unter

NCVS

National Center for Voice and Speech

University of Iowa • Denver Center for the Performing Arts
University of Wisconsin-Madison • University of Utah

Status and Progress Report

Volume 8/July 1995

Editorial and Distribution Information

Editor, Ingo Titze
Production Editors, Julie Lemke 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

NCVS Status and Progress Report
Volume 8/July 1995

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.

Part I

Research papers submitted for peer review in archival journals

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
Department of Preventive Medicine and Environmental Health

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

Memphis State University
Department of Audiology and Speech Pathology

The University of Illinois
Department of Speech and Hearing Science

NCVS Personnel

Administration

Central Office
Ingo Titze, Director
Erich Luschei, Deputy Director
Julie Ostrem, Program Assistant
Julie Lemke, Secretary

Area Coordinators
Research - Ingo Titze
Training - Erich Luschei
Continuing Education - Julie Ostrem
Dissemination - Cynthia Kintigh

Institutional Coordinators
University of Iowa - Ingo Titze
Denver Center for the Performing Arts - Ronald Scherer
University of Wisconsin-Madison - Diane Bless
University of Utah - Steven Gray

Investigators, Affiliates and Support Staff

Fariborz Alipour, Ph.D.
Patricia Benjamin, B.S.
Bridget Berning, B.S.
Florence Blager, Ph.D.
Diane Bless, Ph.D.
James Brandenburg, M.D.
Gayle Brosovic, B.A.
Geron Coale, M.A.
Stefanie Countryman, M.A.
Jeanne Delaney, B.A.
Heather Dove, M.A.
Christopher Dromey, Ph.D.
Wendy Edwards, B.A.
Kimberly Fisher, Ph.D.
Charles Ford, M.D.
Steven Gray, M.D.
Chwen-Geng Guo, M.S.
Elizabeth Hammond, M.D.
Susan Hensley, M.M.
Marilyn Hetzel, Ph.D.
Margaret Hoehn, M.D.
Henry Hoffman, M.D.
Heather Hughes
Kathy Ives
Bruce Jafek, M.D.
Antonia Johnson, M.A.
Karen Johnson, B.S.
Joel Kahane, Ph.D.
Judith King, Ph.D.
Cynthia Kintigh, M.A.
David Kuehn, Ph.D.
Jennifer Lehnherr
Jon Lemke, Ph.D.
Julie Lemke
Russel Long, M.S.
Erich Luschei, Ph.D.
Sharon Maclay, M.A.
Kathryn Maes, Ph.D.
Amy Mast, B.S.
Brian McCabe, M.D.
Arlen Meyers, M.D.
Martin Milder, B.S.
Paul Milenkovic, Ph.D.
Kenneth Moll, Ph.D.
Jerald Moon, Ph.D.
John Nichols, B.A.
Lorraine Olson Ramig, Ph.D.
Carrie Ocken, B.S.
Julie Ostrem., B.S.
Namrata Patil, M.D.
Annette Pawlas, M.A.
Kathe Perez, M.A.
Mark Peters
Donald Robin, Ph.D.
Robin Samlan, M.S.
Paula Saraceno
Ronald Scherer, Ph.D.
Elaine Smith, Ph.D.
Marshall Smith, M.D.
Nancy Pearl Solomon, Ph.D.
Brad Story, Ph.D.
Edie Swift, M.S.
Christina Taskoff, M.S.
Laetitia Thompson, Ph.D.
Ingo Titze, Ph.D.
Nancy Tye-Murray, Ph.D.
Vern Vail, B.S.
Katherine Verdolini, Ph.D.
Patricia Ward
Jennifer Weinstein, B.S.
Darrell Wong, Ph.D.
Raymond Wood, M.D.
Patricia Zebrowski, Ph.D.

Doctoral Students

Todd Brennan, M.S.
Eileen Finnegan, M.A.
Mark Leddy, M.S.
Phyllis Palmer, M.A.
Annie Ramos, M.S.
Nelson Roy, M.S.
Alice Smith, M.A.
Kenneth Tom, M.A.
Elease White, M.S.

Postdoctoral Fellows

Michael Bayerl, M.D.
David Berry, Ph.D.
Katsuhide Inagi, M.D.
Aliaa Khidr, Ph.D.
Robert Lange, Ph.D.
Kristin Larson, Ph.D.

Visiting Scholars

Hanspeter Herzel, Ph.D., Germany
Tzu-Yu Hsiao, M.D., Taiwan

Advisory Board

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

Contents

Editorial and Distribution Information....ii
Sponsorship....iii
NCVS Personnel....iv
Forward....viii

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

Velocity Distributions in Glottal Models
Fariborz Alipour, Ronald Scherer and James Knowles....1
Combined Simulation of Two-Dimensional Airflow and Vocal Fold Vibration
Fariborz Alipour and Ingo Titze....9
Bifurcations in Excised Larynx Experiments
David Berry, Hanspeter Herzel, Ingo Titze and Brad Story....15
Coupling of Neural and Mechanical Oscillators in Control of Pitch, Vibrato, and Tremor
Ingo Titze....25
Mechanisms of Jitter-Induced Shimmer in a Driven Model of Vocal Fold Vibration
Darrell Wong, Robert Lange, Ingo Titze and Chwen Geng Guo....33
Immunocytochemical Study of Proteoglycans in Vocal Folds
Agnieszka Pawlak, Thomas Hammond, Elizabeth Hammond and Steven Gray....43
The Intermediate Layer: A Morphologic Study of the Elastin and Hyaluronic
Acid Constituents of Normal Human Vocal Folds
Thomas Hammond, Ruixia Zhou, Elizabeth Hammond, Agnieska Pawlak and Steven Gray....49
Midbrain Regions for Vocalization Identified by Electrical Stimulation in Anesthetized Dogs
Kang Liu, Nancy Pearl Solomon and Erich Luschei....55
Sulcus Vocalis: A Rational Analytical Approach to Diagnosis and Management
Charles Ford, Katsuhide Inagi, Diane Bless, Aliaa Khidr and Kennedy Gilchrist....65
Autologous Collagen Vocal Fold Injection: A Preliminary Clinical Study
Charles Ford, Paul Staskowski and Diane Bless....75
Longitudinal Phonatory Characteristics After Botulinum Toxin Type A Injection
Kimberly Fisher, Ronald Scherer, Chwen Guo and Ann Owen....81
Electroglottographic Tracking of Phonatory Response to Botox
Kimberly Fisher, Ronald Scherer, Paul Swank, Cheryl Giddens and Dava Patten....97
Formant Trajectory and Segmental Characteristics of Males with Parkinson Disease
Antonia Johnson and Lorraine Olson Ramig....113
Electroglottographic Measures of Vocal Fold Adduction in Parkinson Disease:
Pre and Post Intensive Speech Treatment
Gayle Brosovic and Lorraine Olson Ramig....131
Sense of Effort and the Effects of Fatigue in the Tongue and Hand
Nancy Pearl Solomon, Donald Robin, Sara Mitchinson, Douglas VanDaele and Erich Luschei....139
Suitability of Minidisc (MD) Recordings for Voice Perturbation Analysis
William Winholtz, Ingo Titze and Robert Lange....153

Part II. Tutorial reports and Updates

Electromyographic Techniques for the Assessment of Motor Speech Disorders
Erich Luschei and Eileen Finnegan.....157
Training Update
Erich Luschei/Patricia Zebrowski.....175
Continuing Education Update
Julie Ostrem.....177
Dissemination of Information Update
Cynthia Kintigh.....181

Forward

This Eighth Status and Progress Report comes at the end of a busy summer that included a general NCVS Conference in Madison, Wisconsin, and a new set of aims and priorities for all investigators. We thank the National Institutes of Deafness and Other Communication Disorders for allowing us to continue in the framework of this research and training center for another five years. We also thank our peers who brought many helpful suggestions to the site visit last February.

As this report shows, many questions are still posed, and answers given, to the complex nature of vocal fold vibration. On the one hand, Alipour, Scherer and company are approaching these questions from a computational point of view, looking carefully at glottal airflow. On the other hand, Herzel, Berry and company are studying vibration patterns on excised larynges, and Gray, Hammond, and colleagues are digging deep to discover the corresponding tissue microstructure.

Clinical studies continue to highlight collagen injections, botulinum toxin injections, and tongue and hand fatigue assessment. For Parkinson's disease, formant trajectory and electroglottographic measures have been added to diagnostic and treatment approaches. Some success is also reported on sulcus vocalis, a disorder that has been poorly understood and managed.

Finally, we have included a tutorial on electromyographic techniques in Part II. Readers at any technical level should find this both informative and entertaining.

Ingo R. Titze, Director
July, 1995

NCVS Status and Progress Report - 8
July 1995, 1-8

Velocity Distributions in Glottal Models

Fariborz Alipour, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa
Ronald Scherer, Ph.D.
Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts
James Knowles, B.S.
Department of Speech Pathology and Audiology, The University of Iowa

Abstract

Understanding phonatory mechanism and acoustics requires the study of the air velocities within the larynx. Velocity distributions within three models of the human larynx, namely, a rigid plexiglas model, an excised canine larynx, and a computational model were investigated. A plexiglas wind tunnel with a glottal constriction was used as a two-dimensional steady flow model to measure velocity upstream and downstream of the glottis. Canine excised larynges were used as a prototype pulsatile flow model to study pressure and velocity variations during phonation. Results of the plexiglas modelling indicated a parabolic laminar velocity profile upstream of the glottal constriction and turbulent and asymmetric velocity profiles downstream of the glottal constriction. The time averaged velocities of the excised larynx had similarities to the plexiglas model results. Velocity instabilities and asymmetries were demonstrated by the computational method. The results suggest that both temporal and spatial velocity asymmetries exist in the glottal flow, and therefore the acoustics of phonation must take them into consideration.

Introduction

The vocal folds within the larynx oscillate during phonation to modulate the airflow from the lungs and create sound. The dynamic motion of the vocal folds depends on their tissue characteristics, their initial positions, and the air pressure distributions within the glottis (the space between the vocal folds). The air pressures on the surface of the vocal folds act as external driving forces to help displace them and dynamically change the shape of the glottis (1-3). Another driving force is the shear force resulting from air viscosity near the tissue surface. The amount of this force is determined by the velocity distribution within the glottis.

Since a vibratory cycle of the glottis will typically result in shapes that are convergent during glottal opening and divergent (or nearly parallel) during glottal closing (4), a number of experimental studies have reported information on the translaryngeal and intraglottal pressure and bulk airflows for the these glottal shapes (5-11). Analytic expressions have been devised in the attempt to predict these pressure drops (11,12), and some of these have been useful in computer models and theories of vocal-fold oscillations that include analogues of the tissue characteristics and acoustic ducts above and below the glottis (13,14).

The glottal flow rate (volume velocity) has been of primary interest in modelling studies of phonatory fluid mechanics because of the ease of its measurement and its usage in the one-dimensional analysis of airflow and acoustics in the larynx and vocal tract. The volume velocity is sometimes referred to as the source of sound in phonation, and is the primary output of the glottis for speech and modelling purposes, particularly in the electric analogs (linear circuits) of vocal tracts (1,14). Under the one-dimensional assumption of airflow in the larynx, the glottal flow rate is considered continuous and mean velocity may be obtained from the continuity of mass and cross-sectional area. However, this one-dimensional analysis suffers from the assumption of uniform velocity profiles in the larynx and vocal tract which has been challenged recently (15,16). Also, shear forces can only be obtained from nonuniform particle velocity distributions.

Since the particle velocity output is the more fundamental measure for the acoustic nature of sound production, any non-uniformities in instantaneous velocity distributions may be relevant to the resulting acoustic output. There have been few theoretical studies of particle velocities of the larynx (17-19), and experimental investigation of the velocity in glottal models is at the beginning stage.

The glottis is a 3-dimensional space. At any instant of time, the airflow within the glottis is not necessarily uniform across any specific cross section. In particular, when the glottis takes a diverging shape, the flow may move toward one of the boundaries, more or less clinging to the wall on that side, but separating from the wall on the other (17). This asymmetry of flow may also give rise to differences in the wall pressures on the two sides of the diffuser (the glottis). Because of the importance of the intraglottal pressures in controlling vocal fold motion, the flow profiles within the glottis need to be studied. Also, if the velocity profile within the glottis is asymmetric, the velocities just downstream of the glottis will also be asymmetric, and thus can be detected and measured with hot wire anemometry. Subtleties of acoustic variation dependent upon velocity distributions just downstream of the glottis depend upon knowledge of those velocity distributions, the topic of this report.

In this study complementary information on particle velocities are obtained within three models of the human larynx, namely, a rigid plexiglas model, excised canine larynx models, and a computational model. The purpose of this study was to obtain and compare velocity profiles upstream and downstream of the glottis within these models. Only the divergent glottis which occurs just prior to glottal closure was considered in the plexiglas and computational models

Methodology

In this section, various methodologies involving the plexiglas model, the excised larynx model, hot wire anemometry, and the computational model will be discussed.

Plexiglas Model

The plexiglas model was constructed with two primary parts, a relatively long duct that acts like a "wind tunnel", and the glottal orifice made of vocal fold pairs and shims that create realistic glottal shapes and diameters.

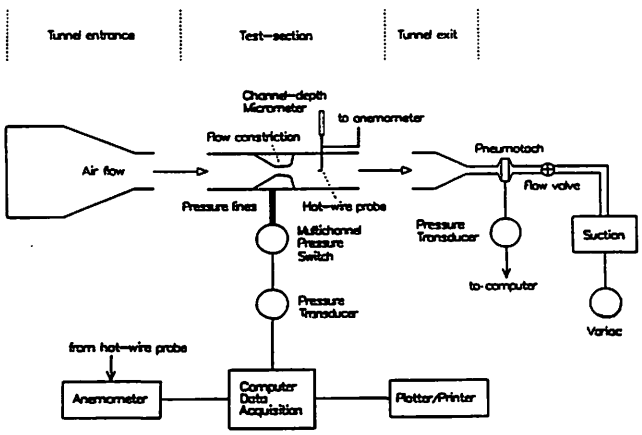


Figure 1. Schematic of the plexiglas wind tunnel model.

The plexiglas wind tunnel (Figure 1) was created by the Iowa Institute of Hydraulic Research at the University of Iowa. The length of the tunnel was approximately 2 meters, with an internal height of approximately 25.2 mm and an internal width of 152 mm, for an aspect ratio of 6:1. This aspect ratio can be decreased by insertion of styrofoam walls (down to 1:1). The tunnel inlet consisted of an extended aluminum honey comb and wire mesh assembly 15.5 cm long followed by a gradual contraction. This construction was used to laminarize the flow. The downstream extension contracted over an extent of 32 cm and attached to a suction source (a 4 hp Power Tank sweeper) to draw the air. The output of the sweeper was controlled by a Variable Autotransformer (Staco Energy Products, type 3PN1010V), with an intervening valve for fine flow control. A pneumotach (F.E. Wiedeman & Sons double core) was used with a Validyne DP103-10 pressure transducer to record the bulk flow through the model.

The glottal flow constriction (Figure 1) was created by the Medical Instrument Shop of the University of Iowa. This section consisted of removable plexiglas plates on the wall of the wind tunnel to which were attached aluminum vocal fold pieces and separation shims. The length of the vocal fold pieces were the same as the width of the flow tunnel (152 mm), creating essentially two-dimensional flow conditions with a glottal length of 152 mm.

Brass and aluminum shims ran the width of the tunnel and allowed a range of glottal gaps (distance between the two constriction pieces). After the fitting of the glottal pieces with the prescribed glottal gap, the gap was measured using feeler gages along the length of the gap. In this study, a uniform glottal gap with diameter of 0.6 mm ± 0.01 mm was used.

Excised Larynx Model

Excised canine larynges acquired from the University of Iowa Hospitals and Clinics were kept in saline solution prior to experimentation. The larynges were trimmed and the false vocal folds removed to expose the true vocal folds. Air of 100% humidity and 37 degrees Celsius was achieved using a Concha-Therm III Servo Control Heater unit (RCI Laboratories). The larynges were mounted on a pseudo-tracheal rigid tube (i.d. of 17.5 mm). The glottis was easily viewed by a camera and accessible by the equipment (Figure 2). Adduction and tension controls were established by connecting cartilages to micrometers with sutures (20).

The mean pressure in the subglottal region 7 cm below the glottis was monitored with a wall mounted manometer (Dwyer No. 1230-8). The mean flow rate was monitored with an in-line flowmeter (Gilmont rotameter model J197) upstream of the Concha-Therm III device. The time varying subglottal pressure was measured with a piezoresistive pressure sensor (Microswitch 136PC01G1) at the same location as the manometer tap. Bandwidth for the pressure transducer was approximately 0-800 Hz.

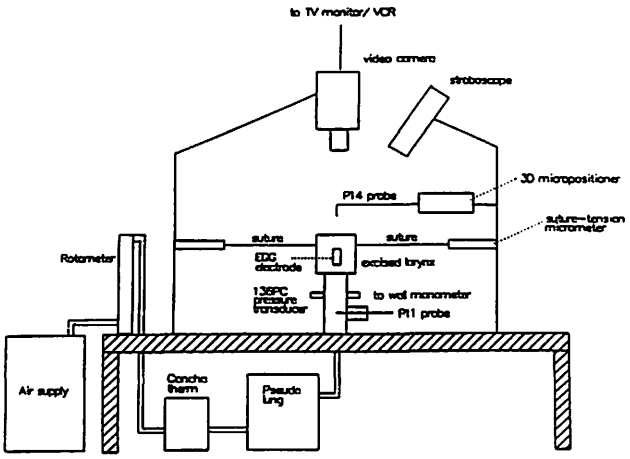


Figure 2. Schematic of the excised larynx experimental setup.

The Dantec 55P11 hot wire probe was placed in the subglottal section 4.5 cm below the pressure tap location, and the 55P14 probe was placed above the glottis over the midline at a height of approximately 10 mm.

During the experiment, analog data from the hot-wire probes, EGG (Synchrovoice), and the pressure transducers were monitored on a digital oscilloscope (Data Precision Data 6000) and simultaneously recorded on a Sony model PC-108M Digital Audio Tape (DAT) recorder. These data were digitized later with a 16-bit analog/ digital convertor and analyzed on VAX-station computers.

