid, geniohyoid and anterior belly of the digastric muscles can be considered primary contributors to the submental surface recording.

Testing the Model

Testing the Model Using Adjusted Correlation Values as Weights

Average correlation values were calculated for tokens 1, 3, and 5 from the correlation analysis. This yielded the following average correlation values for each of the five muscles: 0.755 for the mylohyoid muscle, 0.738 for the anterior belly of the digastric muscle, 0.750 for the geniohyoid muscle, -0.056 for the genioglossus muscle, and 0.349 for the platysma muscle. To eliminate negative values, the average value calculated for the genioglossus muscle was adjusted to zero. Utilizing intramuscular EMG from tokens 2 and 4, these adjusted correlation values served as the weights used to generate a submental waveform. The calculated output was correlated with the actual submental recording. An error value was calculated to reflect the difference between the actual and the calculated submental activity.

Table 6 compares group average correlation values and average error sum per data point for the tokens used to develop the weights (tokens 1, 3, 5) and the tokens used to test the weights (tokens 2, 4). The difference between the averages from the two sets of tokens is also included in this table. Figure 2 contains a sample of the actual and calculated submental waveform.

Table 6.
Average Correlation Values (Avg Corr) and Error Sum Per Data Point (Avg Error) for Tokens used to Develop Weights (Tokens 1, 3, and 5) and Tokens Used to Test the Weights (Tokens 2 and 4) for the Three Analysis Approaches

	Develop Weights		Test Weights	
	Avg Corr	Avg Error	Avg Corr	Avg Error
Correlation Analysis	0.84	0.056	0.85	0.052
Numeric Analysis	0.86	0.016	0.86	0.016
Analytic Analysis	0.86	0.016	0.86	0.016

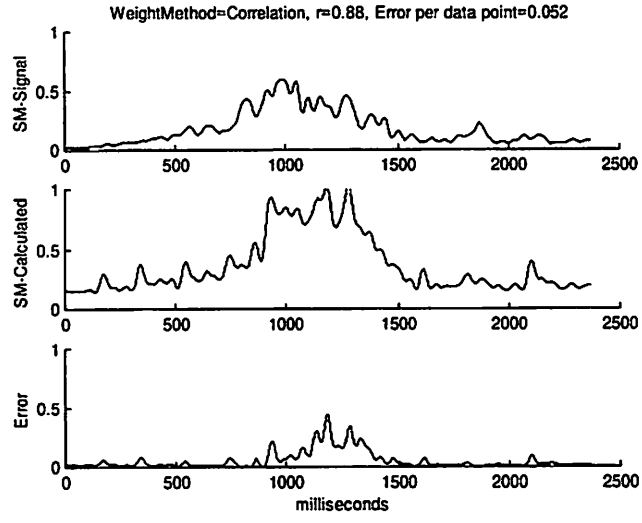


Figure 2. A sample (token ssas3) of an actual (SM-Signal) and calculated (SM-Calculated) submental waveform obtained when correlation values were used to calculate weights in the model (Equation 1). Error is presented in the bottom waveform.

Table 7.
Average Correlation Values (Avg Corr) and Average Error Sum Per Data Point (Avg Error) for the Five Swallow Tasks for Weights Developed With Correlation Values

	Develop Weights (a)		Test Weights (b)		Difference Values (a-b)	
	Avg Corr	Avg Error	Avg Corr	Avg Error	Diff Corr	Diff Err
3 ml Water	0.87	0.049	0.88	0.048	-0.01	0.001
10 ml Water	0.87	0.062	0.87	0.049	0.00	0.013
20 ml Water	0.87	0.055	0.89	0.052	-0.02	0.003
3 ml Applesauce	0.81	0.049	0.82	0.042	-0.01	0.007
3 ml Peanut Butter	0.87	0.044	0.80	0.070	0.07	-0.026

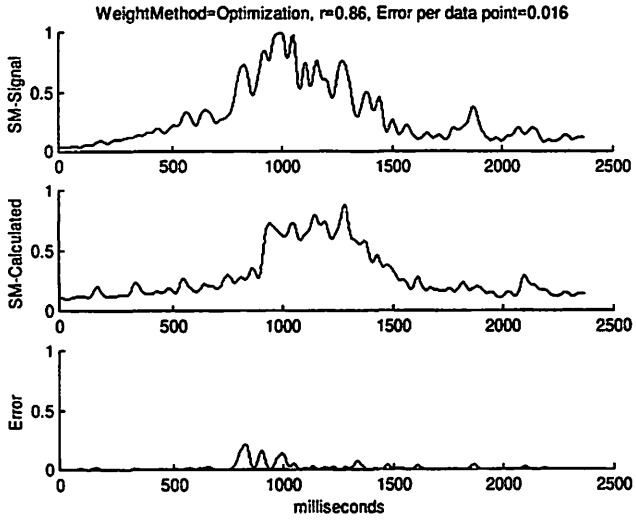


Figure 3. A sample (token ssas3) of an actual (SM-Signal) and calculated (SM-Calculated) submental waveform obtained when optimizaiton weights were used to calculate weights in the model (Equation 1). Error is presented in the bottom waveform.

Correlation values obtained for the two sets of tokens did not differ statistically. Across the seven subjects the average correlation values for the set of tokens used to develop the weights ranged from 0.76 to 0.91 with an average of 0.84. The average correlation values for the set of tokens used to test the weights ranged from 0.77 to 0.91 with an average of 0.85. The average difference between the correlation values for the two sets of tokens is 0.01.

For the subgroup of tokens used to develop the weights, the processed EMG data resulted in an average error sum per data point ranged from 0.012 to 0.137, with an average of 0.056. For the subgroup of tokens used to test the model, average error sum per data point ranged from 0.011 to 0.124, with an average of 0.052. The difference between the error sum per data point for the two subgroups of tokens averaged 0.004. The two sets of tokens did not produce statistically different error values.

For correlation values the primary task difference between the two sets of tokens occurred for the peanut butter swallows (average difference = 0.07) (Table 7). The set of tokens used to develop the weights produced a higher correlation for the peanut butter swallow (0.87) than the set of tokens used to test the weights (0.80). This difference was also noted for average error sum per data point, which ranged from a difference of 0.001 for 3 ml water swallows to 0.026 for peanut butter swallows. For peanut butter swallows, the average error sum per data point for the tokens used to test the weights was greater (0.070) than the average error sum per data point for the subgroup of tokens used to develop the weights (0.044).

Testing the Model Using Optimization Values as Weights in the Model

Weight values used to test the model were calculated by averaging values from tokens 1, 3 and 5 of the optimization analysis, which yielded average values of 0.295 for the mylohyoid muscle, 0.273 for the anterior belly of the digastric muscle, 0.380 for the geniohyoid muscle, 0.073 for the genioglossus muscle, and 0.103 for the platysma muscle. Table 6 shows a comparison of group correlation values and error sum per data point for the tokens used to develop the weights and the tokens used to test the weights in the model. A difference between average correlation and average error values calculated from the two sets of tokens is also included in the table. Figure 3 shows a sample of an actual and calculated submental waveform.

The two sets of tokens resulted in similar average correlation values. Across subjects, average correlation values for the sets of tokens used to develop the weights ranged from 0.79 to 0.85, with an average of 0.81. For the tokens used to test the weights, average correlation values ranged from 0.78 to 0.86, with an average of 0.82.

For the tokens used to develop the weights, average error sum per data point ranged from 0.012 to 0.022, with an average of 0.016. For the sets of tokens used to test the model, average error sum per data point ranged from 0.011 to 0.022, with an average of 0.016. There was no difference between the error sum per data point for the two subgroups of tokens.

Testing the Model Using Analytically-Derived Values as Weights in the Model

Weight values used to test the model were calculated by averaging values from tokens 1, 3 and 5 of the analytical analysis. This yielded an average weight of 0.285 for the mylohyoid muscle, 0.295 for the anterior belly of the digastric muscle, 0.373 for the geniohyoid muscle, 0.057 for the genioglossus muscle, and 0.102 for the platysma muscle. Table 6 shows a comparison of group correlation values and error sum per data point for the tokens used to develop the weights and the tokens used to test the weights in the model. A difference between average

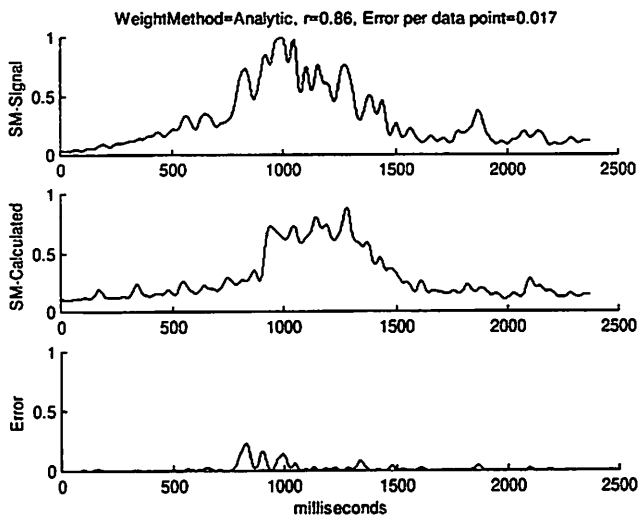


Figure 4. A sample (token ssas3) of an actual (SM-Signal) and calculated (SM-Calculated) submental waveform obtained when analytic values were used to calculate weights in the model (Equation 1). Error is presented in the bottom waveform.

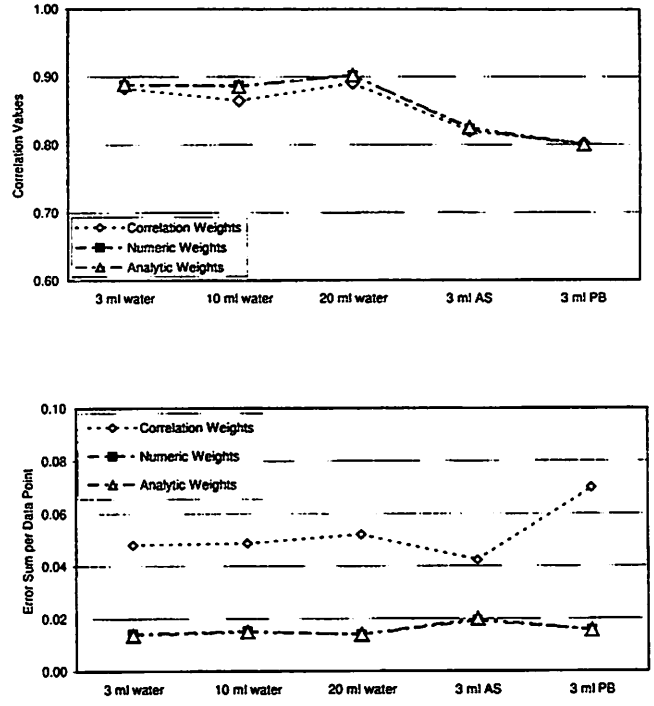


Figure 5. Comparison of the three approaches used to develop weights for the model. The top graph represents the correlation values between the actual and calculated submental recording. The bottom graph represents the error sum per data point.

correlation and average error values calculated from the two sets of tokens is also included in the table. Figure 4 shows a sample of an actual and calculated submental waveform.

The two sets of tokens resulted in similar average correlation values. For all subjects, average correlation values for the tokens used to develop the weights ranged from 0.78 to 0.85, with an average of 0.82. For the tokens used to test the weights, average correlation values for all subjects ranged from 0.78 to 0.86, with an average of 0.82. For the tokens used to develop the weights, average error sum per data point ranged from 0.012 to 0.022, with an average of 0.015. For the sets of tokens used to test the model, average error sum per data point ranged from 0.011 to 0.022, with an average of 0.016. The difference between the average error sum per data point for the two subgroups of tokens averaged 0.001.

Comparison of the Three Approaches

No statistically significant differences were noted between the three approaches with regard to correlation values (Figure 5). However, the error sum per data point generated from the three analysis methods differed statistically ($p < 0.01$). Overall, error sum per data point was greater when correlation values were used as weights to test the model (Figure 5).

Discussion

Typically, when EMG is used in the investigation of swallowing disorders, electrodes are placed on the submental skin's surface. The measured activity is thought to reflect the activity of muscles in the floor-of-mouth region, specifically the mylohyoid, geniohyoid, and anterior belly of the digastric muscles. This experiment evaluated this assumption by simultaneously recording EMG from the submental surface and individual muscles: mylohyoid, anterior belly of the digastric, geniohyoid, genioglossus and the platysma muscles. Three analysis methods were performed to estimate individual muscle contributions: correlation, numeric and analytic. In the correlation analysis individual muscles that showed a high positive correlation with the recording from the submental surface were considered potential contributors to the surface recording. A linear model was developed to represent the hypothesized relationship between the individual muscles and the surface recording. This model was used in both the numeric and the analytic analyses. For the numeric analysis, the linear model was used to optimize weights such that the best fit between the submental surface recording and the sum of the weighted muscle activity was obtained. Using an analytic approach, linear equations were used to derive a matrix of equations. Muscles that were assigned a high optimization value or a high analytic value were considered potential contributors to the submental recording.

A secondary goal was to test the linear model that was developed. In the analysis of the model, weights were assigned to each of the five individual muscles and used to

calculate a submental waveform. This calculated waveform was correlated with the actual submental recording. An error value between the two waveforms was calculated. It was assumed that if the two waveforms had a high correlation and a low error value then the defined model and the weights served as a good estimate of the degree of individual muscle contributions.

Contributors to the Submental Recording

The Primary Contributors

Regardless of the analysis method, primary muscle contributions to the submental surface recording were generally observed from three of the five muscles evaluated: the mylohyoid, anterior belly of the digastric, and the geniohyoid muscles. For some subjects, the primary contributions were not equal across the three muscles. In these cases, a major contribution was noted from at least one of the three muscles and a minor contribution was noted from the remaining two muscles.

The subject-specific patterns noted in this investigation are not uncommon in the swallowing literature. Previous investigations have identified subject-dependent muscle activation patterns during swallowing. Hrycyshyn and Basmajian (1972) evaluated firing patterns of muscles during swallowing. They reported that approximately two-thirds of their subjects had high agreement in the firing sequence observed across all swallows. The remaining subjects had a low agreement with the majority of the group, although they did have high intrasubject agreement. These findings were corroborated in a similar investigation. Palmer, Perlman, McCulloch, VanDaele, and Luschei (1995) evaluated patterns of muscle activation during swallowing and found that intrasubject pattern agreement was higher than intersubject pattern agreement.

While these previous investigations did not address the issue of contributions of individual muscles to the submental surface, they did report variability in the behavior of the muscles of swallowing. Considering that there is variability in the temporal firing pattern of the individual floor-of-mouth muscles, one would expect some variability to be evident in the pattern of contributions from individual muscles.

The Genioglossus and Platysma Muscles

Based on data obtained from all three analysis methods, the other two muscles investigated, the genioglossus and the platysma muscles, had lower contributions to the submental surface recording. Thus both of these muscles may be considered minimal contributors to the submental surface recording.

Task Effects on the Contribution

Data from previous investigations regarding the effect of bolus volume on the submental surface recording have been inconsistent. Some investigators have reported that the submental recording is sensitive to alterations in the bolus volume (Cook et al, 1989; Ertekin et al., 1997). Others did not report bolus volume effects (Dantas &Dodds, 1990; Miller & Watkin, 1996). In this experiment bolus volume and viscosity were altered. Increased bolus volume did not statistically alter the contribution of the individual muscles to the submental recording.

While it is possible that bolus volume does not alter the contribution of the individual muscles to the submental surface, limitations in the methodology used to evaluate bolus volume should also be considered. Perhaps the variability in anatomy and the idiosyncratic nature of swallowing imply that different guidelines should be considered when defining changes in bolus volume. For example, Adnerhill, Ekbreg and Groher (1989) asked subjects to swallow a comfortable amount of liquid to determine the normal bolus size for thin liquids. They reported an average bolus size of 21 ml. However, normal bolus size ranged from approximately 10 ml to 30 ml across all the subjects. This suggests that a regulated bolus volume such as 15 ml could be perceived as either a small bolus or a large bolus depending on the subject. Defining small, medium and large boluses by determining an absolute bolus size to be used across all subjects may dilute the potential bolus volume effect. An alternate method of defining bolus volume should be considered. By measuring a person's average bolus size, incremental increases and decreases from that average could serve as smaller and larger boluses for that subject. Instead of using absolute bolus volume categories such as 3 ml, 10 ml and 20 ml, bolus volumes could be labeled as average bolus size, large bolus size and small bolus size, yet the absolute volume that coincides with each of those categories would be subject specific.

Although changes in bolus volume did not alter the contribution of individual muscles to the submental recording, alterations in bolus viscosity yielded statistically significant differences in the contribution of the individual muscles to the submental recording. As viscosity increased so did the peak of the EMG as well as the area under the curve of the EMG signal. Based on the relationship between EMG output and pressure (Jaffe, McCulloch, Palmer & Luschei, 1996), the greater EMG signals associated with more viscous substances implies that more pressure is being exerted on the bolus for transport.

The increased activity that was noted in this current investigation is consistent with previous investigations that have shown increased intra-bolus pressure with increased viscosity (Dantas, Kern et al., 1990) and increased magnitude and duration of the recorded submental

signal (Dantas & Dodds, 1990; Ertekin et al., 1997; Reimers-Neil et al., 1994). Although, we did not analyze duration of EMG activity in this current study, peak values and integration of the filtered submental waveforms increased with viscosity. These changes observed in the submental recording were also seen in the data from the individual muscles.

Modeling the Submental Recording

Equation 3 was used as a model to represent the relationship between the surface recording and the individual muscles. This linear equation required weight values for each of the five muscles represented in the model. Although the model was tested three times using three different sets of weights, all sets biased the system such that the mylohyoid, anterior belly of the digastric and the geniohyoid muscles were more heavily weighted than the genioglossus and the platysma muscles. Strong correlation values were obtained between the actual and calculated submental recording regardless of which of the three sets of weights were used. Thus, based on correlation values between the actual and calculated submental recording, the model provided an adequate representation of the contribution of individual muscles to the submental surface recording.

Although correlation values were high regardless of which of the three sets of weight values were used, the three sets of weights resulted in somewhat different error values. Weights derived from the numeric and analytic values yielded lower error values than those obtained when weights were derived from the correlation analysis. These findings are not surprising considering that the values used to develop the weights for the model were optimized to keep error at a minimum. This parameter was not considered in the correlation analysis. While it is possible that the lower error values are a function of the analysis approach used to develop the weights, it is also possible that the error values implicate either limitations in the model or contributing muscles that are not accounted for in the model. Thus, when error values are regarded as important criteria, the weights developed from the numeric and analytic methods provided a better representation of the contribution of the muscles to the submental surface recording.

In summary, based on the high correlation values between the actual and modeled submental waveform, the model provided an adequate estimate of the contribution of the five muscles. Taking into account that the error values were significantly lower when weights were derived from the optimization or analytic analysis method, these weights can be considered better estimates of the actual contribution of the individual muscles to the submental surface recording.

Acknowledgments

The first author is sincerely grateful to the other authors as well as to Ingo Titze, Brad Story, Jerry Moon, and Kevin Rosenberg for their guidance throughout this project. This project received funding from (1) the National Center for Voice and Speech (NIH grant number P60 DC00976), and (2) the University of Iowa Hospitals and Clinics, Department of Otolaryngology, Head and Neck Surgery.

References

Adnerhill, I., Ekbreg, O., and Groher, M.E.(1989). Determining normal bolus size for thin liquids. *Dysphagia, 4*, 1-3.

Barofsky, I. (1995). Surface electromyographic biofeedback and the patient with dysphagia: clinical opportunities and research questions. *Dysphagia, 10*, 19-21.

Bouisset, S., & Maton, B. (1972). Quantitative relationship between surface EMG and intramuscular electromyographic activity in voluntary movement. *American Journal of Physical Medicine, 51*, 285-295.

Bryant, M. (1991). Biofeedback in the treatment of a selected dysphagic patient. *Dysphagia, 6*,140-144.

Cook, I.J., Dodds, W.J., Dantas, R.O., Kern, M.K., Massey, B.T., Shaker, R., & Hogan, W.J. (1989). Timing of videofluoroscopic, manometric events, and bolus transit during the oral and pharyngeal phases of swallowing. *Dysphagia, 4*, 8-15.

Dantas, R.O. & Dodds, W.J. (1990). Effect of bolus volume and consistency on swallow-induced submental and infrahyoid electromyographic activity. *Brazilian Journal of Medical and Biological Research, 23*, 37-44.

Dantas, R.O, Kern, M.K., Massey, B.T., Dodds, W.J., Kahrilas, P.J., Brasseur, J.G., Cook, I.J., & Lang, I.M. (1990). Effect of swallowed bolus variables on oral and pharyngeal phases of swallowing. *American Journal of Physiology, 258*, G675-681.

deLarminat, V., Montravers, P., Dureuil, B., & Desmonts, J.M. (1995). Alterations in swallowing reflex after extubation in intensive care unit patients. *Critical Care Medicine, 23*, 486-490.

Doty, R.W. & Bosma, J.F. (1956). An electromyographic analysis of reflex deglutition. *Journal of Neurophysiology, 19*, 44-60.

Ertekin, C., Aydogdu, I., Yuceyar, N., Pehlivan, M., Ertas, M., Uludag, B., & Celebi, G. (1997). Effects of bolus volume on oropharyngeal swallowing: an electrophysiologic study in man. *American Journal of Gastroenterology, 92*, 2049-2053.

Gay, T., Rendell, J.K., & Spiro, J. (1994). Oral and laryngeal muscle coordination during swallowing. *Laryngoscope, 104*, 341-349.

Hirose, H. (1971). Electromyography of the articulatory muscles: Current instrumentation and technique. *Status Report on Speech Research: Haskins Laboratories, SR25/26*, 73-86.

Hrycyshyn, A. & Basmajian, J. (1972). Electromyography of the oral stage of swallowing in man. *American Journal of Anatomy, 133*, 335-340.

Huffman, A.L. (1978). Biofeedback treatment of orofacial dysfunction: a preliminary study. *American Journal of Occupational Therapy, 32*, 149-154.

Jaffe, D.M., McCulloch, T.M., Palmer, P.M., & Luschei, E.S. (1996). Quantitative contributions of the middle, lateral and posterior genioglossus to intraoral pressure generation. Paper presented at the fifth annual meeting of the Dysphagia Research Society, Aspen, Colorado.

Koole, P., de Jongh, H.J., & Boering, G. (1991). A comparative study of electromyograms of the masseter, temporalis, and anterior digastric muscles obtained by surface and intramuscular electrodes: Raw EMG. *The Journal of Craniomandibular Practice, 9*, 228-240.

Lehr, R.P., Blanton, P.L., & Biggs, N.L. (1971). AN electromyographic study of the mylohyoid muscle. *Anatomical Records, 169*, 651-660.

Miller, J.L. & Watkin, K.L. (1996). The influence of bolus volume and viscosity on anterior lingual force during the oral stage of swallowing. *Dysphagia, 11*(2), 117-124.

Palmer, P.M., Perlman, A.L., McCulloch, T.M., VanDaele, D.J., & Luschei, E.S. (1995). The effect of bolus volume on electromyography of select oral, laryngeal, and pharyngeal muscles [Abstract]. *Dysphagia, 10*, 139.

Palmer, P.M., Perlman, A.L., VanDaele, D.J. & McCulloch, T.M. (1993). Electromyography of select laryngeal and pharyngeal muscles during swallowing, valsalva and phonation. Poster session presented at the annual meeting of the Dysphagia Research Society, Lake Geneva, Wisconsin.

Perlman, A., Luschei, E.S., & Dumond, C.E. (1989). Electrical activity from the superior pharyngeal constrictor during reflexive and nonreflexive tasks. *Journal of Speech and Hearing Research, 32*, 749-754.

Perry, J., Easterday, C.S., and Antonelli, D.J. (1981). Surface versus intramuscular electrodes for electromyography of superficial and deep muscles. *Physical Therapy, 61*, 7-15.

Person, R.S. (1963). Problems in the interpretation of electromyograms: 1. Comparison of electromyograms on recordings with skin and needle electrodes. *Biofizika, 8*, 82-89.

Pouderoux, P. & Kahrilas, P.J. (1995). Deglutitive tongue force modulation by volition, volume, and viscosity in humans. *Gastroenterology, 108*, 1418-1426.

Reimers-Neils, L., Logemann, J., & Larson, C. (1994). Viscosity effects of EMG activity in normal swallow. *Dysphagia, 9*, 101-106.

Ren, J., Shaker, R., Zamir, Z., Dodds, W.J., Hogan, W.J. & Hoffmann, R.G. (1993). Effect of age and bolus variables on the coordination of the glottis and upper esophageal sphincter during swallowing. *The American Journal of Gastroenterology, 88*, 665-669.

Shaker, R., Milbrath, M., Ren, J., Campbell, B., Toohill, R., & Hogan, W. (1995). Deglutitive aspiration in patients with tracheotomy: effect of tracheostomy on the duration of vocal cord closure. *Gastroenterology, 108*, 1357-1360.

Sauerland, E.K., Sauerland, B.A.T., Orr, W.C., & Hairston, L.E. (1981). Non-invasive electromyography of human genioglossal activity. *Electromyography Clinical Neurophysiology, 21*, 279-286.

Sukthankar, S.M., Reddy, N.P., Canilang, E.P., Stephenson, L., & Thomas, R. (1994). Design and development of portable biofeedback systems for use in oral dysphagia rehabilitation. *Medical Engineering and Physics, 16*, 430-435.

NCVS Status and Progress Report - 13
April 1999, 83-92

Thyroarytenoid Muscle Activity Associated with Hypophonia in Idiopathic Parkinson Disease and Aging

Kristin K. Baker, Ph.D. CCC-SP
Denver Center for the Performing Arts, Wilbur J. Gould Voice Research Center
Lorraine Olson Ramig, Ph.D. CCC-SP
Speech Language Hearing Sciences Department, University of Colorado - Boulder
Denver Center for the Performing Arts, Wilbur J. Gould Voice Research Center
Erich S. Luschei, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa
Marshall E. Smith, M.D.
University Medical Center, The University of Utah - Salt Lake City

Abstract

Objective: This paper compared electromyographic (F-MG) amplitude levels of the thyroarytenoid (TA) muscle in young and aged individuals and individuals with idiopathic Parkinson disease (IPD) under conditions of known vocal loudness (sound pressure level). **Background:** Although a voice disorder is frequently associated with aging and the hypokinetic dysarthria of IPD, it is still unclear how laryngeal muscle physiology is affected by these processes and how potential changes in laryngeal muscle activity result in characteristic changes of the voice. **Methods:** *Absolute and relative* (to maximum) EMG amplitudes of the TA muscle were compared during phonatory, pre-phonatory and non-phonatory tasks. Corresponding sound pressure level (SPL) measures were obtained for the phonatory tasks. **Results:** *Absolute* TA amplitudes (expressed in microvolts) were consistently the highest in the young individuals (1 0. 81 - 29.73), lowest in the individuals with IPD (3.23 - 9.76) and intermediate in the aged individuals (6.72 - 16.87). *Relative* TA amplitudes (expressed as a per cent of maximum) were generally the highest for the young individuals. (9.11 - 21.56), lowest for the aged individuals (3.66 - 11.27) and intermediate (5.08 - 15.27) for the individuals with IPD. SPL findings showed the young individuals produced speech tasks with the highest SPLS. The aged individuals and individuals with IPD showed comparable SPLs for most of the phonatory tasks and these levels were consistently lower than the young individuals. **Conclusions:** These findings suggest reduced levels of TA muscle activity may contribute to the characteristic hypophonic voice disorders that frequently accompany aging and IPD.

Voice disorders are often a primary characteristic of the hypokinetic dysarthria associated with idiopathic Parkinson disease (IPD). Eighty nine per cent of the 200 individuals with IPD studied by Logemann et al.,[1] presented laryngeal dysfunction as one of several dysarthric symptoms, and approximately half of those individuals presented a voice disorder as the only symptom. Perceptual descriptors of this hypophonic voice disorder typically include reduced loudness[2], monoloudness[2-3], monotonicity[2-3] and a hoarse, breathy voice quality[1]. Studies investigating acoustic correlates have shown variable results, but in general indicate there are reductions in: voice sound pressure level (SPL)[4], fundamental frequency range and variability[5] and phonatory stability[6] in individuals with IPD compared to non-disordered individuals. Laryngeal imaging methods have identified changes in vocal fold adduction[7-8], such as vocal fold bowing and an open phase configuration, and changes in vocal fold stability[8], such as phase asymmetries and laryngeal tremor, as some of the more common laryngeal movement abnormalities that occur in IPD.

Although these findings provide valuable descriptive information about the phonatory disorder observed in IPD, a comprehensive understanding of the pathophysiology underlying this voice disorder is lacking. Very few

studies have directly investigated laryngeal muscle function in individuals with IPD. Of these studies, higher laryngeal muscle activation levels[9] and co-contraction of the thryoarytenoid and cricothyroid muscles[10] were observed during speech and maximum gestures in individuals with IPD compared to age-matched controls. These laryngeal electromyographic (EMG) observations are comparable to limb electrophysiologic findings in individuals with IPD that are indicative of the pathophysiology of rigidity[11]. Thus, it has been suggested that stiff or rigid laryngeal muscles may contribute to the etiology of the voice disorder associated with IPD[8-10]. However, this interpretation seems inconsistent with the hypophonic voice characteristics that frequently accompany IPD. Furthermore, it has been hypothesized that alterations of cortico-basal ganglionic activity result in reduced facilitation of motoneuron pools[12] and hypokinetic movements in IPD. These movements are characterized by reduced peak and overall EMG amplitudes[13] and reduced EMG magnitudes at movement onset[14]. Considering these findings, it is possible that reduced levels of laryngeal muscle activity contribute to the voice symptoms in some individuals with IPD. Additional investigations would clarify how laryngeal muscle function is affected by IPD and how changes in muscle function result in characteristic changes of the voice.

This paper compared *absolute and relative* thyroarytenoid (TA) muscle amplitude levels in young and aged individuals and individuals with IPD under conditions of known SPL. investigating the TA was of interest because this laryngeal muscle is involved in the control of related variables that are affected by hypophonia, such as, vocal fold adduction[15-16], SPL and vocal loudness[17]. Young and aged comparisons were included because many individuals with IPD are aged and muscular[18] and nonmuscular[19-20] changes have been observed in the laryngeal mechanism as a function of age. These changes may influence vocal fold movement and acoustic voice parameters and result in perceptual changes that are similar to the disordered voice characteristics observed in individuals with IPD[21]. Without group comparisons, it would be difficult to determine if potential changes in TA amplitude levels were related to aging alone or a combination of aging and a neuropathology.

Methods

Participants

Four young (mean age, 27.7 years), four aged (mean age, 73.5 years) and five individuals with IPD (mean age, 71 years) participated. The young and aged individuals reported good health and no neurological conditions. The individuals with IPD were optimally medicated as determined by his or her neurologist and medication was taken within 60 to 90 minutes of all procedures.

Prior to participation, a screening, which consisted of a voice, speech and background history and a rigid (Kay Elemetrics SN1077) or flexible (Olympus ENF-P3) video laryngostroboscopic exam (Kay Elemetrics RLS 9100) of the vocal folds was performed on each individual to determine if they met certain inclusion criteria. This procedure was performed by an impartial speech-language pathologist with 14 years of clinical experience. The screening inclusion criteria for the young individuals was normal voice, speech and vocal fold function. The screening inclusion criteria for the aged individuals and the individuals with IPD was: a) the presence of hypophonic voice characteristics commonly associated with aging and IPD, specifically, reduced loudness (volume) and a breathy voice quality and b) some level of vocal fold incompetence. Additionally, these criteria had to be met in the individuals with IPD while they were experiencing the effect of their medication ("on" state). Descriptive information for the participants is summarized in Tables 1 and 2.

Data Collection and Analysis

Each person was seated comfortably in an IAC sound treated booth and acoustic, respiratory kinematic and laryngeal EMG data were collected. The voice was recorded with a head-mounted microphone (AKGC410) positioned 8 cm from the participant's mouth. Voice intensity was measured at 30 cm with a Bruel and Kjaer Type 2230 sound level meter (SLM). Respibands were placed on the rib cage just below the armpits and on the abdomen centered over the umbilicus. Laryngeal EMG data were collected using slightly modified bipolar hooked-wire electrodes[22]. Each electrode consisted of a pair of stainless steel wires (50 μm in diameter) that were insulated and glued together along their

Table 1.
Descriptive Information for the Young and Aged Participants

Pt.	Age (yr.mo.)	Sex	VFI
1	29.5	M	None
2	28.2	M	None
3	26.1	F	None
4	27.1	F	None
5	79	M	Mild bowing
6	72.1	M	Mild-moderate bowing
7	68.3	F	Posterior gap
8	74.5	F	Moderate bowing

Note. VFI=vocal fold incompetence.

Table 2.
Descriptive Information for the Participants with Idiopathic Parkinson Disease

Pt.	Age (yr.mo.)	Sex	VFI	H & Y	YPO	Medication
1	78.2	M	Moderate bowing	ND	7	Sinemet CR, Permax
2	71.7	M	Mild-moderate bowing	3	5	Sinemet, Permax, Eldepryl
3	65.1	F	Moderate-severe bowing	3.5	4	Sinemet, Amantadine
4	78.1	F	Severe bowing	4	5	Sinemet, Sinemet CR
5	59.9	M	Mild bowing	ND	5	Sinemet, Benzotropine

Note. VFI=vocal fold incompetence; YPO=years post disease onset; H & Y=Rating from Hoehn and Yahr scale[26]; ND=not done.

length (bifilar configuration). The bifilar wires were thread through a 25 gauge 1.5 inch hypodermic needle and approximately 2 mm of the insulation was removed from the end of each wire. The bared wire at the electrode tip was then bent back to form a hook. After assembly, each electrode was steam sterilized. Electrode insertion was performed by an otolarygnologist. First, the laryngeal area near the cricothyroid membrane was anesthetized and then an electrode was directed through the medial cricothyroid space and angled upward and laterally into the TA muscle[23].

After positioning of the transducers, at least three repetitions of sustained vowel phonations[24], dry swallows and effortful glottal closures (Valsalva maneuvers)[22,25] were performed to verify TA electrode placement. Once electrode placement was confirmed, each individual produced several speech and nonspeech tasks. The speech tasks included three comfortable and three maximum sustained phonations of the vowel /ɑ/, ten repetitions of the sentence "Pop took his socks off', one reading of a standard passage (Rainbow passage) and one 30 second monologue describing a "happy day." This allowed TA amplitudes to be examined for speech tasks that required relatively simple (sustained vowel phonations) and more complex (connected speech) levels of laryngeal motor control. Pre-phonatory TA activity levels were also measured for these tasks because previous studies have shown that individuals with IPD have difficulty producing appropriate levels of force at the onset of certain limb movements[13, 26-27]. Measuring prephonatory TA activity made it possible to determine if a similar type of movement initiation problem affected the laryngeal mechanism. The non-speech task consisted of two minutes of quiet rest breathing. This task was performed to determine if the TA muscle showed increased levels of tonic resting and background EMG activity, an observation that has been noted in the lip[28-29] and jaw[30] muscles of individuals with IPD. All tasks were performed by each individual with the exception of the comfortable sustained vowel phonations and the monologue. One of the aged individuals and one individual with IPD did not produce the comfortable sustained vowel phonations and one young individual did not produce the monologue.

The EMG, acoustic and respiratory signals were recorded to a Sony PC-108M eight-channel digital audio tape recorder and simultaneously digitized (5 KHz sample

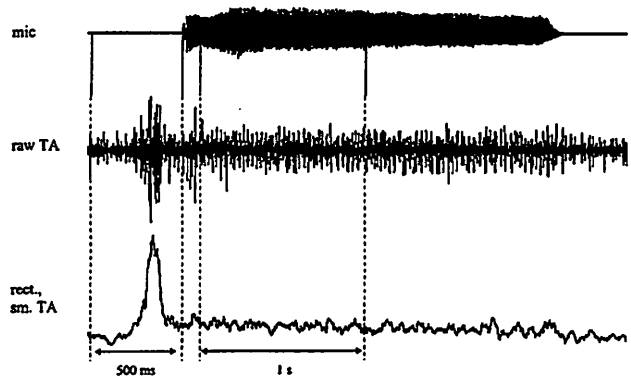


Figure 1. Example of a microphone (top trace), raw TA EMG (middle trace) and rectified and smoothed (200 ms) TA EMG signal (bottom trace) for an aged individual's comfortable sustained vowel phonation. Cursor placements that quantified phonatory and pre-phonatory EMG activity are indicated by the vertical lines. The 500 ms pre-phonatory EMG activity are indicated by the vertical lines. The 500 ms pre-phonatory segment is indicated by the first and second vertical lines and the 1 s segment 100 ms after phonation onset, is indicated by the third and fourth vertical lines.

rate) to a Hewlett Packard Pentium/200 MHz computer using WINDAQ acquisition software (DATAQ Instruments Inc). Before the EMG signals were recorded and digitized, they were amplified (20 Hz - 5 KHz) and anti-aliased filtered (2.5 KHz). The EMG signals were then zero-meaned, full wave rectified, smoothed (moving average constant of 200) and converted to microvolts. The sound level meter data were converted from volts to sound pressure level (dBSPL). This processing was performed using WINDAQ and CODAS software (DATAQ Instruments Inc).

EMG Measurements

Two EMG measures were obtained from the smoothed and converted TA signals: *absolute and relative* TA amplitudes. These measures were acquired by manually cursoring out the following segments of the speech and nonspeech tasks. A 1 second segment, 100 ms after phonation onset was measured for the comfortable and maximum sustained vowel phonations and the segment between phonation onset and offset was measured for the sentence repetitions, reading and monologue. Prephonatory segments for the comfortable and maximum sustained vowel phonations, the reading and monologue were measured 500 ms before phonation onset and for the sentence productions, 100 ms before phonation onset. A 100 ms versus 500 ms prephonatory segment was measured for the sentences because they were produced consecutively with only a brief intersentence interval. Cursor placement for phonation onset and offset was determined by using the microphone signal. Figure 1 shows an example of how phonatory and prephonatory TA amplitudes were obtained for a comfortable sustained vowel phonation. For quiet rest breathing, a 5 second segment of TA activity was measured 60 seconds after the individual was instructed to relax quietly in the booth.

Because *absolute* amplitude measures can be affected by methodological factors, such as electrode placement within the muscle[24], *relative* TA amplitudes were also examined for the above segments. *Relative* TA amplitudes were normalized to the maximum TA activation level for each individual using the following formula:

(absolute TA mean for each task/maximum TA mean) * 100

TA maximum was determined by scrolling through each TA record and visually identifying the task that produced the highest TA amplitude level. Once identified, a 100 ms window was manually segmented and the mean TA value for that interval was calculated. Tasks that generated the highest TA levels were variable across individuals and included dry swallows (aged, IPD), swallows of thin and solid consistencies (young, IPD), Valsalva maneuvers (aged, IPD), hard glottal attacks (young), prephonatory gestures (aged) and extremely loud phonation (young). Any gesture identified during this process as being contaminated with artifacts was not measured.

SPL Measurements

SPL measures for the comfortable and maximum sustained vowel phonations and sentence repetitions were obtained for the same speech segments as TA amplitude measures. However, the microphone signal was not used to segment the SPL signal. Rather, a SPL criterion level (dB_{cl}) was calculated for each vowel phonation and sentence repetition and the points where the dB_{cl} intersected the SPL signal determined cursor placement for SPL onset and offset. The criterion level was computed as follows,

$$dB_{cl} = 0.76*dB_{max} + 0.70*std_{SPL}-2.1$$

where dB_{max} referred to the SPL maximum and std_{SPL} referred to the SPL standard deviation for a given task.

For the reading and monologue tasks, there were several instances where the SPL signal dropped below dB_{cl}. These drops below dB_{cl} were reflective of pauses in the speech sample. in order to exclude these segments from the average SPL measure, a duration parameter (300 ms) that appropriately differentiated speech and non-speech segments was defined (see Figure 2). After multiple segments from each individual's reading and monologue were measured, the overall mean SPL for these tasks was computed by averaging the means of the multiple segments. Data obtained using the dB_{cl}, method were compared to data obtained on-line by recording the digital output of the SLM[4]. Similar patterns were observed across individuals and tasks for both methodologies.

Statistics

Group means and standard errors were calculated from the raw data for each individual and task. In addition,

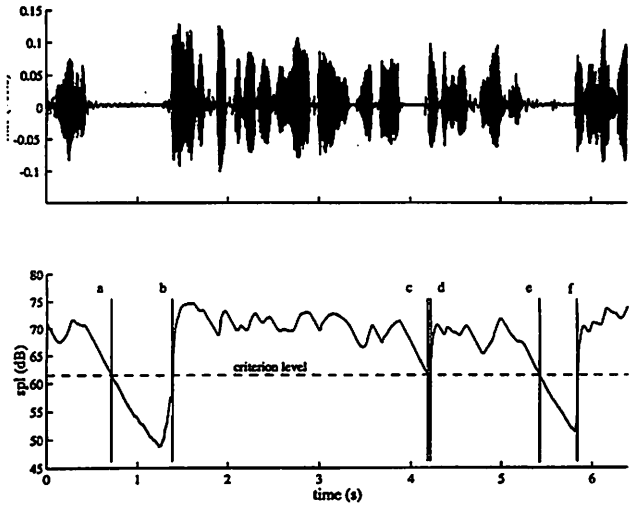


Figure 2. Microphone (top trace) and SPL (bottom trace) signals from a portion of the reading passage (Rainbow passage). The horizontal dashed line indicates dB_{cl}, which was 61.5 dBSPL. The three sets of vertical lines indicate three instances where the SPL signal dropped below dB_{cl}. The segments from (a) to (b) and (e) to (f) were excluded from analysis because both were more than 300 ms and corresponded to pauses, while segment (c) to (d) was included in the analysis because it was less than 300 ms and corresponded to a component of the speech sample. A mean was calculated for the segment from (b) to (e) and averaged with means of other segments from this passage to obtain the overall mean SPL.

TA amplitude and SPL data for the three groups were compared using an analysis of variance (ANOVA) or a repeated measures ANOVA, as appropriate[31]. The former analysis was used to examine measures for the reading and monologue tasks, the pre-phonatory segments of these tasks, quiet rest breathing and maximum amplitude gestures. The latter analysis assesses variability between individuals as well as variability between multiple observations on the same individual and therefore was used to examine the comfortable and maximum sustained vowel phonation and the sentence repetition measures as well as the corresponding pre-phonatory measures. Both types of analyses were completed with SAS statistical software (SAS/IML, Version 6.12, SAS Institute, Inc., Cary, North Carolina, 1996) and a significance level of $p \leq 0.05$ was established for all measures. If the F statistic was significant for a particular task, *a priori* contrasts were performed for each group (IPD versus aged, IPD versus young, and aged versus young).

Intra- and inter-rater reliability for the TA amplitude and SPL measures was assessed by remeasuring twenty percent of the data. Correlation coefficients ranged between .95 and 1.0 for the TA amplitude measures and between .97 and 1.0 for the SPL measures.

Results

Absolute and relative TA Amplitudes

Consistent directional differences were observed across groups for the phonatory, pre-phonatory and quiet

Table 3.
Group Means, Standard Errors (in parenthesis) and p Values for the Main Effect and Multiple Comparisons for *Absolute* TA Amplitude Measures

	Mean (standard error) in microvolts			Main effect p value	Multiple comparisons p values		
Task	Young	Aged	IPD		IPD v. aged	IPD v. young	aged v. young
comfortable vowel	20.42 (0.21)	12.45 (0.25)	6.98 (0.36)	< 0.0001	< 0.0001	< 0.0001	< 0.0001
maximum vowel	23.60 (0.54)	14.29 (0.51)	9.11 (0.45)	< 0.0001	< 0.0001	< 0.0001	< 0.0001
sentence repetition	25.55 (0.22)	14.77 (0.22)	9.43 (0.20)	< 0.0001	< 0.0001	< 0.0001	< 0.0001
reading	21.13 (6.09)	15.25 (3.03)	8.48 (1.56)	= 0.10	ND	ND	ND
monologue	19.81 (8.97)	14.63 (3.52)	7.91 (1.54)	= 0.21	ND	ND	ND
pre-comfortable vowel	29.16 (0.51)	11.83 (0.59)	5.89 (0.59)	< 0.0001	< 0.0001	< 0.0001	< 0.0001
pre-maximum vowel	29.73 (1.70)	16.87 (1.60)	8.70 (1.43)	< 0.0001	< 0.0008	< 0.0001	< 0.0001
pre-sentence repetition	21.10 (0.26)	11.18 (0.26)	9.76 (0.24)	< 0.0001	< 0.0001	< 0.0001	< 0.0001
pre-reading	16.49 (5.15)	11.98 (1.51)	5.89 (1.07)	= 0.07	ND	ND	ND
pre-monologue	18.67 (9.65)	13.15 (2.84)	7.22 (1.27)	= 0.26	ND	ND	ND
rest breathing	10.81 (6.46)	6.72 (2.85)	3.23 (0.65)	= 0.38	ND	ND	ND
TA maximum	147.24 (41.65)	209.86 (69.63)	65.59 (9.03)	= 0.10	ND	ND	ND

Note: ND=not done.

rest breathing mean *absolute* TA amplitude measures (see Table 3). Statistical analyses revealed significance only for the tasks that had multiple repetitions. Specifically, significant differences were observed for the comfortable and maximum sustained vowel phonation and sentence repetition measures as well as for the pre-phonatory measures for these tasks. A Scheffe multiple comparison analysis revealed differences were statistically significant for each group comparison, with the young group consistently producing the highest means (range: 10.81 - 29.73 microvolts), the aged group showing intermediate mean amplitudes (range: 6.72 - 16.87 microvolts) and the IPD group producing the lowest means (range: 3.23 - 9.76 microvolts). Although this directional difference was maintained for the reading, monologue, pre-reading, pre-monologue and quiet rest breathing measures, statistical significance was not observed. Maximum *absolute* values for TA amplitudes also differed among the three groups. The mean maximum TA amplitude was 147.24 (se = 41.65), 209.86 (se = 69.63) and 65.59 (se = 9.03) microvolts for the young and aged individuals and individuals with IPD respectively. These differences were not statistically significant (F = 0.10).

Group means and standard errors for the *relative* TA amplitudes are shown in Table 4. Similar to *the absolute* TA amplitudes, statistical significance was observed only for the tasks with multiple repetitions, the comfortable and maximum sustained vowel phonation, sentence repetition, precomfortable and maximum sustained vowel phonation and pre-sentence repetition measures. Group comparisons for each of these measures were also statistically significant. Statistical significance was not observed for the reading, monologue, pre-reading, pre-monologue or quiet rest breathing measures. Similar directional differences, regardless of statistical significance, were observed for all but three measures. This pattern showed the young individuals with the highest means (range: 9.11 - 21.56 per cent of maximum), the individuals with IPD having intermediate means (range: 5.08 - 15.27 per cent of maximum) and the aged individuals with the lowest mean amplitudes (range: 3.66 - 11.27 per cent of maximum).

Table 4.
Group Means, Standard Errors (in parenthesis) and p Values for the Main Effect and Multiple Comparisons for *Relative* TA Amplitude Measures

	Mean (standard error) in microvolts			Main effect p value	Multiple comparisons p values		
Task	Young	Aged	IPD		IPD v. aged	IPD v. young	aged v. young
comfortable vowel	16.41 (0.21)	10.58 (0.24)	12.11 (0.50)	< 0.0001	< 0.004	< 0.0001	< 0.0001
maximum vowel	18.49 (0.43)	9.02 (0.40)	13.92 (0.36)	< 0.0001	< 0.0001	< 0.0001	< 0.0001
sentence repetition	18.13 (0.15)	11.27 (0.15)	14.53 (0.13)	< 0.0001	< 0.0001	< 0.0001	< 0.0001
reading	15.83 (3.28)	10.05 (3.20)	12.77 (0.97)	= 0.33	ND	ND	ND
monologue	13.28 (2.81)	9.31 (2.82)	11.90 (1.42)	= 0.51	ND	ND	ND
pre-comfortable vowel	21.26 (0.42)	12.17 (0.48)	10.62 (0.48)	< 0.0001	< 0.04	< 0.0001	< 0.0001
pre-maximum vowel	21.56 (1.01)	11.01 (0.95)	13.70 (0.85)	< 0.0001	< 0.05	< 0.0001	< 0.0001
pre-sentence repetition	16.48 (0.24)	9.43 (0.24)	15.27 (0.21)	< 0.0001	< 0.0001	< 0.0004	< 0.0001
pre-reading	13.98 (4.60)	8.93 (3.54)	8.89 (1.14)	= 0.46	ND	ND	ND
pre-monologue	12.47 (3.37)	12.65 (3.00)	10.87 (1.44)	= 0.84	ND	ND	ND
rest breathing	9.11 (5.43)	3.66 (0.93)	5.08 (1.13)	= 0.46	ND	ND	ND

Note: ND=not done.

Deviations from this pattern were observed for three of the pre-phonatory measures: the comfortable sustained vowel phonation, reading and monologue. For the pre-phonatory comfortable sustained vowel phonation measure, the young individuals showed the highest mean (21.26 per cent of maximum), the aged individuals showed an intermediate mean (12.17 per cent of maximum) and the individuals with IPD showed the lowest mean amplitude (10.62 per cent of maximum). For the prephonatory reading measure, the aged individuals and individuals with IPD showed similar mean amplitudes at 8.93 and 8.89 per cent of maximum respectively and these means were lower than the mean amplitude of the young individuals (13.98 per cent of maximum). Finally, for the pre-phonatory monologue measure, the young and aged individuals (young: 12.47; aged: 12.65 per cent of maximum) showed similar mean amplitudes that were higher than the mean amplitude for the individuals with IPD (10.87 per cent of maximum). Although slight directional variations occurred for these three measures, the individuals with IPD still showed either intermediate or low *relative* TA amplitudes. Figure 3 illustrates the patterns that occurred for most of the *absolute and relative* TA amplitude measures. Only results for phonatory measures are displayed.

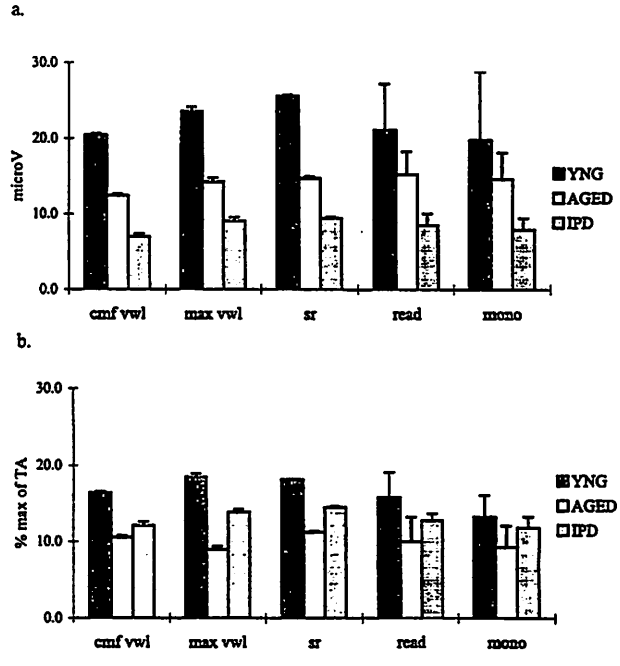


Figure 3. Directional differences that were observed for most of the absolute *(a) and* relative *(b) TA amplitude measures, cmfvwl=comfortable sustained vowel phonation; max vwl=maximum sustained vowel phonation; sr=sentence repetition of "Pop took his socks off"; read=reading of a standard passage; mono-monologue of "happy day".*

Table 5.
Group Means, Standard Errors (in parenthesis) and p Values for the Main Effect for SPL Measured at 30 cm

	Mean (standard error) in dBSPL			Main effect p value
Task	Young	Aged	IPD	
comfortable vowel	78.30 (3.22)	74.20 (3.72)	74.13 (3.22)	0.61
maximum vowel	84.02 (3.51)	73.14 (3.51)	72.98 (3.51)	0.10
sentence repetition	70.92 (3.27)	68.33 (3.27)	68.18 (3.27)	0.80
reading	72.26 (4.39)	69.00 (1.55)	68.78 (1.99)	0.61
monologue	71.68 (4.38)	71.23 (1.32)	68.96 (1.23)	0.76

SPL Observations

The young individuals produced speech tasks with higher SPLs (range: 70.92 - 84.02 dB at 30 cm) than both the aged individuals (range: 68.33 - 74.20 dB at 30 cm) and individuals with IPD (range: 68.18 - 74.13 dB at 30 cm). The latter two groups produced most of the speech tasks with similar SPLS. Statistically significant differences were not observed for any of the SPL measures (see Table 5).

Discussion

This study examined TA muscle activity in young and aged individuals and individuals with IPD in an attempt to gain more information about the underlying pathophysiology of the hypophonic voice disorder associated with IPD. In general, the individuals with IPD and the aged individuals showed TA amplitude levels that were lower than the young individuals. This pattern was observed fairly consistently across a variety of tasks. These findings may be explained physiologically by reductions in TA muscle mass, reductions in synaptic input to laryngeal motoneuron pools or a combination of these conditions.

Certain models for IPD suggest that the loss of dopamine in the substantia nigra pars compacta creates a neurochemical imbalance in both direct and indirect pathways of the cortico-basal ganglionic motor circuit[12]. The dopamine depletion produces either increased excitation or decreased inhibition of the globus pallidal (internal segment) and substantia nigra pars reticula neurons. Either effect (increased excitation or decreased inhibition) results in increased inhibition of thalamic nuclei and subsequent decreased excitation of motor cortical areas. Reduced excitation of motor cortical areas ultimately diminishes input to brainstem and spinal motoneuron pools and effects movement parameters, such as amplitude[13] and velocity[32].

The present findings of reduced TA muscle activity and low levels of tonic background EMG activity lend support to these theoretical models. They are also consistent with descriptions of hypokinesia (reduced movement) and may indicate bradykinesia (slowed movement). Reduced TA amplitudes at the onset of laryngeal movements (prephonatory measures) may increase overall movement time and thereby produce TA movements that are reduced in speed. These movement abnormalities that define hypokinesia and bradykinesia reportedly affect other speech subsystems. Reduced mean amplitudes and peak velocities have been observed in the lip[33] and jaw[34-35] of individuals with compared to individuals without IPD. Also, previous findings of longer voice onset times support the idea that onset parameters of laryngeal movements are affected by IPD[36]. Thus, it appears that in some individuals with IPD, reductions in muscle amplitudes and movement speeds, potentially caused by neurochemical suprabulbar imbalances, may be a primary factor affecting the laryngeal mechanism. Of course, it could be argued that age-related changes, such as neuromuscular denervation or muscle atrophy, affected TA muscle function. However, if age-related changes occurred in the individuals with IPD, they likely occurred in combination with disease-related changes because the individuals with IPD consistently showed *absolute* TA amplitudes that were lower than the aged individuals.

At first *relative* TA amplitude findings for the individuals with IPD may seem to contradict the idea that laryngeal motoneuron pools receive less input as a result of the neuropathological changes associated with IPD. However, it is important to recognize that *relative* TA amplitudes are influenced by TA maximums, such that the lower the divisor or maximum, the higher the quotient or *relative* measure. The findings of intermediate rather than low *relative* TA amplitudes for the individuals with IPD is related to the lower TA maximums that were observed for these individuals compared to the aged individuals. The reduced muscle activation ranges caused by these low TA maximums also indicate that the individuals with IPD used more of the TA's operational range, but very likely did so with similar or lower levels of neural drive to laryngeal motoneuron pools when compared to the aged individuals.

That the aged individuals showed the highest TA values for maximum gestures seems inconsistent with the notion that their TA muscle was affected by age. One explanation for this finding is that with increased motoneuron recruitment, abnormal muscle fibers, for example, denervated/reinnervated muscle fibers, may become activated and generate higher EMG amplitudes. Also, when the aged individuals produced maximum gestures, they may have attempted to overcome peripheral degeneration or denervation of other laryngeal muscles. In doing so, a less efficient, movement pattern that required greater TA activation may have been used. Because neuromuscular demands likely differ for non-maximum and maximum tasks, a similar type of compensation may not have been used for the production of non-maximum tasks.

SPL findings showed that the young individuals consistently produced higher SPLs and TA amplitudes than the other two groups. This is in agreement with results that have shown SPL can be controlled by increasing laryngeal muscle contraction and medial vocal fold adduction. Interestingly, SPLs were similar, but TA amplitudes were different for the aged individuals and the individuals with IPD. These findings indicate that the differences in TA amplitude levels were due to changes in physiology (reduced TA amplitudes) rather than changes in vocal loudness (effort). It also suggests that the individuals with IPD may compensate for physiologic changes by using a SPL control mechanism that differs from the aged individuals. Potential compensatory mechanisms include using: higher levels of the TA's operational range, other laryngeal adductors, such as the lateral cricoarytenoid, increased pulmonary effort[37] or different vocal tract shapes[38]. Without additional measures from other speech subsystems, the compensatory mechanism(s) used by the individuals with IPD to control SPL remains speculative.

Before concluding, methodological issues concerning electrode placement, performance variability, statistical analysis and medication state should be addressed. Because electrode placement differed for each individual, motor units of variable size were recorded[24]. This factor clearly impacts *absolute* EMG amplitude measures. However, this effect probably randomized itself out across individuals and did not contribute much to the group differences that were observed. Performance issues presumably did not influence the results either because the data were collected by the same examiner and consistent instructions and feedback were provided in an attempt to elicit the best possible performance from each individual. Results of statistical analyses showed that only the tasks with multiple repetitions were significant. This appears related to sample size. Statistical significance likely would have been obtained if the connected speech, quiet rest breathing and maximum tasks were repeated, particularly because most of these tasks showed directional differences that were similar to the sustained vowel phonation and sentence repetition tasks. Another possibility for the lack of statistical significance may be related to the observation that the connected speech, quiet rest breathing and maximum tasks showed more within group variability than the sustained vowel and sentence repetition tasks, particularly for the young and aged individuals. Finally, it should be noted that even though the individuals with IPD were experiencing the effect of their medication ("on" state), they still presented hypophonic voice symptoms and reduced TA amplitudes. These observations suggest that, despite potential medication effects, reduced TA amplitudes may contribute to the characteristic voice disorder associated with IPD.

In conclusion, evidence for excessive levels of TA activity was not observed. Rather, these findings point to decreased TA activity as a possible cause for some of the hypophonic voice abnormalities observed in IPD. This reduced muscle activity may result from neurochemical imbalances and subsequent reduced activation of laryngeal motoneuron pools.

Acknowledgements

Supported (in part) by research grants R01 DC-01 150 and P60 DC00976 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.

References

1. Logemann JA, Fisher HB, Boshes B, Blonsky ER. Frequency and cooccurrence of vocal tract dysfunction in the speech of a large sample of Parkinson patients. J Speech Hear Disord 1978;43:47-57.

2. Darley FL, Aronson AE, Brown JB. Motor speech disorders. Philadelphia: W Saunders, 1975.

3. Darley FL, Aronson AE, Brown JB. Differential diagnostic patterns of dysarthria. J Speech Hear Res 1969; 12:246-269.

4. Fox CM, Ramig LO. Vocal sound pressure level and self-perception of speech and voice in men and women with idiopathic Parkinson disease. Am J Speech-Lang Pathol 1997;6:85-94.

5. Ludlow CL, Bassich CJ. Relationships between perceptual ratings and acoustic measures of hypokinetic speech. In: McNeil MR, Rosenbek JC, Aronson AE, eds. The dysarthrias: Physiology, acoustics, perception, management. San Diego: College-Hill, 1984; 163-192.

6. Ramig LO, Scherer RC, Titze IR, Ringel SP. Acoustic analysis of voices of patients with neurologic disease:Rationale and preliminary data. Ann Otol Rhinol Laryngol 1988;97:164-172.

7. Perez KS, Ramig LO, Smith ME, Dromey C. The Parkinson larynx: Tremor and videostroboscopic findings. J Voice 1996; 10(4):341-361.

8. Hanson DG, Gerratt BR, Ward PH. Cinegraphic observation of laryngeal function in Parkinson's disease. Laryngoscope 1984;94(3):348-53.

9. Ludlow CL, Gallena S, Lou G, Menendez L. Laryngeal muscle activation abnormalities in parkinsonism. Presented at the American Speech-Language-Hearing Association; November, 1995; Orlando.

10. Gracco C, Marek K. Laryngeal electromyographic findings in Parkinson's disease. Neurology 1996;46(suppl 1):378.

11. Lee RG. Pathophysiology of rigidity and akinesia in Parkinson's disease. Eur Neurol 1989;29 (suppl 1):13-18.

12. Wichmann T, DeLong MR. Functional and pathophysiological models of the basal ganglia. Curr Opin in Neurobiol 1996;6:751-758.

13. Hallett M, Khosbin S. A physiologic mechanism of bradykinesia. Brain 1980;103:310-314.

14. Godaux E, Koulischer D, Jacquy J. Parkinsonian bradykinesia is due to depression in the rate of rise of muscle activity. Ann Neurol 1992;31:93-100.

15. Baker K, Scherer R, Smith M. Laryngeal muscle control in phonatory adduction. Presented at the Motor Speech Disorders Conference. February, 1996; Amelia Island.

16. Berke GS, Hanson DG, Gerratt BR, Trapp TK, Macagba C, Natividad M. The effect of air flow and medial adductory compression on vocal efficiency and glottal vibration. Otolaryngol Head Neck Surg 1996; 102:212-218.

17. Ramig LO, Dromey C. Aerodynamic mechanisms underlying treatment-related changes in vocal intensity in patients with Parkinson disease. J Speech and Hear Res 1996;39:798-807.

18. Cooper D. Maturation, character, and aging of laryngeal muscles. Presented at the Pacific Voice Conference; September, 1993; San Francisco.

19. Sato K, Hirano M. Age-related changes of elastic fibers in the superficial layer of the lamina propria of vocal folds. Ann Otol Rhinol Laryngol 1997; 106(1):44-8.

20. Sato K, Hirano, M. Age-related changes of the macula flava of the human vocal fold. Ann Otol Rhinol Laryngol 1995;104(11):839-44.

21. Mueller PB. The aging voice. Semin Speech Lang 1997;18(2):159-168.

22. Hirano M, Ohala J. Use of hooked-wire electrodes for electromyography of the intrinsic laryngeal muscles. J Speech and Hear Res 1969;12:362-373.

23. Hirano, M. Clinical examination of voice. In: Arnold GE, Winkel F, Wyke BD, eds. Disorders of human communication. New York: Springer-Verlag Wien, 1981; 11-24.

24. Ludlow, CL. Neurophysiologic assessment of patients with vocal motor control disorders. In: Assessment of speech and voice production: Research and clinical applications. Bethesda: National Institute on Deaftiess and Other Communication Disorders, National Institutes of Health, 1991;161-171.

25. Hirano M, Kakita Y. Cover-body theory of vocal fold vibration. In: Daniloff RG, eds. Speech Science: Recent advances. San Diego: College-Hill Press, 1985; 1-46.

26. Evarts EV, Teravainen H, Calne DB. Reaction time in Parkinson's disease. Brain 1981; 104: 167-186.

27. Flowers, K. Ballistic and corrective movements on an aiming task: Intension tremor and parkinsonian movement disorders compared. Neurology 1975;25:413-421.

28. Hunker CJ, Abbs JH, Barlow SM. The relationship between parkinsonian rigidity and hypokinesia in the orofacial system: A quantitative analysis. Neurology 1982;32:749-754.

29. Leanderson R, Meyerson BA, Persson A. Lip muscle fimction in Parkinsonian dysarthria. Acta Otolaryngol 1972;74:350-357.

30. Moore CA, Scudder RR. Coordination of jaw muscle activity in Parkinsonian movement: Description and response to traditional treatment. In: Yorkston KM, Beukelman DR, eds. Recent advances in clinical dysarthria. Boston: College-Hill, 1989; 147-163.

31. Winer BJ. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1971.

32. Montgomery EB, Nuessen J, Gonnan DS. Reaction time and movement velocity abnormalities in Parkinson's disease under different task conditions. Neurology 1991;41:1476-1481.

33. Caligiuri MP. Labial kinematics during speech in patients with Parkinsonian rigidity. Brain 1987; 110: 1033-1044.

34. Forrest K, Weismer, G. Dynamic aspects of lower lip movement in Parkinsonian and neurologically normal geriatric speakers' production of stress. J Speech Hear Res 1995;38:260-272.

35. Conner NP, Abbs JH. Task-dependent variations in parkinsonian motor impairments. Brain 1991;114:321-332.

36. Forrest K, Weismer G, Turner GS. Kinematic, acoustic and perceptual analyses of connected speech produced by Parkinsonian and normal geriatric adults. J Acoust Soc Am 1989;85(6);2608-2622.

37. Stathopoulous ET, Sapeinza CM. Respiratory and laryngeal fimction of women and men during vocal intensity variation. J Speech Hear Res 1993;36;64-75.

38. Titze I. Principles of Voice Production. Englewood Cliffs: Prentice Hall, 1994.

39. Hoehn MM, Yahr MD. Parkinsonism: Onset, progression and mortality. Neurology 1967;17:427-442.

NCVS Status and Progress Report - 13
April 1999, 93-104

Induced Fatigue Effects on Velopharyngeal Closure Force

David P. Kuehn, Ph.D.
Department of Speech and Hearing Sciences, The University of Illinois at Urbana-Champaign
Jerald B. Moon, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa

Abstract

The purpose of this investigation was to study the effects of induced velopharyngeal fatigue in adult speakers with normal mechanisms. A force sensing bulb was placed in the velopharynx to measure velopharyngeal closure force and intramuscular electrodes were inserted into the levator veli palatini muscle to sample muscle activation levels. The subjects' task was to repeat the syllable /si/ 100 times while an external load was placed on the velopharyngeal mechanism. The external load consisted of various levels of air pressure (0 as a control, 5,15,25,and 35 cm H2O relative to atmospheric pressure) delivered to the nasal passages via a tube and nasal mask assembly. Fatigue was defined as a declination of force across the series of syllables within a pressure condition and was depicted as the slope of a linear regression line that was fit to the data. The more negative the slope, the greater the rate of fatigue. Within each experimental pressure condition, small cyclic variations in force were noted about each regression line that corresponded to individual breath groups. This type of declination, within breath groups, has been reported in the literature previously. Overall declination in force over an entire series of syllables and over several breath groups is a new finding. It was possible to induce such fatigue in most subjects and greater rates of fatigue generally occurred at the higher levels of external loading, i.e., at 25 and 35 cm H2O. Two subjects, one male and one female, reached exhaustion. The female could not perform the syllable repetition task at 25 cm H2O and the male could not complete the task at 35 cm H2O. Three subjects, one female and two males, exhibited virtually no force declination even at the highest level (35 cm H2O) of external loading. There were no discernable differences in patterns of fatigue nor in initial velopharygeal closure force values between the male and female subjects.

It is generally believed that non-impaired individuals use muscle force levels for speech that are far below the maximum levels that they are capable of generating. Thus, Hinton and Arokiasamy (1997) and Thompson, Murdoch, and Stokes (1997) found that normal adult speakers use only about 10% of their maximum interlabial pressure for the production of /p/ in a conversational context. This is consistent with other reports suggesting that speech is produced with less than 20% of maximum orofacial muscle force (Amerman, 1993; Barlow & Abbs, 1984; Barlow & Netsell, 1986; Barlow & Rath, 1985; DePaul & Brooks, 1993; Muller, Milenkovic, & MacLeod, 1985; Netsell, 1982).

With regard to the velopharyngeal mechanism, Kuehn and Moon (1994) found that most normal subjects in their study used levator veli palatini activation levels for speech that are nearer the lower end of their total range of activity determined by a nonspeech task. This finding is consistent with the results reported for labial activity in normal subjects, again suggesting that normal individuals expend relatively little energy in producing speech. In contrast, in a follow-up study, Kuehn and Moon (1995) found that a group of subjects with cleft palate exhibited levels of levator veli palatini activity that were at the higher end of their total range of levator activity. This suggests that individuals with cleft palate might expend more muscle energy in producing speech, at least in relation to velopharyngeal activity, compared to individuals with normal intact palates.

Kuehn and Moon (1995) hypothesized that, although speakers with cleft palate may have to expend more energy in activating the levator muscle to achieve velar elevation for speech compared to individuals without cleft palate, speakers with cleft palate may limit muscle activity so as to remain just below a level of neuromuscular fatigue.

In this fashion, velopharyngeal closure may not be sufficiently tight and speech may be somewhat hypernasal during conversational speech. However, if the threshold of fatigue were exceeded and the rate of fatigue was sufficiently rapid, speech might quickly deteriorate because of neuromuscular exhaustion and become excessively hypernasal. Therefore, speakers with cleft palate who are somewhat hypernasal might tend to avoid overexertion and exhaustion of the muscles of velopharyngeal closure by avoiding excessive rates of neuromuscular fatigue.

Possible fatigue effects are frequently discussed in the clinical context of cleft palate related speech disorders. Anecdotal reports often indicate that a person's speech is more hypernasal when the individual is tired or when he or she "does not try hard enough." Therefore, an understanding of fatigue is important in addressing the possible deleterious effects on speech production.

A large literature exists pertaining to general aspects of neuromuscular fatigue (e.g., Asmussen, 1979; Enoka & Stuart, 1992; Gandevia, Enoka, McComas, Stuart, & Thomas, 1995). Many definitions of neuromuscular fatigue have been proposed but all have in common reduction in force. Bigland-Ritchie and Woods (1984) defined neuromuscular fatigue as "any reduction in the force generating capacity of the total neuromuscular system regardless of the force required in any given situation." Fatigue often is framed in relation to "endurance" or the ability to maintain a given level of force for an extended period of time. Luschei (1991) defined endurance as "resistance to fatigue." Several types of fatigue have been identified in relation to site of transference of energy within the central nervous system and the peripheral effector system (Bigland-Ritchie & Woods, 1984; Edwards, 1984). Although fatigue often is considered as an undesirable by-product of failure at various levels of the neuromuscular chain, it might also be considered in a positive vein as a protective mechanism to prevent an energy crisis (de Haan & Koudijs, 1994; Sargeant, 1994).

The bulk of the literature dealing with fatigue involves studies of the limb and torso musculature in humans as well as in animals. Cerny, Panzarella, and Stathopoulos (1997) reviewed some of the literature pertaining to respiratory strength and endurance conditioning. Several studies have been published that deal with head and neck muscles and, thus, are more directly related to speech phonation and articulation.

Robin, Luschei, and their colleagues have published a number of studies dealing with tongue and hand strength and endurance in normal subjects, "supranormal" subjects (trumpet players and debaters), and subjects with various disorders (Robin, Goel, Somodi, & Luschei, 1992; Robin, Somodi, & Luschei, 1991; Solomon, Lorell, Robin, Rodnitzky, & Luschei, 1995; Solomon, Robin, Mitchinson, VanDaele, & Luschei, 1996; Somodi, Robin, & Luschei, 1995; Stierwalt, Robin, Solomon, Weiss, & Max, 1996). Their results suggest that there may be differences in endurance among these groups. Thus, Robin et al. (1992) found that trumpet players and high school debaters ("supranormal" subjects), who presumably use their tongues more frequently and perhaps at higher activity levels than the matched normal control subjects, exhibited significantly longer endurance times, especially at the level of 25% of maximum voluntary contraction.

Robin et al. (1991) found greater tongue endurance times for normal children compared to children diagnosed with developmental apraxia of speech. Murdoch, Attard, Ozanne, and Stokes (1995) also found significantly longer tongue endurance times for normals compared to a group of subjects with developmental apraxia of speech. Stierwalt et al. (1996) found longer tongue endurance times for normal subjects compared to a matched group of subjects following traumatic brain injury. These latter results are consistent with findings for TBI subjects reported by McHenry, Minton, Wilson, and Post (1994), at least at the 2N target force level.

The studies cited above suggest that tongue muscle endurance might increase (fatigue decrease) with activity levels that exceed typical values and may decrease (fatigue increase) with certain neural or muscular disorders. However, reduced tongue endurance might not occur in all disorders affecting motor behavior. In that regard, although Solomon et al. (1995) found reduced tongue strength (maximum voluntary contraction) in a group of subjects with Parkinson's disease compared to matched control subjects, there was no difference in tongue endurance between the groups.

It is interesting to note that Somodi et al. (1995) and Solomon et al. (1996) found that the tongue is more fatigable than the hand in normal subjects. Their findings indicated that the hand exhibits twice the endurance time at a predetermined perceptual effort level that was comparable between the tongue and hand.

Fatigue has been studied in head and neck structures in addition to the tongue. Several fatigue studies have been published for the lips and jaw (see Solomon et al., 1996 for a partial listing; see also Jow & Clark, 1989; Lyons, Rouse, & Baxendale, 1993; Mao, Stein, & Osborn, 1993; Miles & Nordstrom, 1995). Cooper and Rice (1990) conducted fatigue studies in relation to canine vocal fold muscles.

Although anecdotal reports frequently mention the possibility of velopharyngeal fatigue in relation to speech, we were not able to find any published experimental studies relating to this activity. However, in relation to the playing of musical instruments, Bless, Ewanowski, and Dibble (1983) reported on their experience involving ten subjects who demonstrated velopharyngeal fatigue effects associated with prolonged and effortful playing of wind instruments. All of the subjects had a history of normal speech. Several measures were obtained from the subjects before and after using their wind instruments. The latter measures were obtained

just after velopharyngeal fatigue was induced by playing the wind instrument. Three of the ten subjects demonstrated evidence of velopharyngeal incompetence for speech following fatigue in that nasal airflow was observed on nonnasal speech sounds. Thus, it appears that in adverse conditions, velopharyngeal fatigue effects might generalize to speech in some individuals. We were interested in exploring this possibility first in a group of normal subjects and subsequently in a group of subjects with impaired velopharyngeal mechanisms.