Hot Wire Anemometry

Particle velocities within the plexiglas model, in the inlet pipe of the excised larynx, and in the exiting jet of the excised glottis were measured with a constant temperature hot-wire anemometer system (Dantec 56C01). In the plexiglas model the hot wire probe (Dantec 55P14) was mounted either upstream or downstream of the glottal region for experimental recordings. The upstream location was 20.2 cm from the glottis (from the minimum cross-section), and the downstream location was 2.1 cm from the glottis outlet. The probe tip was traversed in the upstream and downstream channels to transduce the velocity profiles transverse to the glottal slit. The hot-wire signal was calibrated at experimental temperatures and expected bulk flows by placing a pitot tube in the center section of the upstream measurement location, and calibrating the resulting velocity estimates with the hot wire system outputs for the same bulk flows. The expected range of bulk flows were used for this calibration. The best fit for velocity calibration was obtained with a least-square polynomial model.

Calibration of the hot wire probe for the excised larynx setup was performed by placing both a pitot tube and subsequently the hot wire probe in the center of the jet exiting a 0.6 cm uniform tube at the glottal bench attachment location. This allowed calibration in steady flow at experimental temperature and humidity conditions for the typical range of velocities.

Computational Method

To obtain a detailed description of the flow field, steady flow through divergent glottal constriction models was solved numerically. A commercially available package (FIDAP[1]) based on the finite element technique was used in computational fluid dynamic (CFD) modelling. The major task in using any CFD package is preparing the input data, which includes grid point coordinates, element topology, and boundary conditions. A total of 1000 elements and 4141 nodes were used in the computation. Based on the wall coordinates and size of the gap, a data file was prepared that included information on the keypoints on the boundaries, boundary conditions, mesh density and the element distributions. The wall coordinates were obtained from the vocal fold pieces in the plexiglas model.

Computations were performed on the NCSA CRAY Y-MP supercomputer. A nine-node quadrilateral element was used with higher mesh density near the constriction. An iterative method with successive substitution was used to solve the nonlinear Navier-Stokes equations. The Stokes solution was used for initial iteration, and convergence was set based on the maximum velocity residual of 1%. As the Reynolds number increased to 100 for the upstream channel (corresponding to 200 based on the hydraulic diameter of the glottis), the convergence took longer, particularly for narrower glottal gaps. The two control parameters, Reynolds number and glottal gap ratio, both had major influence on the convergence and outcome of the solution which will be discussed later.

Results and Discussion

Results will be discussed relative to the subglottal and supraglottal velocity profiles for the plexiglas, excised and computational models.

Velocities in the Plexiglas Model

Figure 3 shows velocity profiles in the upstream section of the plexiglas model. The distance between the walls was 25.2 mm, and the closest hot wire probe placement was about 0.6 mm from either bottom or top of the tunnel. The six profiles shown represent Reynolds numbers (based on the upstream dimensions) ranging from 1000 to 2000. As Reynolds number increased, the profiles expectedly altered from a parabolic shape to a shape with a relatively flat portion near midline. This is due to the fact that at higher Reynolds number the tunnel length required to achieve the parabolic profile is longer. There is some minor asymmetry noted for the intermediate Reynolds number profiles. The velocity signal from the hot wire probe was quite stable (noise free) in the upstream section of the model.

1 Fluid Dynamics International, Inc., 500 Davis St., Suite 600, Evanston, IL 60201

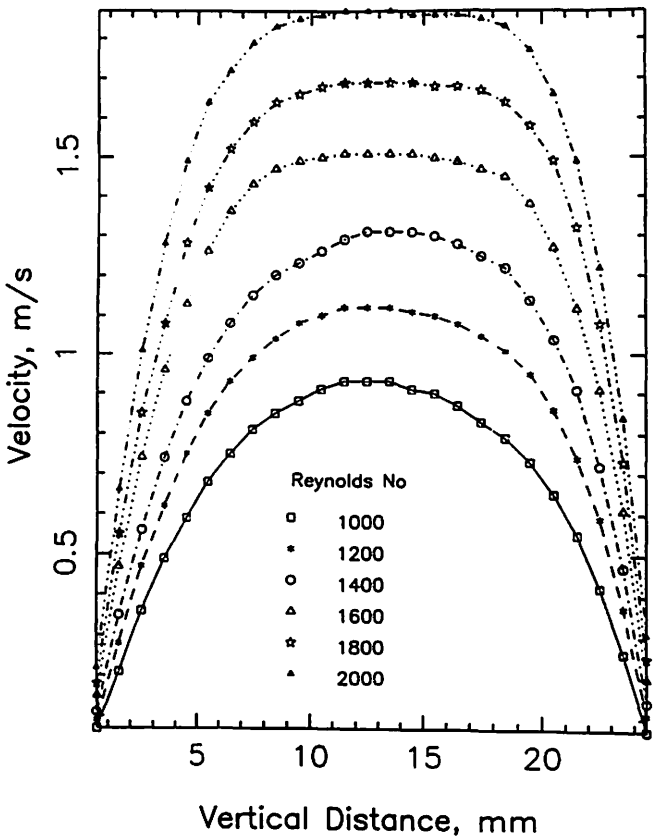


Figure 3. Velocity distributions upstream of the plexiglas model.

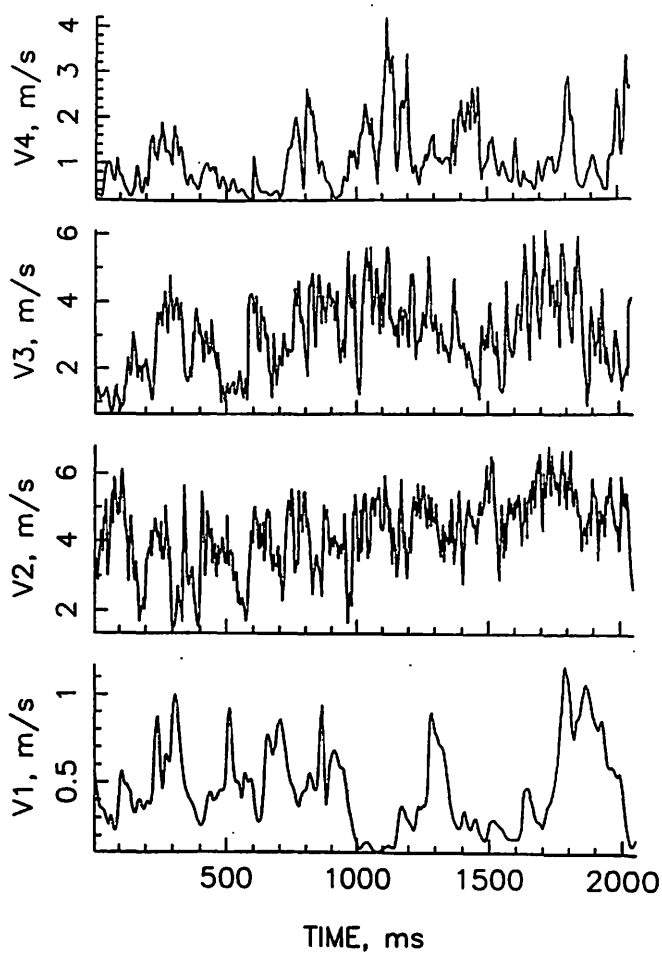


Figure 4. Velocity waveforms recorded 2.1 cm downstream of the plexiglas glottis model with hot-wire positions of 1, 2.5, 4, and 7 mm from the wall (traces from bottom to top, respectively). The glottis shape was divergent; glottal diameter was 0.6 mm.

Unlike the upstream, the downstream velocity signal was changing with time at almost any location. Thus the hot-wire signal was recorded on a microcomputer with a digitizing board. Figure 4 shows velocity time waveforms recorded at four positions 2.1 cm downstream of the glottis. The four waveforms (V1, V2, V3, and V4) correspond to transverse locations that are approximately 1, 2.5, 4.0, and 7.0 mm from the upper wall, respectively, where 7 mm is 28% of the channel height. Turbulence is noted for all four locations, with the waveform V2 showing the largest absolute velocity values, and V3 showing the next highest values. The location closest to the wall shows the least amount of turbulence and the smallest absolute values. In this example, the glottal shape was divergent with an included angle of 30 degrees, and a minimal gap of 0.6 mm. The Reynolds number at the glottal minimal gap, using the hydraulic diameter, was approximately 2000 which corresponds to a Reynolds number of 1000 in the upstream section.

Figure 5 shows the average velocities for a two-second interval at various vertical probe locations in the channel. The profiles are transverse to the constriction gap and were measured 1.70 cm downstream of the glottal exit. The bottom of the channel corresponds to the right on the horizontal axis. The non-uniformity of the velocity values just discussed can be seen in this figure. This asymmetry was probably caused by the instability of the jet. The jet flows are very unstable and the critical Reynolds number for them could be as low as 4 (21). The instability caused flow to cling to one side of the glottal duct, and separate from the other, causing flow to be faster toward one side of the diffuser exit. Further, the instability caused a time dependent (turbulent) flow downstream of the constriction as was demonstrated in Figure 4. This will be demonstrated more clearly in the computational modelling below.

Velocities in Excised Larynx Model

Figure 6 shows a portion of the time varying signals for an excised canine larynx (male, 27 kg, 18.0 mm vocal fold length). In this figure the EGG signal, subglottal pressure (Ps), and subglottal velocity (Vs) for three different probe positions and jet velocity (Vj), are shown. The case corresponds to a mean flow rate of 470 ml/s. About four cycles are shown. The EGG signal waveshape suggests that there was normal vocal fold contact during each period. The frequency spectrum examination of the Ps waveform indicates the existence of the first three partials in the waveshape with a fundamental frequency of 111.3 Hz. The three

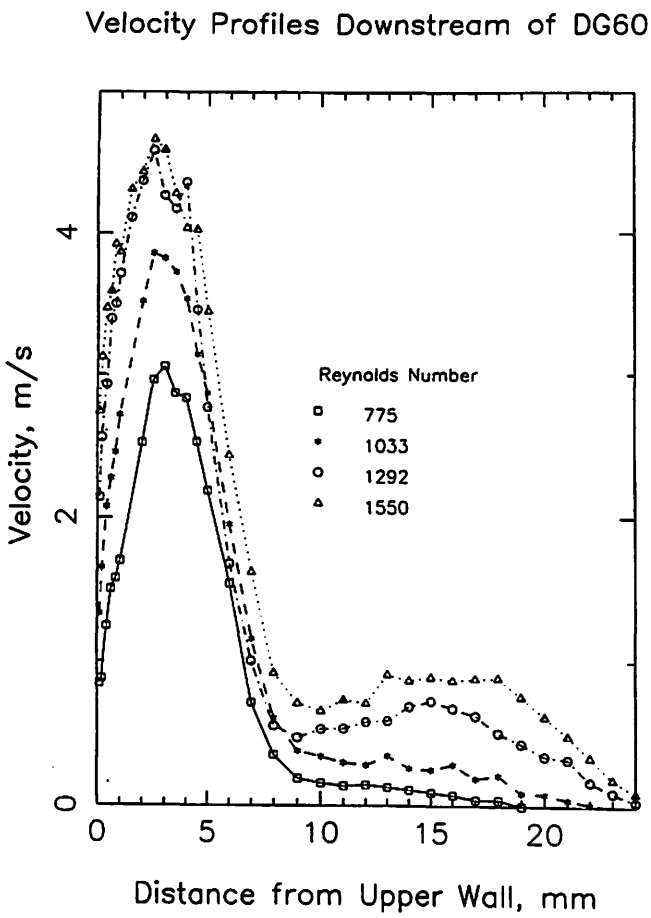


Figure 5. Velocity profiles downstream of the plexiglas model with divergent glottis and 0.6 mm gap.

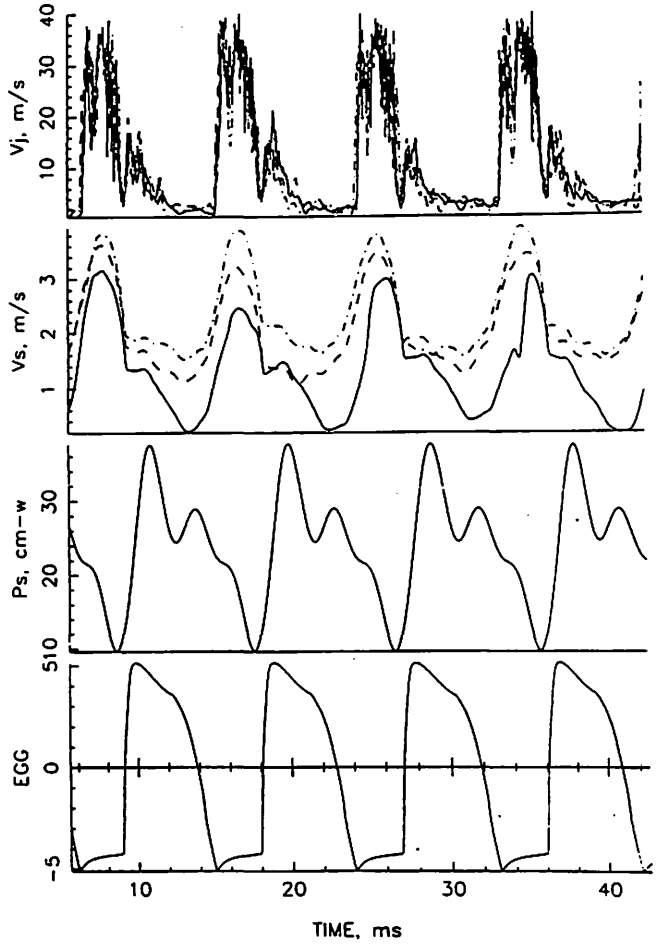


Figure 6. The waveforms of a typical excised larynx result. The signals are: electroglottograph EGG, subglottal pressure Ps, velocity in the trachea Vs, and supraglottal jet velocity Vj.

subglottal velocity waveforms were taken at the 2, 4 and 8 millimeter locations from the tracheal wall within approximately a one minute interval. Although these waveforms were not recorded simultaneously, the three subglottal velocity profiles are similar in shape, with fluctuations in the velocity waveform corresponding to the tracheal pressure changes (20). The flow fluctuations are not completely (phase) synchronized, however, suggesting different spatial effects. It is expected that the velocity measured closest to the wall (solid lines) might have greater viscous effects than the other two locations. The discrepancies within the subglottal velocity profiles across the same cross section warrants deeper investigation. The output velocities Vj will be discussed below.

Figure 7 (following page) shows velocity profiles at the subglottal section 11.5 cm below the glottis. The profiles are foreshortened on the right side of the figure because of cautious placement slightly away from the wall on that side. The subglottal tubing in the hot-wire section was about 20 mm in diameter. Ten cycles of each subglottal velocity waveform were averaged at a location of 0.6 cm; the average waveform is plotted under the profiles. Seven locations on the cycles (1 and 7 at the two ends when flow is minimum, and 4 at the peak) were marked on each waveform and extracted data were plotted as a function of distance from one wall. The velocity profile changes in shape and magnitude with time and has some inconsistent asymmetry and skewness to the side. The average waveform is similar to the glottal volume velocity with skewness to the right.

Figure 8 (following page) shows four profiles for time averaged velocities at the subglottal cross section in the excised larynx setup for four different bulk flows. The data were obtained from the same larynx that was used for Figures 6 and 7. The lowest flow rate of 292 ml/s (100Hz, 13.0 cm H_2O subglottal pressure) had a relatively flat average velocity profile compared to the other higher flow conditions. The subglottal setup included approximately 3 meters of helical plastic tubing to simulate a pseudo lung, attached to one foot of straight tubing with imbedded honey comb to recondition the air particle velocity for the hot wire velocity measurement. The other three profiles correspond to consecutively increased mean flow rates (470, 558, and 680 ml/s, respectively), with corresponding frequencies of 111, 166 and 146 Hz, and subglottal pressures of 21.0, 13.2,

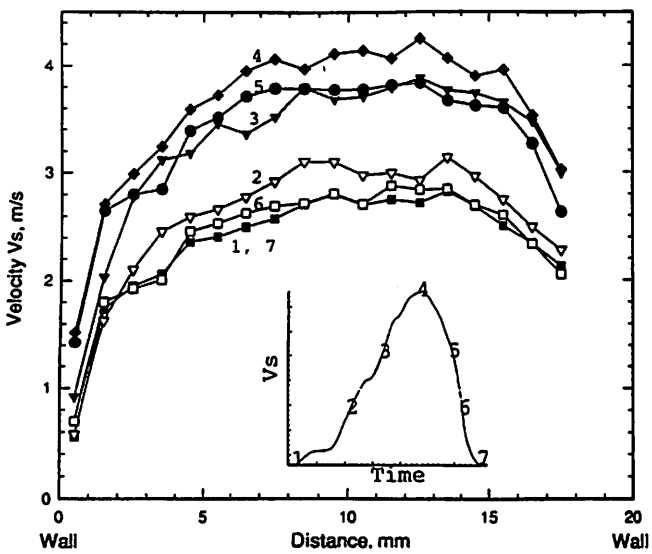


Figure 7. Velocity profiles in subglottal section of the excised model. These velocity profiles were selected at different phase of the oscillation cycle.

and 18.1 cm H_2O. Thus, these cases refer to different frequency, subglottal pressure, and adduction conditions. The four curves show increasing slopes of velocity change away from (but near to) the wall, which is consistent with shear stress dependence on the frequency and mean flow for oscillating flow (21).

The supraglottal jet velocity (Vj) measured 10 mm above and on the midline of the glottis is shown Figure 6. The figure shows significant turbulence in the flow. The flow undoubtedly exits the glottis in the form of a relatively free field jet, giving rise to turbulence. The shape of the velocity waveform depends on the location of the probe (20), and may be quite different from the subglottal velocity shape. It is noted that there was no supraglottal duct to create a significant resonance structure that might have affected the particle velocity waveshape.

Figure 9 shows time averaged and maximum particle velocities taken transversely across the glottis. The data for this figure were obtained from a male canine larynx of 29 kg with a vocal fold length of 18.1 mm. The flow was 510 ml/s, fundamental frequency of 235 Hz, and a subglottal pressure of 14.0 cm H_2O. The peaks of the maximum and time averaged particle flows were offset to the right, and thus did not occur in the midline of this glottis. This suggests that the glottal jet was displaced due to either asymmetric motion of the vocal folds or asymmetric aerodynamics of the glottis. Another possibility might be due to slightly asymmetric mounting of the larynx on the bench or deviation of the supporting tracheal tube from vertical position. Although stroboscopy helps, it is not completely clear in this case if the glottal motion is suspected. The aerodynamic possibility suggests that the divergence of the glottal shape near maximum flow and at the beginning of closure may influence this asymmetric output velocity seen in Figure 9. The data in the

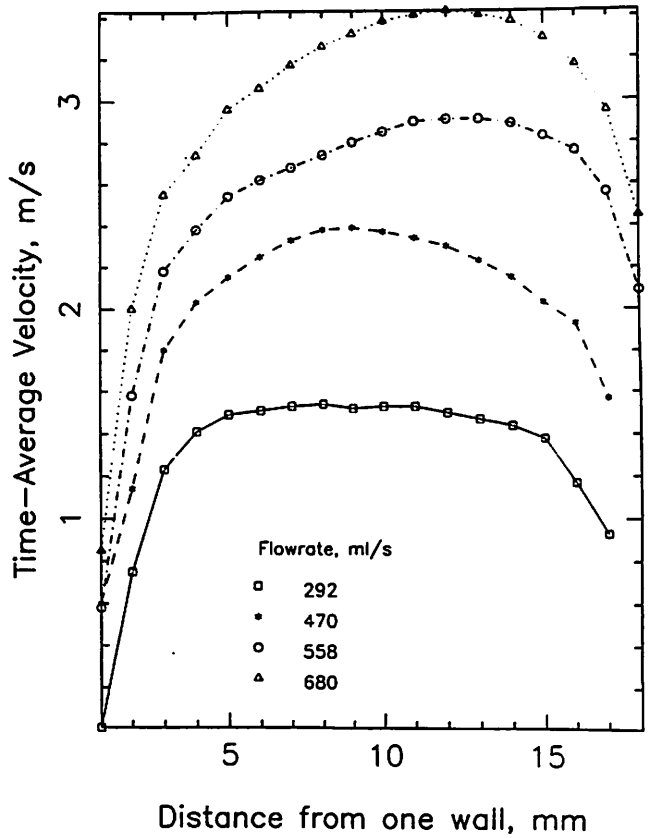


Figure 8. Velocity profiles upstream of excised model. Data were obtained from time-averaging of velocity waveforms measured in the subglottal section at various flow rates.

excised larynx study show that both temporal and spatial velocity asymmetries in the glottal flow are possible.

Velocities in Computational Model

Figure 10 shows the computational results for a divergent glottis model at a Reynolds number of 50 and a gap size of 1 mm. In the top figure, velocity profiles within the glottal region are enlarged to provide greater detail. Parabolic velocity profiles enter the glottis with increasing values, and separate near the minimum glottal cross sectional area. The flow then exits as a jet which maintains its speed for a certain length and gradually spreads towards the walls. At both sides of the jet, vortices are formed and circulate in opposite directions. The shape of the profiles and the jet are symmetric for low Reynolds numbers and wider gaps. As Reynolds number increases or as gap size decreases, this symmetry may not hold. As seen in Figure 11, at the top the glottal flow pattern for the divergent model with a gap size of 0.73 mm and Reynolds number of 100 is symmetric, but at the bottom, the flow pattern for the same Reynolds number and a smaller gap size of 0.32 mm has become asymmetric and the jet attaches to the wall. The flow in the vocal tract past the glottis model becomes so unstable that ordinary steady flow laminar CFD codes are

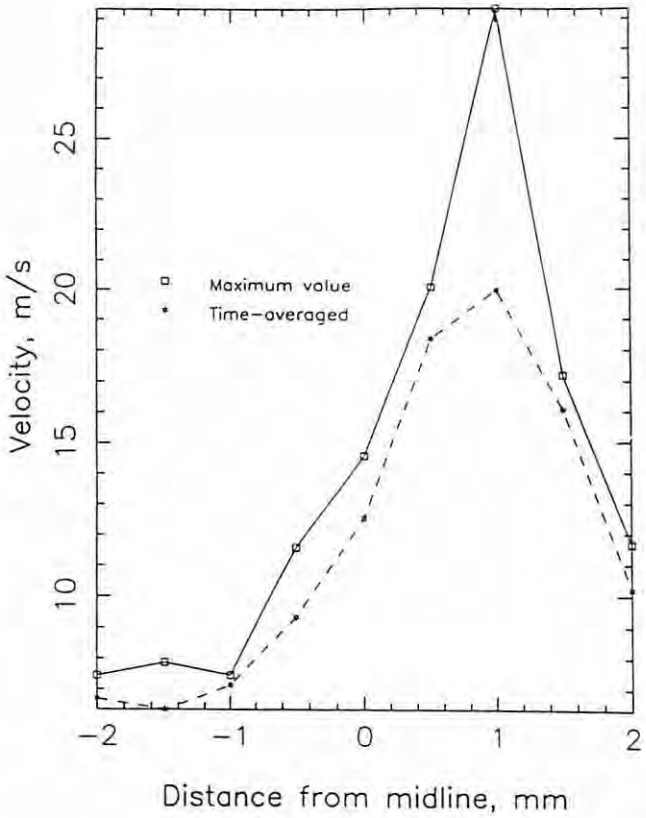


Figure 9. Velocity profile in supraglottal jet of excised model. Data were obtained from waveforms measured 10 mm above the glottis. Measurement were taken transverse to glottal slit.

unstable that ordinary steady flow laminar CFD codes are not capable of resolving any flow details.

Concluding Remarks

The results of this study have emphasized asymmetric velocity profiles obtained from the different glottal models. The steady flow experiment with the plexiglas glottal model clearly demonstrated turbulence and flow instability downstream of the glottal constriction. The flow upstream was laminar and predictable. The exiting jet from the plexiglas glottal model may flow to one side and create an uneven velocity profile. The glottal jet in the excised larynx is also turbulent and distributed unevenly across the glottis. The findings for the computational model are consistent with the experimental data. These findings suggest the necessity of assessing the acoustic effect due to flow asymmetries within the vocal tract and also the importance of turbulence modelling in future analyses of laryngeal flow.

Acknowledgments

The authors would like to thank Dr. Ingo Titze for helpful comments and the National Institute on Deafness and Other Communication Disorders for financial support of this work under Grant No. DC00831-03, and the National Center for Supercomputing Applications (NCSA) for granting the computing time

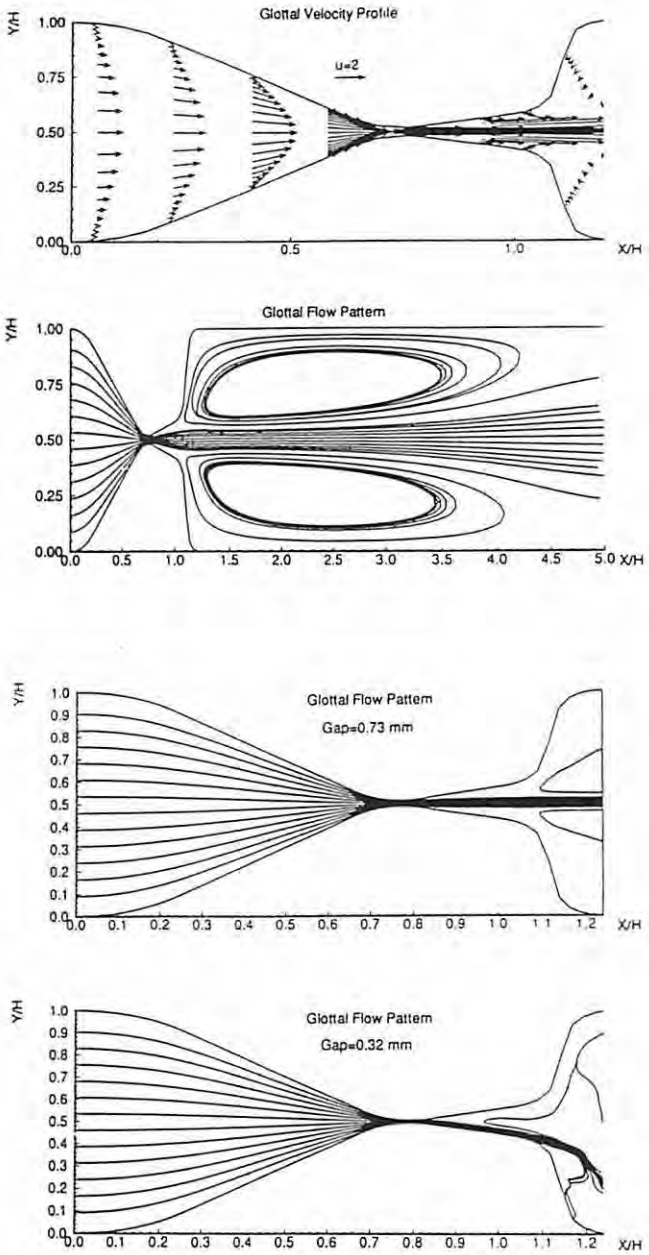


Figure 10 (top). Velocity profiles and flow pattern in a divergent computational model at low Reynolds number. Figure 11 (bottom). Flow pattern in a divergent computational model at Reynolds number of 100 and two gap sizes of 0.73 and 0.32 mm.

References

1. Ishizaka K, Flanagan JL. Synthesis of voiced sounds from a two-mass model of the vocal cords. *Bell Syst. Tech. J.* 1972;51:1233-1268.

2. Titze IR, Talkin DT. A Theoretical Study of the Effects of Various Laryngeal Configuration on the Acoustics of Phonation. *J. Acoust Soc. Am.*, 1979; 66(1):60-74.

3. Alipour F, Titze IR. A finite element simulation of vocal fold vibration. In: LaCourse JR, ed. *Proceedings of the Fourteenth Annual Northeast Bioengineering Conference.* 1988:186-189.

4. Hirano M. Structure and Vibratory Behavior of the Vocal Fold. In: Sawashima M, Cooper FS eds. *Dynamic Aspects of Speech Production.* University of Tokyo Press, Tokyo 1977:13-30.

5. van den Berg Jw, Zantema JT, Doornenbal PJr. On the air and the Bernoulli effect of the human larynx. *J. Acoust. Soc. Am.*, 1957;29(5):626-631.

6. Scherer RC. *Laryngeal Fluid Mechanics: Steady Flow Considerations Using Static Models.* Ph.D. thesis, University of Iowa, Iowa City 1981.

7. Scherer RC, Titze IR. Pressure-flow relationships in a model of the laryngeal airway with a diverging glottis. In: Bless DM, Abbs JH, eds. *Vocal Fold Physiology: Contemporary Research and Clinical Issues.* College-Hill Press, San Diego 1983:179-193.

8. Scherer RC, Titze IR, Curtis JF. Pressure-flow relationships in two models of the larynx having rectangular glottal shapes. *J. Acoust. Soc. Am.*, 1983;73(2):668-676.

9. Binh N, Gauffin J. Aerodynamic measurements in an enlarged static laryngeal model *Technical Phoniatrics.* STL- QPSR 1983.

10. Scherer RC, Guo CG. Laryngeal modeling: translaryngeal pressure for a model with many glottal shapes. In: *Proceedings of the 1990 International Conference on Spoken Language Processing*, The Acoustical Society of Japan, Japan, 1990;1:3.1.1 - 3.1.4.

11. Scherer RC, Guo CG. Generalized translaryngeal pressure coefficient for a wide range of laryngeal configurations. In: Gauffin J, Hammarberg B, eds. *Vocal Fold Physiology: Acoustic, Perceptual, and Physiological Aspects of Voice Mechanisms*, Singular Publishing Group, San Diego, 1991:83-90.

12. Scherer RC, Guo CG. Effect of vocal fold radii on pressure distributions in the glottis. *J. Acoust. Soc. Am.* Suppl. 1, 1990;88:S150.

13. Ishizaka K, Matsudaira M. Fluid mechanical considerations of vocal cord vibration. *SCRL-Monograph.* Speech Communication Research laboratory, Santa Barbara 1972;8.

14. Titze IR. The physics of small-amplitude oscillation of the vocal folds. *J. Acoust Soc. Am.*, 1988;83(4):1536-1552.

15. Alipour F, Patel VC. Numerical simulation of laryngeal flow. In: Vanderby R. ed. *1991 Advances in Bioengineering.* BED-Vol. 20, 1991:111-114.

16. Alipour F, Patel VC. Steady flow through modeled glottal constriction. In: Meghdari A, ed. *Proceedings of the Intternational Conference on Engineering Application of Mechanics.* 1992;1:93-100.

17. Liljencrants J. Numerical simulation of glottal flow. In: Gauffin J, Hammarberg B, eds. *Vocal Fold Physiology: Acoustics, Perception, and Physiological Aspects of Voice Mechanisms.* Singular Publishing Group Inc., San Diego 1991:99-104.

18. Iijima H, Miki N, Nagai N. Glottal impedance based on a finite element analysis of two-dimensional unsteady viscous flow in a static glottis. *IEEE Trans. on Signal Processing.* 1992;40(9):2125-2135.

19. Guo CG, Scherer RC. Finite element simulation of glottal flow and pressure. *J. Acoust. Soc. Am.* 1993;94(2):688-700.

20. Alipour F, Scherer RC. Pulsatile airflow during phonation: An excised larynx model. *J. Acoust Soc. Am.* (in press).

21. White FM. *Viscous Fluid Flow.* McGraw-Hill, New York 1974.

NCVS Status and Progress Report - 8
July 1995, 9-14

Combined Simulation of Two-Dimensional Airflow and Vocal Fold Vibration

Fariborz Alipour, Ph.D.
Ingo Titze, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa

Abstract

Self-sustained oscillation of the vocal folds was simulated by combining two-dimensional laryngeal airflow with multi-dimensional tissue movement. A finite-element model was used for the solution of viscoelastic waves in the tissue and a finite volume method was used in the solution of Navier-Stokes equations for the airflow. A so-called 'shadow method' simulated the glottal constriction in the flow model to avoid the complexity of grid movement. The two-dimensional flow equations were solved in an iterative manner until the given transglottal pressure was approximated. The flow solution was then used in the estimation of the aerodynamic forces on the tissue, required in the finite element solution of tissue movement. The results indicate that inlet velocity profiles to the glottis are almost parabolic at any instant of time. The glottal velocity increases to its maximum at glottal exit at the center of a jet, with values exceeding 40 m/s for 0.8 kPa of lung pressure. The jet velocity waveform is similar to that of an excised larynx and the pressure profiles are similar to those of steady flows in physical models. Also, the displacement of the inferior portion of the vocal folds leads the superior position in phase.

Introduction

Voice research has benefitted vastly from technological advancements in the last few decades. Among those advancements are the new generations of workstations that, combined with computer networks put computing power for simulations of complex physical systems within reach of almost every researcher. One particular form of the speech simulation that received a lot of attention in recent years is the biophysical modeling of voice production. In such a models, physical laws of tissue and air movement are involved in every aspect of voice production. Although the models initially may include a set of rules built on empirical relationships obtained from a speaker, gradual refinement and validation ultimately favors physical laws over speaker specific rules.

A biophysical model of voice production requires at least three submodels. These submodels describe tissue mechanics, laryngeal aerodynamics, and vocal-tract acoustics. Tissue mechanics dictates the vibrations of the vocal folds and provides the geometry of the glottis, displacements, and stress and strain distributions within the vocal folds at every instant of time. Laryngeal aerodynamics handles the calculations of pressure and velocity fields within the glottis and the estimation of the external forces on the vocal fold tissues. Vocal tract acoustics brings about the acoustic pressure field within the vocal tract and radiation from the mouth, nose, and skin surface. These submodels can sometimes work independently, but generally need to work interactively. The complete biophysical model involves the integration of these submodels to run simultaneously.

Tissue mechanics has evolved from discrete low dimensional models [1,2] to a finite element method (FEM) continuum model [3]. Among the advantages of continuum models are greater accuracy and ability to incorporate the effects of complex tissue properties, such as anisotropy, inhomogeneity, and mixed boundary conditions. Also, continuum models provide a more detailed description of the movement of fleshpoints that facilitate visual presentation of vocal fold oscillations (animation). The price paid for these benefits is an increase in the computation time and the need to specify more parameters.

Laryngeal aerodynamics can be the most complicated part of the biophysical modeling process. This complexity arises from the nonlinearity of the governing equations and the boundary interactions between tissue movement and airflow at oscillation rates of over 100 Hz. Many

simplified steady flow models have been used to extrapolate to pulsatile flow. Empirical pressure-flow relationships obtained from experimental works [4-7] have been used in computer simulations of speech production with some limitations and caution. Computational studies of airflow through the static models of glottis [8-12] have reported plausible flow patterns and pressure distributions within the glottis, but these results still need to be extrapolated to pulsatile flow.

Pulsatile flow in the larynx has been studied experimentally using the technique of hot-wire anemometry [13-15]. These studies demonstrated the nonuniformity of the glottal jet and the possibility of turbulence above the glottis. The detailed descriptions of pressure and velocity distributions in the subglottal and supraglottal sections of the larynx given by Alipour and Scherer [14] are very useful in the understanding of laryngeal aerodynamics, but they are not easy to implement in biophysical simulations. Thus, it seems necessary to model the pulsatile flow computationally.

To answer this need, a computer program has been developed that solves the pulsatile flow with simultaneous wall movement. This program was first used and tested in a forced-oscillation paradigm [16], but will now be applied to this self-oscillation study. Our goal is to demonstrate that the solution of Navier-Stokes equations for airflow, combined with the finite element solution of tissue displacement, can produce self-sustained oscillations.

Modeling Procedure

Initial Setup

The modeling procedure starts with the setup of the initial configuration for the glottis and the build-up of lung pressure. The prephonatory glottal configuration requires an adjustment of glottal adduction, tissue stress, and vocal fold geometry. Once the configuration is set, the equilibrium geometry of the tissue remains constant. Both the flow domain and the tissue domain are then discretized into a finite number of elements for numerical solutions.

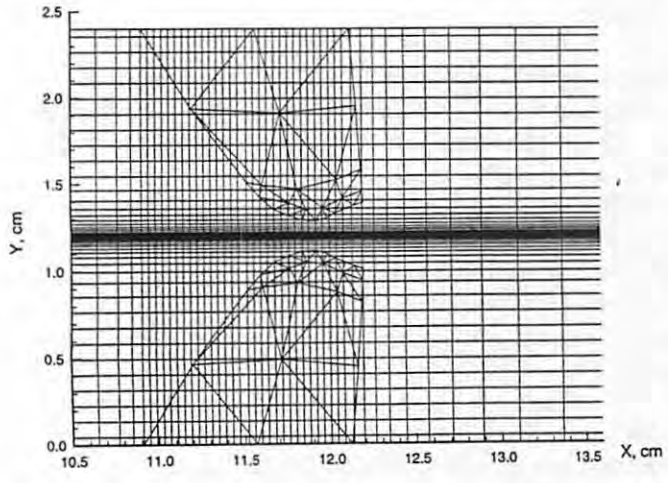


Figure 1. Combined flow and tissue domain grids.