The purpose of this investigation was to obtain information about physiologic fatigue in relation to velopharyngeal closure for speech in a group of normal individuals. Specifically, we sought to answer the following research questions:

1. Are normal adult speakers without velopharyngeal impairment demonstrably susceptible to the effects of induced physiologic fatigue?
2. Are there differences among normal adult speakers with regard to tolerance for induced physiologic fatigue?
3. Are there differences between male and female normal adult speakers with regard to tolerance for induced physiologic fatigue?

It is expected that the data obtained from this study of normal subjects will be useful in serving as a comparison for subsequent studies involving subjects with cleft palate and other conditions involving velopharyngeal dysfunction for speech.

Method

Subjects

Five normal adult males and five normal adult females served as subjects, the oldest of whom was 51 years at the time of this study. The subjects had no reported history of speech, language, hearing, or velopharyngeal impairment and were judged by the investigators to have normal speech at the time of data collection. All were speakers of the General American Dialect, typical of the midwestern United States.

Variables Measured

Velopharyngeal closure force was transduced using a silastic bulb device described in detail by Moon, Kuehn, and Huisman (1994) and Kuehn and Moon (1998). The bulb is teardrop-shaped, flattened anteriorly and posteriorly, and measures 5 mm front-to-back and 10 mm side-to-side. The bulb is rather easy to insert transnasally in most normal individuals because the bulb is soft, pliable, and collapsible upon insertion. The advantages of the size and shape of the bulb have been discussed previously (Moon et al., 1994; Moon, Kuehn, & Huisman, 1995). The bulb is attached to a silastic tube with a 3 mm outside diameter and the tube is connected to a Honeywell Microswitch (model 162PC01D) transducer. The transducer output was amplified using a Biocommunications Electronics (model 205) amplifier. The bulb was calibrated externally from the subjects by using applications of gram weights and, therefore, all force measures will be reported in grams.

Levator veli palatini muscle activity was recorded using stainless steel electrodes, 110 micrometers in diameter. The EMG signals were amplified using Biocommunications Electronics preamplifiers (model 301) and amplifiers (model 205). The audio signal from a dynamic microphone was amplified using a Nakamichi preamplifier and Tascam tape recorder (model 22-4) amplifier.

A light spray of 2% lidocaine topical anesthetic was applied to the more patent nostril for force bulb insertion and to the oral cavity for EMG electrode insertion. The bulb was placed at the most sensitive location within the velopharynx as the subject repeated the sound /s/. Velopharyngeal contact was monitored during bulb insertion in reference to the pressure trace displayed on an oscilloscope as the subject produced /s/ repeatedly. The hooked-wire electrodes were inserted perorally, using disposable hypodermic needles, within the dimple of the velum at an angle following the course of the levator muscle. Details of these placement procedures have been described previously (Kuehn & Moon, 1998).

Bulb force, EMG activity, and the audio signals were monitored on an oscilloscope (Tektronix model 5111A) and recorded on a Sony digital instrumentation recorder (model PC108M). EMG signals were full-wave rectified and low-pass filtered with a 40 ms time constant. Voice and force bulb signals were digitized at 1000 Hz sampling rate using a laboratory computer and commercially available analog-to-digital conversion software, CSpeech (Read, Buder, & Kent, 1990). Closure force signals were then digitally low-pass filtered at 30 Hz. Data were then displayed and analyzed using custom graphics and analysis routines. Levator EMG activation levels were normalized within each subject. This was accomplished by identifying the maximum peak value across the entire data set for a given subject and setting that value at 100%. All other EMG values recorded for that subject were referenced to the maximum value.

Experimental Procedure

After placement of the electrode wires in the levator muscle and the force bulb within the velopharynx, the tube leading outward from the bulb through the nasal passage was threaded through a small hole in a nasal mask. The mask was then positioned over the face to cover the nose so that air pressure could be delivered by an external source to the nasal cavities. The outer end of the bulb tube that extended from the mask was then reattached to the transducer.

The subjects' task was to produce the syllable /si/ 100 times for each of five experimental conditions. A few extra syllables were sometimes produced. Subjects were instructed to produce the syllables in a conversational manner using normal rate and loudness and to quickly inhale between breath groups.

A reverse-flow commercial vacuum cleaner was used to provide the external constant air pressure source that was delivered to the nasal passages. A variable transformer was used to control the voltage delivered to the vacuum cleaner and thus the level of pressure output to the subject. The air pressures delivered to the subjects were as follows: 0 (control condition with vacuum cleaner turned off but with nasal mask on), 5, 15, 25, and 35 cm H2O.

Pilot work showed that the 35 cm H2O was within safe limits. This was done by testing the tolerance level of each of the investigators by gradually increasing the nasal pressure level up to 35 cm H2O. Also, we measured the pressure generated by three adult males and two adult females during nose blowing. The males exhibited intranasal air pressures in the 50-60 cm H2O range and the females in the 40-50 cm H2O range. Therefore, we concluded that, although the 35 cm H2O level might be uncomfortable, it was tolerable and safe for the relatively brief duration of the experiment. This conclusion was warranted based on the subsequent performance of the subjects.

Each subject performed the experimental conditions in the same increasing sequence, from the 0 pressure control condition to the highest pressure condition, rather than randomizing across pressure conditions. This was done to avoid possible latent fatigue effects from influencing the results for lower level pressure conditions, thus giving a false or exaggerated impression of fatigue at lower pressure levels. A 1.5 minute rest interval was used between each pressure condition. This was done in an attempt to restore the neuromuscular system to a nonfatigued state and thus to prevent possible latent fatigue effects from influencing the subsequent condition. Following the 1.5 minute rest and just prior to the onset of the next condition, the pressure for the next condition was set. This was done by instructing the subject to prolong an /s/ sound while the investigators adjusted the transformer dial and observed the change in nasal mask pressure on an oscilloscope. This adjustment required about 30 sec.

There did not appear to be a latent fatigue effect of one experimental condition upon subsequent conditions for most subjects. This follows from the fact that for nine of the ten subjects, the initial velopharyngeal closure force was either similar to or actually increased relative to the initial closure force of the previous experimental condition (see Figures 6 and 7).

Data Analysis and Reliability

Force and EMG measures were obtained during the vowel /i/ rather than during the consonant /s/. This ensured that the force measures obtained were related to active muscle activity in holding the velum up and back rather than possibly related to an indirect effect of heightened pressure within a nearly-closed chamber with the tongue tip making nearly complete oral closure during the /s/ consonant. That is, within a closed chamber, it is possible that the bulb could have functioned as a pressure sensor for heightened intraoral air pressure for /s/ in addition to contact pressure in relation to velum to pharyngeal wall contact. This was especially true once fatigue began to occur and tight contact between the velum and posterior pharyngeal wall was lost or reduced. This would not be a factor for the vowel /i/ because intraoral air pressure for vowels is at or near atmospheric pressure and thus the forces sensed by the bulb would be expected to be due solely to velum-pharyngeal wall contact. This contact would be generated by the muscles of velopharyngeal closure, including that of the levator veli palatini.

Intra- and inter-measurer reliability were determined for similar data in a previous study (Kuehn & Moon, 1998) by re-measuring the entire data sets for force and EMG activity for one subject. All Pearson r values exceeded 0.98.

Results

Figure 1 shows an example of the force data for one subject for the 25 cm H2O pressure condition. Figure 2 shows the accompanying data for the levator veli palatini muscle for the same subject and pressure condition. Linear

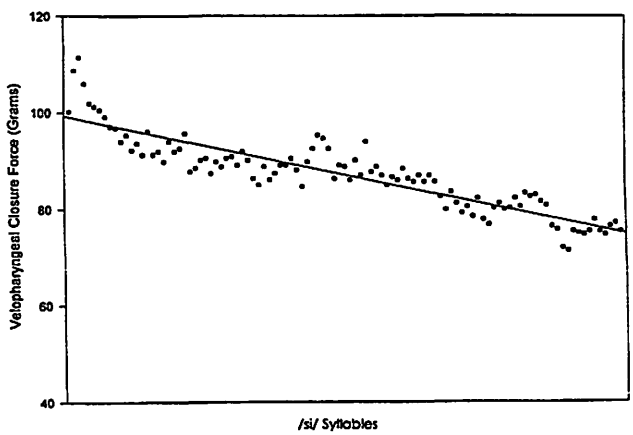


Figure 1. Velopharyngeal closure force values in grams across 105 syllable repetitions for one subject. Each data point represents the force value obtained during the vowel of the syllable /si/. Nasal cavity air pressure = 25 cm H2O.

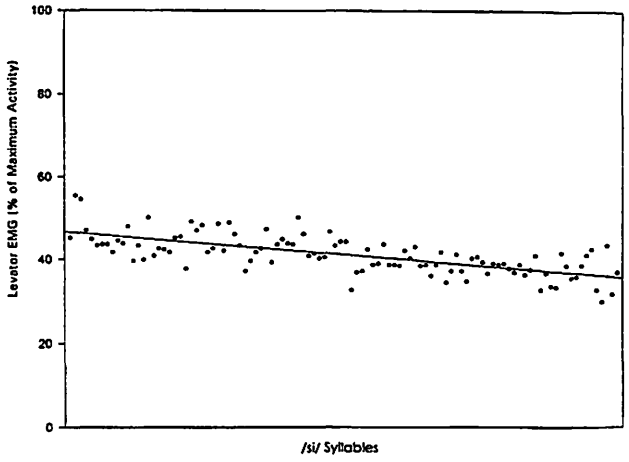


Figure 2. Levator veli palatini EMG values for the corresponding velopharyngeal closure force values in Figure 1. EMG values are expressed as percentages of maximum activity for the individual subject shown. Each data point represents the levator EMG value obtained during the vowel of the syllable /si/. Nasal cavity air pressure = 25 cm H2O.

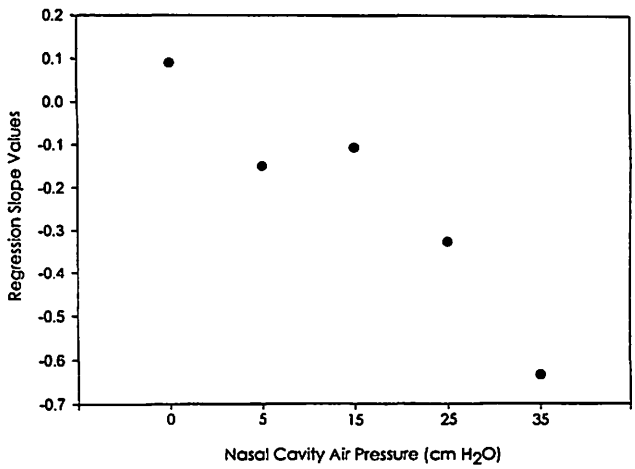


Figure 3. Regression slope values relating velopharyngeal closure force to repeated syllable production for one subject across the five experimental conditions. Experimental conditions are nasal cavity air pressures in cm H2O. The greater the negative value per experimental condition, the greater the force declination across syllables for each experimental condition.

regression lines are included on each of these figures that indicate the slope of the function of velar closure force (Figure 1) and levator EMG activity (Figure 2) as a function of syllable number. Cyclic activity about the regression lines can be observed, especially for the force data. This cyclic activity appeared to be related to breath groups in that there tended to be a slight increase followed by a slight decrease in force associated with each breath group. Thus in Figure 1, approximately eleven breath groups (about nine syllables per breath group on average) are depicted for the 105 syllable data points.

In addition to the short term declinations of force within breath groups, it is obvious, judging from Figure 1, that there was an overall declination of force across the course of syllable production and across breath groups as well. This longer term force declination is reflected in the negative slope of the linear regression line. A slight overall declination in levator activity is also apparent across the course of syllable production (Figure 2). We interpret these overall declinations as depicted by the negative slopes of the regression lines, especially that of force (Figure 1), as a sign of neuromuscular fatigue.

The prevailing level of muscle activity or change in that level is not consistently associated with force declination and, therefore, with fatigue, as defined in the literature. That is, with a decrease in force, accompanying muscle activity might also decrease, stay relatively constant, or increase. Clearly, there is an interaction between peripheral and central mechanisms that dictate whether changes in muscle activity will occur and the exact nature of that change if it occurs. Results of the present study reflected this lack of systematic relation between force change and muscle activity in that levator activity sometimes decreased, stayed fairly constant, or increased over the course of syllable production. No obvious pattern was discernible in this regard. Therefore, the remainder of the results presented are those associated with the force data. The exact relation between muscle activity and force declination will be the subject of a subsequent study in our laboratory.

Because there is no standard method of portraying fatigue in the literature, other than demonstrating a declination in force over time, we chose to focus on the slope of the regression line relating force to syllable number as an indication of fatigue. The more negative the slope, the greater the rate of fatigue involved for a given condition. Figure 3 shows an example of slope values for one subject. For this subject, there was actually a slight overall increase in force over the course of syllable production for the control condition (control condition = 0 cm H2O). The slope value in this case was 0.1 which is close to zero, that is, a constant function. Figures 4 and 5 show that all 10 subjects exhibited slope values near zero for the control condition. Because the slope was near zero for the control condition, this indicates that the subjects did not exhibit fatigue or that the rate of fatigue was very gradual as they repeated syllables with no heightened pressure introduced into their nasal cavities.

Figure 3 shows that the slope values remained near zero for the 5 and 15 cm H2O conditions for this subject, again, indicating that this subject exhibited very gradual rates of fatigue at these pressure levels. However, when nasal air pressure was increased to 25 and 35 cm H2O, slope values became progressively more negative indicating that greater rates of fatigue occurred as nasal air pressure increased.

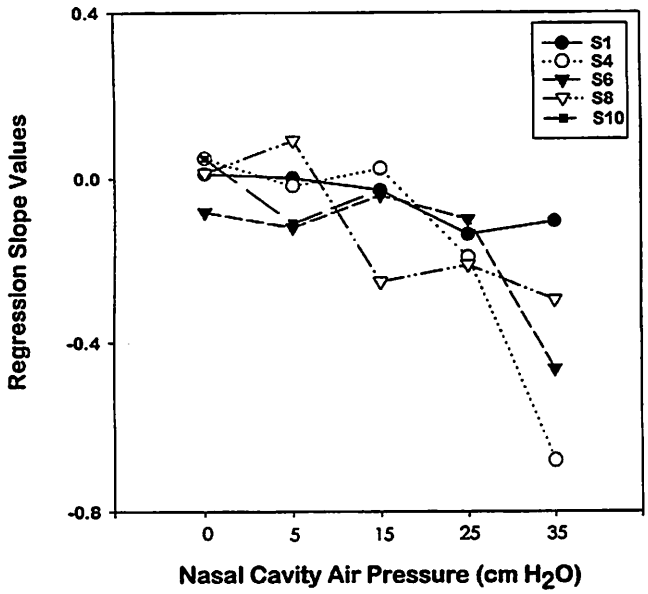


Figure 4. Regression slope values for the female subjects across the five experimental conditions. Experimental conditions are nasal cavity air pressures in cm H2O.

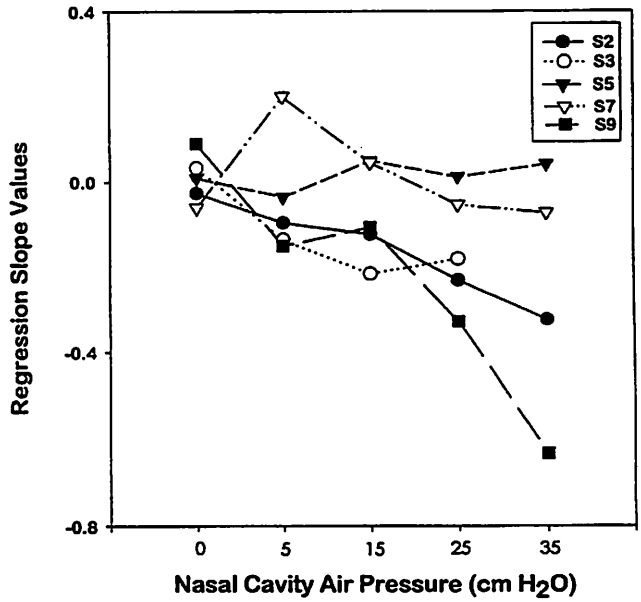


Figure 5. Regression slope values for the male subjects across the five experimental conditions. Experimental conditions are nasal cavity air pressures in cm H2O.

Figures 4 and 5 show composite results for fatigue slope values for female and male subjects respectively across the experimental conditions. In general, both male and female subjects appear to be fairly resistant to the effects of induced fatigue for the control condition (0 cm H2O) and the two lowest pressure conditions of 5 and 15 cm H2O.

One female subject, Subject 10, had difficulty sustaining syllable repetition at the 25 cm H2O level. Intranasal pressure was reduced to 20 cm H2O for this subject. However, she could not sustain syllable repetition at this level either and reached a point of exhaustion during which no velopharyngeal closure force could be generated at the level of 20 cm H2O. Thus, 15 cm H20 was the highest sustainable level obtained for this subject.

For the 25 cm H2O condition, the remaining four female subjects were remarkably similar in their slope values. All four demonstrated declination of force but not extensive. The males, as a group, were more variable than the females for this condition. Two males, S5 and S7, exhibited essentially no fatigue effects whereas the other three males did demonstrate more substantial force declination.

At the highest level of pressure, 35 cm H2O, the subjects were quite variable in their ability to withstand the effects of fatigue, as shown in Figures 4 and 5. Female subject S10 did not participate at 35 cm H2O (and did not complete the task at 25 cm H2O) because, as indicated above, she reached the level of velopharyngeal exhaustion at 20 cm H2O. One male subject, S3, reached exhaustion at the 35 cm H2O level and was not able to complete 100 syllable repetitions at that level. The remaining four female and four male subjects were quite diverse with regard to their slope values. Female subject S1 and male subjects S5 and S7 demonstrated essentially no fatigue effects even at this highest level of pressure. Female subjects S6 and S8 and male subject S2 exhibited moderate levels of fatigue effects whereas female subject S4 and male subject S9 demonstrated more marked rates of fatigue.

Figures 6 and 7 show initial velopharyngeal closure force values for each experimental condition for the female and male subjects. The individual data points are averages over the first 15 syllable productions per subject for each experimental condition. It can be seen that, as a group, the females were more similar to each other in initial force values than the males for nasal cavity air pressure conditions 0 (the control condition) and 5 cm H2O. Across all five experimental conditions, the males, as a group, exhibited a greater range of variability compared to the females. However, the initial force values dispersed rather widely across the female subjects as well for the conditions 15, 25, and 35 cm H2O. At 35 cm H2O, the lowest starting value was 28 grams for female subject S6 and the highest starting value was 124 grams for male subject S7. In comparing the data in Figures 6 and 7, it can be seen that there is considerable overlap between males and females suggesting that there is no systematic difference between the two groups in the initial force of velopharyngeal closure for syllable repetition of /si/ regardless of the experimental condition.

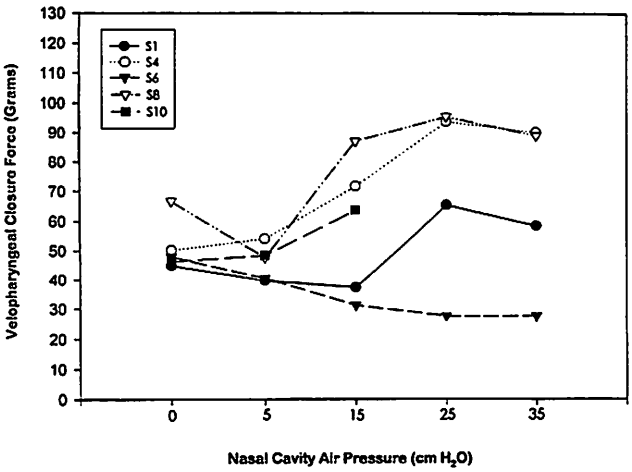


Figure 6. Initial velopharyngeal closure force values across the five experimental conditions for the female subjects. Each data point is an average across the first 15 syllables in a sequence for each experimental condition per subject.

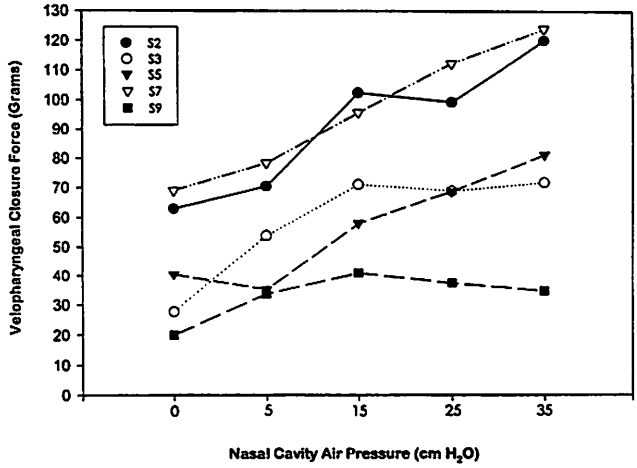


Figure 7. Initial velopharyngeal closure force values across the five experimental conditions for the male subjects. Each data point is an average across the first 15 syllables in a sequence for each experimental condition per subject.

With regard to individual patterns, only one subject, female subject S6, showed a trend of decreasing initial velopharyngeal force as the nasal cavity air pressure increased (Figure 6). All other subjects had similar or actually increasing initial closure forces with increasing nasal cavity air pressure. This suggests that for most subjects in this study, there was not a latent fatigue effect of one fatiguing stimulus upon subsequent experimental series. In other words, it appears that the rest period between experimental conditions was sufficient to reset and presumably to restore velopharyngeal force control.

Discussion

Short Term Force Declination

Both short term as well as longer term velopharyngeal force declination were observed in this study. Short term cyclic force variation occurring within breath groups was observed regularly and is consistent with declination within breath groups reported by other investigators. Specifically, Krakow, Bell-Berti, and Wang (1995), using a Velotrace, reported declination of velar position from the beginning to the end of sentences. Presumably, declination of velar positions, as observed by Krakow et al. would be correlated with declination of velopharyngeal closure force, as observed in the current study. Krakow et al. related their findings to previous reports of declination involving respiratory, laryngeal, and other supralaryngeal structures during speech. Thus, short term declinations occurring within breath groups appear to be rather pervasive within the speech apparatus in general.

Considering expenditure of energy associated with velopharyngeal control, Figure 1 of the current study suggests that short term force declinations occurring within individual breath groups are relatively minor in comparison to the longer term, potentially much greater overall force declination that can occur across several breath groups. Therefore, importantly, the results of this study demonstrate that short term force declinations are accommodated easily by the normal velopharyngeal mechanism and can be concatenated for long stretches, much like a modulating signal, without failure of the velopharyngeal mechanism to achieve functional closure.

Long Term Force Declination

Rate of force declination and total time over which force declines are both important variables to consider in assessing the effects of physiologic fatigue. For most episodes of syllable production in this study, rate of force declination, as determined by the regression slopes, was gradual enough so that the subject could continue to achieve velopharyngeal closure. However, for two subjects, S3 and S10, the rate of force declination was rapid at the higher external loads and exhaustion was reached within the time limits imposed in this experiment.

The relation between rate of force declination and time in reaching exhaustion is depicted in the stylized graph of Figure 8. In reference to this figure, we define "fatigue" as any declination in velopharyngeal force that occurs steadily and continuously over time. The lighter gray region, between level c (starting force value for an utterance) and level b (force value at threshold of exhaustion), is considered as the zone of fatigue. "Exhaustion" is defined as the end product of fatigue at which level the neuromuscular system cannot maintain sufficient force to accomplish the intended task. The darker gray region, between level b and level a (zero force) on Figure 8, is considered as the zone of exhaustion.

In the case of the velopharyngeal mechanism, exhaustion would be manifested as the lack of a functional separation between the oral and nasal cavities. In the present study, exhaustion was signaled by the inability of the velopharyngeal mechanism to counteract the opposing pressure head of the external air pressure load. The result was that the external pressure source opened the velopharyngeal port in a retrograde manner (i.e., from the nasal side) and a snorting sound was produced. The snorting sound was not under the control of the subject and was caused by the air pressure vibrating the soft palate presumably in a nasal-to-oral direction. It is possible that an individual could fail to achieve sufficient velopharyngeal closure to provide a functional separation between oral and nasal cavities simply because of insufficient strength to provide the needed force. We would consider this as a type of neuromuscular "failure." Therefore, neuromuscular failure to achieve sufficient velopharyngeal closure might occur due to two different underlying causes according to our terminology: 1) exhaustion, as an end product of fatigue or 2) insufficient strength.

Figure 8 shows several regression lines each of which is meant to depict force over time for an individual speaker. In the strictest sense, the line that remains horizontal at level c is an example of nonfatigue, that is, no force declination, and all other solid lines that are diagonal are examples of fatigue regardless of how slight the negative slope value might be.

The results of the current study indicated that the subjects varied in their ability to resist the effects of fatigue, but resistance depended on the level of the external air pressure load. All subjects were fairly resistant to fatigue up to a point. Both female and male subjects tended to cluster at the regression slope value of zero for the control condition and also for the lower nasal air pressure conditions of 5 and 15 cm H2O. This suggests that for reiterative speaking tasks, normal individuals tend not to demonstrate rapid rates of fatigue even when some external loading is imposed on the velopharyngeal mechanism. These findings appear to be consistent with the results of Webb, Starr, and Moller (1992) who observed little or no change in nasality in a group of eight normal subjects, aged 15-22, after speaking continuously for a period of 40 minutes.

It is possible that individuals with impaired velopharyngeal mechanisms might exhibit more rapid rates of force declination such as that labeled x and y in Figure 8. In these hypothetical examples, the two speakers reach exhaustion as shown in the figure, with x reaching exhaustion sooner than y. Once in the zone of exhaustion, the ability to generate sufficient force to accomplish the given task likely continues to fall off very rapidly as suggested by the dashed lines.

Figure 9 is a stylized graph that depicts the theoretical relation between time to exhaustion versus percent maximal voluntary contraction (MVC). The latter is a common measure reported in the physiology literature. MVC is difficult to assess for the velopharyngeal muscle system because it is difficult to contract the muscles of velopharyngeal closure as forcefully as possible on command. This is in contrast, for example, to a hand grip task. However, estimates, at least in a gross fashion, might be derived from our previous work (Kuehn & Moon, 1994) in which normal subjects tended to use levator veli palatini activation levels for speech that were at the lower end of the total range of levator activation levels as determined in relation to a blowing task. In reference to that work, it is likely that normals use roughly about 20% MVC or less during speech.

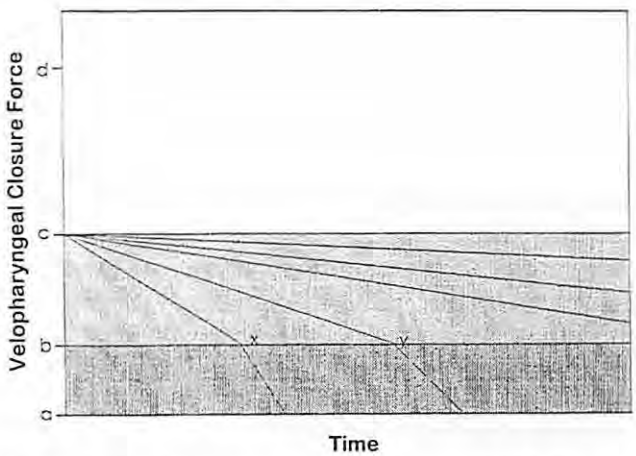


Figure 8. Velopharyngeal closure force versus time. Each line on the graph from starting point c represents force generation for an individual subject. a = zero force level; b = force level at threshold of exhaustion, c = beginning force level for a given utterance string; d = maximum possible force. a-b, the darker gray region = zone of exhaustion; b-c, the lighter gray region = zone of fatigue. x and y are shorter and longer times to exhaustion respectively.

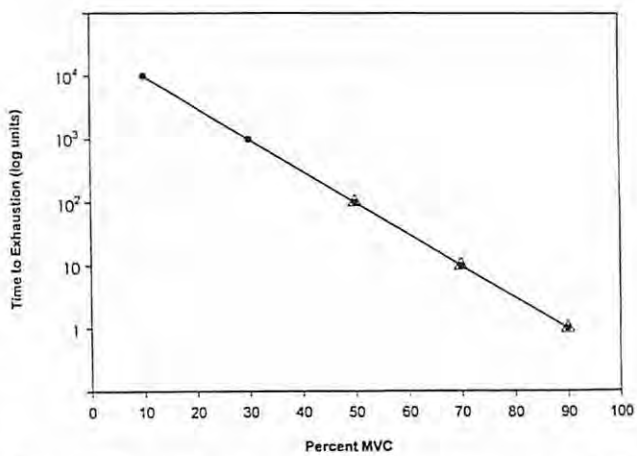


Figure 9. Time to exhaustion in arbitrary log units versus percent MVC (maximal voluntary contraction).

Figure 9 demonstrates that the higher in one's MVC that a task is performed, the shorter will be the time at which exhaustion is reached. In the case of normal velopharygeal closure, and assuming that one uses a level of activation that is at 20% MVC or less, one could speak for a very long time without reaching exhaustion. Again, this is consistent with the results of the no-load and small-load conditions of the current study and with the results of the Webb et al. (1992) study discussed above.

However, if an individual exerts more energy and uses a higher percent MVC for various reasons, the time to exhaustion might become a significant factor and could become very short. In the current study, several subjects increased their starting force values with increased external air pressure loads (especially Subjects 2,4,5,7,8,10). This can be deduced by comparing Figure 4 versus 6 and Figure 5 versus 7. Presumably, these subjects used a higher percent MVC as the external load increased. It is likely that the time to exhaustion for these subjects would have decreased correspondingly and would have been reached if the experiment had been prolonged in time. Indeed, two subjects, S3 and S10, did reach exhaustion within the time allotted to perform the syllable production task.

Another important aspect to consider in relation to the possible effects of fatigue on velopharyngeal closure is the level of force required to achieve closure and where in one's total force range that level of force lies. Figure 10 addresses this issue. Presumably, normal individuals achieve velopharyngeal closure with relatively little expenditure of energy (Kuehn and Moon, 1994). In reference to Figure 10, we suggest that speakers without velopharyngeal impairment would achieve velopharyngeal closure with as little as 20% MVC and perhaps touch closure with as little as 10% MVC. Contraction effort above about 20% MVC typically would be unnecessary in most speaking situations and would be overly fatiguing in the long term. That is, in reference to Figure 9, increasing percent MVC reduces time to exhaustion owing to the more rapid rate of fatigue and, presumably, more rapid depletion of energy. Contraction much under 20% MVC would be insufficient to achieve consistent velopharyngeal closure (Figure 10) and logically would be avoided in typical situations.

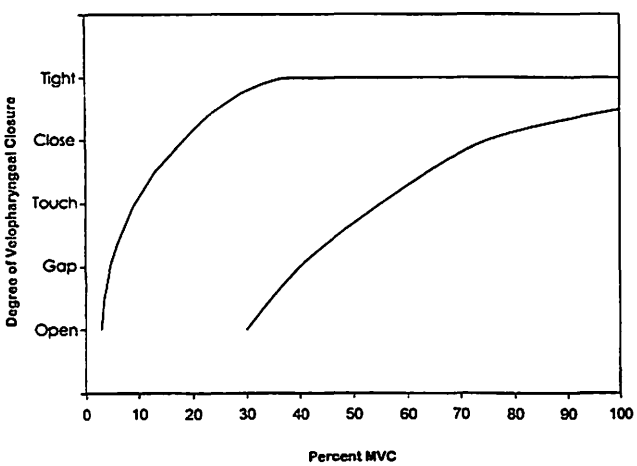


Figure 10. Degree of velopharyngeal closure versus percent MVC (maximal voluntary contraction).

It is interesting to speculate about speakers with impaired velopharyngeal mechanisms in reference to Figure 10. For subjects without impairment, a cluster of curves near the one shown in Figure 10 would be expected. But for subjects with cleft palate, for example, the curves might be expected to be shifted downward and to the right. That is, even with greater levels of effort (increased percent MVC), speakers with cleft palate might still have lower degrees of velopharyngeal closure than speakers without impairment. This suggestion is consistent with the results of a previous study (Kuehn & Moon, 1995) in which it was found that speakers with cleft palate, exhibiting some degree of hypernasality, tend to use levator veli palatini activation levels that are at the higher end of their muscle's total activation range as determined in relation to a blowing task. Thus, if these individuals used even higher muscle activation levels (higher percent MVC) to achieve a greater degree of velopharyngeal closure (Figure 10) they might increase rate of fatigue markedly (Figure 8) and quickly reach a level of exhaustion (Figures 8 and 9).

Sex Comparison

With regard to possible differences between males and females, Figures 4 versus 5 and Figures 6 versus 7 indicate few if any systematic differences in either velopharyngeal fatigue patterns or initial velopharyngeal closure force values. The latter is consistent with a previous report in which no difference in absolute values of velopharyngeal closure force for speech was found between normal adult males and females when the data were grouped (Kuehn & Moon, 1998).

Individual Differences and Implications for Future Research

In reference to the highest loading provided in this study, 35 cm H2O, there was great variability among the normal subjects with regard to fatigability of the velopharyngeal mechanism (Figures 4 and 5). In typical circumstances, such loading would not occur. However, for those subjects who are more susceptible to the effects of induced fatigue (i.e., Subjects 2,4,6,8,9, and especially Subjects 3 and 10) it is possible that velopharyngeal functioning for these subjects might be affected if the velopharyngeal mechanism is taxed. Examples of such taxing phenomena might be playing a musical wind instrument extensively, overall bodily fatigue, trauma, neurologic insult, or perhaps

combinations thereof. Of course, if the velopharyngeal mechanism itself is frankly impaired, such as with cleft palate, the effects of physiologic fatigue might be amplified in these individuals.

It is interesting to note in reference to Figures 6 and 7 that several subjects actually increased their levels of velopharyngeal closure force in relation to increasing levels of nasal air pressure. This suggests that latent effects of fatigue did not build up from one experimental condition to the next for these subjects and, on the contrary, that muscle potentiation (Botterman, 1995) actually might have occurred at least at the onset of each new series of syllable productions. It is possible that the subjects reset their force output control in anticipation of greater load demands on the mechanism. It should be noted that the subjects were not given feedback about their velopharyngeal force output but they were expecting increasing external loads (i.e., increasing nasal air pressure) in subsequent conditions.

The possibility of a potentiation effect would be interesting to pursue given that it might be under control of the central nervous system and not solely reliant on peripheral muscle mechanics. Thus, even if the musculature is deficient, it might be possible to compensate for that deficiency, at least to some degree, by tapping into central control mechanisms.

Our goal in studying physiologic fatigue of the velopharyngeal mechanism is to apply this knowledge in treating individuals with velopharyngeal impairment. A greater understanding of individual and group differences might aid in treatment regimens, especially those that utilize muscle resistance approaches such as that of continuous positive airway pressure, CPAP (Kuehn, 1991). We are currently studying the effects of physiologic fatigue on the velopharyngeal mechanism in a group of subjects with cleft palate.

Acknowledgments

This study was supported in part by PHS Research Grants DC-00976 from the National Institute on Deafness and Other Communication Disorders and DE-10436 from the National Institute of Dental Research. We would like to acknowledge Wendy Edwards for her assistance in data analysis and graphics.

References

Amerman,J.D.(1993). A maximum-force-dependent protocol for assessing labial force control. Journal of Speech and Hearing Research,36,460-465.

Assmussen,E.(1979). Muscle fatigue. Medicine and Science in Sports,11,412-420.

Barlow,S.M.,& Abbs,J.H.(1984). Orofacial fine motor control impairments in congenital spasticity:Evidence against hypertonus related performance deficits. Neurology,34,145-150.

Barlow,S.M.,& Netsell,R.(1986). Differential time-force control of the upper and lower lips. Journal of Speech and Hearing Research,29,163-169.

Barlow,S.M.,& Rath,E.M.(1985). Maximum voluntary closing forces in the upper and lower lips of humans. Journal of Speech and Hearing Research,28,373-376.

Bigland-Ritchie,B.,& Woods,J.J.(1984). Changes in muscle contractile properties and neural control during human muscular fatigue. Muscle & Nerve,7,691-699.

Bless,D.M.,Ewanowski,S.J.,& Dibble,D.(1983). Velopharyngeal valving problems related to extensive playing of wind instruments. XIX Congress of the International Association of Logopaedics and Phoniatrics. University of Edinburgh, Scotland, United Kingdom.

Botterman,B.R.(1995). Task-dependent nature of fatigue in single motor units. In S.C.Gandevia, R.M.Enoka, A.J.McComas, D.G.Stuart, & C.K.Thomas. Fatigue: Neural and muscular mechanisms (pp. 351-360). New York: Plenum Press.

Cerny,F.J.,Panzarella,K.J.,& Stathopoulos,E.(1997). Expiratory muscle conditioning in hypotonic children with low vocal intensity levels. Journal of Medical Speech-Language Pathology,5,141-152.

Cooper,D.S.,& Rice,D.H.(1990). Fatigue resistance of canine vocal fold muscle. Annals of Otology, Rhinology & Laryngology,99,228-233.

deHaan,A.,& Koudijs,J.C.M.(1994). A linear relationship between ATP degradation and fatigue during high intensity dynamic exercise in rat skeletal muscle. Experimental Physiology,79,865-868.

DePaul,R.,& Brooks,B.R.(1993). Multiple orofacial indices in amyotrophic lateral sclerosis. Journal of Speech and Hearing Research,36,1158-1167.

Edwards,R.H.T.(1984). New techniques for studying human muscle function, metabolism, and fatigue. Muscle & Nerve,7,599-609.

Enoka,R.M.& Stuart,D.G.(1992). Neurobiology of muscle fatigue. Journal of Applied Physiology,72,1631-1648.

Gandevia,S.C.,Enoka,R.M.,McComas,A.J.,Stuart,D.G.,& Thomas,C.K.(1995). Fatigue: Neural and muscular mechanisms. New York: Plenum Press.

Hinton,V.A.,& Arokiasamy,W.M.C.(1997). Maximum interlabial pressures in normal speakers. Journal of Speech, Language, and Hearing Research,40,400-404.

Jow,R.W.,& Clark,G.T.(1989). Endurance and recovery from a sustained isometric contraction in human jaw-elevating muscles. Archives of Oral Biology,34,857-862.

Krakow,R.A.,Bell-Berti,F.,& Wang,Q.E.(1995). Supralaryngeal declination: Evidence from the velum. In F.Bell-Berti,& L.J.Raphael (Eds.),Producing speech: Contemporary issues (pp.333-353). New York: American Institute of Physics Press.

Kuehn,D.P.(1991). New therapy for treating hypernasal speech using continuous positive airway pressure (CPAP). Plastic & Reconstructive Surgery,88,959-966.

Kuehn,D.P.,& Moon,J.B.(1994). Levator veli palatini muscle activity in relation to intraoral air pressure variation. Journal of Speech and Hearing Research,37,1260-1270.

Kuehn,D.P.,& Moon,J.B.(1995). Levator veli palatini muscle activity in relation to intraoral air pressure variation in cleft palate subjects. Cleft Palate-Craniofacial Journal,32,376-381.

Kuehn,D.P.,& Moon,J.B.(1998). Velopharyngeal closure force and levator veli palatini activation levels in varying phonetic contexts. Journal of Speech, Language, and Hearing Research,41, 51-62.

Luschei,E.S.(1991). Development of objective standards of nonspeech oral strength and performance: An advocate's views. In C.A.Moore,K.M.Yorkston,& D.R.Beukelman (Eds.),Dysarthria and apraxia of speech: Perspectives on management(pp.3-14). Baltimore: Paul H. Brookes Publishing Co., Inc.

Lyons,M.F.,Rouse,M.E.,& Baxendale,R.H.(1993). Fatigue and EMG changes in the masseter and temporalis muscles during sustained contractions. Journal of Oral Rehabilitation,20,321-331.

Mao,J.,Stein,R.B.& Osborn,J.W.(1993). Fatigue in human jaw muscles: A review. Journal of Orofacial Pain,7,135-142.

McHenry,M.A.,Minton,J.T.,Wilson,R.L.,& Post,Y.V.(1994). Intelligibility and nonspeech orofacial strength and force control following traumatic brain injury. Journal of Speech and Hearing Research,37,1271-1283.

Miles,T.S.,& Nordstrom,M.A.(1995). Fatigue of jaw muscles and speech mechanisms. In S.C.Gandevia, R.M.Enoka, A.J.McComas, D.G.Stuart, & C.K.Thomas. Fatigue: Neural and muscular mechanisms (pp. 415-426). New York: Plenum Press.

Moon,J.B.,Kuehn,D.P.,& Huisman,J.J.(1994). Measurement of velopharyngeal closure force during vowel production. Cleft Palate-Craniofacial Journal,31,356-363.

Moon,J.B.,Kuehn,D.P.,& Huisman,J.J.(1995). Author response: Letter to the editor. Cleft Palate-Craniofacial Journal,32,263.

Muller,E.M.,Milenkovic,P.H.,& MacLeod,G.E.(1985). Perioral tissue mechanics during speech production.In C.Delisi & J.Eisenfield (Eds.), Proceedings of the Second IMAC International Symposium on Biomechanical Systems Modelling,1,363-371.

Murdoch,B.E.,Attard,M.D.,Ozanne,A.E.,& Stokes,P.D.(1995). Impaired tongue strength and endurance in developmental verbal dyspraxia: A physiologic analysis. European Journal of Disorders of Communication,30,51-64.

Netsell,R.(1982). Speech motor control and selected neurologic disorders.In S.Grillner,B.Lindblom,J.Lubker,& A.Pearson (Eds.), Speech motor control (pp.247-261).New York:Pergamon Press.

Read,C.,Buder,E.H.& Kent,R.D.(1990). Speech analysis systems: A survey. Journal of Speech and Hearing Research,33,363-374.

Robin,D.A.,Goel,A.,Somodi,L.B.,& Luschei,E.S.(1992). Tongue strength and endurance: Relation to highly skilled movements. Journal of Speech and Hearing Research,35,1239-1245.

Robin,D.A.,Somodi,L.B.,& Luschei,E.S.(1991). Measurement of strength and endurance in normal and articulation disordered subjects. In C.A. Moore, K.M. Yorkston, & D.R. Beukelman (Eds.), Dysarthria and apraxia of speech: Perspectives on management (pp. 173-184). Baltimore: Paul H. Brookes Publishing Co., Inc.

Sargeant,A.J.(1994). Human power output and muscle fatigue. International Journal of Sports Medicine,15,116-121.

Solomon,N.P.,Lorell,D.,Robin,D.A.Rodnitzky,R.L.& Luschei,E.S.(1995). Tongue strength and endurance in mild to moderate Parkinson's disease. Journal of Medical Speech-Language Pathology,3,15-26.

Solomon,N.P.,Robin,D.A.,VanDaele,D.J.,& Luschei,E.S.(1996). Sense of effort and the effects of fatigue in the tongue and hand. Journal of Speech and Hearing Research,39,114-125.

Somodi,L.B.,Robin,D.A.,& Luschei,E.S.(1995). A model of "sense of effort" during maximal and submaximal contractions of the tongue. Brain and Language,51,371-382.

Stierwalt,J.A.G.,Robin,D.A.,Solomon,N.P.,Weiss,A.L.,& Max,J.E.(1996). Tongue strength and endurance: Relation to the speaking ability of children and adolescents following traumatic brain injury. In D.A.Robin, K.M.Yorkston, & D.R.Beukelman (Eds.),Disorders of Motor Speech: Assessment, Treatment, & Clinical Characterization (pp. 241-256). Baltimore: Paul H. Brookes Publishing Co., Inc.

Thompson,E.C.,Murdoch,B.E.,& Stokes,P.D.(1997). Interlabial contact pressures during performance of speech and nonspeech tasks in young adults. Journal of Medical Speech-Language Pathology,5,191-199.

Webb,M.,Starr,C.D.,& Moller,K.(1992). Effects of extended speaking on resonance of patients with cleft palate. Cleft Palate-Craniofacial Journal,29,22-26.

NCVS Status and Progress Report - 13
April 1999, 105-114

Vocal Fold Bulging Effects on Phonation Using a Biophysical Computer Model

Fariborz Alipour, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa
Ronald C. Scherer, Ph.D.
Department of Communication Disorders, Bowling Green State University

Abstract

Glottal adduction is a primary laryngeal variable determining glottal configuration and phonatory output. Greater adduction of the vocal folds manifests as a reduction of the arytenoidal vocal process gap as well as medial bulging of the membranous vocal fold surface. This study examined phonatory effects due to changing the degree of bulging using a computational model. Bulging was modeled as a quadratic surface, and the bulging parameter B was the coefficient (0 to 1) to the active muscle stress used in our simulation studies. Results indicated that bulging had a significant effect on glottal flow resistance, maximum glottal width and area, and on mean glottal volume velocity. The results are discussed relative to clinical issues of hyperfunction.

Introduction

Glottal adduction is a primary laryngeal posturing gesture. There are three local points of adduction which govern the whole of normal adduction, namely the posterior portion of the arytenoid cartilages, the vocal processes of the arytenoid cartilages, and along the mid-membranous portions of the glottis.

The mechanisms for glottal adduction are not completely clear. It is generally assumed that the interarytenoid (IA) muscles approximate the posterior portions of the arytenoid cartilages due to their direct attachments for such action[1]. Adduction of the vocal processes is more complicated. Nasri et al[2] used an in vivo canine model and found that vocal process adduction forces were greatest via lateral cricoarytenoid (LCA) muscle contraction, followed by thyroarytenoid (TA) muscle contraction and then by IA contraction. These results are consistent with those of Hirano et al[1]. In contrast, Farley[3] used a complex theoretical biomechanical model and found that LCA action contributed relatively little to vocal process adduction whereas IA action played a primary role, with further contributions from TA and cricothyroid (CT) muscles. Contraction of the posterior cricoarytenoid (PCA) muscles abduct and elevate the vocal processes[1], but may also add fine control to the level of vocal process adduction during phonation.[4] The TA muscle appears to have a primary role in creating additional medial adduction of the vocal folds by forming a medial bulge when contracted sufficiently.[1,2,5] This latter topic, adductory effects due to bulging of the vocal folds, a result of increased contraction of the TA muscle, is the subject of this report.

This study is related to two earlier adduction studies,[6,7] in which the effects of arytenoid movement were examined with excised canine larynges and a human subject, respectively. In those studies, the relation between mean subglottal pressure and mean glottal flow for each value of vocal process gap was reasonably represented by a straight line with a slope that increased with greater levels of adduction. The present study explored the same relation, and others, for parametrically increased medial surface bulging. Instead of a tissue model, however, a computer simulation model was used.

Modeling Procedures

The computer simulation model used in this study includes tissue mechanics, glottal aerodynamics, and acoustics of the vocal tract. The tissue mechanics were modeled with a finite element technique.[8] The glottal aerodynamics were modeled with a finite volume solution of the Navier-Stokes equations.[9] The vocal tract acoustics were modeled using a wave reflection and transmission technique.[10,11]

None of these modeling techniques will be discussed in detail here, but a brief synopsis will be given.

Tissue Mechanics

The tissue mechanics included the determination of the pre-phonatory shape of the vocal folds, estimation of the mechanical properties of the vocal fold tissues, and the calculation of vocal fold vibration in the time domain. Vocal fold vibration was modeled with a finite element technique. The three-dimensional geometry of the vocal fold was divided into thin layers along its length (see Figure 1). When the layers are thin enough, the displacement field does not change much across the longitudinal thickness of each layer. Thus, a two-dimensional finite element solution with plane strain (laterally and vertically) was used in each layer. As shown in Figure 2, within each layer, the geometry and the mechanical tissue properties were quantified with triangular elements. The figure shows a typical meshing with muscle (dark gray region), mucosa (gray region), and ligament (white region) as major tissue components.[12] The solution of the tissue mechanics can be outlined as follows:[8]

1. The pre-phonatory shape of the vocal fold surface in the vertical direction was defined by a quadratic expression,[13] with arytenoidal adduction controlled.
2. Viscoelastic properties of each tissue component were read in from an experimentally oriented database.[14,15]
3. The static nodal coordinates in each layer were calculated from an automatic custom-developed mesh generation routine.
4. The mass, stiffness, and damping matrices for each element were calculated and assembled into the corresponding global matrices.
5. Aerodynamic pressures were applied to the vocal fold surfaces.
6. The nodal forces for each element were calculated from the pressure distribution and assembled into a global force vector.
7. The matrix differential equation of motion resulting from the finite element formulation was solved for each time step (the global displacement vector was obtained).

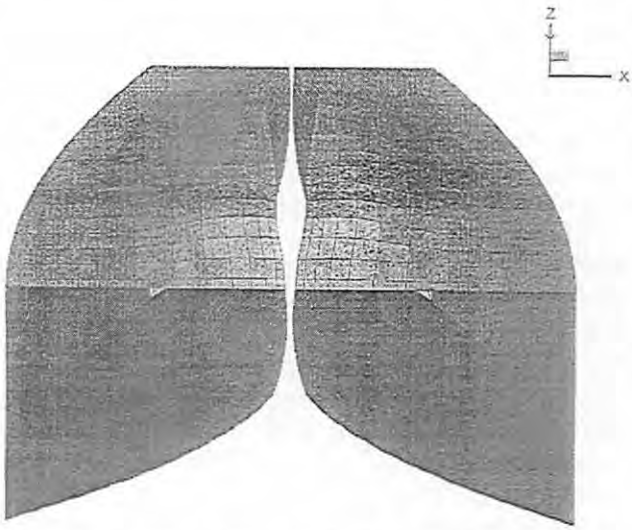


Figure 1. The vocal fold geometry in the biophysical model.

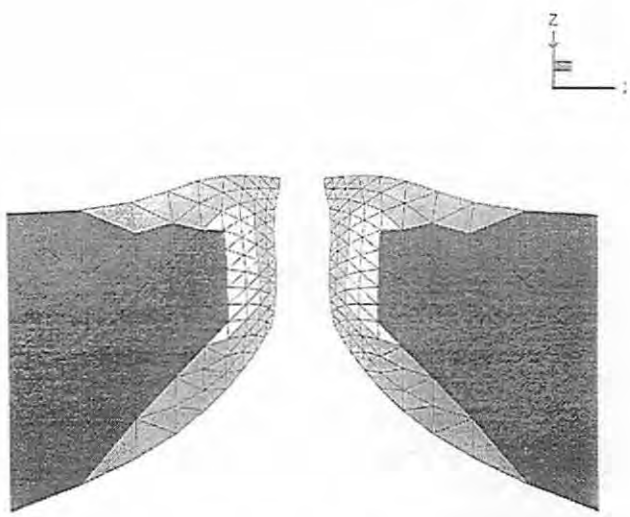


Figure 2. The vocal fold mesh in a medial layer during oscillation. The dark gray represents the vocal fold body, light gray represents the cover, and white region represents ligament.

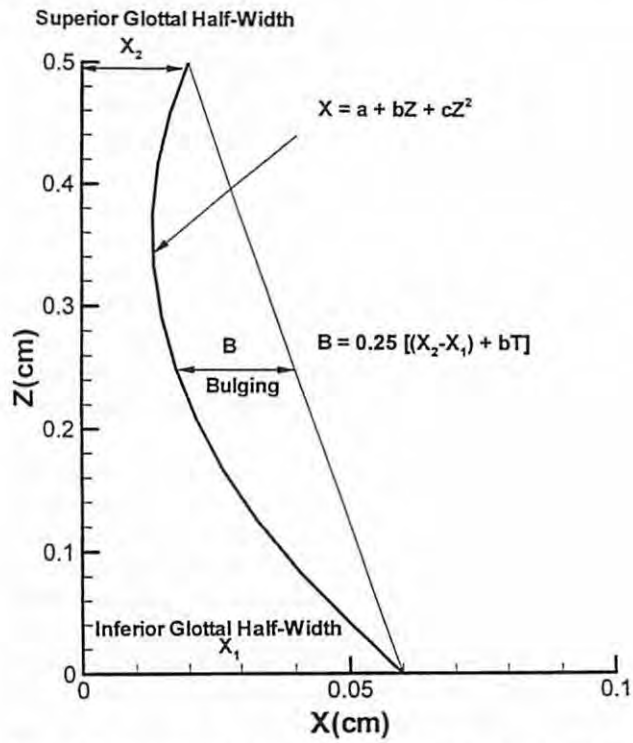


Figure 3. Portion of the glottal wall with parameters that define the bulging and adduction.

8. The nodal coordinates were updated by adding the dynamic displacement vector to the previous equilibrium coordinates, and the boundary conditions were verified.

9. The updated geometry was used in other modules (aerodynamics and acoustics).

The parameter of primary interest in this paper, medial surface bulging, was modeled by a quadratic expression[16,17] between glottal entry and exit, as shown in Figure 3. In the figure, X_1 is the inferior glottal half-width and X_2 is the superior glottal half-width, which were kept constant at values of 0.06 cm and 0.02 cm, respectively, in this study. The curved line is the quadratic fit between the upper and lower glottis. As seen on the figure, the bulging parameter B was defined as a function of the following: the upper and lower glottal diameters, the thickness T of the glottis (0.5 cm), and the value of b in the quadratic function. The bulging parameter was placed midway between the upper and lower edges of the glottis. The value of bulging varied from zero to one.

Bulging was related to the degree of tension in the vocal fold tissue elements corresponding to thyroarytenoid muscle contraction. The bulging value acts as a coefficient to the active muscle contraction stress as determined from physiological data[18] and used in prior simulation studies.[3,19] It is assumed that this approach is sufficiently valid to pursue the phonatory effects of the bulging parameter, but recognize that further refinement is necessary.

Glottal Aerodynamics

The glottal aerodynamics module calculated the pressure and flow in the glottis. There were two procedures for this process. First, the lossless Bernoulli equation was combined with an exit pressure recovery term[20] and acoustic reflection pressure just above the glottis.[21] This provided a first order solution for the airflow and pressure, and was fast and easy. Second, the solution of the Navier-Stokes equations, which are more accurate but require the flow-field discretization and a numerical scheme, was used. Although this second procedure was more time consuming, it provided more information than was possible with the first, specifically the complete pressure and velocity field within the glottis. The flow rate obtained from the first procedure was applied to the Navier-Stokes equations, and a solution was obtained for the entire flow field. The combination of both procedures was highly efficient. The following steps were taken to determine the glottal aerodynamics.

1. Glottal area was calculated from the tissue finite-element solutions.

2. A two-dimensional equivalent of the glottis (with the same transverse cross-sectional areas) was calculated for use with the two-dimensional Navier-Stokes equations.

3. The flow domain was divided into a non-uniform staggered grid (see Figure 4).

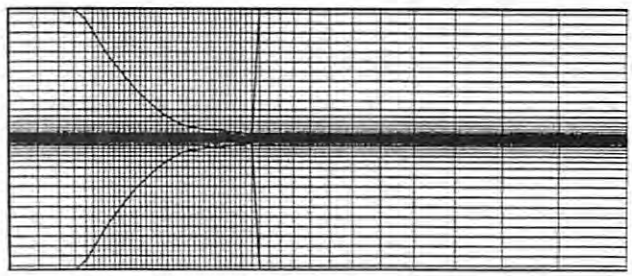


Figure 4. Grids in the flow domain.

4. The glottal flow was obtained from the transglottal pressure and the glottal area using the first procedure (Bernoulli) discussed above.

5. The Reynolds number for the current time step was calculated from the flow and glottal area.

6. The time dependent Navier-Stokes equations were solved for pressure and velocity distributions throughout the entire vocal tract with the finite-volume numerical method.

7. Because the flow field (grid) extended through the vocal folds (see Figure 4), a large source term was used in the discretized Navier-Stokes equations for all cells within the space taken up by the vocal folds. This kept all flow within the airway and set the air velocity essentially to zero on the vocal fold surface.

8. The SIMPLER method[22] for pressure correction was iterated until a converged solution was obtained for that time step.

9. Transglottal pressures and nodal pressure values were updated.

Vocal Tract Acoustic

The acoustics module computed one-dimensional wave propagation in the airways (trachea, pharynx, mouth and nose) on the basis of transmission-line theory[10]. In this model, the vocal tract area function for a particular vowel was obtained from published data.[11,23] The procedure for the solution of the one-dimensional acoustic wave propagation was as follows:

1. The vocal tract was divided into a number of cylindrical sections. For the vowel /a/ the model employed 8 sections in the pharynx, 10 sections in the mouth, 12 sections in the nose (not used in these simulations), and 16 sections in the trachea.

2. The reflection and loss coefficients at the section boundaries were calculated.

3. Using the aerodynamic pressure profile in the glottis and the volume velocity, a vector of the forward pressure downstream of the glottis was calculated.

4. Reflected and output pressures at the mouth and nose were calculated.

5. The aerodynamic pressures in the subglottal and supraglottal sections were updated. The instantaneous

transglottal pressure was strongly influenced by the acoustic properties of the subglottal and supraglottal ducts during phonation[21].

The independent variables of interest in this study were the bulging of the medial vocal folds and the lung pressure. The dependent variables to be reported here (with limited examples) are the mean flow rate, mean subglottal pressure, the mean transglottal pressure, mean glottal flow resistance and impedance, fundamental frequency, maximum glottal area, and the maximum glottal amplitude. The vocal fold length (2.0 cm) and the adduction (0.02 cm) at the vocal processes did not vary.

Results

Figure 5 shows an example of four output variables and two derived variables for the simulation case for a bulging value of 0.4 and a lung pressure of 8.0 cm-H_2O for the /a/ vowel. The fundamental frequency was 113 Hz. The lowest trace is the subglottal pressure (Ps) signal. The variations in the pressure correspond mainly to the first subglottal formant frequency of approximately 500 Hz. The trace just above the subglottal pressure is the glottal volume velocity (Ug). The waveform near the baseline suggests only a brief segment near zero flow. The next trace (Ag) is the projected glottal area, which appears approximately triangular, with again a brief moment of zero area. The fourth trace is the (maximum) glottal width (Gw) calculated between the two vocal folds. It also is somewhat triangular in shape, but with distinct variations. The next trace is the transglottal pressure (Tgp), measured by the simple subtraction of the instantaneous subglottal pressure and the supraglottal pressure values. The double peaks are dependent upon the first two pressure peaks of the subglottal pressure. Notice that the transglottal pressure, which drives flow through the glottis as well as indicates pressure differences between the inferior and superior surfaces of the vocal folds, can be determined whether or not there is flow through the glottis. The top trace is the glottal flow impedance (Imp), obtained by the simple division of the instantaneous transglottal pressure by the instantaneous glottal volume flow. For glottal flow values less than 0.01 ml/s, the impedance value was calculated by interpolating neighboring points. The largest impedance values occur near glottal closure when the transglottal pressures are highest and the flows are lowest.

Instead of presenting each of the signals for the different conditions studied, we will emphasize the glottal volume flow here, since it (or its derivative) is considered to be the source of sound in phonation. Figure 6 shows four glottal volume velocity waveforms corresponding to bulging conditions of 0.4 and 0.8 and lung pressure levels of 8

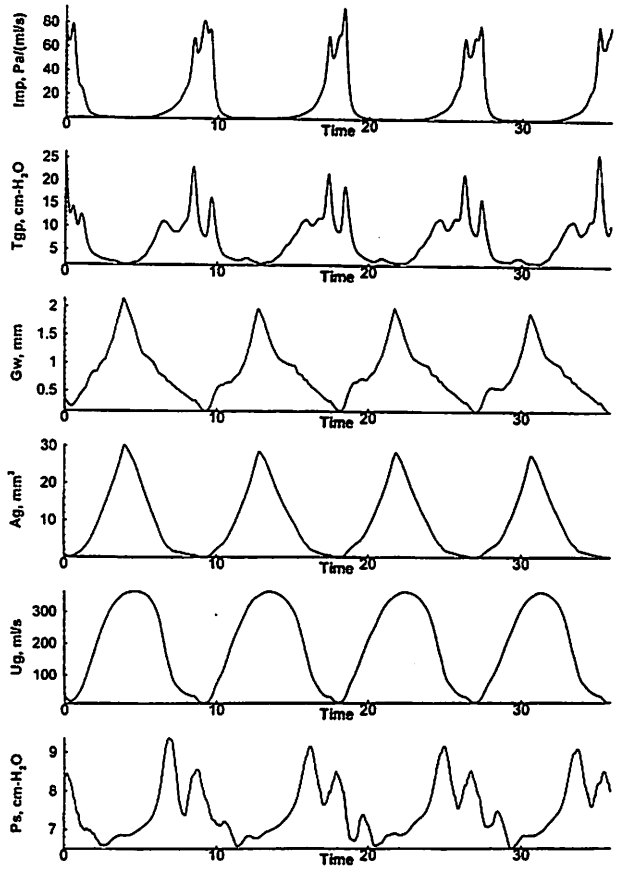


Figure 5. Typical glottal signals in the model.

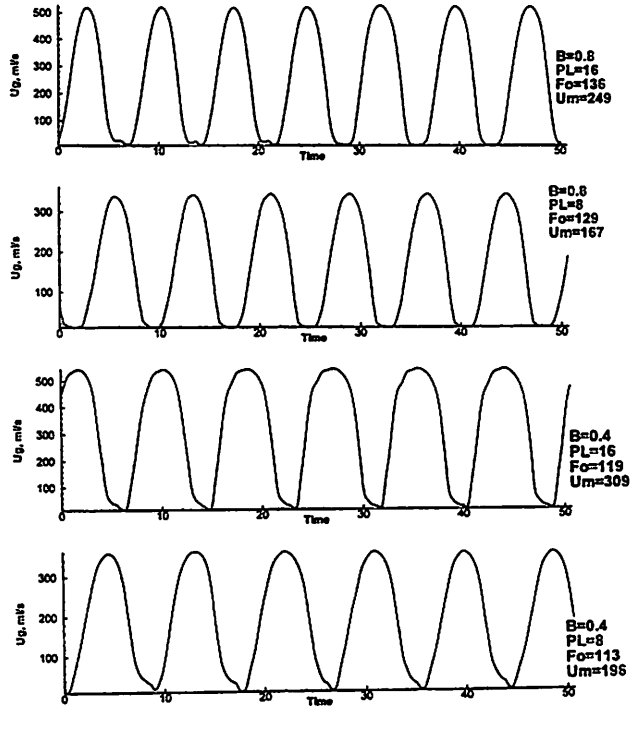


Figure 6. Glottal flow signals for various lung pressures and bulging factors.

and 16 cm-H_2O. When lung pressure was increased from 8 to 16 cm-H_2O for a constant bulging level equal to 0.4, peak flow increased from 364 ml/s to 546 ml/s and the maximum flow derivative increased from approximately -300 l/s^2 to -600 l/s^2. For the same increase of pressure from 8 to 16 cm H_2O for an increased constant bulging level equal to 0.8, peak flow increased from 346 ml/s to 530 ml/s and the maximum flow derivative increased from approximately -300 l/s^2 to -500 l/s^2. These results suggest that increasing the bulging value at constant lung pressure has much less effect on the peak glottal flow and maximum negative flow derivative compared with increasing the lung pressure at a constant bulging value. When the bulging level increased from 0.4 to 0.8 at a constant lung pressure level, there was a longer closed time of the glottis (relatively longer zero flow), suggesting greater adduction and flow resistance, and an increase in fundamental frequency.

Figure 7 shows the results comparing the mean transglottal pressure and the mean glottal flow with the bulging level as the parameter. These results can be compared to similar plots of mean subglottal pressure and mean flow for different levels of arytenoidal adduction for excised canine larynges[6]. In the excised larynx, due to the lack of a vocal tract, the subglottal pressure was the same as the transglottal pressure and the relationship was found to be linear for each level of adduction and essentially non-overlapping between the levels. Here, the relationships are not linear, the slope of the curves tending to vary with higher levels of pressure and flow. The data indicate that greater levels of bulging (closer prephonatory vocal folds) result in less mean glottal flow for the same level of mean transglottal pressure. This implies that greater flow resistance accompanies bulging. This is seen more clearly in Figure 8. The mean glottal flow resistance (defined as the mean subglottal pressure divided by the mean glottal flow) increased as the bulging increased for any level of subglottal pressure. There are a few reversals (which need deeper investigation), but the general trend is strong. The closer vocal folds together with the increased tissue tension account for this expected finding.

Both Figure 7 and Figure 8 appear to suggest that there is a difference in the spread of results at different levels of subglottal pressure. The least spread of flow rate in Figure 7, and the least spread of flow resistance in Figure 8, across bulging levels, is shown at approximately 8 cm-H_2O. The greatest spread of values is seen near 14 cm-H_2O in both figures. This suggests an interesting finding, that bulging may create the least sensitive control of flow resistance at the most normal values of subglottal pressure (near 8 cm-H_2O), whereas the most sensitive control of flow resistance may occur for loud speech (near 14 cm-H_2O). The variation of mean flow resistance at 8 cm-H_2O is approximately 3.6 to 4.6 Pa/(ml/s), a change of 28%, whereas the variation at 14 cm-H_2O is approximately 3.2 to 5.4 Pa/(ml/s), a change of 69%.

Glottal flow impedance, a time dependent quantity defined as the ratio of the transglottal pressure signal and the flow signal, was also obtained in the modeling study (see Figure 5). It was noticed in Figure 8 that glottal

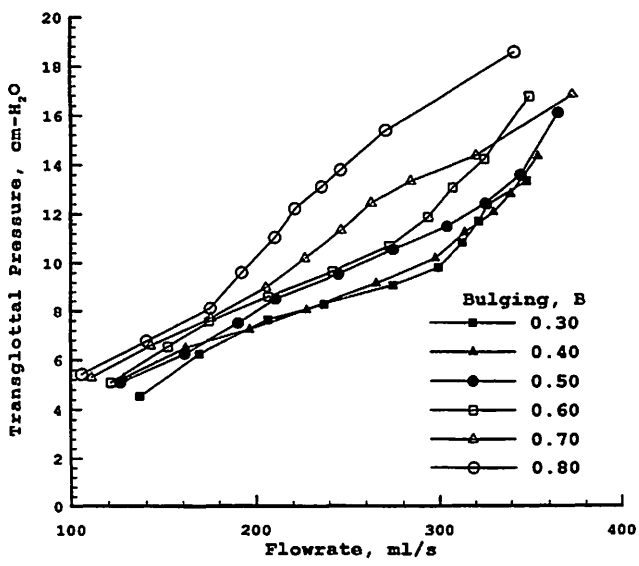


Figure 7. Effects of bulging factor on the pressure-flow relationships.

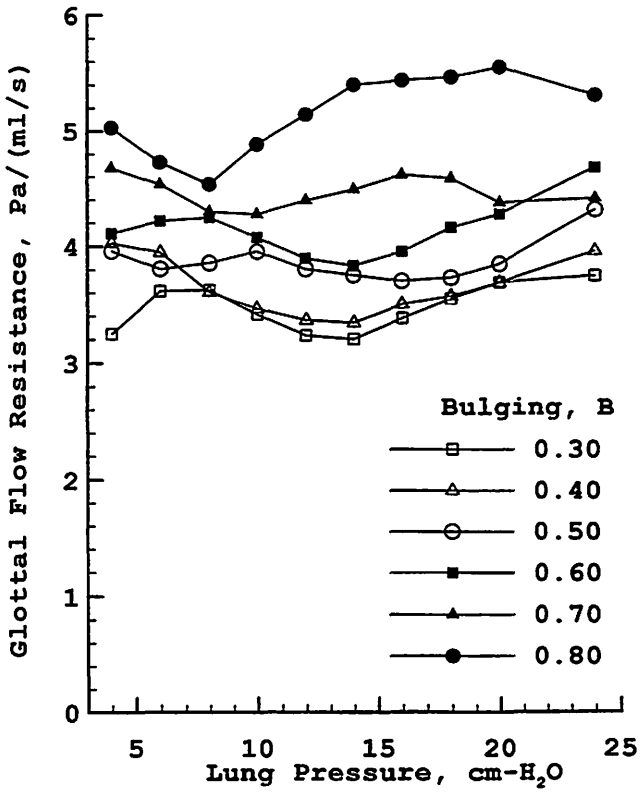


Figure 8. Effects of bulging factor on the mean glottal resistance.

resistance was nearly independent of mean lung pressure for a constant bulging value. Figure 9 indicates that average glottal flow impedance, on the other hand, shows essentially increasing monotonic relationships across lung pressure values. The significance of this and the practical difference between the mean glottal resistance and the mean glottal impedance is yet to be determined.

Increasing lung pressure for specific levels of bulging should increase both the maximum glottal amplitude and area due to the increased external forces pushing upon the vocal fold surfaces. However, for a given value of lung pressure, increasing bulging should reduce the glottal amplitude and area because of the increased adduction and stiffness. These expectations are consistent with the results shown in Figures 10 and 11. Both the glottal area contours and the glottal amplitude contours show increasing values toward the upper left of the figures. For the range of pressures and bulging levels used, both the amplitude and area nearly double. Also, the contours are nearly diagonal, suggesting that there is significant influence on the glottal area and amplitude from both the lung pressure and the vocal fold bulging. For example, doubling the bulging value decreases the maximum glottal amplitude by approximately 20%, and doubling the lung pressure increases the amplitude by about 30% (20% to 39% with greater changes at higher bulging values). These results suggest

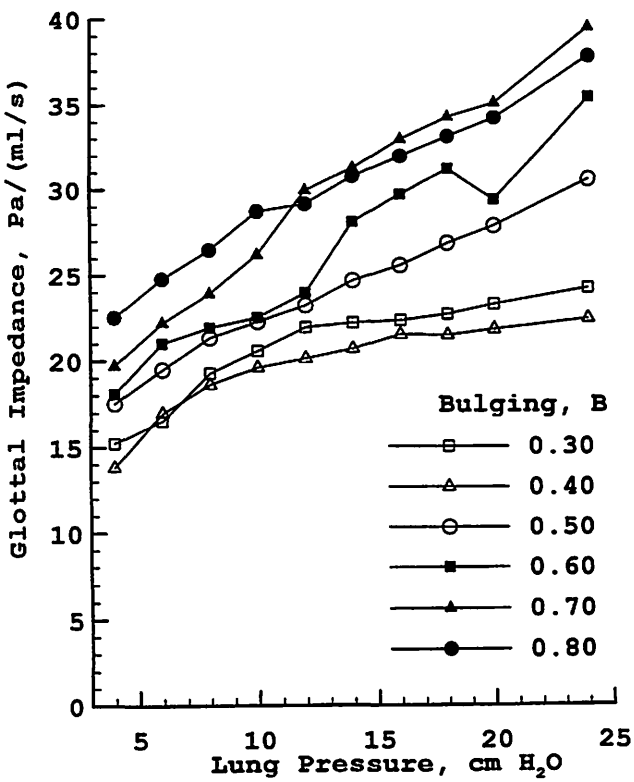


Figure 9. Effects of bulging factor on the mean glottal impedance.

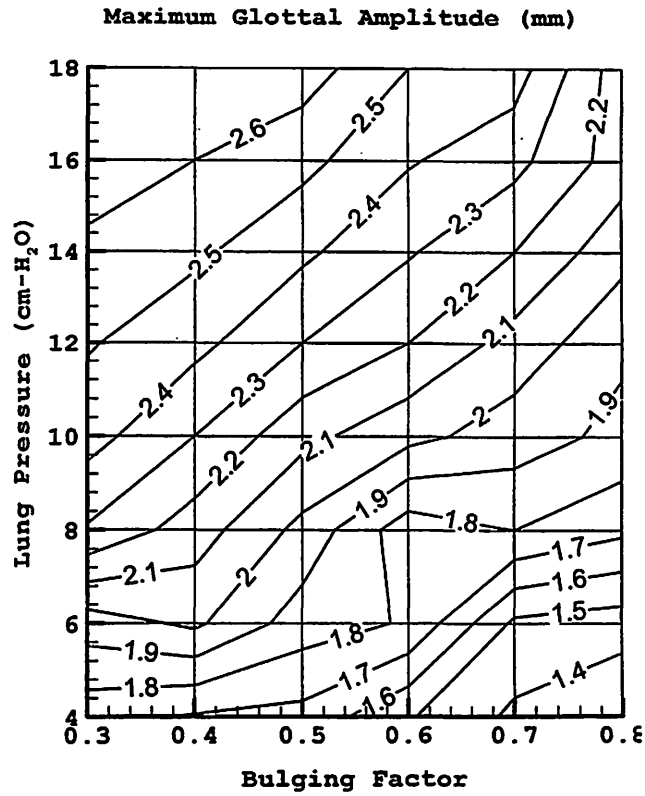


Figure 10. Maximum glottal amplitude as a function of lung pressure and bulging factor.

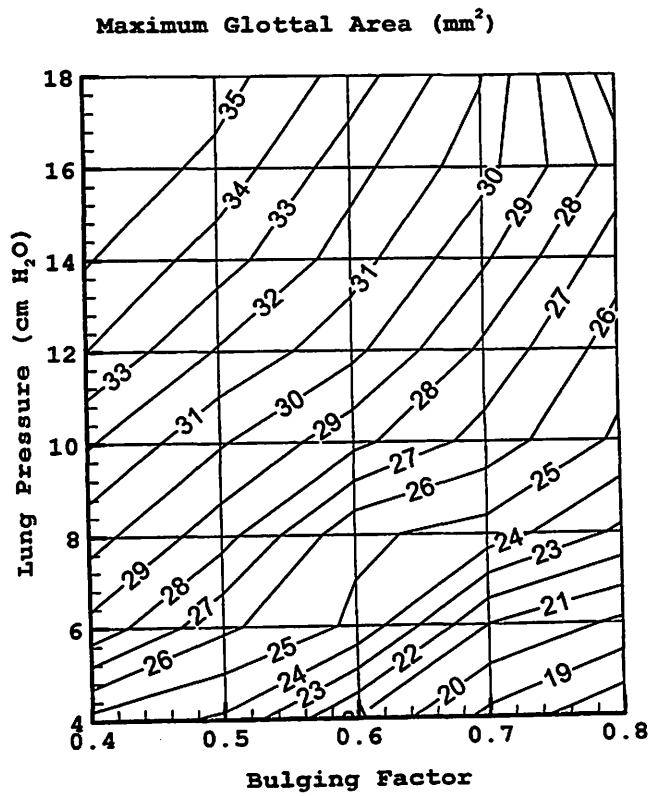


Figure 11. Maximum glottal area as a function of lung pressure and bulging factor.

that bulging may have a strong controlling influence over glottal dynamics.