Computational Grids

The computational grids include a tissue grid and a flow grid. The choice of a grid is crucial in any computational fluid dynamic analysis. Since the flow domain is assumed to be two-dimensional, a nonuniform rectangular grid was selected such that regions of higher velocity and larger pressure gradient contain more grid points. Figure 1 shows a portion of the grid for the flow domain. Since the glottal gap changes during each cycle of oscillation, a logarithmic distribution of grids was designed to ensure the presence of a number of grid points in the region near closure. The glottis that is used in the two-dimensional flow is calculated by averaging the glottal gap in the longitudinal direction (along the vocal fold length).

The tissue domain is three-dimensional, anisotropic, and inhomogeneous. The best method that handles such a domain is the finite-element method. To avoid the complexity of three-dimensional elements, a pseudo three-dimensional method was used in which two-dimensional tissue layers along the vocal fold length are coupled with elastic strings (fibers). The more detailed analysis of tissue mechanics and formulation of the equation of the tissue motion with finite elements are given by Alipour and Titze [3] and in a forthcoming book by Titze and Alipour [17]. For this discussion it suffices to say that a two-dimensional triangular finite element mesh pattern is used in all layers as shown in Fig. 1. This pattern is used in automatic mesh generation at the initial setup. The finer tissue elements are distributed in the regions of the larger displacements for greater resolution and accuracy.

Numerical Solutions

The solutions of tissue mechanics, aerodynamics, and vocal tract acoustics are obtained simultaneously. At every time step (50 μs intervals), the airflow solution is first carried out, then the tissue mechanics is solved, followed by the solution of wave propagation in the vocal tract. These solutions exchange data at every time step, and there is a short start-up period (transient period) before the periodic solution obtained. The solution of wave propagation is based on the wave reflection method [18,19] and will not be discussed here.

The governing equations for airflow are the continuity law and the Navier-Stokes (N-S) equations for unsteady two-dimensional incompressible laminar flow [9,12]. These equations are solved with a finite volume (control volume) method with staggered grids. In this method the Navier-Stokes equations are integrated numerically over every finite control volume of the flow domain, resulting in a set of algebraic equations to be solved. The N-S equations

are elliptic and can be solved only with a specified velocity at inlet. This necessitate the knowledge of flow rate *a priori*. However, in the biophysical model the transglottal pressure is known and flow rate is to be solved for. Thus an iterative method is used that solves N-S equations with a given transglottal pressure. In this method, a Bernoulli solution is first used to approximate the initial flow rate and the N-S equations are then solved; transglottal pressure drop is compared with given values and inlet flow rate is modified iteratively until the pressure profile matches the given transglottal pressure. The details of the numerical solution are given in Alipour, Fan and Scherer [16].

The governing equations for tissue mechanics are cast into a variational form and integrated in space with the finite element method. This results in a system of linear second order differential equations with nodal displacement vector as dependent variables and time as independent variable. The time integration is performed with a Crank-Nicholson finite difference method and nodal displacements are obtained at every time step.

Three different tissue materials were considered in the model of vocal folds, corresponding to the vocal fold muscle, a non-muscular mucosa and the vocal ligament. As mentioned before the vocal fold was divided into nine thin layers parallel to the coronal plane, such that a two-dimensional solution could be applied to each layer. Horizontal and vertical displacements of nodal points were calculated as a function of time. The coordinates of each nodal point is updated using its equilibrium coordinates and its displacement vector.

Boundary Conditions

The Navier-Stokes equations were solved with the appropriate pressure and velocity boundary conditions at every time step. These included inlet and outlet conditions and the no-slip condition on the walls. The pressure boundary condition is generally met by specifying the inlet and outlet flux. Due to the motion of the glottal wall, however, the no slip condition on the this wall was replaced with the moving wall boundary conditions (equating the velocity of the air and tissue at every wall grid point). For the outlet boundary condition (60 cm from the glottal exit), a fully developed flow was specified. Also, the outlet pressure (at the end of vocal tract) was set to zero. The outlet location was far enough from the glottis for these conditions to apply.

The mechanical boundary conditions are met by enforcing the known degrees of freedom to each node on the vocal folds. This includes fixed nodes on the cartilages and air-driven nodes on the glottis surface. The displacements of fixed nodes are always set to zero and the displacement of air-driven nodes are calculated from the pressure distributions within the variational formulation. During vocal fold collision, the nodes are limited from crossing the midsagittal plane.

Results and Discussion

Results were obtained for lung pressure values of 0.4-1.2 kPa. The computation was performed initially on a DECstation 5000 and later on a Silicon Graphics Indy2 machine. Results include the velocity profiles, pressure profiles, and some glottal waveforms that are discussed below.

Figure 2 shows a typical flow pattern in the laryngeal model at a lung pressure of 0.8 kPa and a fundamental frequency of 123 Hz. At the early stage of the opening, streamlines are parallel. As the vocal folds opens from a convergent glottal shape, the jet of air exits from the glottis with vortices around it. The vortex bubbles grow and move toward the end of the channel. When glottis is closing, these bubbles join together and make a larger bubble. This process is repeated in each cycle. Vortex shedding seems to be a major mechanism of energy transfer in the larynx and may play an important role in the acoustics of vocal tract.

Figure 3 shows the instantaneous velocity profiles across the glottis at glottal inlet, just before the contraction of the glottis. The data correspond to a lung pressure of 1 kPa and fundamental frequency of 124 Hz. The solid line corresponds to the moment of glottal opening and indicates a small flow. Different profiles at various phases of the cycle have their peak located in the center and reach a maximum almost in the middle of cycle.

In Fig. 4, velocity profiles are given at the halfway thickness (vertically) of the vocal folds. The profiles are flatter in shape and their peak reaches up about 14 m/s for a lung pressure of 1 kPa. The width of the profiles is smaller than the inlet because of the convergence of glottis.

The jet velocity profiles at exit of the glottis are shown in the Fig. 5. The peak jet velocity initially goes down due to the opening of the vocal folds, but then increases and reaches its maximum of about 44 m/s towards the end of the

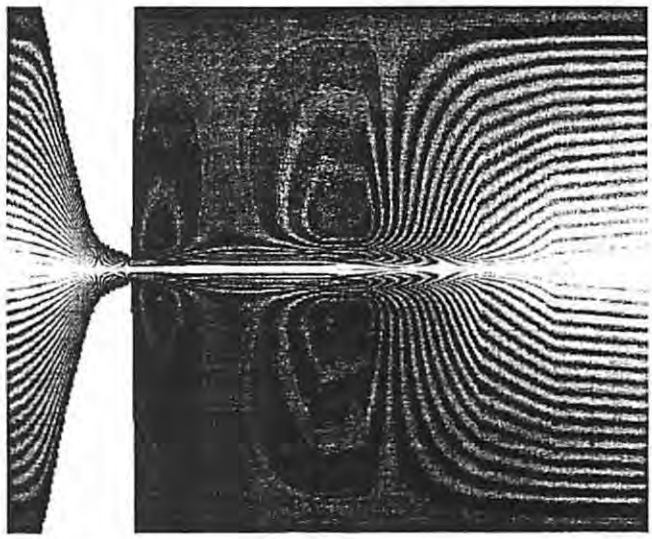
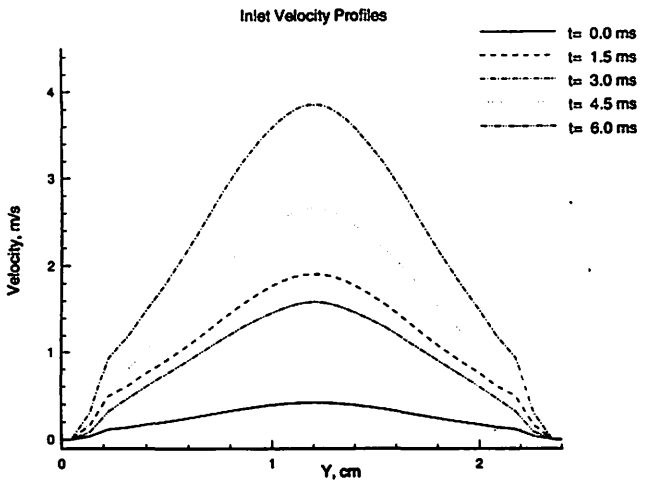


Figure 2. A typical flow pattern in the laryngeal model.



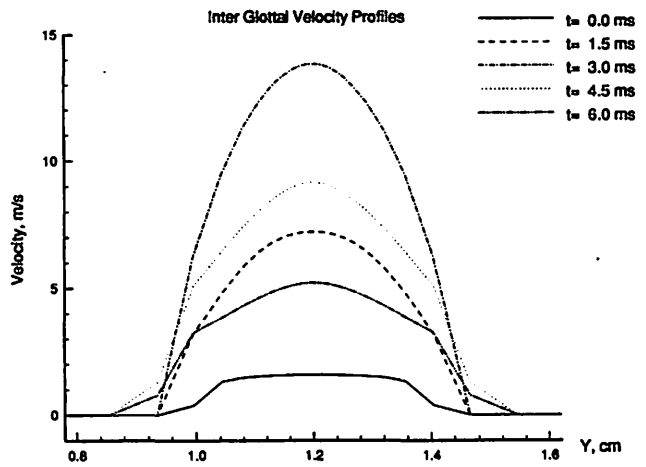


Figure 3 (top). Inlet velocity profiles entering the glottis at various phases of a cycle. Figure 4 (bottom). Intraglottal velocity profiles within the glottis.

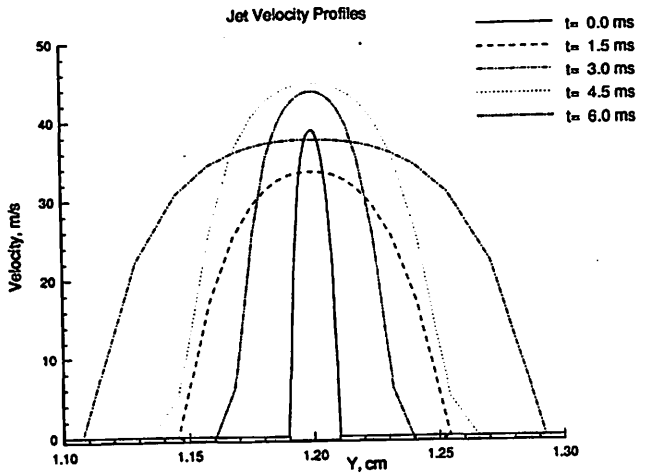


Figure 5 (above). Jet velocity profiles at the glottal exit.

cycle. This predicted maximum velocity is close to the measured peak velocity in the excised larynx with hot-wire anemometry [14].

The effect of wall movement can be seen as the jet width changes. Examination of jet velocities at other cross sections shows that the peaks of the jet velocities stay unchanged for a couple of centimeters away from the glottis, but they pulsate as a function of time.

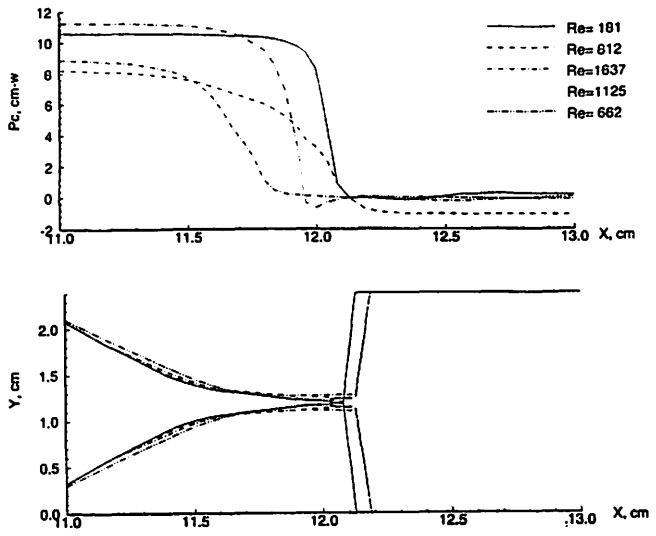
The instantaneous centerline pressure profiles of the simulation case with lung pressure of 1 kPa are plotted along the channel length in Fig. 6. The geometry of the vocal fold wall is shown below. As vocal folds separates, the pressure distribution is built-up in the glottis and a sharp pressure gradient, similar to that of steady flow, controls the laryngeal airflow. During the oscillation of the vocal folds, the amount of the transglottal pressure and the shape of the pressure profile changes. Once the glottis is closed, the pressure becomes almost uniform on each side of the glottis. The pressure drops on the walls (surface of the vocal folds) may be a little higher than at centerline because of the longer particle path on the walls [12].

Figure 7 shows the typical time-waveforms of a simulation at lung pressure of 0.8 kPa. The waveforms from top to bottom are glottal Reynolds number, jet velocity at the glottal exit, and glottal flow rate (volume velocity). The Reynolds number is essentially proportional to glottal flow rate and can exceed 1300 during a cycle. The flow rate signal differs from that found by Bernoulli's solution, but its maximum is about the same. The jet velocity has a big jump from zero at glottal opening and shows some fluctuations that could be due to the iteration method of the solution.

Figure 8 shows overlaid waveforms of glottal width and glottal area for the lung pressure values of 0.4-1.2 kPa. The profiles show skewness, a well known phenomena in the phonation [20]. As the lung pressure increases, the peak of these waveforms increases and moves to the left.

Conclusions

The results of this study have shown that flow-induced self-oscillation is obtainable by solution of the Navier-Stokes equations in a two-dimensional glottal flow duct. the calculations are tedious and require iterations within a time step. The benefit, however, is a better description of the velocity and pressure profiles during all parts of the glottal cycle. This detailed information about the velocity and pressure distribution should complement experimental data from excised larynx in the development of comprehensive theory of aerodynamics of phonation. Vortex shedding and jet formation are now predictable for given specific glottal geometries, even for time-varying cases. From these results, we hope to provide correction factors for simplified Bernoulli flow equations when computation speed is a major criterion.



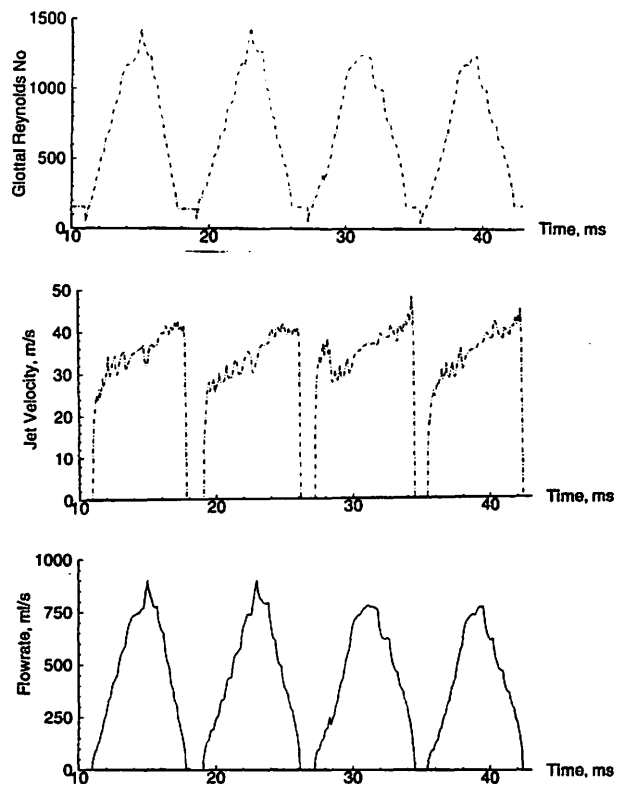


Figure 6 (top). Instantaneous centerline pressures at various phases of a cycle plotted on top of the corresponding two-dimensional wall. Figure 7 (bottom). Waveforms from combined simulation with lung pressure of 8 cm- water. The graph includes from the top to bottom, glottal Reynolds number, glottal jet velocity, and glottal flow rate (volume velocity).

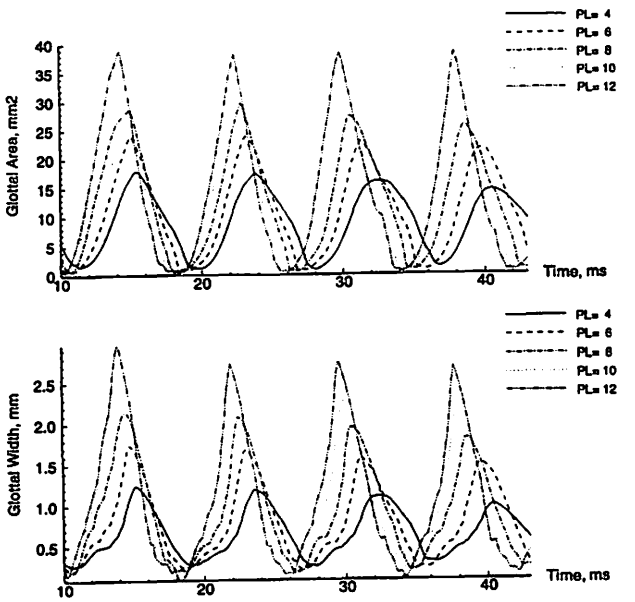


Figure 8. Waveforms from the combined simulation at various lung pressures. The top graph shows the glottal area waveforms at lung pressure values of 4 to 12 cm water. In the bottom graph glottal width are shown for the similar lung pressure values.

Acknowledgments

The authors would like to thank Mr. Chenwu Fan for his assistance in computer programming. This work was supported by the National Institute on Deafness and other Communication Disorders, Grant No. DC00831-04.

References

1. Ishizaka, K. and Flanagan, J. L. "Synthesis of voiced sounds from a two-mass model of the vocal cords" *Bell System Technical Journal*, **51**(6), 1233-1268 (1972).

2. Titze, I.R. and Talkin, D.T. "A theoretical study of the effects of various laryngeal configuration on the acoustics of phonation" *Journal of the Acoustical Society of America*, **66**(1), 60-74 (1979).

3. Alipour, F. and Titze, I.R. "A Finite element simulation of vocal fold vibration" in *Proceedings of Fourteenth Annual Northeast Bioenginering Conference*, Ed. J.R. LaCourse (IEEE #88-CH2666-6, New York 1988) pp. 186-189.

4. Ishizaka, K. and Matsudaira, M. (1972). *Fluid mechanical considerations of vocal cord vibration*. SCRL-Monograph 8, (Speech Communication Research laboratory, Santa Barbara 1972).

5. Scherer, R.C. *Laryngeal Fluid Mechanics: Steady Flow Considerations Using Static Models*, Ph.D. Dissertation, Department of speech Pathology and Audiology, The University of Iowa, Iowa City, Iowa, 1981.

6. Scherer, R.C., Titze, I.R. and Curtis, J.F. "Pressure-flow relationships in two models of the larynx having rectangular glottal shapes" *Journal of the Acoustical Society of America*, 73(2), 668-676 (1983).

7. Binh, N. and Gauffin, J. "Aerodynamic measurements in an enlarged static laryngeal model" *Technical Phoniatrics*, (Department of Gas Dynamics, Royal Institute of Technology, Stockholm, STL- QPSR 1983).

8. Liljencrants, J. "Numerical simulation of glottal flow" in *Vocal Fold Physiology: Acoustics, Perception, and Physiological Aspects of Voice Mechanisms* Ed. J. Gauffin and B. Hammarberg (Singular Publishing Group Inc., San Diego 1991) pp. 99-104.

9. Alipour, F. and Patel, V.C. "Numerical simulation of laryngeal flow" in *1991 Advances in Bioengineering*, Ed. R. Vanderby (ASME, New York 1991) pp. 111-114.

10. Iijima, H., Miki, N. and Nagai, N. "Glottal impedance based on a finite element analysis of two-dimensional unsteady viscous flow in a static glottis" *IEEE Transaction on Signal Processing*. 40(9), 2125-2135 (1992).

11. Guo, C.G. and Scherer, R.C. "Finite element simulation of glottal flow and pressure" *Journal of the Acoustical Society of America*. 94(2), Pt. 1, 688-700 (1993).

12. Alipour, F. and Patel, V.C. "Steady flow through modeled glottal constriction" *Journal of Engineering, Islamic Republic of Iran*, 7(1), 13-18 (1994).

13. Berke, G.S., Moore, D.M., Monkewitz, P.A., Hanson, D.G., and Gerratt, B.R. "A preliminary study of particle velocity during phonation in an in vivo canine model" *Journal of Voice*. 3(4), 306-313 (1989).

14. Alipour, F. and Scherer, R.C. "Pulsatile airflow during phonation: an excised larynx model" *Journal of the Acoustical Society of America*. 97(2), 1241-1248 (1995).

15. Alipour, F., Scherer, R.C. and Patel, V.C. "An experimental study of pulsatile flow in canine larynges" *Journal of Fluids Engineering*, (in press).

16. Alipour, F., Fan, C. and Scherer, R.C. "A numerical simulation of laryngeal flow in a forced-oscillation glottal model" *Journal of Computer Speech and Language* (in review).

17. Titze, I.R. and Alipour, F. *The Myoelastic-Aerodynamic Theory of Phonation*. in preparation.

18. 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.

19. Story, B.H. *Physiologically-Based Speech Simulation using an Enhanced wave-Reflection Model of the Vocal Tract*. Ph.D. Dissertation, Department of speech Pathology and Audiology, The University of Iowa, Iowa City, Iowa, 1995.

20. Rothenberg, M. "Acoustic interaction between the glottal wave source and the vocal tract" in *Vocal Fold Physiology*, Ed. K.N. Stevens and M. Hirano (University of Tokyo Press, Tokyo 1981) pp. 305-323.

NCVS Status and Progress Report - 8
July 1995, 15-24

Bifurcations in Excised Larynx Experiments

David Berry, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa
Hanspeter Herzel, Ph.D.
Institute for Theoretical Physics, Berlin, Germany
Ingo Titze, Ph.D.
Brad Story, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa

Abstract

Bifurcation analysis is applied to vocal fold vibration in excised larynx experiments. Phonation onset and vocal instabilities are studied in a parameter plane spanned by subglottal pressure and asymmetry of either vocal fold adduction or elongation. Various phonatory regimes are observed, including single vocal fold oscillations. Selected spectra are shown to demonstrate the correspondence between these regimes and vocal registers noted in the literature. To illustrate the regions spanned by the various phonatory regimes, two-dimensional bifurcation diagrams are generated. Many instabilities or bifurcations are noted in the regions of coexistence, i.e., the regions where the phonatory regimes overlap. Bifurcations are illustrated with spectrograms and fundamental frequency contours. Where possible, results from these studies are related to clinical observations.

Introduction

Given the nonlinearities associated with aerodynamically coupled oscillations, a nonlinear dynamics approach is essential in any study of vocal fold vibrations. While nonlinearities are routinely implemented in the *simulation* of vocal fold movement, it is less common to use a nonlinear approach in the *analysis* of vocal fold vibrations. The bifurcation diagram (to be explained shortly) is a crucial element in nonlinear analysis, especially when the system is expected to be governed by low-dimensional dynamics.

In relation to vocal fold dynamics, a bifurcation is a sudden qualitative change in the vibratory pattern of the folds, usually induced by a small change of some parameter such as lung pressure, vocal fold tension, or asymmetry. One common bifurcation (induced by gradually increasing the subglottal pressure) is phonation onset, an initialization of self-sustained vocal fold oscillations from a state of rest (1-3). Other common bifurcations include frequency jumps to new periodic oscillations or to subharmonic oscillations ($F_0/2$, $F_0/3$, etc.), the onset of low-frequency modulations into the vibration pattern, and jumps to irregular oscillations associated with chaos (4-9).

The purpose of the bifurcation diagram, especially the two-dimensional bifurcation diagram explored in this paper, is to document the regions where distinct vibration patterns occur. Some regions which might have physiological relevance include the parameter space spanned by subglottal pressure and vocal fold tension, or the region spanned by subglottal pressure and an asymmetry parameter, such as the ratio of tensions between the left and right fold.

For nonlinear systems like the vocal folds, the regions of different vibratory regimes may overlap. This overlap between regions poses serious consequences for vocal control: wherever an overlap occurs, more than one vibration pattern may result from the same vocal fold configuration. In such regions, involuntary, spontaneous jumps from one vibration pattern to another may occur. In fact, all the examples of bifurcations just listed may occur in such regions. Consequently, bifurcation diagrams have immediate clinical relevance for pin-pointing the regions where instabilities, voice breaks, and other transitions may occur. They are also useful for documenting the overall range and capabilities of a particular voice.

A two-dimensional bifurcation diagram that has often been used in connection with voice training and voice analysis is the Voice Range Profile (VRP). In the plane spanned by fundamental frequency and intensity, the VRP maps out the region where phonation is possible. Consequently, a central focus of the VRP is the Hopf bifurcation, the bifurcation associated with phonation onset. In generating the VRP, usually no attempt is made to maintain a particular voice type or vocal fold configuration. Rather, just one cumulative region is noted to encompass the entire vocal range of the subject.

In a recent study by Steinecke and Herzel (3), extensive bifurcation diagrams were generated for an asymmetric two-mass model of the folds. While the results of that study have implications for vocal fold paralysis, they need to be substantiated with models more directly related to the vocal fold morphology. Thus, in this investigation, bifurcations will be generated with excised larynges in an attempt to narrow the gap between biomechanical simulations and clinical observations.

In this paper, we first describe the methods for generating phonation from the excised larynges, and the protocol for data collection for the bifurcation diagrams. Next, an illustration of Hopf bifurcations (bifurcations associated with phonation onset) is presented. A qualitative description is given of the vibration patterns observed in the investigations, as well as a comparison of these patterns to registers discussed in the literature. In another section, bifurcations of relatively symmetric folds are presented. Finally, two bifurcation diagrams are shown for bifurcations induced by both adduction and elongation asymmetries. All observations are discussed and compared with clinical observations and with results from biomechanical simulations.

Methods

Five larynges of large mongrel dogs (on the order of 25 kg) were obtained from the cardiac unit of the University of Iowa Hospitals and Clinics. The animals were sacrificed for other purposes, but the tissue was made available to us post mortem. Each larynx was dissected and trimmed as described elsewhere (10). The thyroid cartilage was bracketed and firmly mounted on a tube supplying heated and humidified air. One suture was attached to the anterior arch of the cricoid cartilage, allowing the vocal folds to be symmetrically elongated before asymmetries were applied, i.e., an upward pulling on the suture would elongate both folds, while a downward pulling would shorten both folds.

To permit asymmetrical adductory adjustments, sutures A_1 and A_2 were attached to the muscular processes of the arytenoid cartilages (see Figure 1). As shown, the sutures coursed anteriorly, medially and inferiorly, to approximate the action of the lateral cricoarytenoid muscle during adduction. Sutures L and S (see Figure 1) were

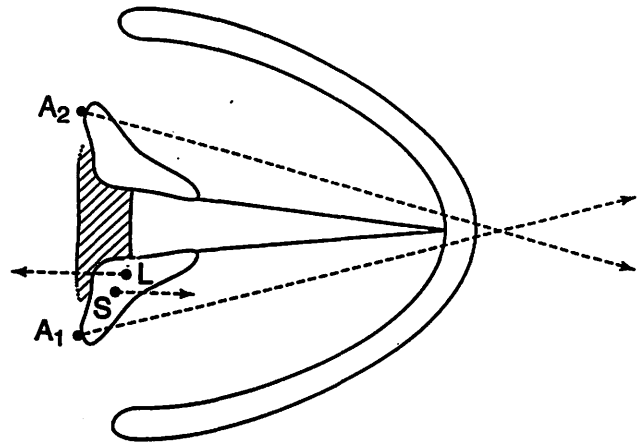


Figure 1. A schematic of the sutures used to induce the asymmetries on the excised larynx.

attached to the left arytenoid cartilage, enabling asymmetrical elongation of the folds. All sutures were attached to micrometers for precise control in manipulating the degree of asymmetry.

All experiments in this study were recorded with a Super VHS color video system, including an audio recording with a Sennheiser-MD441U3 microphone, displaced from the glottal opening by approximately 6 inches vertically and 4 inches horizontally. Selected portions of the audio signals were later digitized with 16-bit resolution and a sampling rate of 22.1 kHz (anti-aliasing filters were set at 10 kHz). In general, the following protocol was used:

* The micrometer settings were adjusted (typically in 1 mm increments) and documented.
* The subglottal pressure was gradually increased from zero to 20-30 cm H_2O. A special note was made of pressures associated with phonation onset, jumps from one vibration pattern to another, instabilities, etc. Where possible, an attempt was made to strobe the various phonatory regimes.
* The subglottal pressure was gradually decreased back to zero. Again, it was noted where phonation stopped, and where instabilities and transitions occurred.

The process was repeated until all desired asymmetries had been investigated on 5 different larynges. No instabilities or bifurcations were noted for the asymmetries applied to Larynx 1 or 2 other than those associated with phonation onset. "Symmetric" bifurcations related to vocal fry-like regimes were noted for Larynx 3, especially at high pressures. For Larynx 4, extensive bifurcations were induced with adduction asymmetries. Finally, bifurcations due to elongation asymmetries were observed in the investigations of Larynx 5. All these findings will be presented in the following sections on Phonation Onset, Symmetric Bifurcations, Bifurcations due to Asymmetric Adduction, and Bifurcations due to Asymmetric Elongation of the Folds.

Phonation Onset

How much pressure is required to initiate and sustain phonation? In a more formal sense, where does the Hopf bifurcation (phonation onset) occur within a given parameter space? This question has immediate clinical relevance for investigating the "ease of phonation" under various conditions, including hydration level (11), prephonatory glottal shaping (12), vocal fold tension or fundamental frequency, and degree of disorder (e.g., to initiate phonation with unilateral paralysis, a relatively large subglottal pressure would be required). For computer models of phonation, related questions have already been carefully documented (2-3,13-14).

It is also of interest to know whether the Hopf bifurcation is supercritical (a soft onset with *no* hysteresis) or subcritical (an abrupt onset *with* hysteresis). For the latter case, two thresholds exist: one associated with phonation onset as subglottal pressure is gradually increased, and one corresponding to the cessation of phonation as subglottal pressure is decreased (15). Examples of subcritical Hopf bifurcations from Larynx 2 are shown in Figure 2. The upper line denoted by "+" symbols indicates the sudden onset of phonation as pressure is increased, whereas the lower line denoted by "◊ " symbols indicates the location where phonation stopped as the pressure was lowered.

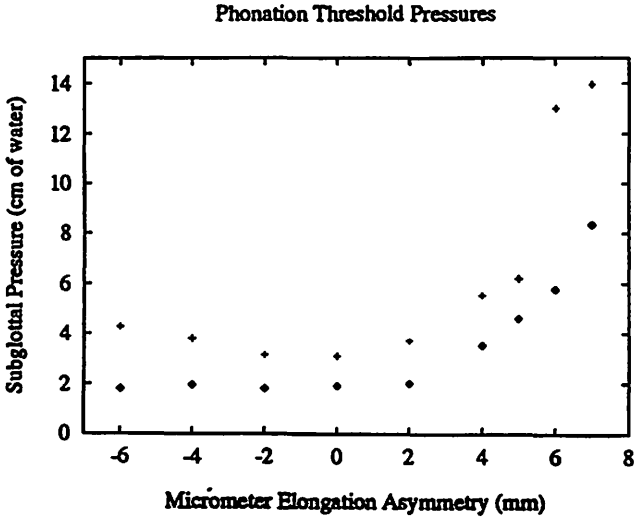


Figure 2. Phonation threshold pressures for Larynx 2. The horizontal axis displays the micrometer readings associated with suture S *for negative elongation values, and suture* L *for positive values.*

Both the upper and lower thresholds shown in the Figure 2 represent the average value of at least two repeated measurements for Larynx 2. The mean difference of repeated measures was approximately 0.35 cm H_2O. Over most of the range of induced asymmetries, there is a 1-2 cm H_2O hysteresis observed in the curves separating aphonia from phonation. However, for large asymmetries, this number increases substantially. Also, as might be suspected, the phonation threshold increases as the elongation asymmetry is increased.

The area between the upper "+" and the lower "◊" symbols is the region where aphonia and phonation coexist for Larynx 2. In this region, small perturbations (i.e., suture adjustments or variations in subglottal pressure) were observed to induce or terminate phonation. Such findings are reminiscent of the involuntary and intermittent aphonia observed in patients with vocal paralysis. Although the details of Figure 2 are not necessarily general observations, all 5 larynges showed some hysteresis in connection with phonation onset.

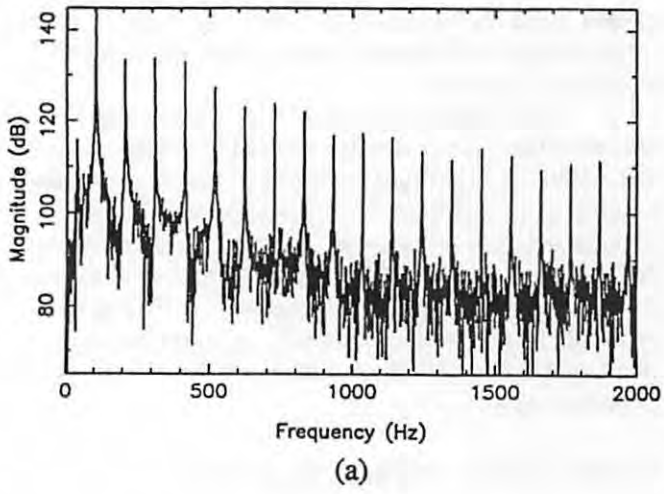
Vibration Patterns and Phonatory Regimes

Through systematic and extensive variations of subglottal pressure and asymmetric adduction and elongation, a wide variety of vibratory regimes have been explored in the canine larynges, reminiscent of the wide range of vocalizations possible in the human voice. Although not always strictly defined, the concept of registers is often used to describe different vibratory regimes (16-20). Some of the registers which are commonly distinguished include:

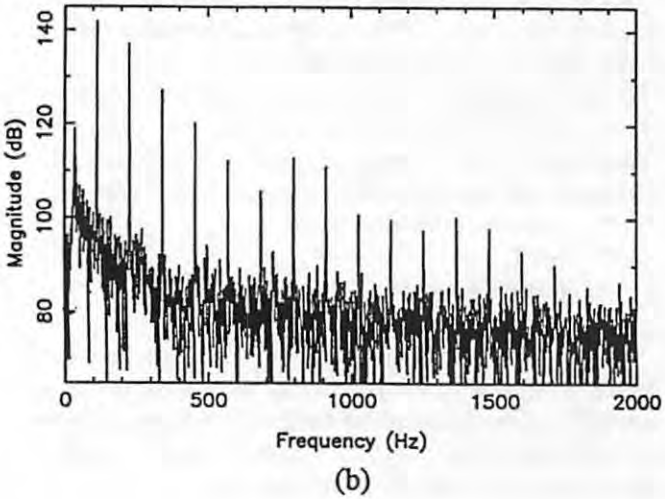
* vocal fry or pulse register
* chest
* falsetto or head
* whistle, flageolet, or flute register.

Below we present regimes which might fit within these general classifications, although no rigorous justification will be given for assigning a particular vibration pattern to a particular register. Rather, assignments will be based on perceptual evaluations and qualitative differences in spectra.

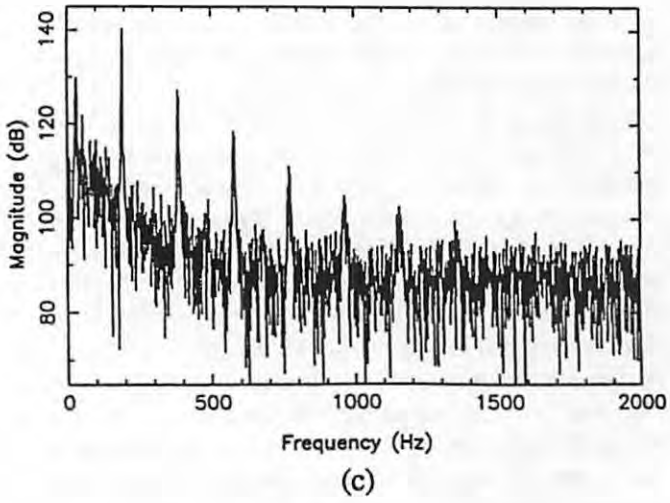
Figure 3 (following page) shows spectra for (a) chest-like phonation, (b) falsetto-like phonation, and (c) flute-like phonation. Spontaneous jumps were observed between (a) and (b), and just a small variation in elongation asymmetry (i.e., a 1 mm increase) induced the jump to (c). The vibration pattern for chest-like phonation revealed large amplitudes of vibration, complete closure, a pronounced mucosal wave, and relatively intense higher harmonics (see Figure 3a). The falsetto-like phonation was characterized by small amplitude vibrations, incomplete closure, weak phase delay between the upper and lower edges of folds, and less intense higher harmonics (see Figure 3b). The flute-like phonation had nearly double the frequency of the other two vibrations patterns (about 200 Hz, instead of 100 Hz), despite similar conditions. It had small vibrations, a weak glottal opening, no visible mucosal wave, and very weak harmonics (Figure 3c). Water on the folds appeared to be rotating synchronously with vocal fold oscillations, giving an indication of vorticity at the same frequency as the vocal fold oscillations. Virtually the same characterization has



(a)



(b)



(c)

Figure 3. Spectra from Larynx 5 for: (a) chest-like vibrations, (b) falsetto-like vibrations, (c) vortex-induced vibrations (related to flageolet, flute, or whistle register).

been given by Keilmann and Michek (18) for stroboscopic observations of whistle register in the soprano voice.