Figure 12 shows the influence of lung pressure and bulging on the resulting mean transglottal pressure. The figure suggests that bulging has relatively little influence over the transglottal pressure, which is dominated by the lung pressure. This holds more for lower lung pressure levels where the contour is relatively flat. At the higher levels of lung pressure, there is some effect that can be observed from the gradient of these contour lines. For example, at a lung pressure of 16 cm-H_2O, changing bulging from 0.6 to 0.7 increases the mean transglottal pressure for about 0.5 cm-H_2O.

Increasing lung pressure should increase the mean glottal volume velocity because of increased forces acting within the airflow, whereas increasing the bulging levels should decrease the mean volume velocity because of the greater membranous vocal fold adduction levels. This is reflected in Figure 13, showing increasing contours of mean volume velocity toward the upper left of the figure. Doubling the bulging value results in greater reductions in mean volume velocity as the lung pressure increases (about 13% reduction for 6 cm-H_2O to about 25% reduction at 14 cm-H_2O).

Discussion

Glottal adduction, subglottal pressure, and vocal fold length are relatively independent variables of the larynx used in producing and controlling phonation. Because true independent control of one without variation in the others is probably impossible in the human, models are necessary.

This study used a biophysical model to look at one aspect of adduction, the medial positioning or bulging of the vocal folds due to contraction of the thyroarytenoid muscle. The results indicate that a relatively large increase of bulging creates significant increases in glottal flow resistance, reductions in maximum glottal amplitude and area, and decreased values of mean volume velocity.

In this study with constant superior and inferior glottal width, the full range of flow resistance across the bulging levels (0.3 to 0.8) was approximately 3.2 to 5.6 Pa/(ml/s) over the range of subglottal pressure used (5 to 25 cm-H_2O). This range can be compared to that obtained in the earlier adduction study[6] using excised canine larynges, which was approximately 3 to 11 Pa/(ml/s). It is noted that the excised canine tissue did not involve bulging because there was no active TA contraction. Examination of figure 5 of the canine study suggests that, over the range of bulging levels studied here, the range of resistance values is equivalent to changing the vocal process gap by approximately 2 mm, a relatively large value. Thus, the bulging parameter may be a relatively strong determinant of the overall glottal flow resistance. This result is plausible only

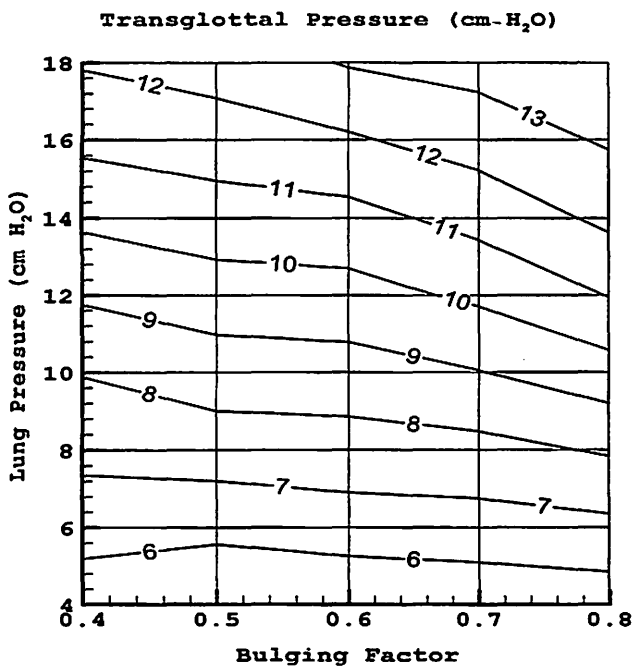


Figure 12. Transglottal pressure as a function of lung pressure and bulging factor.

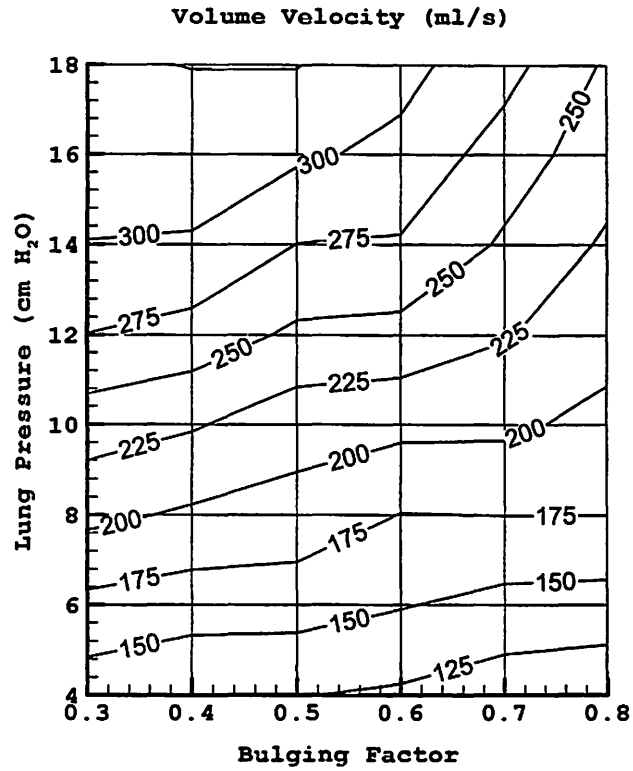


Figure 13. Mean volume velocity as a function of lung pressure and bulging factor.

if the human actually can or does vary bulging over the 0.3 to 0.8 range studied here. The human subject study mentioned earlier[7] appeared to be more consistent with the canine study, wherein the adduction value represented by the vocal process gap was a strong determinant of the relation between a given subglottal pressure and the glottal flow. The scatter in the human data, however, may be due in part to inherent TA contraction and the related bulging. The range of flow resistance in the human subject study, which included a range of adduction and voice quality from breathy to pressed, was approximately 0.5 to 11.5 Pa/(ml/s) over a range of subglottal pressure of 5 to 15 cm-H_2O. The results here suggest that glottal flow resistance is shared between two sources, arytenoidal adduction and TA activation that rounds the vocal fold medial surface and stiffens the vocal fold. Although a significant change in flow resistance is possible with TA activation, it is yet unknown under what circumstances people use it to control flow resistance.

The increased values of glottal flow resistance due to increased bulging appear to be an accessible method to augment or perhaps fine-tune the flow resistance produced by decreasing the vocal process gap. Both types of increases of adduction (vocal process gap, vocal fold bulging) may decrease the open quotient (ref. Figure 6), suggesting both aerodynamic and acoustic control through both methods of adduction.

The clinical problem of tension dysphonia may be directly related to the results of this study. A subset of those with generalized laryngeal tension may combine both types of adduction (complete arytenoidal adduction and vocal fold bulging) to compress the glottis so that only large amounts of subglottal pressure can be used to create vocal fold oscillation and the subsequent low flows (pressed phonation). Those with the combination of breathy-strained voice may use a glottal configuration with an open posterior glottis but compressed vocal folds[24], suggesting the presence of relatively high TA activation (bulging) and the antagonistic action of the LCA-INT-PCA complex (for vocal process adduction) with dominating PCA. Cases in which the vocal processes are touching but with a lack of touching of the middle of the membranous vocal folds suggests that the TA activation may be set at a low level or there is anatomical bowing of the vocal folds. Techniques to counteract vocal fold bowing would include the attempt to create high TA activation as well as increased vocal process adduction, as in the Lee Silverman Voice Treatment approach[25] for louder voicing by Parkinson disease patients.

In addition to these possible clinical cases, too much bulging may create a restricted location on the vocal fold (the most medial portion of the bulge) that may receive the primary impact forces when the vocal folds collide. Thus, too much TA activation (bulging) may promote the creation of vocal fold swelling midmembranously. Contrary to bulging that potentially may contribute to biological inefficiency (tissue damage) is the possibility that some bulging will make oscillation easier by lowering the required phonation threshold pressure. This is consistent with Titze's theoretical suggestion[26] that a closer lower edge of the vocal folds lowers the oscillation threshold pressure. Here, bulging creates closer vocal folds not in the lower glottis but superior to that, effectively accomplishing a similar contour change below the superior portion of the glottis.

Conclusion

In this study a biophysical computer model incorporating both air and vocal fold tissue properties was used to investigate the phonatory effects of medial vocal fold bulging. The bulging was characterized by a quadratic function and was motivated by contour shaping due to contraction of the thyroarytenoid muscles.

The results suggest that vocal fold bulging associated with thyroarytenoid muscle contraction may have significant influences on phonation:

1. Vocal fold bulging increases glottal flow resistance.
2. Flow resistance is more sensitive to bulging for lung pressure values above 10 cm-H_2O (Fig. 8).
3. Lung pressure and bulging both have significant but opposite effects on the maximum glottal amplitude and area (Figs. 10 and 11).
4. Increasing vocal fold bulging decreases the mean glottal volume velocity, with greater reductions at higher lung pressures.

The relative influence of glottal flow resistance due to vocal process adduction and to vocal fold bulging needs further investigation. Results of such studies may be highly relevant to voice therapy and training strategies related to adduction, vocal control in general, and to refined physiological-biomechanical models of phonation.

Acknowledgments

The authors thank Ingo Titze for his editorial advice. This work was supported by NIDCD grant No DC03566-01.

References

1. Hirano M, Kiyokawa K, Kurita S. Laryngeal muscle and glottic shaping. In: Fujimura O, ed. *Vocal Physiology: Voice Production, Mechanisms and Functions*. New York, NY: Raven Press 1988:49-65.

2. Nasri S, Sercarz JA, Azizzadeh B, Kreiman J, Berke GS. Measurement of adductory force of individual laryngeal muscles in an in vivo canine model. *Laryngoscope*. 1994;104:1213-18.

3. Farley GR. A biomechanical laryngeal model of voice F0 and glottal width control. *J Acoust Soc Am.* 1996;100:3794-3812.

4. Choi HS, Berke GS, Ye M, Kreiman J. Function of the thyroarytenoid muscle in a canine laryngeal model. *Ann Otol Rhinol Laryngol.* 1993;102:769-776.

5. Choi HS, Berke GS, Ye M, Kreiman J. Function of the posterior cricoarytenoid muscle in phonation: in vivo laryngeal model. *Otolaryngology - Head & Neck Surgery.* 1993;109:1043-1051.

6. Alipour F, Scherer RC, Finnegan EM. Pressure-flow relationship during phonation as a function of adduction. *Journal of Voice.* 1997;11:187-194.

7. Scherer RC, Alipour F, Guo CG. Glottal flow resistance. A paper presented at the Fourth Biennial Phonosurgery Symposium, 1996, Madison, WI.

8. Alipour-Haghighi F, Titze IR. A finite element simulation of vocal fold vibration. In: LaCourse JR, ed. *Proceedings of the Fourteenth Annual Northeast Bioengineering Conference.* Durham, NH: IEEE;1988:186-189.

9. Alipour F, Fan C, Scherer RC. A numerical simulation of laryngeal flow in a forced-oscillation glottal model. *J. Computer Speech and Language.* 1996;10:75-93.

10. Liljencrants, J. *Speech Synthesis with a Reflection-Type Line Analog.* Doctoral Dissertation, Department of Speech Communications and Musical Acoustics, Royal Institute of Technology, Stockholm, Sweden, 1985.

11. Story BH. *Physiologically-Based Speech Simulation Using an Enhanced wave-Reflection Model of the Vocal Tract.* Ph.D. thesis, The University of Iowa, Iowa City, Iowa 1995.

12. Hirano M, Kakita Y, Ohmaru K, Kurita S. Structure and mechanical properties of the vocal fold. *Speech and Language.* 1982;7:271-297.

13. Titze IR, Talkin DT. A theoretical study of the effects of various laryngeal configurations on the acoustics of phonation. *J Acoust Soc Am.* 1979;66:60-74.

14. Alipour-Haghighi F, Titze IR. Stress relaxation in vocal fold tissues. In: Goldstein SA, ed. *Advances in Bioengineering,* ASME 1990;BED#17:21-24.

15. Alipour-Haghighi F, Titze IR. Elastic models of vocal fold tissues. *J Acoust Soc Am.* 1991;90:1326-1331.

16. Titze IR. A four-parameter model of the glottis and vocal fold contact area. *Speech Communication.* 1989;8:191-201.

17. Titze IR, Alipour F. *The Myoelastic-Aerodynamic Theory of Phonation,* in preparation.

18. Alipour-Haghighi F, Titze IR, Perlman AL. Tetanic contraction in vocal fold muscle. *J Speech Hear Res.* 1989;32:226-231.

19. Cooper DS, Partridge L, Alipour F. Biomechanics, muscle energetics, vocal efficiency. In: Titze IR, ed. *Vocal Fold Physiology.* San Diego, CA, Singular Publishing, 1993:37-92.

20. Ishizaka K, Matsudaira M. Fluid mechanical considerations of vocal cord vibration. *SCRL-Monograph.* Speech Communication Research laboratory, Santa Barbara, 1972;8.

21. Titze IR. Parameterization of the glottal area, glottal flow, and vocal fold contact area. *J Acoust Soc Am.* 1984;75:570-580.

22. Patankar SV. *Numerical Heat Transfer and Fluid Flow,* Hemisphere Publishing Corp., McGraw-Hill, New York, 1980.

23. Fant, G. *Acoustic Theory of Speech Production.* Mouton, The Hague 1960.

24. Brown WS, Vinson BP, Crary MA. *Oraganic Voice Disorders, Assessment and Treatment.* Singular Publishing, San Diego, 1996.

25. Ramig LO, Bonitati CM, Lemke JH, Horii Y. Voice treatment for patients with Parkinson disease: development of an approach and preliminary efficacy data. *Journal of Medical Speech Language Pathology,* 1994;2:191-209.

26. Titze IR. The physics of small-amplitude oscillation of the vocal folds. *J Acoust Soc Am.* 1988;83:1536-1552.

NCVS Status and Progress Report - 13
April1999,115-118

Intraoperative Decision-Making in Medialization Laryngoplasty

Charles N. Ford, M.D.
Division of Otolaryngology-Head and Neck Surgery, Department of Surgery, The University of Wisconsin

One of the major advantages of laryngeal framework surgery is that it affords the surgeon an opportunity to obtain feedback during the procedure so that appropriate modifications can be made intraoperatively to achieve an optimal result. In order to maximize the opportunity for intraoperative adjustments, it is necessary to perform these procedures under local anesthesia. This way, the patient can phonate while the surgeon monitors the response perceptually and observes the alterations in surface anatomical configuration by flexible fiberoptic endoscopy. Medialization thyroplasty is the most predictable of the framework procedures and allows the surgeon to create measured contour alterations of the vocal fold. Routine decisions must be made regarding exposure, placement of incisions, and shim designs. The decision to employ adjunctive procedures or alternative materials is often made during surgery. Above all, the surgeon must be prepared to alter the planned procedure when faced with operative complications and inadvertent tissue injuries that threaten the patient's well being and the surgical result.

Exposure Decisions

Adequate exposure always can be attained through a horizontal cervical incision eccentrically placed so that it is centered on the operative side but crosses the midline sufficiently to allow for flap elevation. Flaps are undermined in the subplatysmal plane and the strap muscles initially preserved. The outer perichondrium of the thyroid cartilage is incised in the midline and elevated to allow visualization of the thyroid lamina from superior to inferior. If there is inadequate posterior exposure, the medial aspect of the sternohyoid and thyrohyoid muscles can be sectioned just inferior to the hyoid. For thyroplasty, it is necessary to identify the inferior border of the thyroid cartilage; a musculoperichondrial flap can be elevated and splayed out, thereby preserving the insertions of the cricothyroid muscle and avoiding injury to this richly vascular muscle (Figure 1). If it is decided to perform adjunctive arytenoid adduction, it will be necessary to cut the inferior perichondrial attachments and remove the insertions of the thyrohyoid and constrictor muscles to allow adequate rotation of the larynx.

Although some surgeons advocate routinely cutting the inner perichondrium, it is an unnecessary step that introduces additional risk to the procedure. In addition to a greater risk of bleeding, there is a slight risk of penetration of the ventricular mucosa. This risk is greatest anteriorly where the ventricle closely approximates the inner surface of the thyroid cartilage. When drilling circumferentially around the window or attempting to separate the cartilage from inner perichondrium, it is best to work first on the inferior, posterior, and superior margins. This preserves the anterior perichondrium so that an anteriorly based inner perichondrial flap can be created, allowing for additional displacement if necessary while protecting the anterior soft tissues (Figure 1).

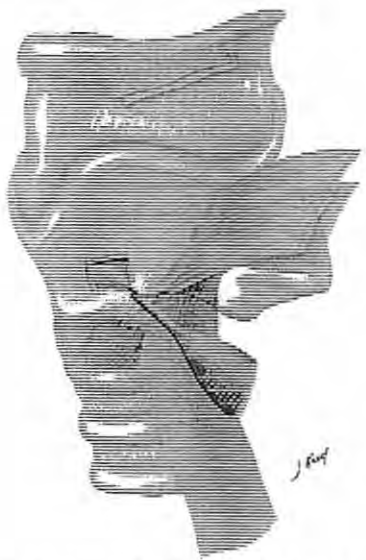


Figure 1. Adequate exposure for thyroplasty is shown with a musculoperichondrial flap elevated preserving the cricothyroid muscle insertions in the thyroid cartilage. A thyroid cartilage window is shown with incisions around three sides of the inner perichondrium preserving the anterior margin as a protective barrier when the entire inner perichondrium cannot be preserved.

Deciding on Proper Placement of Cartilaginous Incisions

There are several formulae and rules that have been used for placement of the cartilaginous window in the thyroid cartilage, but size and placement decisions must be guided by intraoperative observations. Precise placement and size of the window depends on the overall size of the larynx, the size and shape of the planned shim implant, and careful attention to the internal laryngeal anatomy to avoid inappropriate displacement and injury. The surface anatomy must be adequately displayed to determine the location of the vocal fold. If the inferior margin is not identified, there is a tendency towards superior placement of the window, resulting in ineffectual displacement of the ventricle and false vocal fold. Common mistakes are: (1) adhering to a formula based on the use of one type of implant when a shim of different design is employed; (2) not adjusting the size of the window to the overall size of the larynx; and (3) placing the window too high and too far anterior.

The exact placement of the thyroid cartilage window incisions should be dictated by the anticipated shape of the displacement shim (1). Once a decision is made to employ a shim of a particular shape, the surgeon must decide how that implant will displace the vocal fold and adjacent soft tissues. If the shim overlaps the thyroid cartilage superiorly, the superior level of the window must be placed inferior to the midpoint of the thyroid lamina to avoid inappropriate medialization of the supraglottis (Figure 2). Ignoring this detail will result in increased supraglottic resistance and decreased vocal efficiency (2). If the shim is designed to overlap anteriorly, then the window needs to be placed more posteriorly. Recognizing that the window is inappropriately sized or placed requires intraoperative sculpting of the implant to adjust to the situation. When in doubt about the placement of the window, it is generally best to err on the side of a more inferior and posterior placement. Leaving a 3 mm inferior strut of thyroid cartilage is sufficient to support most implants.

Placing the window too far anteriorly can have a detrimental effect on the voice and can promote extrusion. Shims extending to less than 5 mm from the anterior face of the thyroid cartilage will distort the anterior third of the vocal fold such that there is excessive homolateral convex anterior vocal fold displacement and stiffness. This can aggravate more posterior glottic insufficiency and restrict effective oscillation of the contralateral vocal fold anteriorly due to excessive contact or impaction by the displaced fold. A more troublesome consequence of anterior placement is the potential to penetrate the ventricular mucosa predisposing to granuloma and extrusion. The extreme proximity of the ventricular mucosa to the inner perichondrium is apparent by studying the axial sectional anatomy (Figure 3). It is safer to displace the posterior aspect of the vocal fold as there is more intervening soft tissue between the implant and laryngeal mucosa.

Deciding on Shim Design

There are several types of preformed displacement shims, made of different materials and in various sizes and shapes. Displacement shims can also be individually constructed from silastic, Gore-Tex, or autogenous cartilage. Regardless of the material selected, it is important to use a shim that interlocks with the thyroid cartilage to prevent motion and deter extrusion. One approach to an interlocking design is to carve notches in the superior and inferior surface of a silastic block such that the shim snaps into place (tongue-in-groove) and is not easily dislodged (Figure 2). The basic design is then modified by trial displacements of the paraglottic soft tissues while the patient is asked to phonate and the endolarynx observed by flexible endoscopy. The implant is sculpted to achieve optimal medial displacement of the vocal fold, on plane with the contralateral fold, avoiding supraglottic displacement and assuring

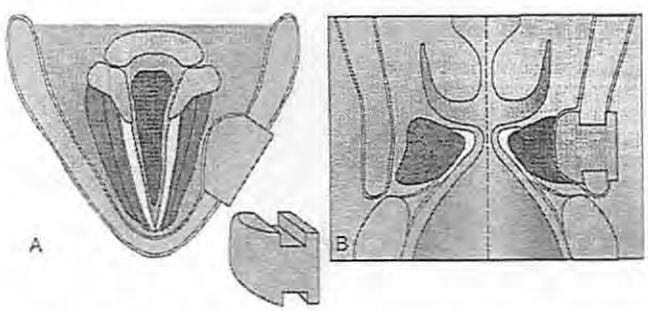


Figure 2. This shows a useful design for a silastic shim that interlocks in a tongue-in-groove fashion with the thyroid cartilage. Since it extends above the cartilaginous window internally, it is desirable to place the superior margin of the window inferior to the estimated plane of the vocal fold.

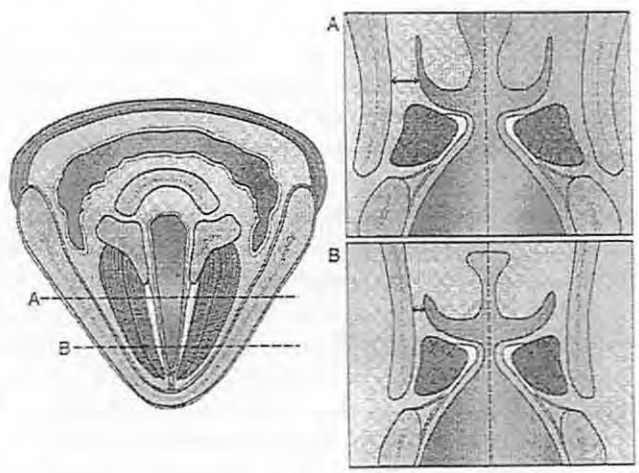


Figure 3. Axial and coronal views demonstrate the relative proximity of ventricular mucosa to the thyroid cartilage anteriorly (B) compared to more posteriorly (A).

glottic and intrafold fullness while avoiding subglottic distortion. When optimal placement cannot be achieved, the surgeon can decide to perform an adjunctive arytenoid adduction. The shim is removed and the larynx rotated after separation of the attached strap and constrictor muscles. This exposure is facilitated by rongeur excision of a posterior rectangular strip of cartilage. Rotation should be gentle to avoid fracture of the cartilage. Passage of the adduction suture is facilitated by direct observation through the thyroplasty window. Satisfactory adduction should be secured before reintroduction of the displacement shim. Further refinement of the shim will often be necessary after the adduction suture is secured.

DecisionsAddressing Intraoperative Complications

It is important to operate with extreme care and meticulous hemostasis to avoid complications and tissue edema that can prevent adequate intraoperative voice assessment. The major cause of operative complications is inadvertent damage to vocal fold structures. Medialization thyroplasty as described by Isshiki (3) preserves the cartilage window and attached inner perichondrium. Once the cartilaginous window is removed, as in the technique described here, there is a risk of tearing inner perichondrium, damaging the thyroarytenoid muscle, and inducing bleeding and edema. Preserving the inner perichondrium reduces these risks, but if damaged, it is helpful to preserve the anterior margin as a protective barrier that still allows for generous displacement but with decreased risk of muscle damage or ventricular mucosal penetration. In the event of inadvertent penetration of mucosa into the laryngeal lumen, it is not advisable to proceed with shim placement since the exposure and contamination will likely result in granuloma and extrusion. A decision to abort the procedure or change to an arytenoid adduction approach is in the patient's best interest.

Figure 4. A flap comprised of an inferiorly based medial slip of the sternohyoid muscle is shown routed on the outer surface of the cricothyroid membrane deep to the thyroid cartilage to fill the paraglottic space, maintain medialization of the vocal fold, and seal any defects present in the ventricular mucosa.

The use of a muscle flap is an alternative that we have used successfully to correct mucosal tear, substitute for a displacement shim, and replace foreign body granulomas associated with extruding prostheses. This approach originally described by Netterville (4) to manage Teflon granulomas entails the use of a medial slip of the sterno-hyoid muscle to fill the space normally occupied by the shim. The inner perichondrium is widely separated from the inner surface of the thyroid cartilage and the elevator is used to create a communication from the window to the external surface of the cricothyroid membrane. A medial slip of sternohyoid muscle is mobilized with a Metzenbaum scissors and passed through the window to create medialization of the vocal fold (Figure 4). While not as precise as a solid shim, it provides sufficient bulk while functioning as a physiologic sealer.

Another intraoperative complication that warrants immediate attention is accidental laryngeal fracture. While this is rarely a problem with thyroplasty, the additional torque often needed to rotate the larynx during arytenoid adduction can lead to unintentional fracture of the thyroid cartilage. When arytenoid adduction is performed in conjunction with thyroplasty, several factors favor fracture: (1) the presence of the thyroplasty window; (2) the possible additional excavation of cartilage from the posterior edge of the thyroid cartilage; (3) extensive cartilage calcification; and (4) a noncompliant larynx requiring extensive rotational force for posterior exposure. In most instances it is possible to continue on with the arytenoid adduction but the thyroplasty will occasionally need to be modified. The interlocking shim described here is often sufficient to stabilize the cartilaginous skeleton. If not, then either sutures or microplates can be used. Gore-Tex can be used alternatively as an implant since it does not require rigid stabilization to be effective.

Summary

In summary, the greatest advantage of medialization thyroplasty is the opportunity afforded the surgeon for intraoperative adjustments to optimize the voice result. To take advantage of this opportunity, it is essential to perform the surgery under local anesthesia so that various soft tissue displacements can be tried, anatomical changes assessed, and intraoperative voice results appreciated. The surgeon should have several alternative approaches for dealing with undesirable results and events. It is often useful to add

arytenoid adduction to achieve optimal vocal fold placement, airway protection, and voice. Intraoperative decision making is critical in both achieving the best result and dealing with operative complications. Bleeding, soft tissue damage, mucosal penetration, and framework fracture are some of the problems that must be addressed during surgery to avoid poor outcomes and assure the very best results.

Acknowledgments

This work was supported in part by the National Center for Voice and Speech through Grant P60 00976 from the National Institute on Deafness and Other Communication Disorders.

References

1. Ford CN, Unger JM, Zundel RS, et al: Magnetic resonance imaging (MRI) assessment of vocal fold medialization surgery. Laryngoscope 105:498-504, 1995.

2. Titze IR: Principles of Voice Production. Englewood Cliffs, NJ, Prentice Hall, Inc., 1994.

3. Isshiki N, Morita H, Okamura H, et al: Thyroplasty as a new phonosurgical technique. Acta Otolaryngol (Stockh) 78:451-453, 1974.

4. Netterville JL, Coleman JR, Ossoff RH, et al: Lateral Laryngotomy for the removal of teflon granuloma. Ann Otol Rhinol Laryngol (in press).

NCVS Status and Progress Report - 13
April 1999, 119-120

Voice Restoration by Tracheo-Tracheolaryngeal Shunt After Laryngotracheal Diversion for Chronic Aspiration

Charles N. Ford, M.D.
Robin Samlan, M.S.
Division of Otolaryngology-Head and Neck Surgery, Department of Surgery, The University of Wisconsin
Joanne Robbins, Ph.D.
Department of Medicine, Section of Geriatrics and Gastroenterology, The University of Wisconsin
Division of Otolaryngology-Head and Neck Surgery, Department of Surgery, The University of Wisconsin
VA GRECC, Madison, Wisconsin

Removal of the larynx is the most definitive procedure for achieving airway protection in patients with life-threatening aspiration. However, it is difficult for surgeons to sacrifice this structure and subject the patient to lifelong alaryngeal communication, especially if there is any chance of neuromuscular recovery. Because of this, a number of techniques that preserve the larynx have been attempted, including adjacent tissue flaps, laryngeal closure, obturators to occlude the laryngeal lumen, and variations on laryngotracheal separation (1). In our experience, the most gratifying results have come from performing tracheoesophageal diversion as described by Lindeman (2). In this approach, a low tracheostomy is performed and the superior segment is anastomosed to the adjacent esophagus such that all oral contents are routed into the esophagus, either directly or via the laryngotracheal segment. Since the innervation is not disturbed, this is a potentially reversible procedure.

In patients who have shown signs of neurological recovery after diversion surgery, we perform a swallow study to determine if there is sufficient laryngeal control to warrant attempted takedown of the diversion (Figure 1). Over the past 4 years, patients who have failed to show sufficient laryngeal functional recovery were offered an alternative method of phonation in which a shunt is created from the trachea at the posterior wall of the tracheal stoma to the superior laryngotracheal segment. This tracheo-tracheolaryngeal shunt (TTLS) affords the patient the opportunity to generate a near-normal voice using a technique similar to the routine tracheoesophageal puncture, but with the added benefit of using a structurally intact larynx. One of the 4 patients who underwent this procedure required a subsequent takedown, and the lessons learned in this case have helped us to modify the technique to enhance success.

Operative Technique

The basic technique of the tracheoesophageal diversion as described by Lindeman (1) is performed to resolve the aspiration problem. Key points that are particu-

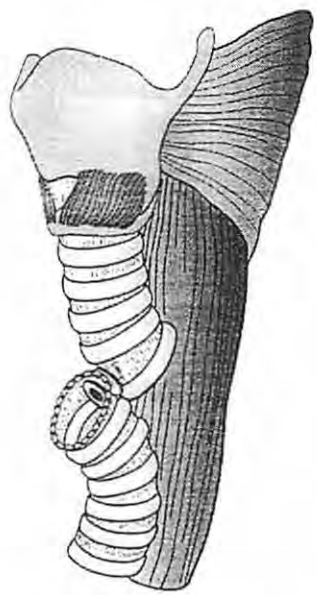


Figure 1. The incision is placed through the posterior wall of the tracheal stoma parallel to the plane of the tracheal rings to avoid cutting across cartilage. Note the gentle curvature of the superior tracheal segment allows the incision to be favorably placed.

Figure 2. This preoperative magnetic resonance imaging study in the sagittal plane depicts the relationship of the air-containing contiguous tracheal segments and the laryngotrachealesophageal passageway is clearly depicted (clear white arrow).

larly relevant to the subsequent performance of the TTLS are: (1) avoiding a high tracheotomy—preservation of 5 rings is desirable; (2) placing the esophageal incision sufficiently low in the neck to avoid substantial kinking of the upper tracheal segment and assuring that the outflow point is most dependent; and (3) making a sufficiently large esophageal opening through a vertical incision to facilitate easy egress of food into the esophagus. Performance of the TTLS should be deferred 6 to 12 months to determine if the patient shows sufficient signs of recovery from their neurological deficit to warrant takedown of the original laryngotracheal shunt and re-establishment of normal laryngeal function.

The TTLS procedure can be performed under local or general anesthesia. An incision is placed in the posterior wall of the inferior segment of the trachea, parallel to the axis of the tracheal rings. It is important to avoid cutting across the rings to facilitate adequate collapse of the soft tissues around the obturator and subsequent sealing around the valved prosthesis. The optimal level for placement of the incision can be judged by assessing a preoperative magnetic resonance imaging (MRI) study. The MRI shows the relationship of the tracheal segments to the esophagus and helps to determine where an incision can be placed in the most horizontal direction to avoid pooling at the site of placement (Figure 2). The incision can be stretched in the horizontal plane with a hemostat and a foley catheter inserted for 5 to 7 days before replacing it with a flanged Blom-Singer prosthesis (In Health Technologies, Carpinteria, CA) (Figure 3).

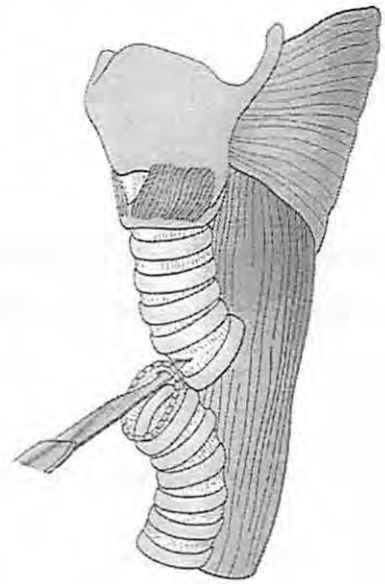


Figure 3. This shows the tracheo-tracheolaryngeal shunt maintained with the flanged valved prosthesis.

Summary and Keys to Avoid Complications

*Place tracheostomy at approximately the interspace between the 5th and 6th tracheal ring.

*Make a generous vertical incision in the esophagus to facilitate passive egress of contents into the esophagus.

*Avoid kinking or dependent sink in laryngotracheal segment of trachea.

*Place incision as near to horizontal plane as possible and avoid cutting across tracheal cartilage.

Acknowledgments

This work was supported in part by the National Center for Voice and Speech through Grant P60 00976 from the National Institute on Deafness and Other Communication Disorders.

References

1. Eisele DW, Yarington CT, Lindeman RC, et al. The tracheoesophageal diversion and laryngotracheal separation procedures for treatment of intractable aspiration. Am J Surgery 157:230-236, 1989.

2. Lindeman RC. Diverting the paralyzed larynx: a reversible procedure for intractable aspiration. Laryngoscope 85:157-180, 1975.

NCVS Status and Progress Report - 13
April 1999, 121-130

Comparison Between Electroglottography and Electromagnetic Glottography

Ingo R. Titze, Ph.D.
Brad H Story, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa
The Denver Center for the Performing Arts
Gregory C. Burnett, Ph.D.
Department of Applied Science, The University of California at Davis
Lawrence Livermore National Laboratory
John F. Holzrichter, Ph.D.
Lawrence C. Ng, Ph.D.
Lawrence Livermore National Laboratory
Wayne A. Lea, Ph.D.
Speech Sciences Institute, Apple Valley, MN

Abstract

Newly developed glottographic sensors, utilizing high frequency propagating electromagnetic waves, were compared to a well-established electroglottographic device. The comparison was made on four male subjects under different phonation conditions, including three levels of vocal fold adduction (normal, breathy, and pressed), three different registers (falsetto, chest, and fry), and two different pitches. Agreement was always found for the movement of glottal closure, but for the general waveshape the agreement was better for falsetto and breathy voice than for pressed voice and vocal fry. Possible differences in the field patterns and scattering mode were addressed. The hypothesis that the electromagnetic sensors could operate in either the forward scattering (diffraction) mode or in the backward scattering (reflection) mode was tested. Results favor the diffraction mode, but reflections may contribute to the signal. Several observations are made on the uses of the electromagnetic sensors for operation without skin contact and possibly in an array configuration for improved spatial resolution within the glottis.

Introduction

Glottography is a measurement of the time variation of the glottis during phonation, where the glottis is defined as the airspace between the vocal folds. Glottography usually involves the transmission of a probe signal from one side of the larynx to the other, with the time variation of the glottis modulating the probe's properties. The modulation is then detected and interpreted in terms of the expected geometry of the glottis, which is formed by laryngeal tissues that are in partial stages of contact during the phonation cycle. The probes used in the past have been electrical current flow, ultrasonic waves, light transmission, and airflow as generated by the speaker. Recently, propagating high frequency electromagnetic waves (EM waves) have been used to measure tissue motions in concert with glottal time variation (Holzrichter et al. 1998). In this paper, the properties of such an EM wave device are compared to those of an electric current flow device.

As in many scanning and imaging techniques, there are two traditional modes of probing an air-tissue interface: the transmission mode and the reflection mode. In the transmission mode the received signal passes through the interface (or a series of interfaces), whereas in the reflection mode, the received signal is scattered (i.e., reflected in the sense of ray optics) off the interface. As discussed below, when the wavelength of the probe becomes comparable or larger than the size of the object being measured, the concepts of transmission and reflection are better described by wave scattering concepts.

Examples are wave diffraction (usually defined as forward direction EM wave scattering), and backward EM wave scattering (due to a reflection from an interface). Generally speaking, the motion and geometry of a tissue interface can be detected in both forward and backward scattering modes. This is true provided there is adequate spatial and temporal resolution of the sendingreceiving probe, that there is sufficient signal relative to the background noise, and that there is a clarity of signal interpretation with respect to the physical phenomena causing the tissue interface modulation.

Both transmission and reflection modes have been attempted in glottography. In almost all attempts in the past, the interface of interest has been the medial surface of the vocal folds. This interface (or pair of interfaces for two vocal folds) is on the order of 1 cm horizontally (front to back) and 0.5 cm vertically (top to bottom) and has significant curvature that varies over the glottal cycle. Thus, any signal reflecting from this interface involves rapid changes in direction. For transmission, there are always at least two interfaces, first a tissue-to-air interface on one vocal fold and then an air-to-tissue interface on the other. When contact between the vocal folds occurs (over all or part of the medial surfaces), scattering from either of the interfaces begins to vanish and transmission approaches unity. The opening and closing of the glottis happens cyclically at frequencies of 100-1000 Hz; thus, instrumentation resolution times are from several hundreds of microseconds to several milliseconds.