A preliminary explanation consistent with the observations would be an interpretation as "vortex-induced vibrations" (21). It is known that in a wide range of conditions between steady laminar flow and highly turbulent flow, a periodic formation of vortices can appear. For example, whistling is related to the excitation of resonances due to such vortex shedding. By means of transverse "lift forces," such periodically generated vortices may induce self-sustained oscillations of vibrating structures. This mechanism may cause the observed vocal fold vibrations and might also be related to the still controversially discussed flute or whistle register. Admittedly, this hypothesis is still somewhat speculative and needs to be confirmed by further measurements and computational modeling.

Finally, we present a fourth regime comparable to vocal fry phonation, taken from Larynx 3 at 19 cm H_2O. The spectrum in Figure 4 displays strong subharmonics at multiples of 50 Hz leading to a rough sounding phonation. Using a synchronization of approximately 100 Hz, a videostroboscopic recording of this regime was also obtained, revealing an alternation of large and small glottal openings, a finding which is often characteristic of vocal fry phonation.

Bifurcations of Symmetric Folds

Although instabilities generated by artificially induced asymmetries were the focus of this study, instabilities were occasionally observed for symmetric folds, especially at large subglottal pressures. Such observations have been noted in previous studies (Scherer & Austin, personal communication). In this section, characteristic bifurcations of Larynx 3 for nearly symmetric conditions are presented. By symmetric folds, we mean that no artificial asymmetry was

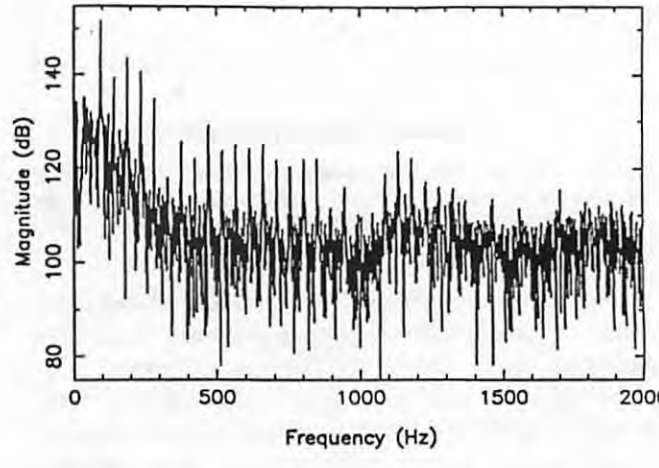
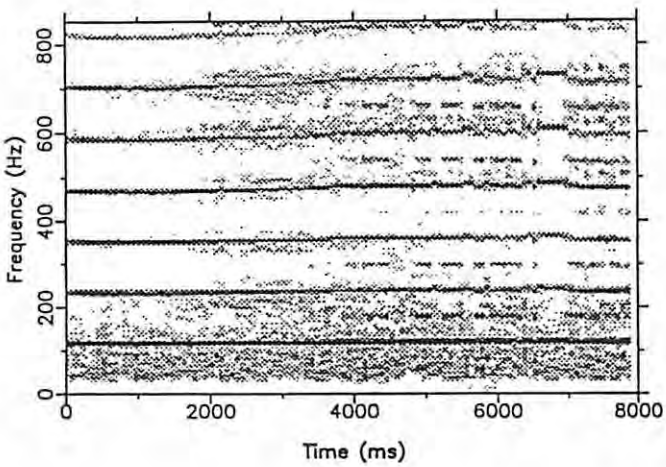


Figure 4. Spectrum from Larynx 3 for vocal fry-like vibrations.

intentionally induced on the folds. It is readily acknowledged that, in practice, true symmetry cannot be achieved.

The spectrogram in Figure 5 shows a transition to a subharmonic regime with peaks at multiples of $F_0/2$. This bifurcation was observed while slowly increasing subglottal pressure around 16 cm H_2O. Qualitatively, this transition might be described as a jump between chest-like phonation and vocal fry-like phonation. A detailed spectrum of the vocal-fry-like phonation was already shown in Figure 4. For even larger subglottal pressures (at about 26 cm H_2O), intermittent jumps between normal phonation and highly irregular phonation occurred, as depicted in Figure 6. The five dark verticle bars (occurring at 0.3 s, 1.1 s, 4.5 s, 5.8 s, 7.2 s) represent irregular phonation.



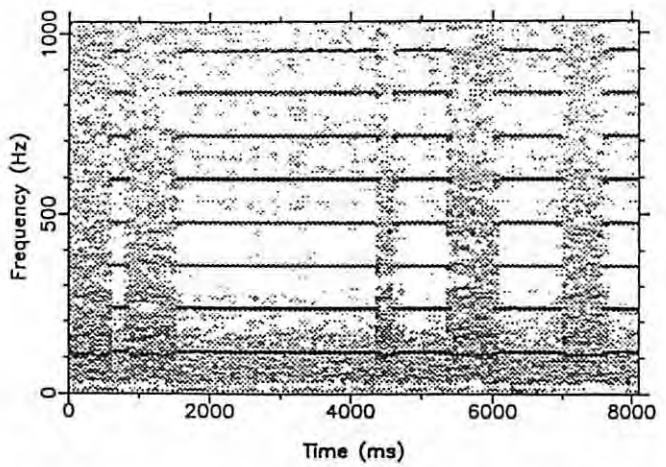


Figure 5(top). Spectrogram from Larynx 3 showing a transition from chest-like regime (1-2 s) to vocal fry-like regime (2-8 s), accompanied by a period-2 subharmonic as indicated by the additional spectral lines at $3F_0/2$ (~180 Hz), $5F_0/2$ (~300 Hz), $7F_0/2$ (~420 Hz), etc. Figure 6 (bottom). Spectrogram from Larynx 3 illustrating intermittent jumps from chest-like vibrations to irregular vibrations.

Bifurcations for Asymmetric Adduction

To systematically vary the asymmetry of adduction, adjustments were made in the micrometer settings of sutures A_1 and A_2 (see Figure 1). Suture A_1 was varied from 2 to 10 mm, while suture A_2 remained constant at 2 mm. The micrometer asymmetry, or the difference between the two micrometer readings, is shown as the horizontal axis in Figure 7. Given the experimental constraints associated with this study (i.e., there was only a limited time to make measurements because of tissue dehydration, and most threshold measurements had to be made at least twice to examine repeatability issues), only a 1 mm resolution was possible along the horizontal axis.

In this introductory paper on bifurcation analysis, no attempt was made to relate micrometer settings to specific glottal geometries. By analyzing the video recordings of these experiments, presumably one could investigate such a correspondence. However, for the purposes of this study, bifurcation diagrams are generated directly from the micrometer readings. Figure 7 shows the locations of various vibratory regimes of Larynx 4 in the parameter plane spanned by adduction asymmetry and subglottal pressure. The observed regimes included (1) chest-like vibrations, (2) irregular vibrations, and (3) periodic single vocal fold oscillations (the other fold appeared to be stationary). The overlap of horizontal and vertical lines, occupying a relatively large portion of the bifurcation diagram for low asymmetry, indicates the region of coexistence of chest-like vibrations and irregular vibrations. The number "8" appears along the upper edge of this region of coexistence, indicating the parameter values of the folds associated with the audio signal analyzed in Figure 8 (following page).

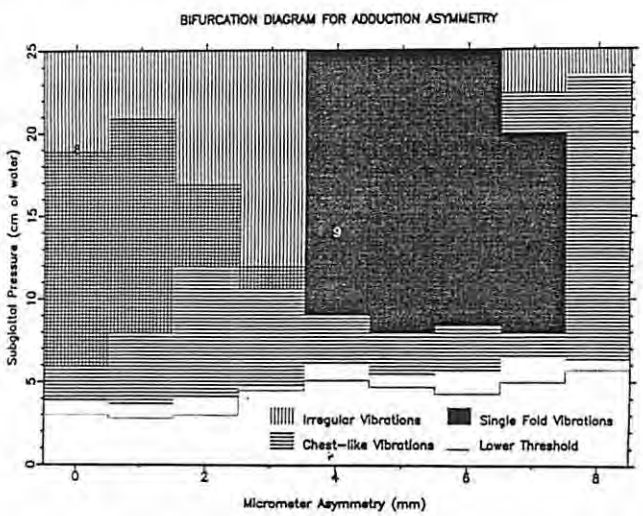


Figure 7. Bifurcation diagram for adduction asymmetries applied to Larynx 4. Square-like patterns indicate the overlap of the regions associated with chest-like vibrations and irregular vibrations. The dotted line indicates the lower phonation threshold for chest-like vibrations. The numerals "8" and "9" are positioned to indicate the laryngeal parameters which correspond to the signals used to generate Figure 8 and 9.

Typical of this region of coexistence, Figure 8a shows the spectrogram of a transition from chest-like vibrations to irregular vibrations. Note that around 1 s, the transition to irregular vibrations is initiated by a period-doubling, as indicated by the spacing of the harmonics. To examine this region more closely, a high precision fundamental frequency contour was generated (see Figure 8b) using GLIMPSES, a speech analysis package (22). The zig-zag pattern found in the contour from about cycle 32 to 52 is characteristic of a period-doubling. Because period-doublings are known to be common precursors of chaos, this gives an indication that the irregular oscillations may be associated with chaos. However, a closer examination of the spectrogram (Figure 8a) shows that the "irregular" vibrations are not totally random. For example, immediately following the transition, one sees hints of a fundamental frequency around 130 Hz (F_0), as well as a period-3 subharmonic, as manifested by the spectral lines at $F_0/3$, $2F_0/2$, and $4F_0/3$.

As indicated by the symbol "9" in the bifurcation diagram (Figure 7), Figure 9 was generated with parameter values associated with single vocal fold vibrations. Although too complicated to be displayed in the bifurcation diagram, the region of periodic single vocal fold vibrations showed considerable overlap with chest-like and irregular vibrations. A transition from single vocal fold vibrations to chest-like is shown in Figure 9. The transition occurs around 2.5 s, as manifested by the sudden drop in fundamental frequency. The single vocal fold vibrations are accompanied by a slight modulation, as depicted by the sidebands surrounding the harmonics during the first 2 seconds.

Bifurcations Due to Elongation Asymmetry

To systematically vary the asymmetry of elongation, adjustments were made in the micrometer settings of sutures *L* and *S*. The employment of suture L was referred to as a positive micrometer asymmetry, and the employment of suture S as a negative micrometer asymmetry. For 3 separate larynges (Larynges 1, 2, and 5), systematic lengthening and shortening of the left vocal folds were performed, with micrometer asymmetries ranging from 13 mm to -5 mm.

In this preliminary paper on bifurcation analysis, no attempt was made to precisely correlate micrometer readings with more kinetic measures such as force or tissue strain. Nevertheless, two microsutures were placed along the mid-length of both folds, allowing strain to be calculated. A few spot computations revealed a general correspondence between micrometer asymmetry and tissue strain. However, for the purposes of this study, only micrometer asymmetries are reported.

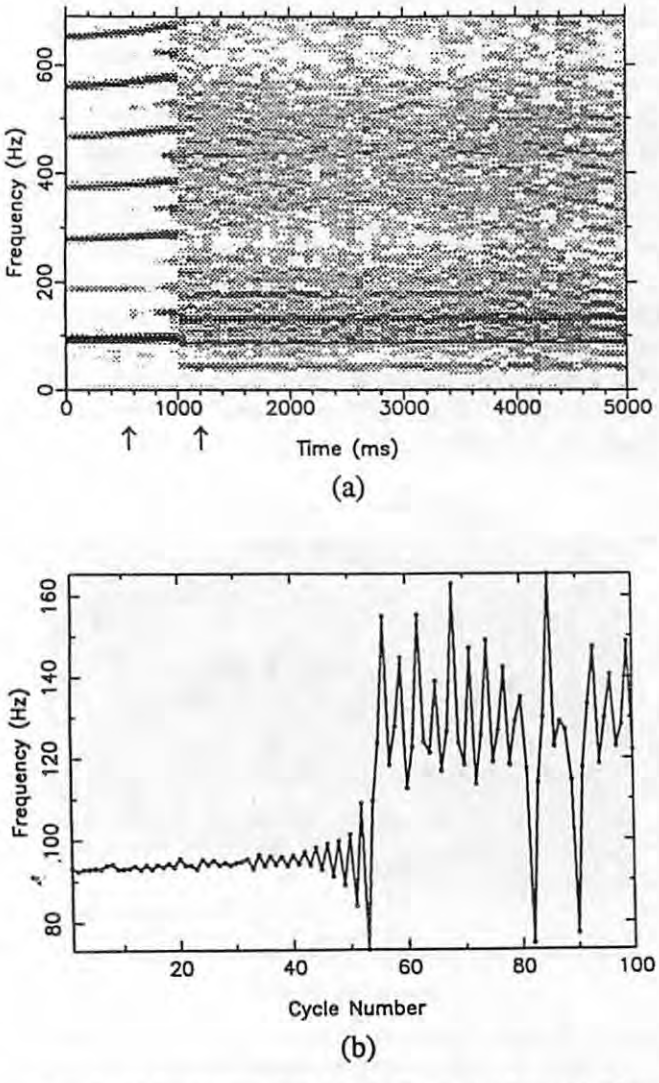


Figure 8. (a) Spectrogram from Larynx 4 illustrating a transition from chest-like vibrations to irregular vibrations, and (b) a magnified view of the fundamental frequency contour near the transition in (a), as indicated by the arrows.

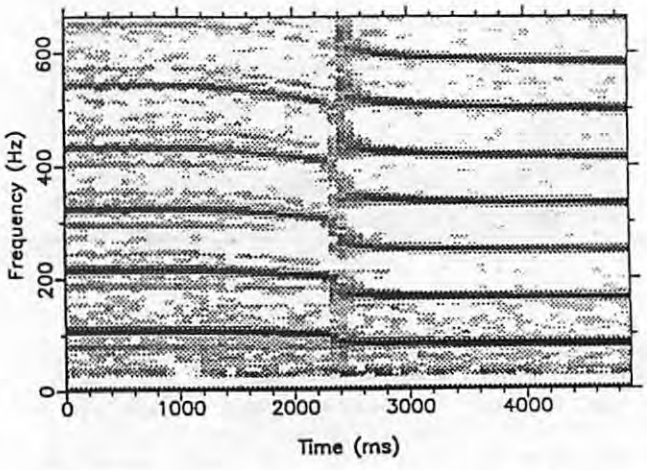


Figure 9. Spectrogram from Larynx 4 illustrating a transition from single vocal fold vibrations (first half) to chest-like vibrations (second half).

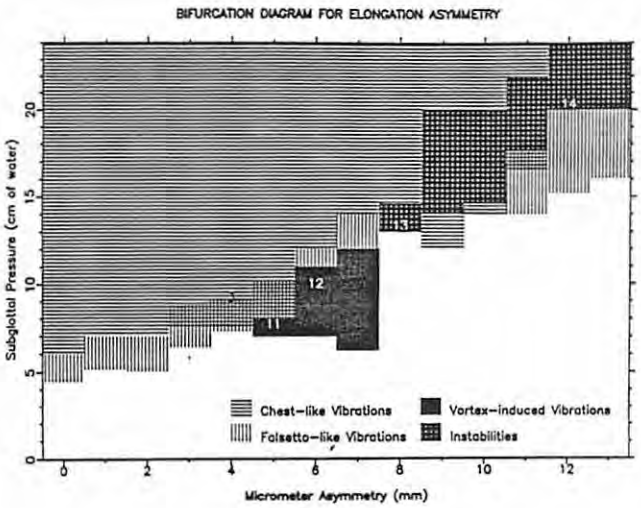


Figure 10. Bifurcation diagram for elongation asymmetries applied to Larynx 5. The numerals "11" and "14" are positioned to indicate the laryngeal parameters which correspond to the signals used to generate Figures 11-14.

With the applied elongation asymmetries, the only instabilities observed for Larynges 1 and 2 were those associated with the transition from aphonia to phonation (and vice versa). In the case of Larynx 1, perhaps the applied asymmetries were simply not great enough to induce further instabilities. For the relatively large asymmetries applied to Larynx 2 (see Figure 2), a significant coexistence region existed between aphonia and phonation. As noted before, small perturbations were observed to both induce and terminate phonation, reminiscent of the intermittent, involuntary aphonia often associated with vocal fold paralysis.

Larynx 5 demonstrated many instabilities which are summarized in Figure 10. Near the number "3" a small coexistence region of chest-like and falsetto-like vibrations is denoted by the overlapping horizontal and vertical lines. In this region, the previously mentioned spontaneous jump from falsetto-like vibrations (Figure 3a) to chest-like vibrations occurred (Figure 3b). Figures 11 (Figures 11 and 3c correspond to the same audio recording) and 12 are taken from the gray region where a whistle or flageolet-like register was observed (possibly related to vortex-induced vibrations). Although an audible high-pitched whistling was present for elongation asymmetries greater than 7 mm, it is not indicated in Figure 10 because no vocal fold vibrations were observed. Although too complex to be shown, the indicated whistling-region has considerable overlap with the falsetto and chest-like regions. The complicated bifurcations shown in Figure 11 are associated with the coexistence of whistle-like phonation (fundamental frequency around 200 Hz) and falsetto-like phonation (fundamental frequency around 100 Hz). Another observation from the same region is depicted in Figure 12. Here the

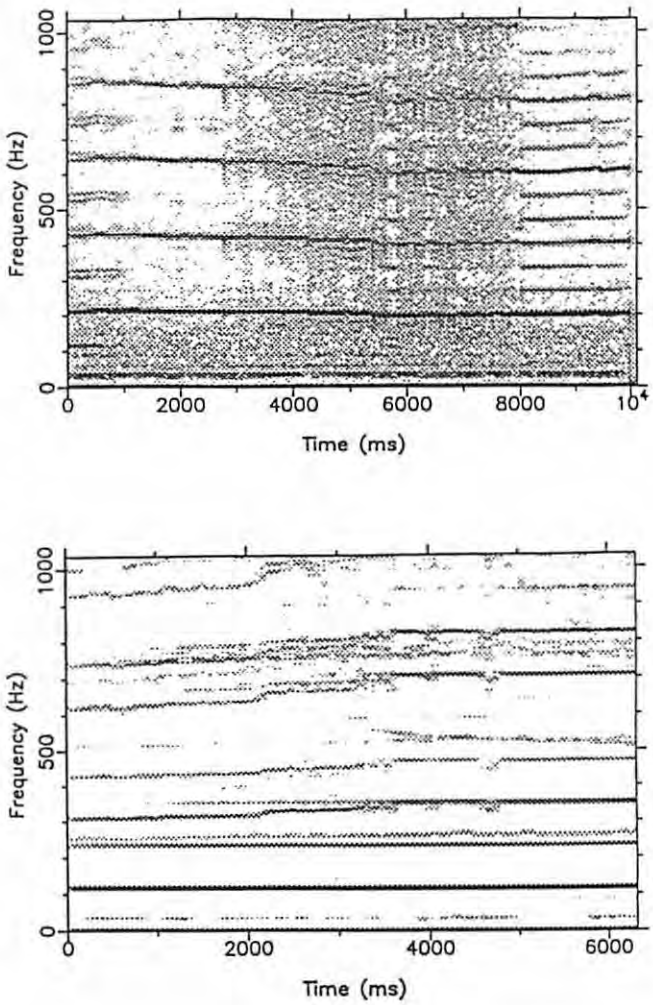


Figure 11(top). Spectrogram from Larynx 5 illustrating jumps from $F_0/2$ to F_0 to irregular vibrations to $F_0/3$. Figure 12 (bottom). Spectrogram from Larynx 5 illustrating a whistle-like register with two independent frequencies, i.e., some spectral lines remain constant, while others increase in frequency.

coexistence induces two independent frequencies, as demonstrated by the two sets of spectral lines, one set remaining constant and the other set gradually increasing.

Although the black region (speckled with white dots) of Figure 10 is just another coexistence region of chest and falsetto-like vibrations, it is specifically denoted as an instability region. This particular region was the most remarkable of all the instability regions studied on the 5 larynges. A few typical examples are shown in Figures 13 and 14.

The modulations shown in Figure 13 appeared as a precursor of the jump from falsetto to chest-like phonation. Although vibrato in a true singer's voice is probably induced by changes in muscle activity (23), our studies demonstrate that comparable modulations may also be induced by bio-

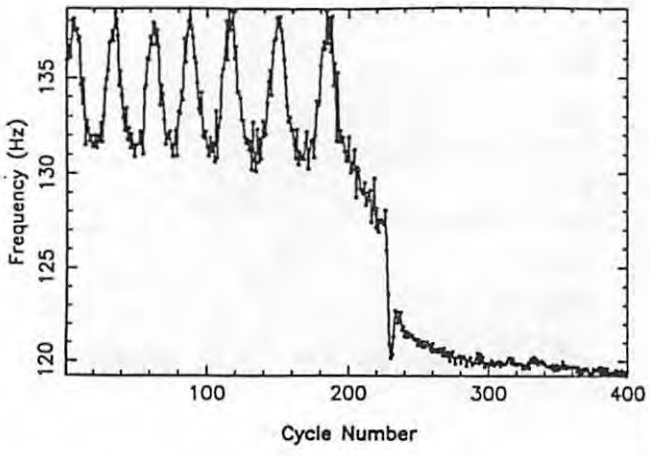


Figure 13. Fundamental frequency contour from Larynx 5 illustrating a spontaneously induced "vibrato" (i.e., a low frequency modulation) for fixed laryngeal parameters, followed by a sudden decrease in fundamental frequency.

mechanical-aerodynamic instabilities. However, the observed modulation frequencies were often well above an acceptable vibrato range. For example, the sidebands in Figures 9 and 14 suggest modulation frequencies of about 30 and 15 Hz, respectively.

At the parameter values marked 14 in Figure 10, more than 20 spontaneous jumps occurred between chest and falsetto-like phonation. A few examples of such spontaneous jumps are shown in Figure 14. The falsetto-like phonation (upper frequency) was accompanied by a low-frequency modulation, as indicated by sidebands in the spectrogram of Figure 14a, and by the F_0 modulation in the fundamental frequency contour of Figure 14b.

Summary and Conclusions

This investigation complements bifurcation analysis in computer models (5,9,24-26) and observations of bifurcations in voice disorders (5-8). In comparison to computer models, excised larynges are more closely related to vocal fold morphology. Furthermore, because the parameters of interest can be carefully monitored and controlled in excised larynx experiments, the researcher can pin-point the sources of irregularity with greater confidence than would be possible in human subjects. Typical nonlinear phenomena observed in earlier studies (subharmonics, modulations, and transitions to irregular phonation associated with chaos) were also observed in the excised larynx experiments. For the most part, instabilities were observed in parameter regions where phonatory regimes coexisted. In these regions, small parameter changes were seen to induce bifurcations. Sometimes jumps to new regimes occurred in these regions without any external intervention. Additionally, it was shown how many of these phonatory regimes might be related to common vocal registers noted in the literature.

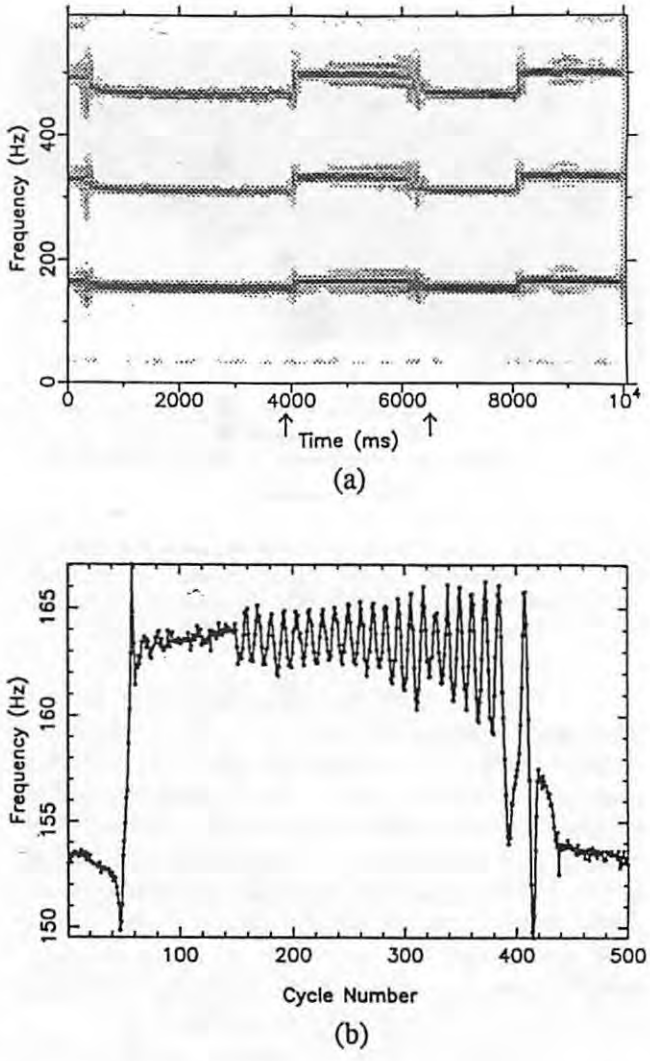


Figure 14. (a) Spectrogram from Larynx 5 illustrating spontaneous jumps from chest-like phonation (lower frequency) to falsetto-like phonation (upper frequency with modulations), and (b) fundamental frequency contour near the mid-portion of the spectrogram, as indicated by the arrows.

Sufficient asymmetric adduction was seen to induce single vocal fold oscillations. Clinically, such observations are found in unilateral vocal fold paralysis. Further asymmetric studies of this type can be used to better quantify the nature of this disorder, and to estimate the amount of surgical correction that might be necessary to restore normal, chest-like phonation under similar dysphonic conditions.

At various elongation asymmetries, several vocal registers were observed including chest, falsetto, and flageolet. In this regard, Figure 10 represented a comprehensive illustration of our most interesting results. In general, falsetto was observed at lower subglottal pressures than chest-like vibrations. For relatively low asymmetries, the coexistence region of chest and falsetto was either very

small or non-existent. At larger elongation asymmetries, the coexistence region became very significant and many spontaneous jumps were observed between chest and falsetto-like vibrations.

Only a small region corresponded to flageolet or whistle register. Although a high frequency whistling sound was present for higher asymmetries as well, vibrations did not appear to be induced at the higher frequencies. This high-pitched whistling might be comparable to "whining" elicited from anesthetized dogs (27). We expect that pure whistling and whistle or flageolet register (i.e., vortex-induced vibrations) are closely related, both being induced by vorticity above the glottis. Indeed, if plotted in a bifurcation diagram such as Figure 10, these two phenomena would occur at neighboring parameter regions.

Although elongation asymmetries were often not visible (they increased the tension of the folds, but usually did not significantly alter the vocal fold geometry), experimental results indicate their profound influence on the overall vocal fold vibration patterns. Thus, in the treatment of unilateral paralysis, the influence of elastic asymmetries needs to be considered in addition to the more obvious geometric asymmetries.

In summary, bifurcation diagrams have proven useful for studying vocal fold dynamics, just as they have for many other nonlinear systems. We envision that the systematic framework demanded in the generation of bifurcation diagrams will guide many future studies of vocal fold oscillations.

Acknowledgments

This work was supported by Grant No. P60 00976 from the National Institute on Deafness and Other Communication Disorders and by the Deutsche Forschungsgemeinschaft. The authors also thank Lucy Karnell for assistance with experimental design, dissection and suturing.

References

1. Ishizaka K, Matsudaira M. Fluid mechanical considerations of vocal cord vibration. In: *SCRL Monogr 8*. Santa Barbara, CA: Speech Commun Res Lab, 1972.

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

3. Steinecke I, Herzel H. Bifurcations in an asymmetric vocal fold model. *J Acoust Soc Am* 1995;97:1874-1884.

4. Mende W, Herzel H, Wermke K. Bifurcations and chaos in newborn cries. *Phys Letters A* 1990;145:418-424.

5. Herzel H, Steinecke I, Mende W, Wermke K. (1991). Chaos and bifurcations during voiced speech. In: Mosekilde E, ed. *Complexity, Chaos, and Biological Evolution*. New York: Plenum Press, 1991:41-50.

6. Baken RJ. Géométrie fractale et évaluation de la voix: application préliminaire á la dysphonie. *Bull d'Audiophonogie Ann Sc Univ Franche-Comté* 1991;7:731-749.

7. Titze IR, Baken RJ, Herzel H. (1993). Evidence of chaos in vocal fold vibration. In: Titze IR, ed. *Vocal Fold Physiology: Frontiers in Basic Science*. San Diego, CA: Singular Publishing Group, Inc., 1993:143-188.

8. Herzel H, Berry DA, Titze IR, Saleh M. Analysis of vocal disorders with methods from nonlinear dynamics. *J Speech Hear Res* 1994;37:1008-1019.

9. Berry DA, Herzel H, Titze IR, Krischer K. Interpretation of biomechanical simulations of normal and chaotic vocal fold oscillations with empirical eigenfunctions. *J Acoust Soc Am* 1994;95:3595-3604.

10. Durham PL, Scherer RL, Druker DG, Titze IR (1987). Development of excised larynx procedures for studying the mechanisms of phonation. In: *Tech Report VABL-1-1987*. Iowa City, IA: Voice Acoustics and Biomechanics Laboratory, Department of Speech Pathology and Audiology, University of Iowa, 1987 (unpublished).

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

12. Titze IR, Schmidt SS, Titze MR. (in press) Phonation threshold pressure in a physical model of the vocal fold mucosa. *J Acoust Soc Am*.

13. Ishizaka K, Flanagan JL. Synthesis of voiced sounds from a two-mass model of the vocal cords. *Bell Syst Tech J* 1972;51:1233-1267.

14. Lucero JC. Dynamics of the two-mass model of the vocal folds: equilibria, bifurcations, and oscillation region. *J Acoust Soc Am* 1993;94:3104-3111.

15. Bergé P, Pomeau Y, Vidal C. *Order within chaos*. New York: Wiley, 1984.

16. Hollien H, Michel J. Vocal fry as a phonational register. *J Speech Hear Res* 1968;11:600-604.

17. Walker SJ. An investigation of the whistle register in the female voice. *J Voice* 1988;2:140-150.

18. Keilmann A, Michek F. Physiology and acoustic analyses of the female whistle voice. *Folia Phoniatr* 1993;45:247-255.

19. Miller DG, Schutte HK. Physical definition of the "flageolet register". *J Voice* 1993;7:206-212.

20. Sundberg J. (in press). Vocal fold vibration patterns and phonatory modes. *Folia Phoniatr.*

21. Langford WF, Zhan K. Dynamics of strong 1:1 resonance in vortex-induced vibration. In: Paidoussis MP, Akylas T, Abraham PB, eds. *Fundamental Aspects of Fluid-Structure Interactions.* New Jersey: The American Society of Mechanical Engineers, 1992.

22. Titze IR, Liang H. Comparison of F_0 extraction methods for high precision voice perturbation measurements. *J Speech Hear Res* 1993;36:1120-1133.

23. Hsiao TY, Solomon NP, Luschei ES, Titze IR. Modulation of fundamental frequency by laryngeal muscles during vibrato. *J Voice* 1994;8:224-229.

24. Ishizaka K, Isshiki N. Computer simulation of pathological vocal-cord vibration. *J Acoust Soc Am* 1976;60:1193-1198.

25. Wong D, Ito MR, Cox NB, Titze IR. Observation of perturbations in a lumped-element model of the vocal folds with application to some pathological cases. *J Acoust Soc Am* 1991;89:383-394.

26. Smith ME, Berke GS, Gerratt BR, Kreiman J. Laryngeal paralyses: theoretical considerations and effects on laryngeal vibration. *J Speech Hear Res* 1992;35:545-554.

27. Solomon NP, Luschei EL, Kang L. (in press). Fundamental frequency and tracheal pressure during three types of vocalizations elicited from anesthetized dogs. *J Voice.*

NCVS Status and Progress Report - 8
July 1995, 25-31

Coupling of Neural and Mechanical Oscillators in Control of Pitch, Vibrato, and Tremor

Ingo Titze, Ph.D.
Department of Speech Pathology & Audiology, The University of Iowa

Introduction

Traditionally, voice production has been conceptualized as the combination of one active source (an oscillator) and several passive systems that modify the sound produced by the source [1]. Oscillation seemed to occur only at the vocal folds while "resonation" occurred in the vocal tract (subglottal, supraglottal, nasal). This simple division between active and passive systems is no longer the most useful conceptualization. First, the vocal folds themselves have resonances (normal modes) that are in principle no different from the resonances (formants) of the vocal tract. Each normal mode of a single vocal fold can be thought of as a distinct oscillator, as each vocal tract formant is an oscillator. Thus, we have a system of many mechanical and acoustic oscillators, some of which are strongly coupled while others are weakly coupled or non-excited.

Overriding the mechanical and acoustic oscillators are the neurologic oscillators. Evidence is growing that vocal vibrato, for example, may be a cultivated vocal tremor [2]. A collection of tremor frequencies in the 4-7 Hz range is assumed to originate in the central nervous system. This collection of frequencies produces a band of low-frequency noise in laryngeal muscle contractions. A peripheral oscillator (we suspect partly mechanical and partly reflexive) is tuned to a specific frequency within this band (around 5.0 Hz). When the peripheral oscillator is excited by the tremor spectrum, a relatively stable vibrato frequency results, typically between 5-6 Hz.

Confusion can arise between neural oscillations and low frequency modulations within the mechanical and acoustic oscillators. These low frequency modulations are often beat frequencies between two oscillators (say, two similar modes of the left and right vocal folds) that have been observed in the 20-30 Hz region [3].

The purpose of this paper is to lay the groundwork for exploring some of the similarities and differences between neural and mechanical oscillations. The work is incomplete, but there is enough data and theory to draw some tentative conclusions.

Review of an Experiment

During the course of an experiment designed to record the electromyographic (EMG) activity of laryngeal muscles in a singer, the opportunity presented itself to probe the natural vibrato with an artificial vibrato [4]. This allowed for some basic hypotheses to be explored about a peripheral resonance phenomenon for vibrato (and possibly vocal tremor) in the larynx.

Figure 1 illustrates the experimental design. We assumed that a natural vibrato begins with a collection of tremor frequencies in the 4-7 Hz region, shown in the upper left portion of the diagram. This collection of frequencies excites both the TA and CT muscles, which are otherwise activated tonically to produce the mean fundamental frequency (F_o). Mechanical rotation and translation takes place around the cricothyroid joint. A resonant system is presumed to exist on the basis of reflex loops between mechanoreceptors and the intrinsic musculature, as the center of the diagram suggests.

In the experiment, a stimulator was used to excite the cricothyroid muscle with a train of pulses (bottom of Fig. 1). The pulse train contained a harmonic spectrum of frequencies. By superposition, each harmonic component could be thought of as a separate excitation of the muscle. The fundamental component of the pulse train varied between 2-8 Hz. Direct muscle stimulation was used with a pair of hooked-wire electrodes placed into the belly of the muscle. This method of stimulation was not necessarily the best method, but seemed highly convenient since the electrodes were already placed for EMG recording. The subject could produce both vibrato and non-vibrato phonations when the stimulator was turned either on or off. Thus, all

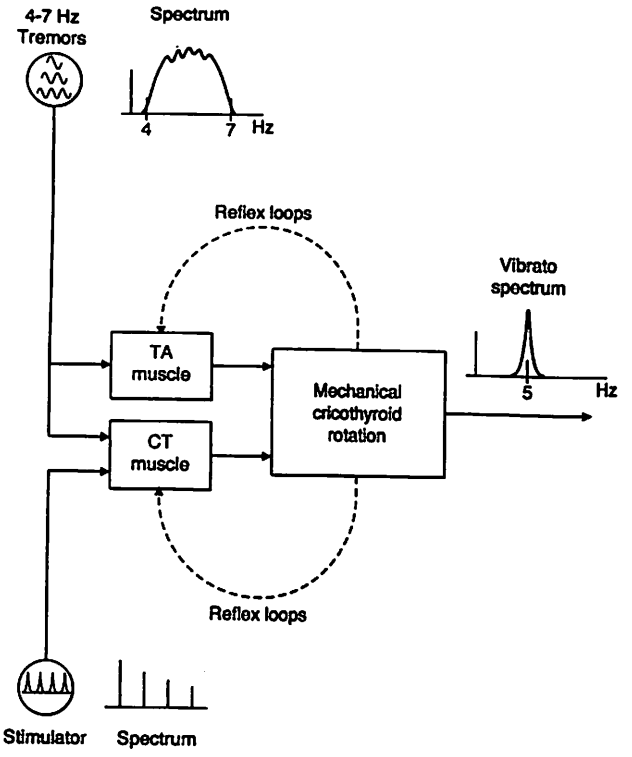


Figure 1. Diagram of an experiment to test the interaction between natural and artificial vibrato.

combinations of natural and artificial vibrato could be probed. We hypothesized that the two vibratos would interfere and that the natural vibrato could be entrained by the artificial vibrato.