Under the highly variable reflection and transmission conditions of the vocal fold structure, the choice of a probe signal is important. As stated above, the classes of signals that have been tried have been electric, ultrasonic, and electromagnetic (in the form of visible light). In addition, flow glottography has been used, which uses the airflow signal that the larynx itself creates in the process of phonation.

With the use of electric current signals (electroglottography, or EGG), there is no scattering from the interfaces because the electric probe current does not propagate as a wave in the tissue; rather, electrons or ions are moving over very short mean free paths. The transmission mode reduces to a conduction mode. Wherever the tissues are in contact, there is conduction across the glottis; wherever there is an air gap, there is no direct conduction. Because the carrier signal is usually chosen to be an alternating current at a high frequency (in the MHz range), the capacitance of the vocal fold air gap does allow some displacement current to be transmitted across the gap. However, unless the gap is very small, the impedance of the gap is much larger than the impedance of the tissue, and the signal is determined by ionic conduction through the contacting tissues that surround the air gap (Titze, 1990).

With propagating electromagnetic and ultrasonic waves, there can be both transmission (e.g., forward scattering or diffraction) and reflection (e.g., backward scattering) from the larynx tissues. Ultrasound waves used for tissue probing usually have wavelengths of 0.1 to 1 millimeters, and follow ray optic trajectories when used in probing glottal structures. Ideally, when used in the transmission probe, they can quantify the contact area; when used in the reflection mode, they can quantify the medial surface motion.Hamlet and Reid (1972) used ultrasonoglottography (UGG) in the transmission mode. The receiving transducer was placed on the other side of the neck. Hamlet had better success with this arrangement than with reflection, but in general the contact area signal was not very stable. The amplitude was very sensitive to head and neck movement, especially vertical movement of the larynx within the neck. In the ultrasound reflection mode, Zagzebski et al (1983) placed a signal transducer on one side of the neck, acting both as a sender and receiver. Owing to the non-parallel orientation of the transducer relative to the tissue-air interface in the glottis, the signal transducer received little of the reflected signal. Most of the wave was reflected away from the incident direction, making the detected interface signal very noisy.

Incoherent, visible light has been used as an electromagnetic probe, and is known as photoglottography (PGG). The measurement has been used for a number of decades (e.g. Lofqvist and Yoshioka, 1980). Although coherent light sources (i.e., lasers) can be used, light scattering in tissues is sufficiently intense that the measurement mode becomes essentially incoherent in nature. Because visible light is highly attenuated by tissue, a signal path has to be chosen that minimizes the loss. Usually, one transducer is placed above the glottis inside the throat (using a fiber-optic cable) while the other is placed outside the throat, usually below the thyroid prominence and, thereby, below the glottis. The signal propagation mode is transmission though the horizontal plane of the glottis and the surrounding vocal fold tissues, from top to bottom of the glottis. This direction of the probe signal is perpendicular to that used by UGG and EGG, which are directed along the horizontal plane of the glottis and perpendicular to the midsagittal plane. Thus, PGG provides information that differs from UGG and EGG, a minimum projected *glottal area* rather than the *contact area* measured by EGG and UGG.

Finally, in flow glottography (FGG) the signal is associated with the stream of air forced through the glottis, from the lungs to the lips. As in all other forms of glottography, the movement of the vocal folds modulates the signal. What is profoundly different in FGG is that the probe is self-generated by the speaker. The signal is a natural part of the phonatory process. The FGG receiver is

a flow mask placed over the mouth and nose, which measures the oral flow, as produced by the glottis and modulated by the vocal tract. From the measurement of oral airflow out of the nose and lips, the actual glottal flow versus time is estimated by inverse filtering, a process that attempts to remove the flow modulations imposed by the vocal tract. The estimated glottal airflow is then further inverse-filtered to obtain the time-varying glottal area by using a nonlinear airflow model that has glottal area as an input (Rothenberg, 1973).

Electromagnetic Glottography

Electromagnetic glottography (EMGG) is a new technique, whereby an electromagnetic wave in the GHz range is propagated and then detected to obtain information on the condition of the larynx tissue interfaces (Holzrichter et al, 1998). A transducer arrangement for combined EMGG and EGG measurements described herein is shown in Figure 1. The EM antennae are placed on the front of the neck near the thyroid prominence (Adam's Apple) and the EGG electrodes are placed on both sides of the neck. The transmitted EM wave is a pulse containing about 10 wave cycles at 2.3 GHz, with a wavelength of 13 cm in air and 1.8 cm in tissue. The average power emitted is about 10 microwatts/CM^2 and the energy per pulse is $< 10^{-9}$ joules. The pulse is transmitted at 2 MHz (see McEwan 1994-1996), and reception is accomplished using a homodyne mode detector, signal integration, and bandpass filtering. The system can detect motion in the near and intermediate-field, with one antenna being used as a transmitting and the second as a receiving element. The transmit and receiver antennas are simple monopoles, about 2 cm long, placed end to end, about 2 millimeters apart, along a common horizontal axis (Figure 2).

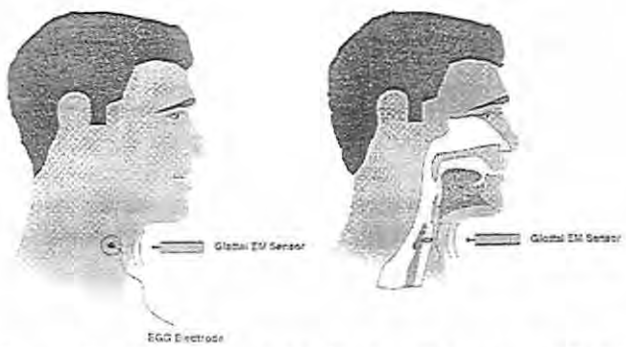


Figure 1. Two right side-views of a subject showing positions of EMGG and EGG sensors. The left image shows the locations of the EGG electrode and the EMGG sensor, and the right image illustrates the vocal tract, along the mid sagittal plane of the neck and head, and the EMGG sensor location. The axis of the transmit and receive antennas of the EMGG is perpendicular to the figure plane. When the sensor is placed against the neck, below the thyroid prominence, the antenna axis is tangent to the neck at the midsagittal plane.

The antenna radiation pattern is toroidal around the horizontally oriented monopole element axis (with the center toroidal radius determined by the antenna conductor), whereupon the electric (E) field oscillates in the poloidal direction, and the magnetic (B) field in the toroidal direction. In the horizontal plane of the sensor, cutting across the neck at the level of the vocal folds, the E field is horizontal, as shown by the arrows near the sending electrode. With the sensor centered on the mid-sagittal plane (0° azimuth), the transmit antenna is located about 1 cm to the right side of the mid-sagittal plane, and the receive antenna about 1 cm to the left. The angle of the sensor axis relative to the midsagittal plane, in the horizontal plane of the vocal folds, defines the azimuthal angle used in experiments described below.

Many other EM sensor modes of operation can be used. For example, for operation at a longer distance from the subject, a far field mode sensor (i.e., radar mode with a directional antenna) can be employed (Skolnik 1990). The EMGG device used for these experiments was operated in the near to intermediate field, where the antenna elements were about one wavelength from the internal glottal structure (about 13 cm in air or 1.8 cm in tissue, where the relative dielectric constant of the tissue is about 50). The 1.8 cm wave length is similar in size to several of the glottal tissue structures and thus both wave scattering and diffraction methods are needed to describe the detailed EM wave trajectories. Numerical modeling validates this assump-

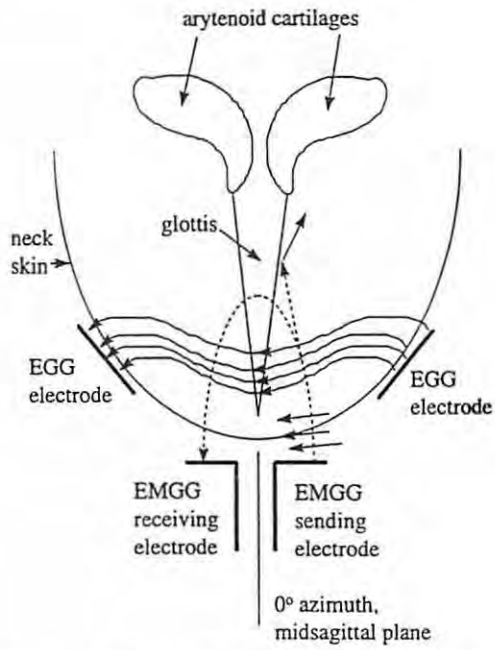


Figure 2. EMGG and EGG electrode placement in a horizontal plane, with a cross section through the neck and larynx Expected electric (E) field patterns are shown for the EGG electrodes from right to left across the glottis and the propagating E field is shown with arrows near the sending electrode.

tion. The sensor transmits an EM wave pulse train toward the interior of the neck, and receives a diffracted or backscattered signal a few nanoseconds or so later. The signal is either backscattered (e.g., reflected) from tissue-air or tissue-tissue interfaces or forward scattered (e.g. transmitted and diffracted) across or around the partially sealed glottal structure.

The signals received from each transmitted pulse are integrated and filtered so that only those tissue interface motions that occur with rates >70 Hz and <7 kHz are detected. Thus, reflections from stationary or slowly moving tissues, such as slow artery blood flow pulsation, are not detected. This high-pass filtered mode is called the "field disturbance" mode and is particularly useful for survey work, where absolute tissue interface locations relative to the antennae are uncertain. The field disturbance mode has the advantage that very small tissue interface motions, with position changes as small as 10 micrometers, are easily detected in the presence of other stationary anatomical structures. However, it has a disadvantage in that it is difficult to accurately identify the absolute locations of the moving, reflecting structures.

The initial experiments described in previous publications (Holzrichter et al 1998, 1996; Burnett et al 1997) showed that the EMGG sensors detect tissue motions that oscillate in phase with the closure and opening of the vocal folds. Because the antenna used with the EMGG was nonfocusing, and because several EM wave cycles were transmitted per pulse, it was not possible to measure which oscillating tissue interfaces were directly responsible for the signals. However the position of the sensors, located directly under the thyroid prominence, together with the longitudinal range limit of the sensor (due to time gating of the received signal) restricted the sight of oscillation to either the vocal folds or the nearby tracheal wall motions. The association of the signal with specific tissue interfaces is ambiguous, as presently understood, because the sensors measure the product of tissue area and vibration amplitude (i.e., a volume displacement). As an example, for the vocal fold medial surface, the area is relatively small, 1-2 cm^2, but the amplitude of motion is relatively large, 0-0.5 cm; whereas for pressure induced tracheal wall motions, the area can be relatively large 5-10 CM^2, with the amplitudes being relatively smaller, perhaps < 0.1 cm. The waveforms for EMGG will be shown to be so consistent with EGG signals that it is "natural" to associate the EM scattering site with changes in the glottal tissue configurations. However, Burnett et al. (1997) stress that EM wave reflection from subglottal tracheal wall motion, induced by pressure buildup as the glottis closes, can lead to similar appearing EMGG signals. Focusing and time gated EM sensor systems are being developed to accurately measure the tissue interface motions of the differing tissues to determine the exact sources of EM wave reflections.

At this point both hypotheses of forward scatter (diffraction) and backscatter (reflection) must be considered. For forward scatter, the waves are diffracted (due to phase cancellations from boundary conditions at or near the electrodes) from the sending electrode through the glottal gap to the receiving electrode, as shown by the bent dashed lines in Figure 2. The anterior part of the glottal gap is the preferred path because the bending (diffracting) angle is less severe and the pathlength is smaller than for waves penetrating the posterior gap. For backscatter, the waves are likely to be reflected first from the vocal fold medial surfaces, toward the dorsal region of the neck, as shown by the straight dashed line and the scattering arrow. There would then have to be secondary reflections from cartilages and other interfaces back to the receiving electrodes. It is not clear which part of the glottis would be favored by backscattering. In any case, a reflected signal from the air-tissue interface should assume more of the shape of the glottal area function (or PGG) than the contact area function.

Note that for both types of scattering, the E field of the EMGG device has the same orientation in the horizontal plane as the E field of the EGG device at the sending electrode (see arrows to the left in Figure 2). At the receiving electrode, the direction of the E field is not clear because of the uncertainty of the scattering mechanism. Further studies are underway at Lawrence Livermore National Laboratories to determine the relative strengths and directions of EM wave scattering from the folds under varying conditions of contact and from the nearby tracheal walls. The experiments presented in this report are designed to compare, under a variety of conditions, well understood EGG signals associated with vocal fold contact to those obtained by an EMGG sensor measuring laryngeal tissue conditions associated with vocal fold action.

Methods

To test the scattering hypotheses, a side-by-side comparison of the EGG and EMGG devices was undertaken. Because of the difference in design and construction, the two devices did not interfere with each other, neither in terms of physical placement nor in terms of carrier frequency and spurious EM noise generation. Their simultaneous use enabled simple one to one comparisons of the two systems' measurements.

At the University of Iowa, four subjects were asked to perform phonatory tasks that produced a range of voice types and transducer placements. The subjects were all young to middle-aged males (ages 28, 31, 55, and 57). The male gender was preferred for this study because males have greater protrusion of the thyroid cartilage than females, making their EGG signals stronger. For this comparative study, we felt that the stronger signal-to-noise

ratio was an advantage to tease out the signal differences. Also, none of the subjects had excessive fat in the neck, the presence of which would have limited the signal to noise ratio of the EGG signal as well (Titze, 1990). The subjects are identified by ascending age with the following symbols: S1, S2, S3, and S4.

The vocal conditions included three different levels of vocal fold adduction (normal, breathy, and pressed), three different registers (falsetto, modal and fry, and two pitches (low and high). It is known that variation in adduction and register cause major changes in the waveform of the EGG signal (Childers et al. 1986). In particular, breathy voice and pressed voice create differences in the vocal fold contact patterns in the anterior and posterior regions of the glottis. These differences should be noticeable in the glottographic signals. All in all, each subject phonated five tokens of a steady /ɑ/ vowel for each of three levels of adduction, three registers, and two pitches.

To test the scattering hypotheses, the EMGG electrodes were moved azimuthally around the neck (Figure 3). The azimuthal rotation around the neck (away from 0°) should *increase* the EMGG signal strength if the transducer acts in a reflection mode because the scattering angle from the medial surface of the vocal folds would be reduced. However, the distance of the total signal path also increases with angle, making the advantage of off-axis placement somewhat less obvious. On the contrary, the azimuthal rotation should clearly *decrease* the EMGG signal strength if the transducer acts in the transmission mode because less contact surface would be in the field path. As seen in Figure 3 for $\theta \approx 45°$ (5 cm right), any near-field diffraction from one electrode to the other would include little of the glottis, whereas for $\theta = 0$ (centered) the anterior glottis is likely to be in the diffraction path.

The signals were recorded using a Sony PC-108M digital tape recorder at the University of Iowa and processed at Lawrence Livermore, with Matlab and other

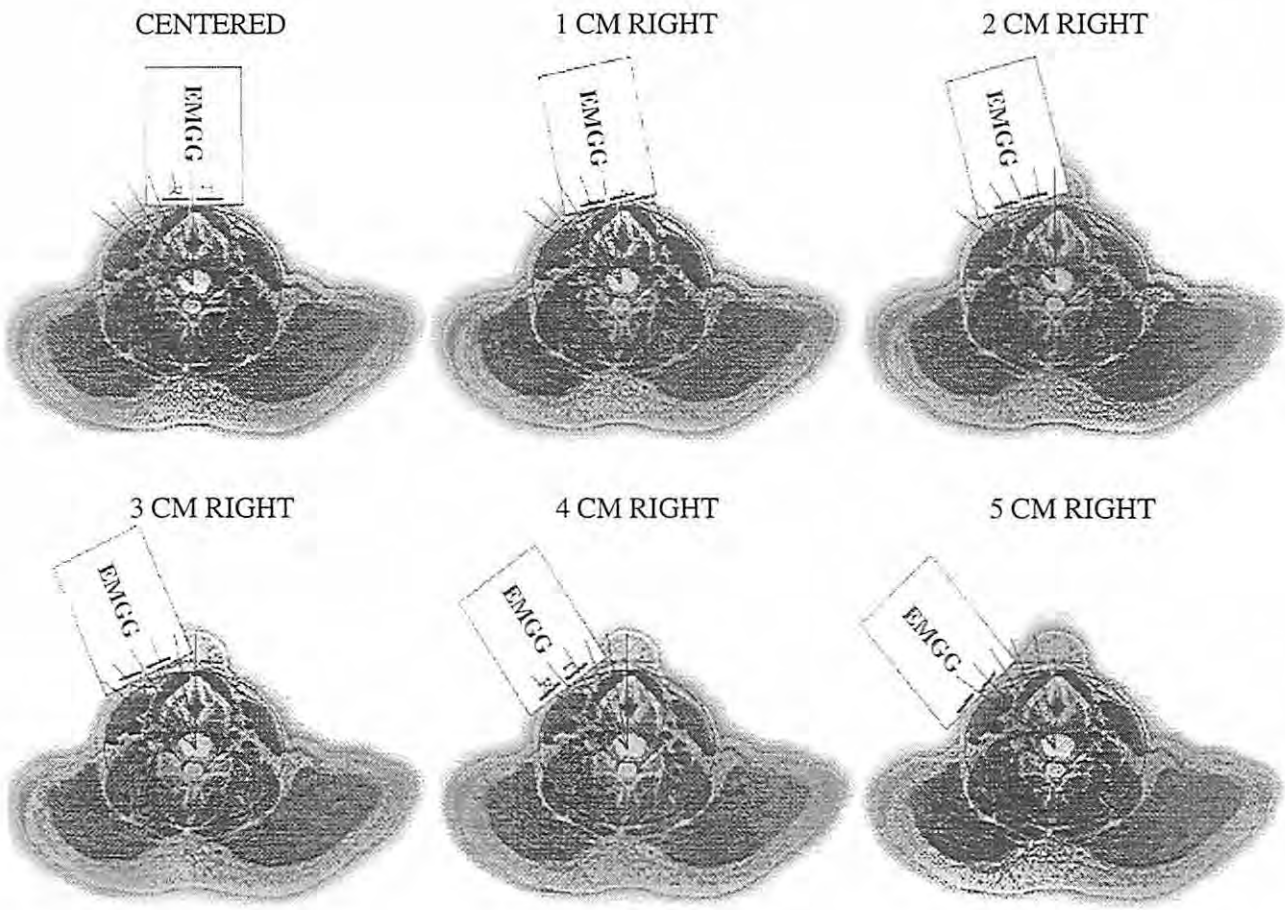


Figure 3. Transverse sections through the neck at the level of the glottis, with the EMGG sensor placed at different locations around the neck.

conventional signal processing software. In all cases, instrumental induced filtering of the signals (i.e., internal high pass and low pass filtering) was removed from the data sets before display in the following figures.

Results

Accurate glottal timing information was obtained from the EGG signal, which is unambiguously associated with the opening and closing of the vocal folds. The EMGG signal corresponded well to the EGG signal for glottal closing, but for some conditions the detailed differences for glottal opening were large.

Figure 4 compares the EMGG signal with the acoustic (audio) signal and the EGG signal for modal (chest) voice, normal adduction, and low pitch (161 Hz) from subject 1. In this example, as in other figures to follow, the acoustic signal was shifted 1.4 ms in time to correct for the relatively slow sound speed from the glottis to the microphone, in contrast to the near instant measurement time of the EGG and EMGG. The negative peak on the acoustic signal is associated with the closure of the glottis. Both sensor signals show agreement on glottal closure and glottal opening, as well as the general shape of the contact pattern. This is a case for best agreement between the two waveforms.

The timing between EMGG and the photoglottography (PGG) have also been validated by employing high speed photography systems, coupled to a laryngoscope (Burnett et al. 1997; Holzrichter, 1996 and 1998; Burnett and Leonard). Results showed that opening and closing events were consistently correlated.

The data in Figure 5 show the EGG and EMGG signals from all four subjects, phonating in modal register.

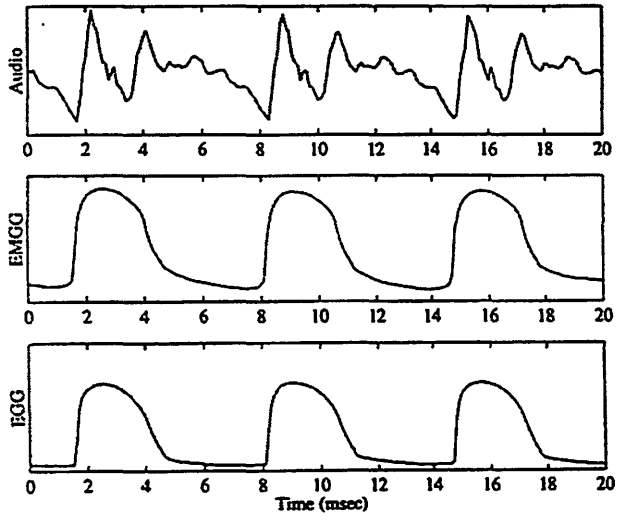


Figure 4. An acoustic (microphone) signal, the EMGG signal, and the corresponding EGG signal measured simultaneously on subject S1 for three pitch periods, modal phonation, and vowel /α/.

The fundamental frequency (F_0) was in the 100 to 120 Hz range for all subjects. Bold lines are for EMGG and fine lines for EGG. The wavefonns have similarity, but there are considerable differences between them for all subjects except S1. Note, for example, the shape differences for subject S2 in the early contact phase. The EMGG does not rise to the same initial height as the EGG, but rises later. A similar effect is seen for S3. For subjects S1 and S4, the differences are greater in the low contact phase. These differences are likely to be the result of different field intensity distributions over different portions of the tissue in contact.

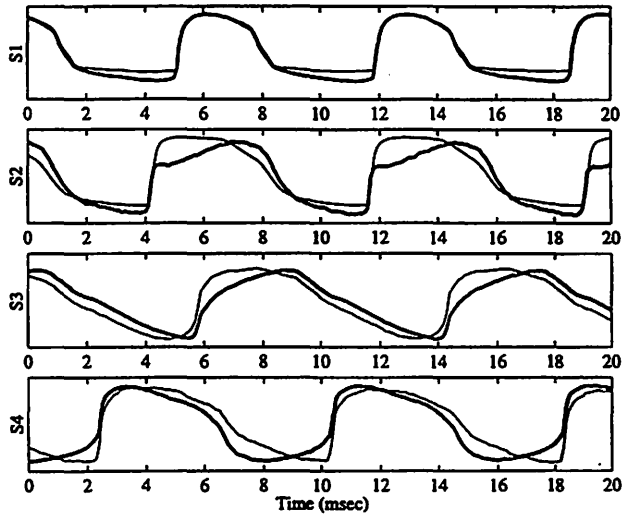


Figure 5. Signals from EGG and EMGG from each of the four subjects for modal phonation for the vowel /α/. Fundamental frequencies ranged between 100-120 Hz across the subjects; EMGG is in bold lines and EGG in fine lines.

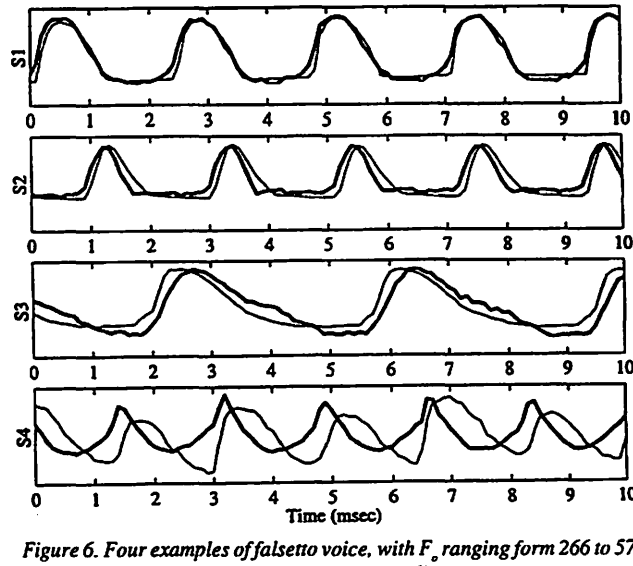


Figure 6. Four examples of falsetto voice, with F_0 ranging form 266 to 578 Hz; EMGG is in bold lines and EGG in fine lines.

Figure 6 shows examples of falsetto voice, with F_0 ranging from 330 to 550 Hz. Waveforms for S1, S2 and S3 are similar, with the main differences being a slight phase offset for S2 and S3, but for subject S4 there is little similarity. The width of the contact pulse is smaller for EMGG than EGG in this subject; since contact is incomplete in falsetto voice, the two systems may be measuring different spatial contact conditions.

Figure 7 shows examples of breathy voice for the four subjects. The fact that the overall agreement between the two sensor systems is best for this phonation type (except for subject S3) speaks in favor of the diffraction (forward scattering) hypothesis. This is because the primary contact region for breathy voice is normally in the anterior glottis (the nearest field region), where maximum diffraction would occur.

Figure 8 shows examples of pressed voice for the four subjects. There are substantial variations in the shapes of the EMGG signals compared to the EGG signals for all individuals. These larger differences again support the diffraction hypothesis because more contact area variation occurs in the posterior glottis during pressed voice than breathy voice, primarily at the bottom of the vocal fold where contact is incomplete and spatially nonuniform. But the reflection hypothesis would also claim less agreement for pressed voice because the tissue-air interfaces are only partially preserved with high degree of contact over the entire vocal fold. In spite of these disclaimers, some of the important timing events, such as glottal closure, are preserved between the EGG and EMGG sensors. There is less agreement between the two signals regarding glottal opening.

Figure 9 shows a modal-fry register transition for subject S2. Measured signals are again audio, EMGG, and EGG. In modal voice, represented by the first six or seven cycles, the waveforms have similar shapes, as previously discussed. After about 70 ms, the EMGG shows a strong negative thrust that corresponds to glottal closure and the upward movement of the EGG. Here a reflected (backward scattered) signal may be starting to mix with the forward scattered signal, which gets smaller as the vibrational amplitude decreases in vocal fry. At about 125 ms, the

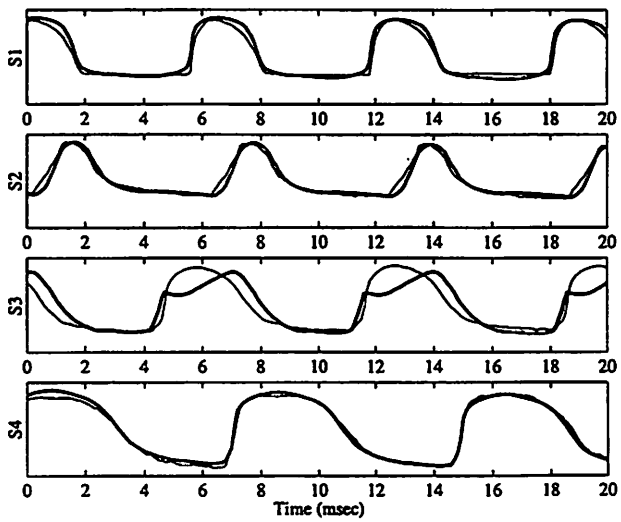


Figure 8. Four examples of "pressed" voice; EMGG is in bold lines and EGG in fine lines.

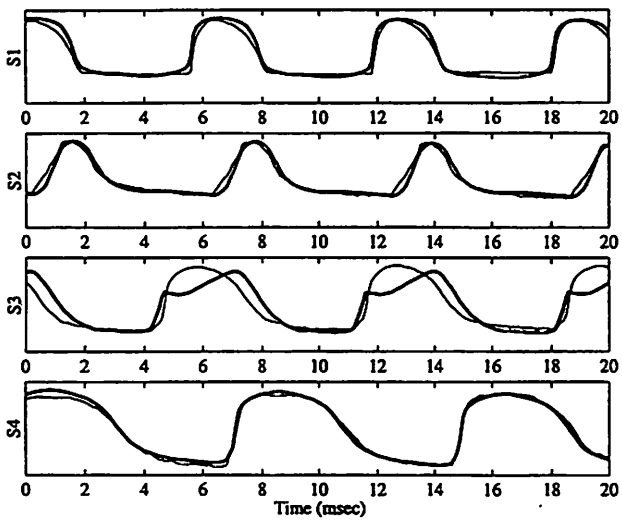


Figure 7. Four examples of "breathy" voice; EMGG is in bold lines and EGG in fine lines.

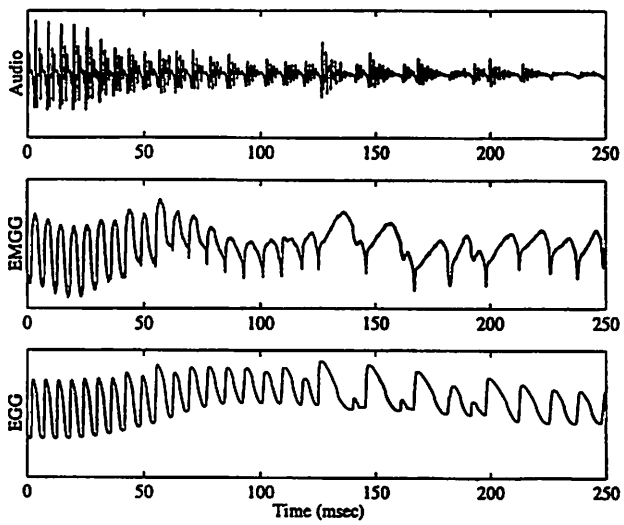


Figure 9. Transition from "modal" speech to vocal "fry", spoken by subject S2. Measured signals are audio, EMGG, and EGG.

downward thrust becomes a spike, apparently cancelling the upward thrust associated with increasing contact area. It is possible that a reflected signal from the expanding tracheal wall at glottal closure can cause this phase cancellation (Burnett et al, 1997). Notice that this is precisely what was beginning to be seen in pressed voice from the same subject in Figure 5, a lower dip prior to closure, followed by a delayed rise.

Figure 10 shows the normalized root mean squared (RMS) amplitude of the EMGG signal as a function of azimuth displacement (measured in cm around the neck in reference to the thyroid prominence). The sensor positions agree roughly with those shown in Figure 3, although the transverse sections through the neck in Figure 3 were not obtained from any of the four subjects. It is clear that the signal strength *decreases* in both directions with azimuth for all four subjects. There is a little ripple in the general maximization of the signal strength near 0, but this is attributed to system and measurement noise. These data are quite compelling in favor of the diffraction mode, for which the 0° azimuth is a clear advantage, as was argued earlier. However, because of the differences in path length, near-field scattering, and multiple moving tissue surfaces (such as the trachea in combination with the medial surface of the vocal folds), the reflection mode hypothesis cannot be discarded.

Conclusions

It appears that electromagnetic glottography, EMGG, may be a viable alternative to electroglottography (EGG) for pitch and vocal fold contact measurement. The EMGG signals are basically similar to those of EGG, but some differences exist as a function of phonation type, subject, and sensor placement. Closer agreement has been found with phonation types for which the anterior glottis is assumed to be the primary probing region, suggesting that the sensor operates more in a diffraction mode than in a reflection mode in the near-field.

A clear advantage of the EMGG approach is that contact with the subject's skin is not needed. The capacity of the EMGG to provide sensing "at a distance" (with distances up to 50 cm measured) leaves space for other sensors to be used simultaneously in clinical applications. In the near field, however, as shown in this study, there may be a confounding between reflected and diffracted waves. Because the EMGG signal has the potential of measuring medial surface movement (and perhaps tracheal movement) under conditions of contact as well as non-contact, it is neither a pure contact area nor a pure glottal area detector. In the future, more spatial discrimination may be attainable with arrays of sensors that could be designed to focus the EM wave onto a specific tissue subsurface to measure glottal positions as a function of both space and time.

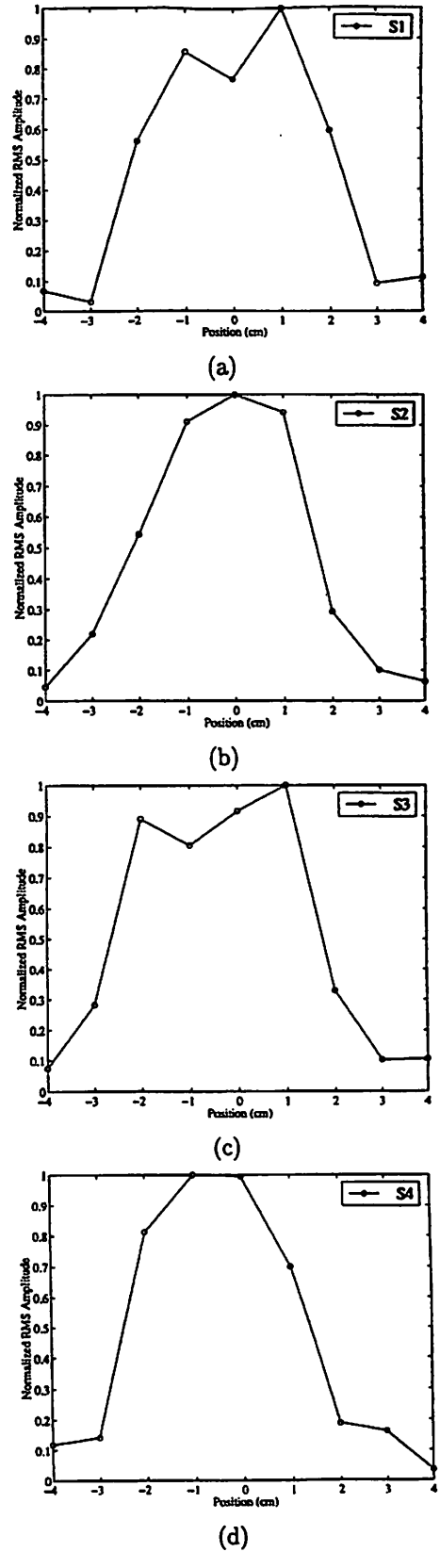


Figure 10. Variation of the root mean square (RMS) value of the EMGG signal as a function of azimuth position relative to the midsagittal plane.

Acknowledgments

This work was supported by Grant #P60 DC00976 from the National Institute on Deafness and other Communication Disorders and under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract W-7405-Eng-48 .

References

Burnett, G.C., Gable,T. G., Holzrichter,J.F., Ng, L.C., (1997), Voiced excitation functions calculated from micro-power impulse radar information. paper 4aSP4. 134th meeting of the Acoustic Society of America, December 1-5, 1997 San Diego. J.Acoust. Soc. Am. Vol. 102 (5)Pt. 2, 3168 (Nov 97) Available at the web site: http://speech.llnl.gov/

Burnett, C.S. & Leonard, R. J. Comparisons of EM sensor signals to high speed video images during speech production; private communication.

Childers, D., Hicks, D., Moor, G., & Alsaka, T. (1986). A model for vocal fold vibratory motion, contact area, and the electroglottogram. *Journal of the Acoustical Society of America,* 80, 1309-1320.

Hamlet, S. & Reid, J. (1972). Transmission of ultrasound through the larynx as a means of determining vocal-fold activity. *IEEE Transactions on Biomedical Engineering, 19*(1), 34-37.

Holzrichter, J.F., Burnett, G.S., Ng, L.C., & Lea,W.A. (1998). Speech articulator measurements using low power EM-wave sensors. *Journal of the Acoustical Society of America, 103* (1) 622.

Holzrichter, J. F., Lea, W. A., McEwan, T. E., Ng, L. C., & Burnett, G.C. (1996). Speech Coding, Recognition, and Synthesis using Radar and Acoustic Sensors. *University of California Report UCRL-ID-123687* (reprints available from the Office of Scientific and Technical Information, P.O.Box 62, Oak Ridge, TN, 37831).

Lofqvist, A., & Yoshoika, H. (1980). Laryngeal activity in Swedish obstruent clusters. *Journal of the Acoustical Society of America,* 68, 792-801.

McEwan, T. E. US Patents 5,345,471 (1994), 5,361,070 (1994), and 5,573,012 (1996). Generalized circuit descriptions are contained in these patents, however a wide range of EM sensor systems can be used for experiments similar to those described above.

Rothenberg, M. (1973). A new inverse filtering technique for deriving the glottal airflow waveform during voicing. *Journal of the Acoustical Society of America,* 53, 1632-1645.

Skolnik, M.I. (1990) *The Radar Handbook, 2nd ed.*, McGraw-Hill.

Titze, I. (1990). Interpretation of the electroglottographic signal. *Journal of Voice* 4(1), 1-9.

Zagzebski, J., Bless, D., & Ewanowski, S. (1983). Pulse echo imaging of the larynx using rapid ultrasonic scanners. In D.M. Bless and J. H. Abbs (Eds.), *Vocal Fold Physiology: Contemporary Research and Clinical Issues* (pp. 210-220). College-Hill Press, San Diego.

NCVS Status and Progress Report - 13
April 1999, 131-140

Multiple Sound Sources of the Vocal Tract (An Analysis of [Imitated Tibetan] Chant)

Michael Edgerton, Ph.D.
The Waisman Center, The University of Wisconsin-Madison
Aliaa Khidr, Ph.D.
The University of Virginia
Diane Bless, Ph.D.
The Waisman Center, The University of Wisconsin-Madison

This paper will discuss multiple sound sources of the vocal tract, focusing specifically on one type of multiple sound source, that of (imitated Tibetan) chant. First I will present a physiological outline of multiple sound sources, in order to better understand chant as just one of the many types of multiple sound sources that composers and performers might utilize. The main body of the paper will then describe a collaborative study between Drs. Khidr, Bless and the author, which is designed to quantify the physiologic and acoustic parameters of chant. By doing so, this paper will attempt to merge current musical compositional imperatives with current scientific rigor and rationale. Ultimately, this cross-disciplinary work will impact both voice science and musical expression as much more needs to be learned regarding the limits of the voice. Compositionally, the value for introducing scientific rigor into the world of new music has been well understood by those associated with electronic and computer music, acoustic modeling and design, or for composers exploring the limits of expression associated with the extensions of performance technique. However, the voice has been treated much with the same musical aesthetic and principles of phonation that existed during the 18th and 19th centuries. So that while instruments have incorporated multiphonics, complex (noise-based) frequency signals, and transient gestures, the voice has remained, for the most part, the carrier of simple melodic formuli. Therefore, this paper will suggest that by knowing the limits of the voice and what is pedagogically safe, contemporary performance/compositional practices will begin to explore and expand upon the current trends of positing simple solutions when approaching the voice.

Multiple sound sources have been identified as important instrumental resources for contemporary, avant-garde instrumental music since the late 1950's, when Bruno Bartolozzi attempted to systematize a theory behind the production of multiphonic sonorities (Bartolozzi, 1969). This conceptual innovation led, in the 1960's, to the development of expanded sound resources for the contemporary voice through important works such as Maxwell Davies's "Three Songs for a Mad King," Ligeti's "Aventures" and "Nouvelle Aventures," and others. All of these works featured expanded vocal resources and were important early explorations for the development of an extended vocal technique.

The difficulty of presenting new techniques for the voice has much to do with the fact that we are living, breathing instruments and cannot be taken apart (and put back together again) nor manually adjusted to the same degree that instruments are. As a result composers have tended to explore contemporary performance techniques utilizing phonetic based articulatory procedures, and have not examined source, resonance or multiphonic principles outside of a few isolated sound poets and/or vocal improvisers.

Between the late 60's and the late 90's a few composers have continued to explore the voice. These explorations have frequently incorporated more precision regarding the acoustic output of complex sound sonorities, though rarely physiological behaviors. It is the purpose of this study to introduce a systematic physiologic and acoustic framework, from which we will present an analysis of one form of voiced and voiced multiple sound source behavior, that of (imitated Tibetan) chant.

There exists a significant body of literature that discusses multiple sound sources, not as a desired musical expression, but rather as part of a dysfunctional makeup. Various forms of multiple sound sources have been described as dysfunctional and functional (Large and Murray, 1979; Nonomura, Seki, et al, 1966; Terrio and Schreibweiss-Merin, 1993; Mckinney, 1982; Kaufman, 1975; Herzel, Berry, Titze, Saleh, 1994). Although the classification of multiple sound sources as dysfunctional is quite the opposite of my intentions, we are able to gather valuable information regarding the physiology and anatomy involved with differing sound sources. As we will see during the course of this paper, multiple sound sources can be well maintained and non-vocally violent, and should be viewed as another legitimate mode of expression. Of particular relevance to our understanding of the physiology of multiple sound sources may be the literature on Diplophonia. Gerratt (Gerratt, Precoda, et al., 1984; Gerratt, Precoda and Hanson, 1987), Ward (Ward, Sanders, Goldman and Moore, 1969) and Marasovich (Marasovich, Gopal, Gerber and Gibson, 1992) used physiological and acoustical data to make some inferences about the etiological bases of diplophonic patterns. They suggest that there are several possible causes of diplophonia, including: unilateral vocal fold polyps causing different vibratory patterns of the two folds; ventricular folds (or false folds) acting as an additional source; congenitally absent or rudimentary vocal folds; assymetrical loading of mucous on the vocal folds; adolescent voice changes; and the voluntary control of assymetrical vocal fold vibration. What this should mean to composers and performers wishing to explore the multiple sound source phenomenon is that many ways exist to produce these sounds and that the ability to produce certain techniques will most likely have significant variation between performers.

In addition to the articles that focus on dysfunctional, multiple sound sources, there is a growing body of literature that suggest that "functional" multiple sound sources are the bases for chanting and various forms of extended vocal techniques (Herzel, 1996; Herzel, 1994; Large, 1979; Anhalt, 1984; Wishart, 1979; Wishart, 1983; Jensen, 1979; Kavasch, 1980; Clark, 1985; Newell, 1970; Barnett, 1972; Chase, 1975).

Multiple sound sources may occur at any level of the vocal tract and include voiced or unvoiced sounds within a variety of simultaneously produced combinations. Whether the multiple sound sources are considered a problem to be corrected, a technique to substitute alternative vibrating sources for aberrant structures, or a viable singing technique, it seems clear that it is necessary to understand the physiology of production and its acoustic result.

The following outline presents numerous multiple sound source (or combinatorial) possibilities. The underlying principle involves a simple additive procedure that will combine two separate sources.

A. Voiced and Voiced

1. chant
   similar to and including Tibetan Chant, with two sources of oscillation (London – Psalm of these Days II)
2. falsetto with 'fry'
   falsetto with vocal fry (Wishart – On Sonic Art)
3. ingressive glottal pitch with ingressive 'fry'
   similar to chant, but performed on inhalation (Edgerton – In America)
4. asymmetrical vocal fold vibration
   left and right vocal folds oscillate at different frequencies, producing true 2 voice polyphony (very rare)
5. subharmonic, voiced, 'gravel' voice
   vocal fold source with sensation of 2nd source being produced below vocal folds (Christi – Instant Reality)

B. Voiced and Unvoiced

1. voice with lip buzz
   (Edgerton – Mountain Songs)
2. voice with whistle (lateral or rounded)
   phonation while whistling (London – Psalm of these Days II)
3. voice with pharyngeal articulation
   (Hirsch – Haiku Lingo)
4. voice with articulation in the oral cavity
   (Monk – Key)
5. voice with nasal articulation
6. voice with salival articulation
   (Edgeron – Syale #1)
7. voice with lingualabial
   phonation with lip-tongue flutter (London – Psalm of these Days II)
8. voice with lingual
   phonation with tongue flutter (London – Psalm of these Days II)
9. voice with air
   (Dutton - Fugitive Forms)
10. vocal fry with air
    (Edgerton – Mountain Songs)
11. voice with velar articulation
    (Dubreuil)
12. voice with uvula
    (Kavasch – The Owl and the Pussycat)

C. Unvoiced and Unvoiced

1. pharyngeal articulation with lip buzz
2. pharyngeal articulation with oral cavity frication
   *(Edgerton - *&%$*)*
3. whistle with sustained oral cavity fricative or approximation
4. whistle with oral cavity articulation
5. whistle with pharyngeal articulation
6. lateral whistle with egressive nasal fricatives or approximations
7. double tongue vibration
   (Wishart – On Sonic Art)
8. lingua with labial flutter
   lip – tongue flutter (London – Psalm of these Days II)
9. salival frication – dental
   (Edgerton – Mountain Songs)
10. salival frication – cheek
    (Stabler - Druber)
11. salival frication – bilabial
    spit roll ("R" – Frank Cox)
12. salival articulation with percussive dental articulations (staccato attacks)
13. egressive nasal fricative or approximation with percussive dental articulation
14. egressive nasal fricative or approximation with bilabial articulation
15. egressive nasal fricative or approximation with percussive alveolar articulation (lingual)
16. egressive nasal fricative or approximation with sustained alveolar or palatal approx or frication

(Continued...)

17. egressive nasal frication or approximation with lip buzz
18 egressive nasal frication or approximation with ingressive lip buzz

D. Three or More

1. Ingressive dental, salival fricative with an egressive nasal fricative/ approximation
2. Ingressive bilabial, salival fricative with an egressive nasal fricative/ approximation
3. glottal pitch, alveolar or palatal fricative and salival fricative between cheek and gum
4. glottal pitch, pharyngeal articulation, tongue vibration
5. glottal pitch, tongue vibration, bilabial fricative (wide perturbation)
6. glottal pitch, tongue vibration, lip buzz (unvoiced pitch)
7. glottal pitch, rear tongue vibration, front tongue vibration
8. glottal pitch, tongue vibration, labialdental frication

(Please hear the work of Wishart, Minton, Dutton, Moss, Reynolds, Kubota, Sutherland, Blonk, Brooks, Stratos, Cobbing, Kermani, Mirana, Newton, Bijma for many of the techniques listed above)

Figure 1. Combinatorial principles with brief description and selected musical citations.

All of the techniques listed above have been used by vocal artists for many years, but composers have been slow, negligent or resistant to using such combinatorial principles in their work. This has made the task of finding these techniques in composed, written scores a difficult task. Without a doubt there are many other artists in many other areas that will be included within such an outline. In an attempt to counterbalance this lack of compositional activity into such sound complexities, this author has begun to develop systematic approaches to introducing extended vocal techniques through a series of pedagogical pieces that are designed for non-specialists, as well as expert performers. For an in-depth description of the combinatorial elements listed above, please see my upcoming book on extended vocal techniques.

However, as stated earlier, this paper will not present a comprehensive analysis of all of the elements listed in figure 1, but will rather focus upon the first combinatorial principle under the voiced and voiced category, that of chant.

Chant

Introduction

The central task to our understanding of chant is to develop an appropriate and meaningful hypothesis of the physiologic and acoustic elements. Generally, voice science classifies phonation according to four central precepts: airflow, source, resonance and articulation. Airflow refers to the fluid upon which molecular pressure will propagate along, across or within a medium; source refers to the physical element responsible for setting up a disturbance within a medium (such as a string, reed, or vocal fold); resonance refers to the acoustic properties of a system (and in our case is central to our understanding of the voice, as the resonating characteristics are a way of tracking the almost constant change of the acoustic tube that occurs during speech and song); articulation refers to the movement of the vocal tract articulators (lips, jaw, tongue, palate, velum, uvula, pharynx and vocal tract walls). Although important, airflow, resonance and articulation are secondary considerations in this discussion of chant, as the source characteristics present the dominant perceptual and performative feature of chant.

Our central question will ask, what constitutes the double source phenomenon of (imitated Tibetan) chant? For the purposes of this discussion, the double source phenomenon consists of: 1) the combination of two elements that produce the low, gravelly, growl-like sound of chant; 2) two elements that seem to involve the combination of a vocal fold pitch with another periodic source; 3) two sources that generally vibrate at a ratio of 2:1, with the vocal folds comprising the higher pitch, while the 'other element' sounds the lower pitch; 4) two sources that combine in a simple additive process, in which it is possible (and pedagogically effective) to begin with a normal vocal fold source, then to add the 'other element' (or vice-versa), making sure to retain both during the radiated signal (see figure 2).

Theoretically, it is vital to conceptualize chant as two independent sources that may be 'played' separately or

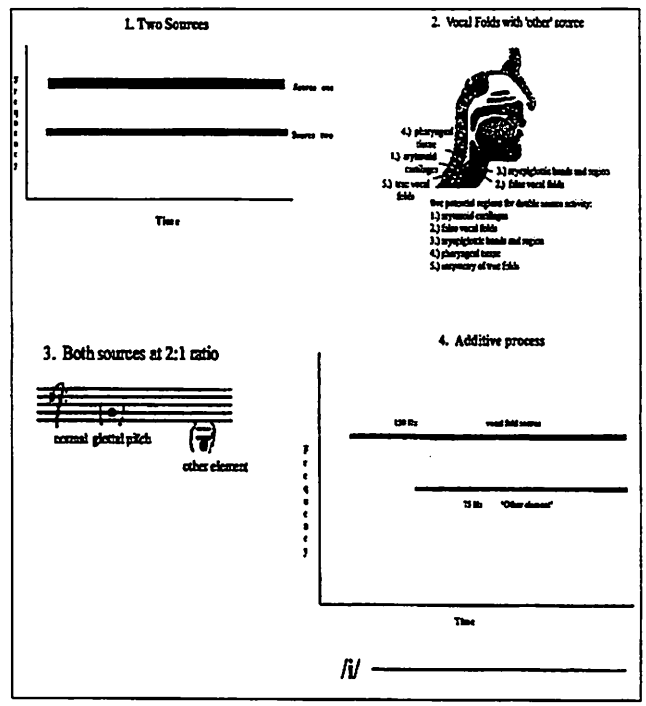


Figure 2. Elements of double source.

together, as this will allow us a greater flexibility of musical expression through the addition of another sound source. There are reported instances when the two sources are tuned, not only to the octave, but that have also occurred at the 12th, 16th, and at non-harmonic intervals. Research into these non-linearities are currently being investigated and will not be reported further in this paper. That the double source phenomenon is capable of producing non-linear and chaotic behavior should suggest exciting potentials for true non-harmonic, multiphonic sonorities that are not stressful or vocally violent behaviors. An example of such behaviors might be notated as such:

Aurally, the precise identification of such low frequencies may be difficult to perceive for the performer.

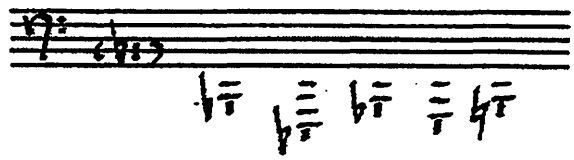


Figure 3. Harmonic and non-harmonic chant notations.

This should be an important consideration for the composer, so that the performer need not necessarily identify specific frequencies of the second source (around 40 to 75 Hz, for example), but might be asked to encompass movement towards or away from whole number multiples or harmonic relationships.

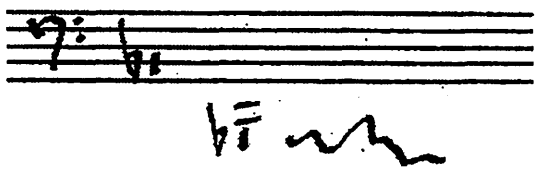


Figure 4. Alternative notation procedures for the lower source.

Methods

Six subjects took part in this study (TC, BD, ME, DH, DM, SS). Their ages were between 25 and 50. Their country of origin included four from the U.S., one from France, one from Tibet. Of the six subjects there were no expert Tibetan chanters. All of our chanters produced a style of imitated Tibetan chant. Our two best chanters were both from the U.S., although all subjects produced the double source characteristic similar to the sound of Tibetan chant. All subjects had the distinct, though subjective, impression of the existence of the second source by identifying a physical sensation of engaging a second oscillator. Aurally and kinetically, this sensation resembled a vocal fry for most.

Two sets of recordings were obtained during both habitual phonatory mode and imitated Tibetan chant. The first set of recordings took place in a sound proof-booth at the University of Wisconsin where all six subjects were seated with a cardioid microphone placed at a distance of 12 inches from each subjects' mouth. Each subject was asked to sustain one vowel (ie. /i/) while switching from 'normal' to 'chant' phonation using his habitual pitch and loudness levels. Each subject was asked to repeat the procedure for other vowels (/a/, /o/, /u/, /e/). A DAT recorder (SONY TCD-D10 PRO II) was used to record four of the subjects, while a reel-to-reel tape recorder (NAGRA LVII, model 1716903) was used for the other two subjects. The acoustic samples were later digitized and analyzed using C-Speech, an acoustic analysis package developed at the University of Wisconsin by Paul Malenkovich. Power spectra and wideband analyses were used to measure resonant characteristics, while narrowband and glottal waveform measures were used to measure the source characteristics.

The second set of recordings took place at the University of Wisconsin's Hospitals and Clinics in the Department of Otolaryngology. Laryngeal behavior was visualized during examinations using 90 and 70 degree rigid endoscopes followed by a nasofiberscope. All were coupled to a Wolfe Endocam Camera and a Bruel and Kjaer Stroboscope light generator. Laryngeal images were recorded on a Hitachi (VT-8A) ½" VHS videocassette recording system for further analysis.

Findings

Visual Observations

Two methods were found to be responsible for the production of the double source phenomena. The first method produced by four subjects, involved assymetries of vocal fold oscillation. The precise methods varied slightly from subject to subject, but all involved assymetries of closure of the vocal folds from anterior to posterior and inferior to superior. The second method of double source production was produced by two subjects and featured medial approximation and/or constriction of the supraglottic structures.

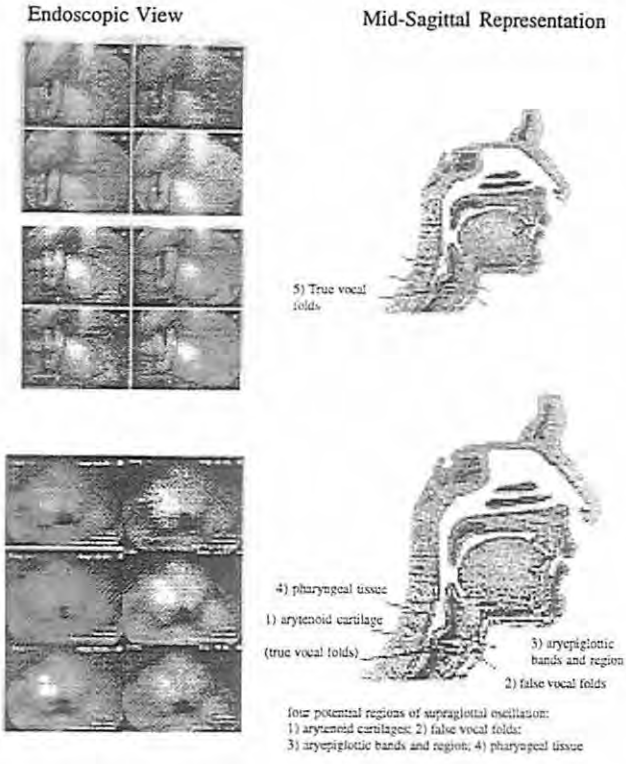


Figure 5. Methods of double source production. Top row – assymetrical vocal fold oscillation; bottom row – supraglottal oscillation.

Although not a significant difference in the sound between the double source produced by assymetrical vocal fold vibration and supraglottal oscillation, there seems to be a subjective rougher, and more resonant quality to supraglottal oscillation. It has been reported elsewhere that the supraglottic structures may serve as an alternative method of phonation for persons with laryngectomies. If this is indeed probable, then it might seem that persons producing chant with supraglottal oscillation will have a greater ability to assume a limited independence from the 2:1 double source coupling than when the double source is produced entirely within the folds themselves. Though not quantified, we might assume that two separate instruments will have greater degrees of freedom of frequency and amplitude for two sounds than for one instrument which will produce two sounds. Additionally, we must consider the construction of the instrument. As the vocal tract is made up of fleshy, pliable material, we must remember that such materials tend to constructively couple, also as we get closer to the original source this coupling effect becomes stronger, so that the coupling within one system will have a stronger tendency to oscillate with each other, while the coupling of two adjacent structures will tend to have more degrees of freedom.

Both methods of this type of double source production are very complex that may have numerous neurological

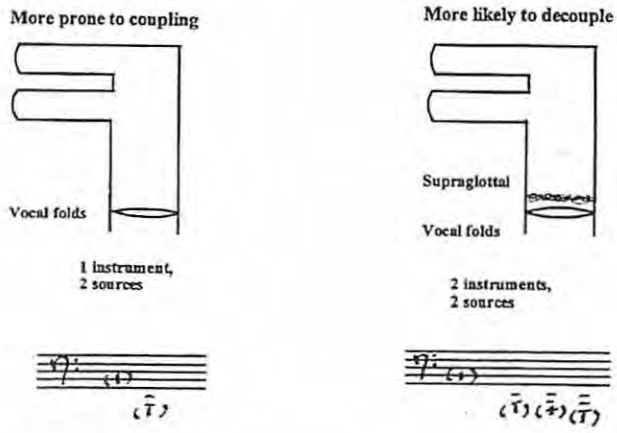


Figure 6. Likely coupling characteristics.

and biomechanical causes, and might very well involve different physiological behaviors from subject to subject, when examining a larger set of subjects and samples.

What we do not know is how the laryngeal physiology differs across larger numbers of subjects who produce a reasonable facsimile of chant. While the arytenoids, false folds and aryepiglottic structures appear to have sufficient vibratory energy, there seems to be skepticism regarding the ability of the supraglottic structures to act as the point of absolute impedance which is necessary to function as the second (or 'other') source of periodic disturbance.

These findings are significant, as the literature on the double source phenomena are few and those that do exist mostly treat this type of phenomena as aberrant behaviors to be avoided. Other studies assume that one type of method is the only way of producing chant, while others are based mostly upon anecdotal evidence with little scientific proof or rationale for their findings. More importantly, these findings give us observations into the complex and little understood world of multiple sound sources, which serves to inform us about the possibilities and limits of such techniques, which immediately serves the artist with imagination and an explorative personality.

Acoustic Measures

Source Characteristics

All subjects showed evidence of a double source signal. Acoustic analyses served to confirm the subjective impression that the additional frequencies functioned as a second regular voiced source. Using a narrowband spectrogram we were able to compare the spectral characteristics from normal to chant. The findings showed that the distance between harmonics halved when moving from normal to chant phonation. This tells us that the vocal fold vibration

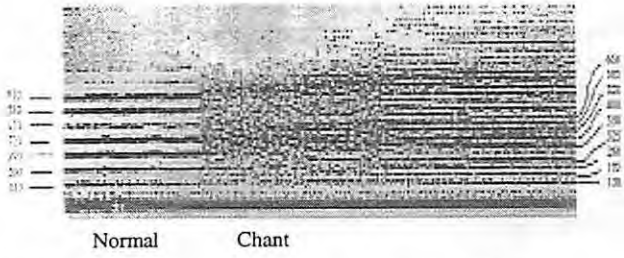


Figure 7. Source doubling; narrowband and waveform analysis.

retained its fundamental frequency and attendant harmonics, while the additional intermediate frequencies were the attendant harmonics of the lower vibrating frequency. Additionally, a glottal waveform provided exact readings of the fundamental period when switching from normal to chant (normal at 130 Hz, chant at 65 Hz). Our findings showed a doubling of the fundamental period when switching from normal to chant phonation.

Resonant Characteristics

Acoustic analysis has been a central tool of voice scientists for many years, as we are limited in our capacity to visualize vocal tract movements. However, scientists have been able to measure vocal tract movements indirectly by analyzing the acoustic output. It is through a separation of the source and resonant characteristics that we are able to analyze the vocal tract. Thus, we understand vocal tract movements through the movements of regions of high and low acoustic pressure. This is a very complex procedure, but is made accessible through widely available acoustic analysis packages which allows us to easily visualize these pockets of energy throughout the vocal tract. Most reliable is the wideband spectrogram window that allows us to pick up the regions of high energy that the tract exhibits at any given point in time. This method of analysis faithfully follows movements of the articulators within the vocal tract to such a high degree of reproduction that we are able to pretty successfully predict large trends of a given speech signal from a given spectral output. Most important for this audience is that certain movements of the formant frequencies will suggest certain articulatory movements and serve to present close approximations for complex questions of area functions and 3-D tube segmentation.

The four subjects with assymetries of vocal fold behavior featured no formant frequency movement (and thus no shift of the regions of high and low pressure) when changing from normal to chant, while retaining the same vowel. This lack of formant frequency movement suggests that the vocal tract retains the same area function when moving from normal to chant. Therefore, the acoustic signal seems to support the visual endoscopic observations that this subject produced the second source at the level of the vocal folds, thus retaining the length of the vocal tract. For this study we will not include interpretations of amplitude and airflow.

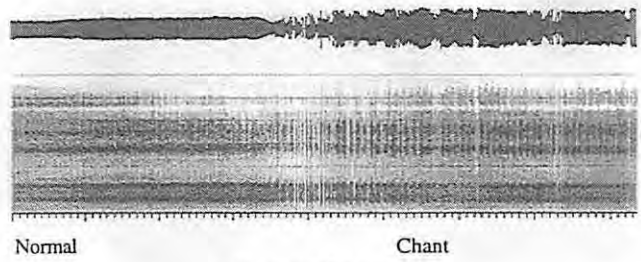


Figure 8. Formant frequency analysis; wideband analysis for assymetrical vocal fold source.

The two subjects who were observed to have supraglottal approximation/closure featured formant frequency movement when changing from normal to chant phonation, while retaining the same vowel. This formant frequency movement, when switching from normal to chant phonation, indicates that a change in the length of the tube has occurred.

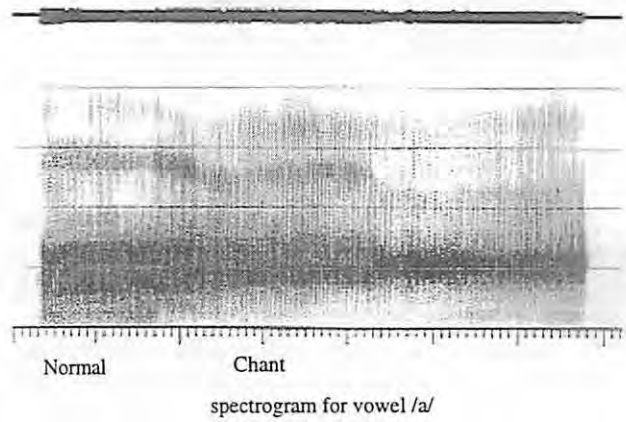


Figure 9. Formant frequency analysis; wideband analysis for supraglottal oscillation.

How do we interpret the formant movements occuring during supraglottal oscillation?

Our conceptual model will be based upon perturbation theory (or how vowels are considered acoustically), and will be calculated following the following seven steps, using a ¼ wavelength principle within a straight and uniform tube:

1) calculate the first three formant frequencies of a normal adult male (17.5 cm)

2) place upon the 17.5 cm tract, the perturbations for vowels /a/, /i/, /o/, /u/ (remember that a constriction in a region of high pressure will raise formant frequencies, while a constriction in a region of low pressure will lower formant frequencies)

3) compare formant frequencies of straight tube with the changes made allowing for perturbation

4) shorten tract to 16.5 (approximating the height of the supraglottal elements)

5) apply perturbation theory for vowels /a/, /i/, /o/, /u/

6) compare 16.5 with 17.5 will they lower, stay the same or raise when switching from normal to chant

7) compare the findings of step 6 with the tracings of formant frequency movement when switching from normal to chant phonation; the 17.5 represents normal phonation, while 16.5 represents the chant phonation.

1. Formant frequency calculations for F1 – F3 for a normal adult male (17.5 cm)

Fr = (2n – 1) * c/4l

Fr = resonant frequency

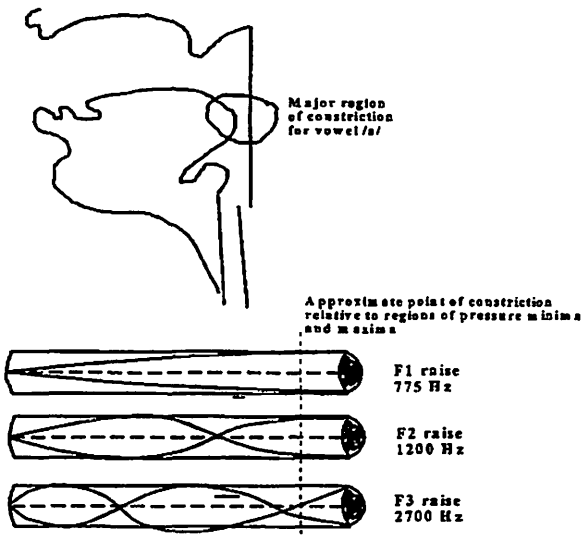
C = constant (speed of air in sound)

L = length of tube

N = resonant frequency

F1 = 480, F2 = 1440, F3 = 2400

2. Perturbations for vowel /a/ for a normal adult male (17.5 cm)



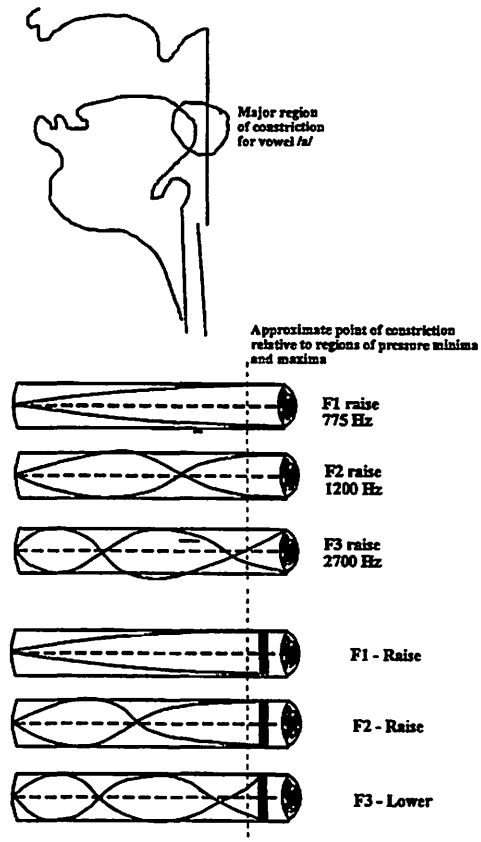
3. Compare uniform tube with perturbation analysis for 17.5 cm tract

	Uniform tube	/a/	/i/	/o/	/u/
F1	480	775	237	452	345
F2	1440	1200	2400	700	818
F3	2400	2700	2990	2400	2560

4. Shorten tract to 16.5 cm, compare normal to chant, will they lower, raise, stay same

Normal (17.5) to Chant (16.5)

	/a/	/i/	/o/	/u/		/a/	/i/	/o/	/u/
F1	775	237	452	345		Raise	Raise	Raise	Raise
F2	1200	2400	700	818		Raise	Raise	Same	Lower/ Same
F3	2700	2990	2400	2560		Lower	Raise	Raise	Raise



5. Compare theoretical model with tracings of our subject

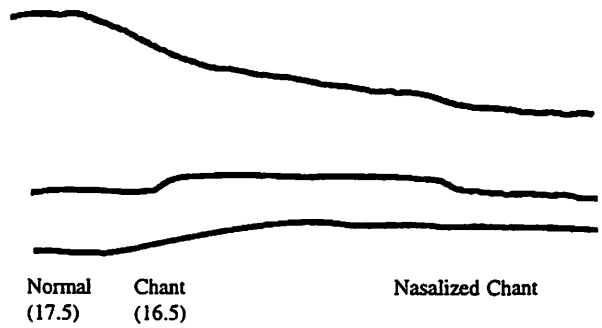


Figure 10. Conceptual model of supraglottal oscillation, steps 1 – 7.

Although these are idealized conceptualizations of a 3-dimensional and highly variable resonant system, the above calculations seem to provide a measure of evidence that the supraglottal oscillators do indeed function as a point of absolute impedance, and thus provides independent confirmation that the vocal tract is indeed shortening when switching from normal to chant for these two chanters.

More analysis and research is needed to determine more precisely the methods and modes of both types of double source phenomenon, but our initial findings seem to strongly indicate that at least two methods of double source production do occur with this type of chant and quite possibly more.

Discussion

Musically, these findings inform us about the construction and capabilities of the vocal instrument. For chant, we have primarily observed tuning between the two sources at the octave, but other relationships do occur and have been reported by other researchers. Generally, chant is mostly produced with the normal glottal pitch in the lower register, but is possible with the glottal voice in a higher, even falsetto (for males) register. In all cases, the non-glottal pitch assumes the lowest member of the dyad, although it would be most welcome to observe cases of the non-glottal pitch oscillating at a higher frequency. This author's observations conclude that amplitude is generally reduced when switching from normal to chant, although a seasoned performer can produce quite a robust sound. Although not vocally violent if done naturally with not a lot of stress, some performers do not seem to be able to sustain chant equal to normal phonation, but this has not been quantified and is purely anecdotal. Lastly, our discussion of chant assumes an outgoing (or egressive) flow of air, but all of the above uses are possible on inhalation (or ingressive airflow) with, potentially, an even greater variation of behavior possible.

Additionally, it was suggested that the two methods of producing chant were mostly similar, but that the supraglottal oscillation had a rougher and more resonant timbre. It was also suggested that the supraglottal, double source oscillation might have more independence, such that the two instruments control the pitches at a ratio of 1:1 and thus carry less of a load, whereas the assymetrical vocal fold oscillation, controlled within one instrument has a ratio of 1:2, in which the single instrument has twice the frequency load and thus taxing the vocal fold system to a much higher degree, which then suggests that in order to produce a periodic source of oscillation, the folds must carry a far higher degree of efficiency with any aberrant behavior potentially serving to disrupt the entire phonatory system. Thus, we might build a working hypothesis that the two instrument system will have the propensity towards a greater, limited independence, while the one instrument system will be effected more highly by the internal stresses and strains upon the delicate vocal fold mechanism.

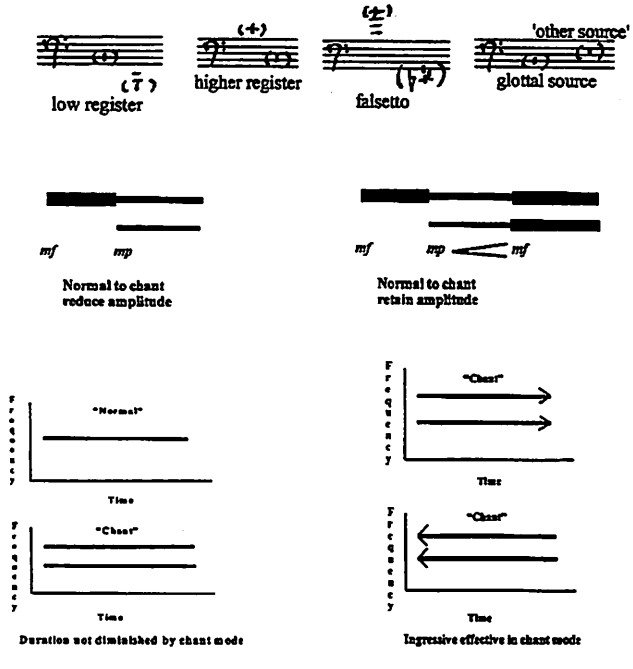


Figure 11. Potential relationships between sources.

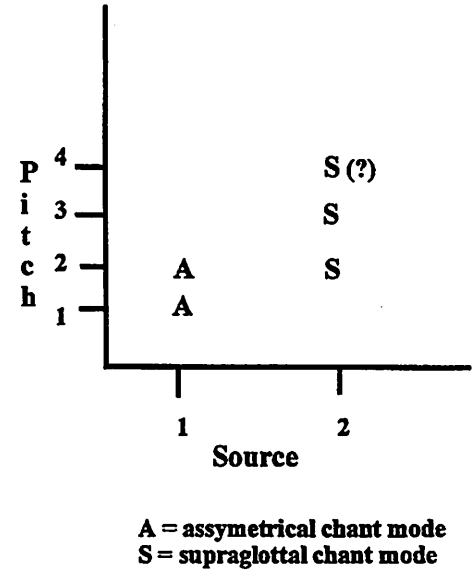


Figure 12. Source to pitch correspondence.

Because of the increased load upon the (1:2) initial integer, it will be harder to break from constructive coupling in this system of fleshy and compromisable material, while the source to frequency ratio of the supraglottal method has a lessor load and may destructively couple the two sources, which to our modern ears and those involved with contemporary art recognize as desired of part of our natural landscape.

Conclusion

This paper is an initial step designed to fill a void within the new music community. For years instruments have had systematic explorations into the extensions of their sounding principles, but the voice has not had that same lineage. As a Research Fellow with the National Center for Voice and Speech my research is designed, in part, to address those needs and concerns of instrument design and development of the contemporary voice. This research will encompass many projects, including chant, ingressive phonation, voiced and unvoiced multiple sonorities, reinforced harmonics, etc. The significance should not be lost on composers as instrument design and its understanding is central to the conceptions of timbre, gesture and the explication of the reproducible, legitimate sound of our contemporary art.

Acknowledgment

This work was supported by grant # P60 DC00976 from the NIDCD. The authors wish to thank Ingo Titze, Diane Bless and Terrence Dolan for their assistance.

Bibliography

Anhalt, I.: Alternative Voices. Toronto: University of Toronto Press, 1984.

Barnett, B.M.: Aspects of Vocal Multiphonics. Masters Thesis: University of California-San Diego, 1972.

Bartolozzi, B.: Metodo per Oboe. Milano: Edizioni Suvini Zerboni, 1969.

Bartolozzi, B.: New Sounds for Woodwinds. London/New York: Oxford University Press, 1967.

Chase, A.M.: Aspects Involving the Performance of Contemporary Vocal Music. Masters Thesis: University of California-San Diego, 1975.

Clark, E.M.: Emphasizing the Articulatory and Timbral Aspects of Vocal Production in Vocal Composition. D.M.A. Thesis: University of Illinois, 1985.

Gerratt, B.R.; Precoda, K.; Hanson, D.G.; Berke, G.S.: Source Characteristics of Diplophonia. Unpublished manuscript, 1984.

Gerratt, B.R.; Precoda, K.; Hanson, D.G.: Diplophonia: Features in the Time Domain. Paper presented at the annual convention of the American Speech-Language-Hearing Association, New Orleans, 1987.

Herzel, H.; Berry, D.; Titze, I.R.; Saleh, M.: Analysis of Vocal Disorders with Methods from Nonlinear Dynamics. Journal of Speech and Hearing Research, 37:1008-1019, 1994.

Herzel, H.; Reuter, R.: Biphonation in Voice Signals. Unpublished manuscript, 1996.

Jensen, K.: Extensions of Mind and Voice. Composer, 2: 13-17, 1979.

Kaufman, W.: Tibetan Buddhist Chant. Bloomington: Indiana University Press, 1975.

Kavasch, D.: An Introduction to Extended Vocal Techniques: Some Compositional Aspects and Performance Problems. Reports from the Center, vol. 1, no. 2. La Jolla: Center for Music Experiment, University of California-San Diego, 1980.

Large, J.; Murry, T.: Studies of Extended Vocal Techniques: Safety. NATS Bulletin 34: 30-33, 1978.

Large, J.; Murry, T.: Observations on the Nature of Tibetan Chant. Journal of Research on Singing, 1979.

Marasovich, W.A.; Gopal, H.S.; Gerber, S.E; Gibson, W.S.: Diplophonia in a Neonate. International Journal of Pediatric Otorhinolaryngology, 25:227-234, 1993.

McKinney, J.C.: The Diagnosis and Correction of Vocal Faults. Nashville: Broadman, 1982.

Newell, R.M.: Writing for Singers in the Sixties. D.M.A. Thesis: University of California-San Diego, 1970.

Nonomura, N., Seki, S.; Kawana, M.; Okura, T., Nakano, Y.: Acquired Airway Obstruction Caused by Hypertrophic Mucosa of the Arytenoids and Aryepiglottic folds. American Journal of Otolaryngology, 17:71-74, 1996.

Smith, H.; Stevens, K.; Tomlinson, R.: On an Unusual Mode of Chanting by Certain Tibetan Lamas. Journal of the Acoustical Society of America, 41: 1262-1264, 1967.

Sundberg, J.: The Science of the Singing Voice. Dekalb: Northern Illinois University Press, 1987.

Sundberg, J.: Articulatory Interpretation of the Singing Formants. Journal of the Acoustical Society of America, 55: 838-844, 1974

Terrio, L.; Schreibweiss-Merin, D. : Acoustic Analysis of Diplophonia: A Follow-up Report. Perceptual and Motor Skills, 77: 914, 1993.

Titze, I.R.; Story, B.H.: Acoustic Interactions of the Voice Source with the Lower Vocal Tract. The Journal of the Acoustical Society of America, 101: 2234-2243, 1997.

Ward, P.H.; Sanders, J.W.; Golman, R.; Moore, G.P.: Diplophonia. Annals of Otology, Rhinology, and Laryngology, 78: 771-777, 1969.

Wishart, T.: The Book of Lost Voices. York: Philip Martin, 1979

Wishart, T.: The Composers View: Extended Vocal Technique. Musical Times, 5:313-314, 1980.

Wishart, T.: On Sonic Art. London: Gordon and Breach, 1983.

Compositions/Artists Listed

Greetje Bijma	various
Jaap Blonk	various
William Brooks	Madrigals
Ellen Christi	Instant Reality (CD)
Bob Cobbing	various
Frank Cox	"R"
Bernard Dubreuil	not commercially available
Paul Dutton	Fugitive Forms
Michael Edward Edgerton	In America.....
	Mountain Songs
	Syale #1
	&%
Shelley Hirsch	Haiku Lingo
Deborah Kavasch	The Owl and the Pussycat
Elise Kermani	various
Nabuo Kubota	various
Edwin London	Psalm of These Days II
Fatima Miranda	various
Phil Minton	various recordings
Meredith Monk	Key
David Moss	various
Lauren Newton	various
Roger Reynolds	Voicespace
Gerhard Stabler	Druber
Demetrio Stratos	various
W. Mark Sutherland	various
Trevor Wishart	On Sonic Art

Part II

Tutorial report

NCVS Status and Progress Report - 13
April 1999, 141-148

Exploring the Human Voice With Computer Simulation

Ingo R. Titze, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa

Introduction

When I think of Carl Seashore, I think of a man who did not compartmentalize science into natural versus behavioral, hard versus soft, clinical versus basic, or any other type of dichotomy. Problems had to be solved, and formal disciplines were just labels on tool boxes that one needed to solve the problems. Clearly, he himself owned a large number of those tool boxes.

The study of human vocalization has thrived on precisely this multidisciplinary approach. Understanding the human voice requires some psychology, some physiology, some linguistics, some physics, some engineering, some music, some medicine, some speech-language pathology, and a lot of open eyes and ears for even more intruders. As I have studied human sound production for the last 20 years here at Iowa, I have found that the larynx is one organ in our body that attracts a multitude of disciplines like a magnet. When we make sound, we vibrate tissues, move air, and radiate acoustic waves. There is a lot of physics and engineering involved in describing these human sound emissions. But as we emit the sound, we also reveal much about our inner selves, not only our anatomy, but our personality, emotional state, health, and (in some cases) our artistic ability. This brings the whole spectrum of arts and humanities into the sphere of investigation.

The Brain and The Larynx

Let's begin with the physiology of voice production. How is the brain involved in making sounds? Figure 1 shows the brain separated from the skull, with the larynx at the bottom. It is connected downward to the windpipe and ultimately to the lungs. Our intent to vocalize is formulated in the cerebral cortex. The speech motor area is then activated and all commands to the larynx go through the nucleus ambiguous, right above the brain stem in the skull section of the diagram. But we must distinguished between vocalization and speech. Speech usually includes vocalization, but vocalization can exist entirely out of the context of speech. Animals vocalize, infants vocalize, speech-impaired or speech-deprived persons vocalize. It is currently believed that the larynx, which is the major sound producing

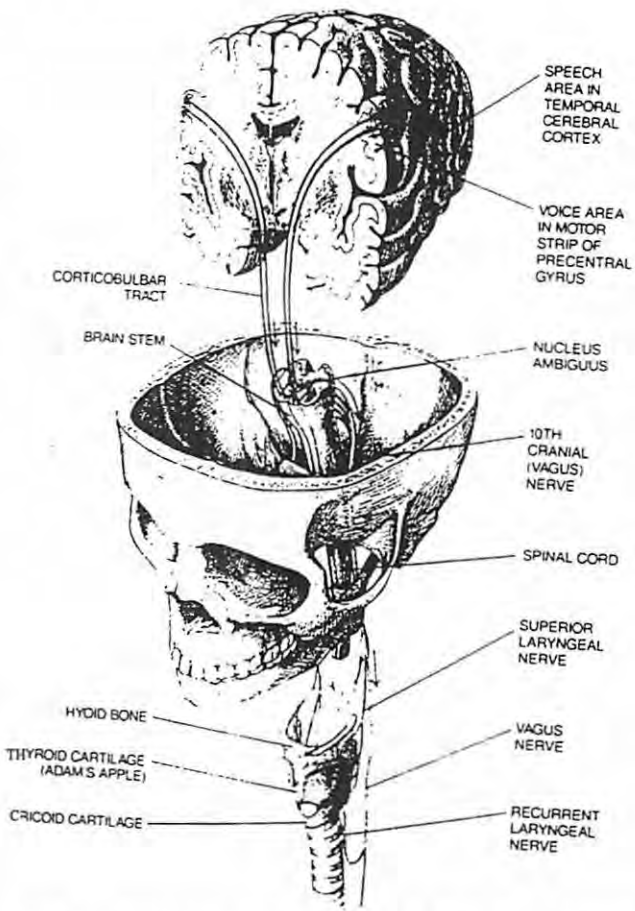


Figure 1. Connections between the brain and the larynx (from Sataloff, 1992, Scientific American; used with permission).

organ in humans is driven by two systems of neural pathways, the speech motor system and the limbic system. The limbic system, activated by emotions and the environment (internal and external to the body) has a strong grip on the larynx and the lungs. Sometimes called the fight or flight system, the limbic system of neural pathways produces motor patterns for cries, grunts, shouts, squeals, moans, and laughter. Thus, we vocalize to locate, to be located, to express anger, fear, joy, and disgust, as do many animals. All of this can be done in or out of the context of speech.

The effect of the limbic system on the larynx is powerful. As a child, I had to walk past a building with a siren on top of it on the way to and from school. I didn't carry a watch, but knew approximately when the siren would go off. I tried never to be at the side of the building, because the sound was so intense and frightening that it caused me to run and shout in an uncontrolled way for several blocks until I finally calmed down. I had little control over my flight responses.

Consider a simple block diagram of the brain-larynx-lung connections as shown in Figure 2. In addition to the limbic system, the speech motor system (top middle) also activates the larynx, but it engages it together with the tongue, the lips, and other speech articulators. It could be hypothesized that the speech motor system is free of emotion and unaffected by the environment, coding only speech gestures into strings of sound. In a newscast, or reading instructions out loud, the system is used that way, a straight line from cortical language and speech centers to vocal output. The main task of the brain is then to coordinate the timing of all the movements. But how detachable is speech from the emotions? Some scientists believe that the speech motor system always engages the limbic system, making optimal use of neural wiring that already exists for primal sounds (Davis et al., 1996). If this is true, emotions and rhythm cycles for breathing and laughter could enhance vocal communication. They become the building blocks for speech. Others believe that the speech motor system is specially wired. Infant cries change to coos, which change to babbles, which change to vowels and consonants, which change to words. Articulatory development, they would say, is part of the human genetic code, perhaps somewhat independent of the limbic system, and unique to humans (Lieberman, 1991). If this is true, emotions may compete with speech for the use of the larynx and the lungs. When you bear spiritual or emotional testimony and you "choke up", your pitch sometimes rises in an uncontrolled fashion. A sob can block your respiration and phonation, and tears and extra mucus can cause you to misarticulate. That's poor speech or poor singing, but extraordinary vocal communication. Some listeners are enraptured by it, others are embarrassed by it. Emotion talks to emotion.

Another part of the nervous system that is deeply involved in vocalization is the spinal reflex system (right side of Figure 2). Breathing is autonomic; it requires no conscious voluntary action. Because the amount of air and carbon dioxide exchanged in the lungs must be regulated, there is an inherent feedback control between the lungs, the larynx, and the spinal cord. The larynx acts as a valve,

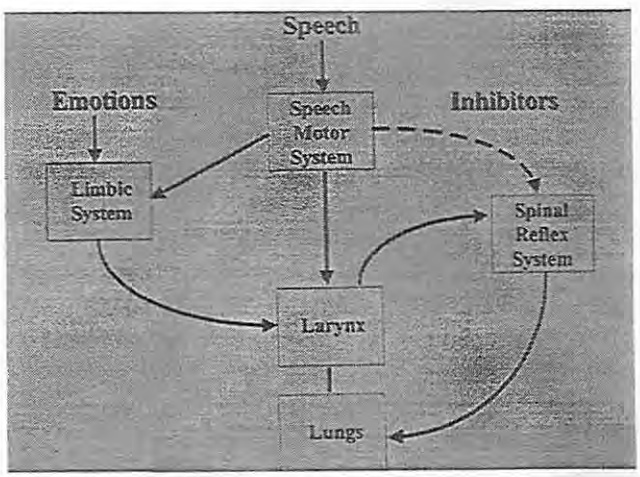


Figure 2. Block diagram of the brain-larynx-lung connections through the limbic and speech motor systems.

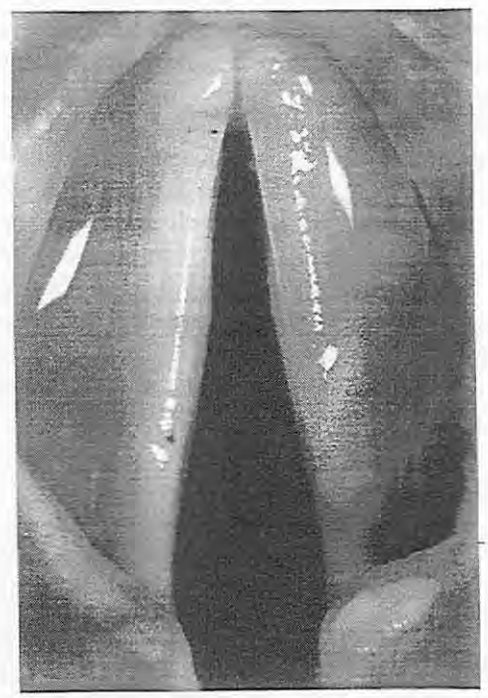


Figure 3. The vocal folds as seen from the top. These folds not only vibrate to produce sound, but they act as a valve to regulate air moving in and out of the lungs.

opening and closing the gate for more or less air to be exchanged (Figure 3). When we vocalize, the breathing rhythm is altered, and the reflex system may get in the way.

Coughing is an example of a respiratory reflex. The slightest touch on the vocal folds with a foreign object will elicit a cough reflex. But why don't we cough when the vocal folds touch each other in collision? One theory is that the speech motor system must inhibit the reflex system (see the dashed line in Figure 2). If this is true, then some voice disorders may result from a malfunction of this inhibition of the reflexes. The reflex system may be triggered too easily and the speech motor system gets tripped up by it. This may be the case in a symptom known as spasmodic dysphonia, where a person speaks with many voice interruptions and much tension.

The main message to be understood here is that the larynx is multi-functional and hard-wired to several major parts of the nervous system. Anatomically, it sits in the middle of the neck, surrounded by all lifelines of the body. It's a busy highway there. The jugular vein and carotid artery pass along-side the larynx, as does the vagus nerve. Food passes directly behind the larynx, and air is drawn through it. Is it any wonder that our voice is a mirror of most of our body functions? Syrus said: "As the man, so is his speech." Voice reflects our age, gender, disease, fatigue and physical fitness. Vocal tremors are like earth tremors, revealing much of what lies under the surface. Shakespeare had Cleopatra utter the following words after learning of the death of Antony: "His legs bestrid the ocean; his rear'd arm crested the world; his *voice* was propertied as all the *tuned spheres*," (Antony and Cleopatra; Act V, Scene I).

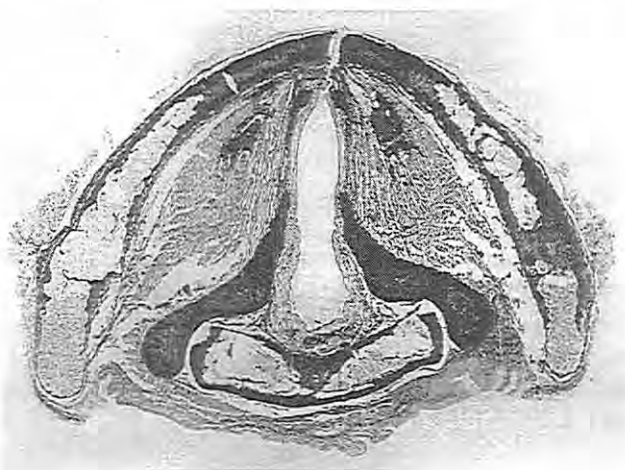


Figure 4. Histological cross-section through the larynx showing the thyroid cartilage (semi-circular boundary around outside), and the arytenoid cartilages (darker areas in the bottom half), and the vocal folds (muscle tissue in upper half).

Shakespeare perhaps recognized that if all the biological systems are functioning synergistically, like pulsating oscillators, human vocal expression can be divine.

Control Dimensions of the Voice

We all know intuitively how to make our voices higher and lower or how to make them louder and softer. But in addition to pitch and loudness, there are at least four other vocal variables we can control: tightness, register, resonance, and roughness.

Consider first how all six of these variables are perceived on a scale:

Perceptual Control Variable	Scale
loudness	soft-loud
pitch	low-high
tightness	breathy-pressed
register	fry-chest-falsetto-whistle
resonance	dull-ringing
roughness	smooth-rough

Most of the scales are continuous, but register is a more quantal (categorical) perception. Listeners perceive one register or another with not much mixture (or overlap) between the perception. Roughness may also be more quantal, but more research is needed on the perception of vocal roughness. We all can distort our voices somewhat, but the mechanisms for control of deliberate distortions are not yet clear.

Loudness Control

The primary control variable for loudness is lung pressure. As lung pressure increases, the aerodynamic power delivered by the lungs to the vocal folds increases; this aerodynamic power is the product of pressure and airflow.

But what is the valve that regulates the amount of flow through the larynx? It is the arytenoid cartilages, shown in cross-section in Figure 4. These cartilages are positioned by various groups of muscles. Adductory muscles move the cartilages together and abductory muscles move them apart. This valving, then, controls the amount of airflow. But more importantly, valving controls the way that airflow is continually interrupted when the edges of the vocal folds vibrate. A sudden interruption of airflow creates sound at high frequency, as in a handclap. Thus, in every cycle of vibration of the vocal folds, airflow must build up, then collapse rapidly, if a wide spectrum of frequencies is to be present in the sound produced.

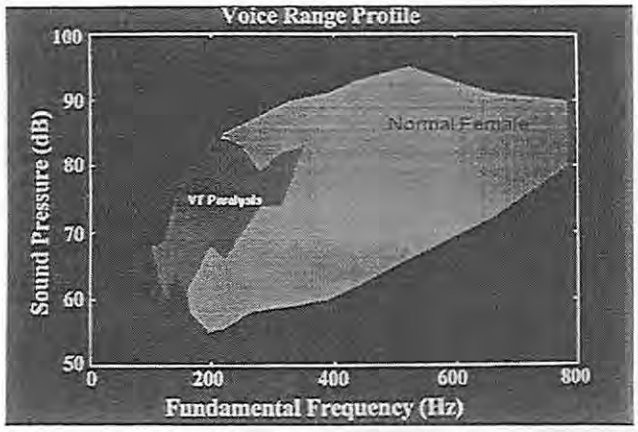


Figure 5. Voice range profile for a normal subject (right side) and a patient with unilateral laryngeal nerve paralysis (left side).

A convenient measure of pitch and loudness range in a human subject is the Voice Range Profile (Figure 5). This graphic representation of intensity on the vertical axis and fundamental frequency on the horizontal axis presents a clear picture of the *range* of phonation, much like an audiogram depicts the *range* of hearing. The lower boundary represent phonation threshold, like the threshold of hearing, and the upper boundary represents the threshold of vocal risk, analogous to the threshold of pain in hearing. Note that the area of the Voice Range Profile is much smaller for a person with unilateral vocal fold paralysis than for someone with normal vocal function.

Pitch Control

Pitch is controlled primarily by two agonist-antagonist muscles in the larynx: the cricothyroid (CT) muscle and the thyroarytenoid (TA) muscle (Figure 6). The TA muscle is the major portion of the vocal fold, hidden behind the thyroid cartilage in the diagram. Cricothyroid action stretches the vocal folds, and thereby tenses them. Without this increased tension, the longer vocal folds would vibrate at a lower pitch, just as in stringed instruments. A long piano string, for example, vibrates at a lower pitch than a short one. But, if the tension of the long string is made to be significantly higher, it is possible to have a higher pitch with a longer string.

In addition to the stretching of the vocal folds by the CT muscle, the TA muscle can contract and stiffen itself internally. In this process, it tends to shorten the vocal folds. In controlling pitch, then, the thyroarytenoid muscle and the cricothyroid muscle are agonists and antagonists, opposing each other.

A third muscular system involved in pitch control is the respiratory system. As lung pressure is increased by the abdominal muscles and other respiratory muscles, air-

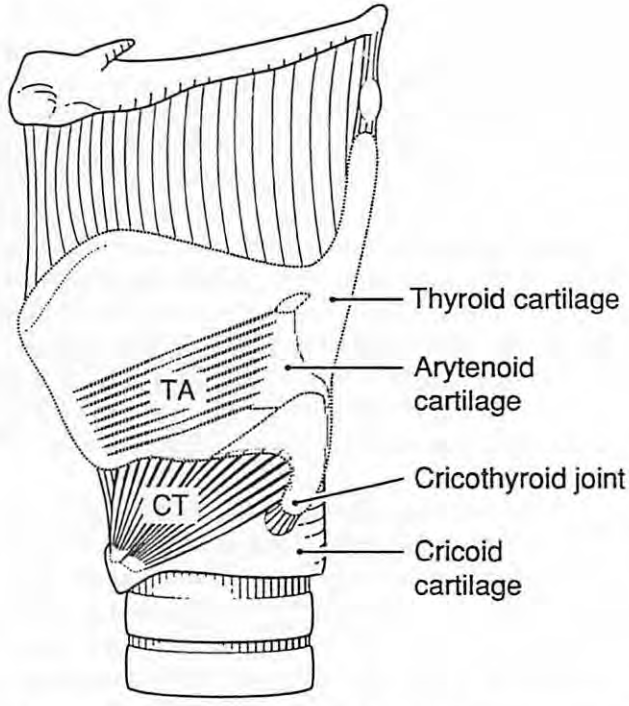


Figure 6. Side view of the framework of the larynx. The thyroarytenoid (TA) muscle and the cricothyroid (CT) muscle are used for control of pitch.

flow increases through the vocal folds. This tends to increase their amplitude of vibration. As amplitude of vibration increases, the average tension of the of the vocal folds also increases. This is the effect you may have noticed when blowing a New Year's Eve noisemaker; as you blow harder into this instrument, the pitch tends to go up. This pitch-loudness interaction reflects the nonlinearity in an acoustic oscillator and can give rise to chaotic behavior. In an infant cry, for example, the frequency of vibration is so dependent on amplitude of vibration that large pitch fluctuations occur with increased lung pressure. Often the vibration patterns becomes very irregular when the infant "bears down".

Control of Tightness

As already stated, the valving action of the arytenoid cartilages brings the vocal folds together or moves them apart. If we use too much force with the adductory muscles, the voice becomes tight or pressed. If we use too little adductory effort, the vocal folds do not approximate tightly enough, and we get a breathy voice.

Tightness and breathiness are symptoms of a number of voice disorders. If the adductor muscles are hyperactive, the voice is judged to be tight, strained, or even strangled. Spasmodic dysphonia has those symptoms; in addition, the tightness appears in spasms rather than in a continuous fashion. In some disorders, the vocal folds are bowed in the middle, but at the endpoints the arytenoid

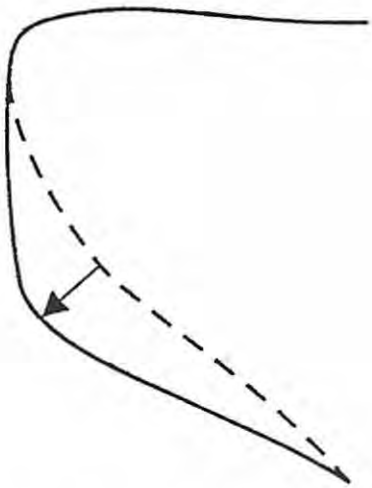


Figure 7. Bulging of bottom of vocal fold for the modal (chest) register production.

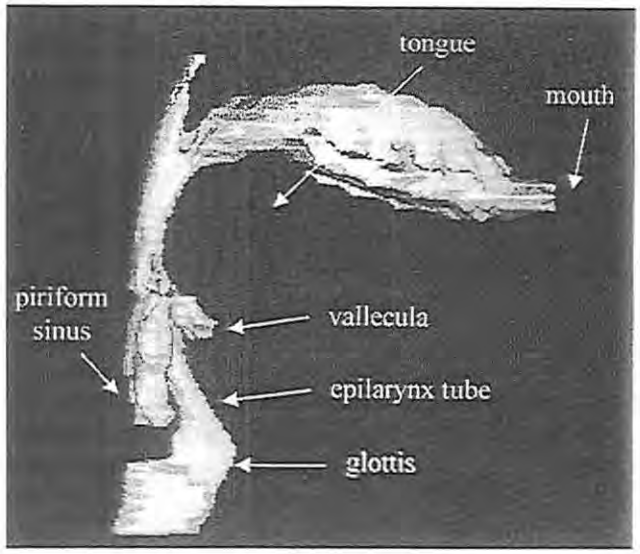


Figure 8. Sagittal (side) view of the human airway as obtained by electron beam computed tomography. The vowel shape is an /a/ as in "father".

cartilages are pressed together tightly. Thus, air can escape in the middle and noise is produced during phonation. Tightness and breathiness may therefore appear together in some disordered voices, but in normal voices the two sounds are generally distinct.

Control of Register and Resonance

A register is a specific vocal quality that can be maintained over some range of pitch and loudness. Examples are modal (or chest) register, falsetto register, fry (or pulse) register and whistle register. The change from modal register to falsetto register is controlled primarily by the thyroarytenoid muscle; that is, the major belly of the vocal folds. As this muscle contracts, the bottom of the vocal fold bulges more medially (Figure 7). As this happens, the bottom of the vocal fold is set into vibration more vigorously by lung pressure. This more vigorous vibration at the bottom of the fold creates an effectively thicker vocal fold and more area of collision over its whole surface. With more collision there is more abrupt disruption of airflow and the excitation of higher frequencies. So, the modal register has a rich timbre, with many harmonic frequencies present, whereas the falsetto register has fewer harmonics and a weaker timbre.

Resonance of the voice (a ringing quality) is primarily controlled by the epilarynx tube. It is a very narrow, short tube just above the vocal folds (Figure 8). This tube functions like the narrow mouthpiece of a trumpet that is directly coupled to the lips. For the trumpet, the mouthpiece and lips create a tight acoustic unit that promotes oscillation of the lips. Similarly, the epilarynx tube and the vocal folds form a tightly coupled oscillating system in the larynx. The narrower the tube, the easier it is for the vocal folds to be set into oscillation. The sound produced when the tube is narrow contains a lot of high frequency energy, which is sometimes called vocal ring or resonant voice.

Operatic singing is characterized by lots of vocal ring, but so is twang, which is heard in some speech dialects and in some country-western singing. In producing twang, the pharynx tends to be narrow also, and there is less low frequency acoustic energy than in opera quality. The nasal port may also play a role in controlling this vocal resonance quality (hence, the term nasal twang), but this involvement of nasality in vocal resonance is still heavily debated.

Voice Analysis

Traditional voice analysis is based on extraction of temporal and spectral features from a microphone signal. This observation-at-a-distance via the airborne signal has limitations, especially if sound production at the larynx and sound transmission through the airways are to be studied separately. The microphone signal is an unfortunate mixture of the combined properties of the source of sound and the propagation of sound through the airways (known as the filter). Thus, speech scientists have been looking for additional information to augment acoustic recordings of human voices. The source-filter theory of voice production can then be studied in greater detail. In particular, the use of fiber-optic viewing of the vocal folds (as in Figure 3) has provided important information about the source, while volumetric imaging of the vocal tract (as in Figure 8) are basically static, affording no temporal resolution of dynamically changing airway structures in speech or singing. But the spatial resolution of these images is so good that even these static shapes help to uncouple the propagation characteristics from the source characteristics. A subject typically

spends many hours in the magnetic resonance imaging machine to map out his or her articulatory space by mimicking the vowels and consonants in speech and holding each steady for several minutes. A complete volumetric scan, with millimeter accuracy, is obtained for each shape, including the oral cavity, the nasal cavity with the sinuses, the pharynx, the larynx, and the trachea. In and around the larynx are small pockets of air that make interesting resonators for sound (the piriform sinus, vallecula, epilarynx tube, and glottis; recall Figure 8).

Analysis of Singing Voice

Recent research has shown that trained singers are able to optimize their source-filter interactions to obtain easier and more efficient voice production. Singers open their mouths wide, particularly at high notes. This has the acoustic benefit of producing a better impedance match from the airway to free space, much like the flare of a trumpet. Impedance (the ratio of acoustic pressure to airflow) is high at the larynx and low in free space, requiring a megaphonelike transformation. A wide mouth also has the benefit of getting the jaw out of the way of the larynx. Often a tense jaw restricts the larynx in finding a position that can be held over a wide range of pitches.

The makeup of a premier singing voice, one that appears only once in a decade and is universally recognized as special, is still speculative; but some hints are beginning to emerge. Naturally great voices tend to have a nonencumbrance of skeletal structures in and around the larynx. This involves, at a minimum, the shoulders, the neck, and the jaw. As a wide range of pitches and loudnesses are accessed, soft and hard tissues do not interfere with each other. There is enough room for expansion, contraction, and linear displacement of the larynx, rib cage, abdomen, diaphragm, and airways as necessary. These organs do not move a lot-on the contrary, a firm equilibrium position (posture) is desirable-but the critical movements must be unencumbered. This applies particularly to the use of opposing muscles (agonist-antagonist pairs). They must not fight each other, but be able to turn on and off gradually (like a dimmer switch) to move structures and change tensions precisely and differentially. Jerky on-off movements are seldom seen in a premier singer. Rather, there is a deathlike calmness on the surface, underneath which huge muscular efforts are expended.

Within the larynx, there are likely to be some morphological differences between ordinary and premier singers, although direct verification by inspection of the organs of deceased singers has not been possible. Scientists have relied on simulation, therefore, to test the optimal structures for sound production. Symmetry between the left and right vocal folds seems to play an important role. In principle, the two vocal folds have their own characteristic modes of vibration (like drums, bells, or strings). These modes depend on the viscoelastic properties of the vocal fold tissues and the boundaries that surround the tissues (the cartilages). If either the boundary structures or the internal tissue properties of the vocal folds are asymmetric, different modes (with different natural frequencies) can be excited. These modes can fight each other. A common airflow between the vocal folds does help to entrain the modes, but there is a limit to this entrainment. If large ranges of pitch and loudness are to be achieved, a highly symmetric pair of vocal folds has a much better chance of avoiding chaotic oscillation.

There is large benefit in having a thick, pliable mucosa as a covering of the vocal folds. This mucosa propagates a surface wave while the vocal folds are vibrating. In fact, it is the surface wave that facilitates the energy transfer from the airstream between the vocal folds to the tissue itself, thereby producing self-oscillation. Highly gifted singers probably have the genetic construct of a thick and pliable vocal-fold mucosa, although direct histological verification is yet pending.

Underneath the loose, pliable mucosa must be a tough ligament that can support large tensions, much like a piano or violin string. For high pitches, this ligament absorbs most of the tension in the vocal folds. The amount and the type of collagen and elastin fibers that make up this vocal ligament may again be genetically determined. Thus, some people may be born with better material properties than others, much like certain woods or metals are more desirable for musical instrument design.

Voice Synthesis

Synthetic speech is beginning to play a role in many facets of our everyday life. Artificial voices are now a familiar sound in phone messages, warning signals, entertainment and computerized self-instruction. There will be many more applications in the future. In particular, we will hear synthesizers with transformed personalities, with exaggerated and diminished emotions, with different dialects, and with voice qualities that will go beyond the dimensions of the human voice.

Speech Synthesis

For many speech synthesis applications, the only measures of success is producing synthetic speech are its intelligibility and naturalness. What comes out of the box is more important than what is in the box. Many coding, storage, and transmission strategies can therefore be used to convert one form of speech to another, or text to speech (Klatt, 1987). There are other applications, however, for which the final output is not the most important product; rather, the principal goal is an improved understanding of the speech production mechanism. What's in the box

matters more than what comes out of the box. The speech may indeed be of lesser quality, but because it helps to illuminate the physical and physiological processes in speech production, it is precisely what the investigator is focusing on.

Clinical Applications of Speech Synthesis

In medicine, the most exciting prospect is predicting the outcome of special types of head and neck surgery; that is, making a model of the entire mechanical or biomechanical process of speech production and then introducing a pathology, such as a vocal fold paralysis (Isshiki & Ishizaka, 1976) or a tissue resection (Fujimura, 1994). The abnormal speech behavior can be quantified in a parametric way. Corrective surgery can then be carried out on the model before any procedures are carried out on humans. In particular, the more radical procedures can be taken to their extremes. Certain parameters in the speech production system can be exaggerated and distorted to find out what lies beyond the limits of normal production. This is very useful in predicting breakdowns of the human speech production system.

With applications to speech training and habilitation, we envision that a vocologist (one who trains and modifies voices) can use simulation to alter any particular voice quality of a client. In addition to giving examples with his or her own phonation, the vocologist can use a synthesizer to play variations of sound qualities, speech dialects and emotions. This will probably be an improvement over the all-human model, because it would allow one variable to be changed at a time. The vocalist will be able to interpret clearly what the acoustic result of a physiologic change in his or her voice will be.

Simulation of the Singing Voice

Although on the surface it would appear that simulation of the singing voice presents a greater challenge than simulation of speech (because of the added artistry), one finds that what is difficult for humans is easy for machines, and what is easy for machines is difficult for humans. For example, humans have great difficulty producing high and loud sounds, that is, sustaining muscle contractions for long periods of time. This is trivial for machines. There is no fatiguing and tiring of muscles. Hence, long, sustained sounds present no difficulty. Furthermore, in singing, much is prescribed by the musical score. Movement from syllable to syllable is coded by meter and rhythm. Thus, it is relatively easy to incorporate articulatory timing into a singing simulation strategy. The melody is also prescribed, whereas in speech the "melody" is free-running (dialect dependent and speaker dependent).

In our work at the National Center for Voice and Speech, we have been reasonably successful in producing some of the sounds of operatic tenors. We have called our singing robot "Pavarobotti" (Figure 9), in honor of our great contemporary tenor, Luciano Pavarotti. Some of the tenor characteristics, such as the vocal ring and the dramatic rich timbre, are surprisingly easy to simulate. Vocal ring is modeled by inserting a small quarterwave resonator above the vocal folds, approximately one-sixth of the length of the vocal tract and one-sixth of the cross-sectional area of the pharyngeal input to the vocal tract (Sundberg, 1987). This quarterwave resonator, tuned to approximately 3000 Hz, mimics the narrow acoustic channel (the epilarynx tube) above the vocal folds. Vibrato can be introduced with sinusoidal frequency modulation of the air pulses emitted from the larynx, but this modulation must also include some neurologically-based randomness.

We have found that the glottal airflow pulse in operatic sounds is not substantially different from that found in normal speech production. Hence, we have concluded that singing is not fundamentally different from speech (acoustically), but it takes great amounts of training and building of muscles to maintain this constancy over a two octave range.

We believe that a future version of Pavarobotti will become a very exciting teaching tool for those who provide voice training and vocal habilitation. Much of what is now done laboriously --- off-line --- will be able to be accomplished in real time. Perhaps a singing robot will be part of every studio and clinic; buttons can be pushed and knobs can be turned to create a variety of sound qualities that vocalists can listen to and evaluate. Beginning with a model of a beginner's voice, attributes of long-range training can be added, one by one, to project the voice of a pupil or client into the future. This is not to say that Pavarobotti will become the teacher, but rather the teacher's aid. By displaying his vocal tract anatomy, there will be some direct visual confirmation of what the system is and what the system does.

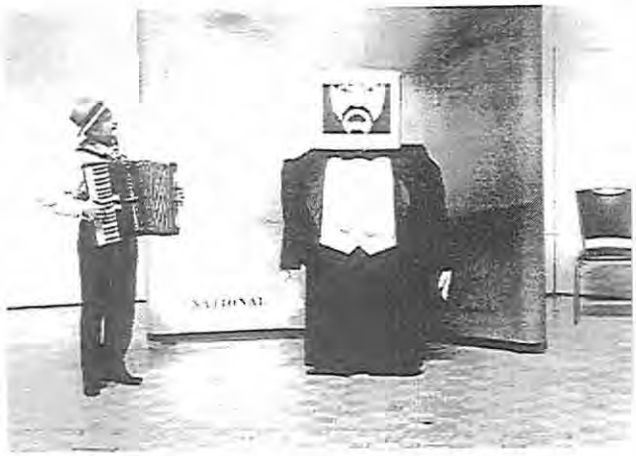


Figure 9. The author with his singing robot, "Pavarobotti".

Summary

Where does this take us in the future? I hope that simulated speech and song will never replace human soundings. Vocalizing, like general body exercise, is a necessary part of life. It is the best emotional and spiritual outlet we have. But with a simulator, we can explore, with little risk and cost to human health, the dimensions of many of our voices. We can explore these dimensions one at a time, slowly and carefully. Even if we take sizes and forces beyond their extremes, we cannot break anything. Furthermore, we can predict some effects of training, surgery, injury, and aging in a matter of minutes. Instantly, we can project a voice into the future, and perhaps recover its past. I have taught Pavarobotti how to sing, but it has also taught me how to sing. The real world and the virtual world will coexist, side by side, for the betterment of all of us.

Acknowledgment

This research was supported by grant # P60 DC00976 from the National Institutes on Deafness and Other Communication Disorders.

Bibliography

Davis, P., Zhang, S., Winkworth, A., and Bandler, R. (1996). Neural control of vocalization: Respiratory and emotional influences. Journal of Voice 10, 1, 23-38.

Fujimura, O. (1994). Tongue prostheses for articulation. Journal of the Acoustical Society of America, 95(5) [Abstract], 3013.

Isshiki, N., & Ishikawa, T. (1976). Diagnostic value of tomography in unilateral vocal cord paralysis. The Laryngoscope, LXXXVI(10), 1573-1578.

Klatt, D. (1987). Review of text-to-speech conversion for English. Journal of the Acoustical Society of America, 82, 737-793.

Lieberman, P. (1991). Uniquely Human: The Evolution of Speech, Thought, and Selfless Behavior. Cambridge MA: Harvard University Press.

Sataloff, R. (1992). The human voice. Scientific American, 108-115.

Sundberg, J. (1987). The Science of the Singing Voice. DeKalb, IL: Northern Illinois University Press.