Figure 2 shows the frequency response curve of the peripheral oscillator when the fundamental frequency of stimulation was varied from 2-8 Hz. The subject phonated an /ɑ/ vowel without natural vibrato at 220 Hz. The excitation level of the stimulator was kept constant as the fundamental component of F_o variation was measured. Note the tuning around 5 Hz, suggesting that there is a resonance in this peripheral system.

In the next part of the experiment, the combination of natural vibrato and artificial vibrato was probed. In Fig. 3, the natural vibrato frequency is plotted as the stimulation frequency (the artificial vibrato) is varied. Note that when the stimulation frequency was considerably below the natural vibrato frequency, e.g., 3-4 Hz, the two frequencies appeared to be independent of each other. In other words, the vibrato frequency remained roughly constant as the stimulation frequency was increased. In the region of 5-6 Hz, however, an entrainment took place between the stimulation frequency and the natural vibrato frequency. The two frequencies locked onto each other. When the stimulation frequency was increased further, e.g., into the 6-8 Hz range, the two frequencies became independent again. This suggested that the frequency of the peripheral oscillator is not hard-wired, but rather can be pulled up or down by about 1 Hz with a more dominant oscillator.

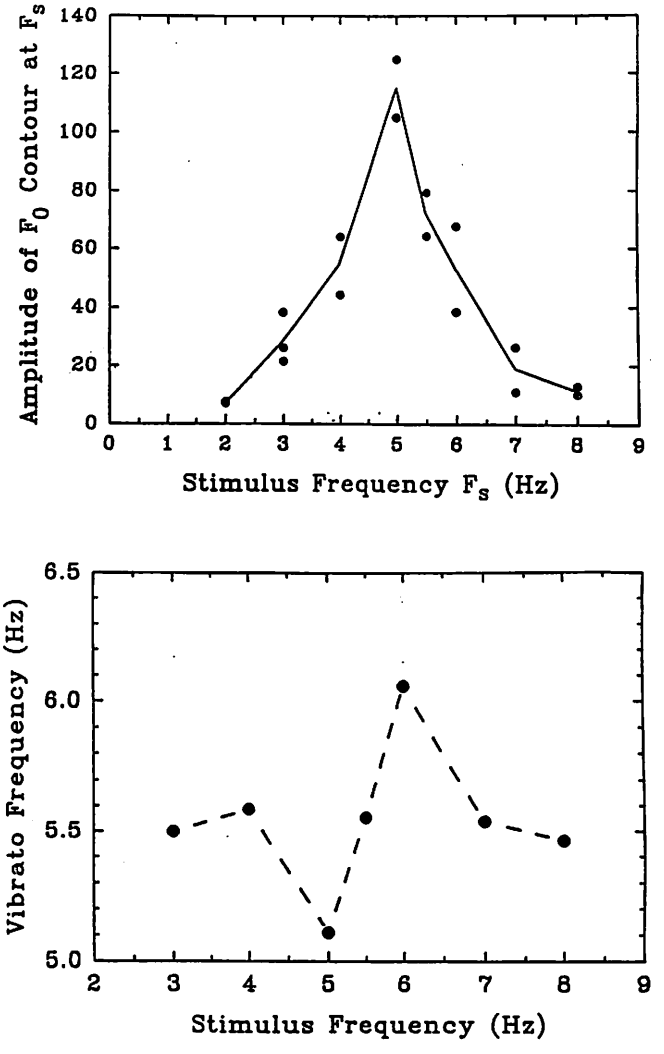


Figure 2 (top). Frequency response curve of a peripheral oscillator apparent in a singer. (From Titze, Solomon, Luschei & Hirano, 1994) Figure 3 (bottom). Vibrato frequency as a function of stimulation frequency, showing entrainment in the 5-6 Hz region. (From Titze, Solomon, Luschei & Hirano, 1994)

The current step in the investigation of vibrato is to determine what some of the mechanical properties of the peripheral oscillator might be, and how they can produce the observed effects.

Characteristics of the Peripheral Oscillator

The mechanism of rotation of the cricoid cartilage relative to the thyroid cartilage is shown in Fig. 4. The force produced by the cricothyroid muscle, F_{CT}, produces a clockwise rotation of the cricoid cartilage about the CT joint. This rotation is quantified by the angle θ. The torque arm for this rotation is w. A counterclockwise torque is produced by the

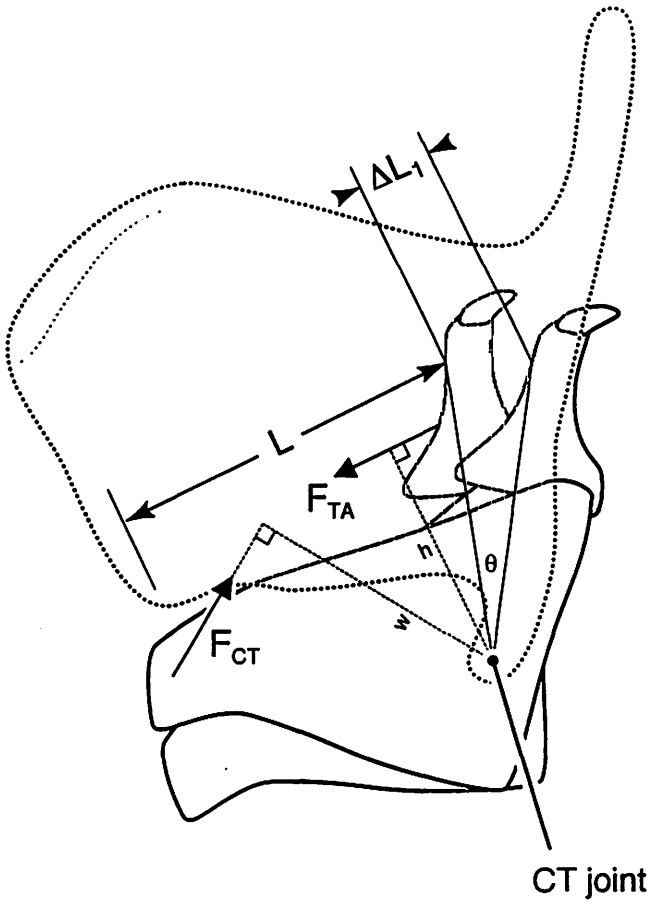


Figure 4. The mechanics of rotation about the cricothyroid joint.

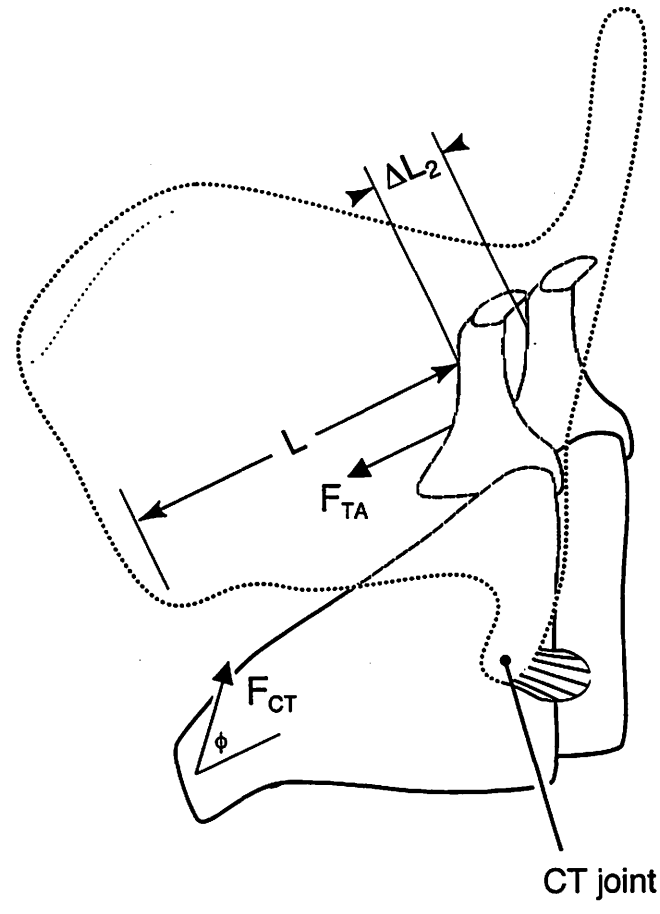


Figure 5. The mechanics of translation (gliding) over the cricothyroid joint.

thyroarytenoid muscle, with a contractile force F_{TA} and a torque arm h. In combination, the two torques produce a length change ΔL_1. The equation of motion for this rotational dynamics can be written as

$$wF_{CT} = hF_{TA} + k_r\theta + d_r\dot{\theta} + I_r\ddot{\theta} \quad . \quad (1)$$

where I_r is the rotational inertia, k_r is the rotational stiffness, and d_r is the rotational damping. The "dot-over" is used to indicate time differentiation. By identifying the rotational strain $\varepsilon_r = h\theta/L_o$, where L_o is the resting length of the vocal folds, equation (1) can be rewritten as

$$wF_{CT} = hF_{TA} + (k_rL_o/h)(\varepsilon_r + t_r\dot{\varepsilon}_r) + (I_rL_o/h)\ddot{\varepsilon}_r \quad (2)$$

where a time constant of the rotational system has been defined as $t_r = d_r/k_r$. This differential equation for the rotational strain ε_r can be solved if the muscle forces F_{CT} and F_{TA} are known.

The translational dynamics of the system are illustrated in Fig 5. We assume that the cricothyroid joint can be displaced from its equilibrium position. In other words, the cricoid cartilage can glide forward or backward on this joint by contraction of either the CT or the TA muscle. The change in vocal fold length produced by this gliding is labeled ΔL_2. The equation of motion can be written as

$$F_{CT}\cos\phi = F_{TA} + k_t(\Delta L_2) + d_t(\dot{\Delta L_2}) + M_t(\ddot{\Delta L_2}) \quad (3)$$

where M_t is the mass, k_t is the stiffness and d_t is the damping for translation. An additional parameter for this mechanical system is now the relative angular orientation $\cos\phi$ between the CT force and the TA force. Defining a second time constraint as $t_t = d/k_t$ and the translational strain as ε_t, the equation becomes

$$F_{CT}\cos\phi = F_{TA} + k_tL_o(\varepsilon_t + t_t\dot{\varepsilon}_t) + M_tL_o\ddot{\varepsilon}_t \quad (4)$$

As before, this equation can be solved for the translational strain if the muscle forces are known.

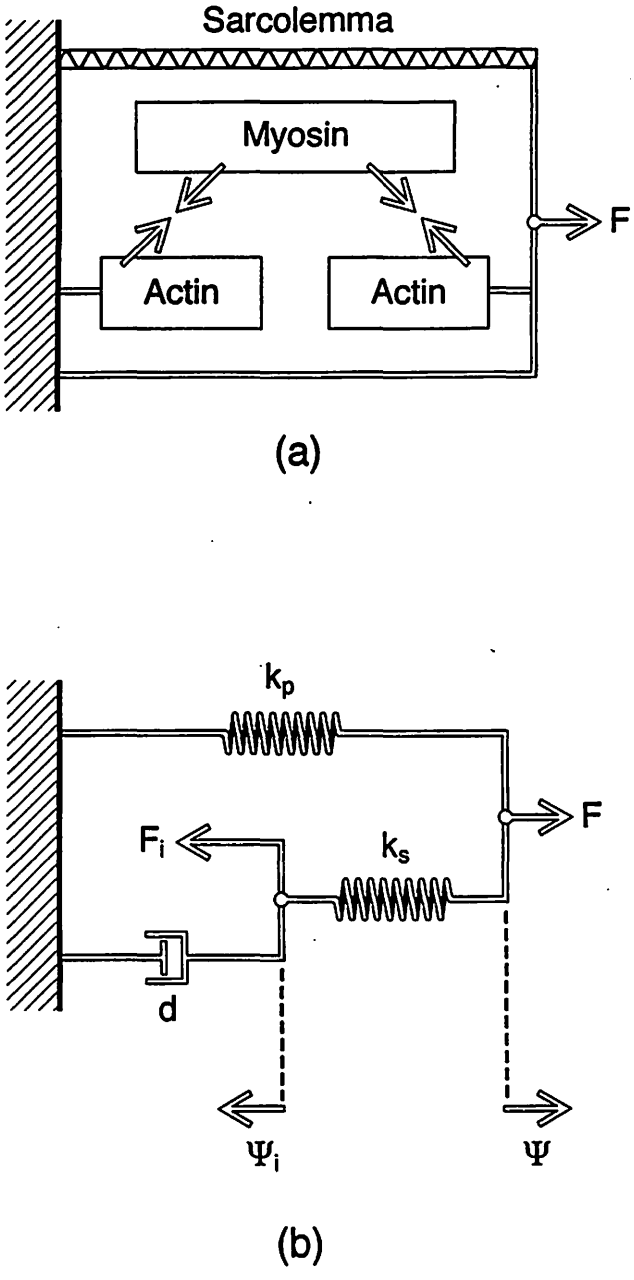


Figure 6. A model of muscle. (a) Sketch of basic sarcomere components, (b) a lumped element mechanical representation.

Lumped Parameter Models for the CT and TA Muscles

A lumped parameter model of muscle tissue is shown in Fig. 6. Part (a) shows a schematic of the sarcomere. According to the sliding filament theory, molecular forces of attraction cause the actin protein molecule to slide in the direction of greater overlap with the myosin molecule, causing a force against the load F. In part (b), the parameter k_p represents the parallel stiffness of sarcolemma, i.e., the connective tissue between muscle fibers; k_s represents the series stiffness within the sarcomere; and d is the internal viscous damping when sliding takes place. The internal contractile force F_i will be discussed later.

Two points are identified in Fig. 6b where Newton's laws of motion can be solved. The first is an internal point, where the displacement of the actin in the sarcomere is ψ_i. At this point the equation of motion is

$$k_s(\psi + \psi_i) + d\dot{\psi}_i = F_i \qquad (5)$$

At the load end of the muscle, the displacement ψ is identified. For this displacement, Newton's law of motion yields

$$k_s(\psi + \psi_i) + k_p\psi = F \quad , \qquad (6)$$

These two equations can be combined by solving for ψ_i in equation 6 and substituting the result into 5. This gives

$$d\frac{k_p + k_s}{k_s}\dot{\psi} + k_p\psi = F - F_i + \frac{d}{k_s}\dot{F} \qquad (7)$$

By defining two new time constants

$$t_p = d\frac{k_p + k_s}{k_p k_s} \qquad (8)$$

and

$$t_s = \frac{d}{k_s} \quad , \qquad (9)$$

equation 7 can be simplified to the following form,

$$k_p(\psi + t_p\dot{\psi}) = F - F_i + t_s\dot{F} \quad . \qquad (10)$$

The equation can now be normalized to include stress and strain variables. If A is the cross-sectional area of the muscle, L_o is the resting length, σ_a is the active stress, and E is the Young's modulus, then

$$F + t_s\dot{F} = A[\sigma_a + E(\epsilon + t_p\dot{\epsilon})] \quad , \qquad (11)$$

where $\varepsilon = \psi/L_o$, $\sigma_a = F/A$, and $E = k_p L_o/A$. Henceforth in the development of the rotational and translational models, this equation will be used to represent the dynamic characteristics of both the CT and TA muscle.

The time course of the activation can be represented by another first-order differential equation,

$$\sigma_a + t_a \dot{\sigma}_a = a\sigma_{am} \quad , \qquad (12)$$

where t_a is the activation time, a is the normalized muscle activity (ranging from 0.0 to 1.0), and σ_{am} is the maximum active stress. In combination, equations 11 and 12 represent a second-order mechanical system for the muscle [5].

To test the model of a muscle, the following paradigms are usually adopted: 1) the stress strain response, 2) the force velocity response, 3) the stress relaxation response after a step strain, 4) the strain creep response after a step stress, and 5) the maximum active isometric stress as a function of strain. We will present only a few of these tests as a validation of the muscle model.

Figure 7 shows the passive stress strain response for a sample of thyroarytenoid muscle (data circles) and for the model (solid curves). Note that this stress strain curve is nonlinear and that the straining process involves energy dissipation, as measured by the area between the rising and falling curves. This hysteresis suggests that the stress is not only a function of strain but also of strain rate. Strain rate is of course quantified by the damping element in the muscle model of Fig. 6b.

Figure 8 shows the responses of the muscle model to step strains and sudden activation. Six different step strains are applied at t = 0.2 s, as shown in part (b). Note how quickly the stress relaxes in part (a). This time constant is less than 0.1 s in this simulation. At t = 2.0 s, the muscle is activated on to 50% of maximum. The response delay is on the order of 0.4 s. Both the relaxation time and the activation time can of course be varied at will with the model parameters.

Table 1 (next page) shows the entire set of parameters selected currently for the CT and TA muscle models and the CTJ dynamics. Note that some of the parameters are constants while some vary with ε, the muscle or vocal fold strain. The most complex parameter so far is the empirical formula for translational strain at the bottom of the table [6]. Also of interest are maximum active stresses and Young's moduli in the CT and TA muscles [7,8]. The constant values are mostly estimates from unpublished data in our laboratory or order-of-magnitude calculations from basic physical principles.

Nonlinear Mechanisms That May Contribute to Oscillation

The major challenge that remains is to determine what nonlinear mechanisms can produce oscillation in this neuromuscular system. The most likely candidate is a slow reflex [9], combined with several mechanical response

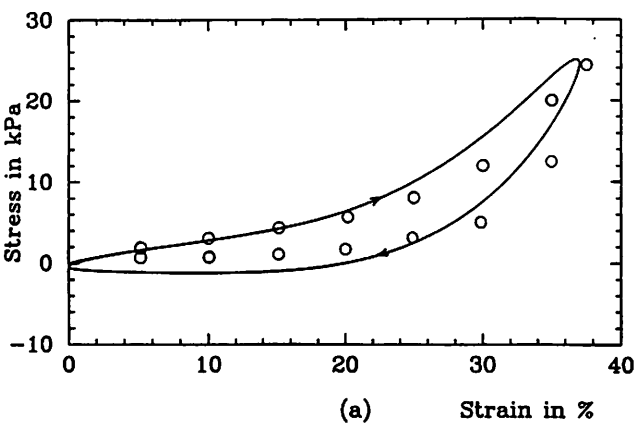


Figure 7. Passive stress-strain curves of muscle, as (a) measured on canine TA, (b) simulated with a model.

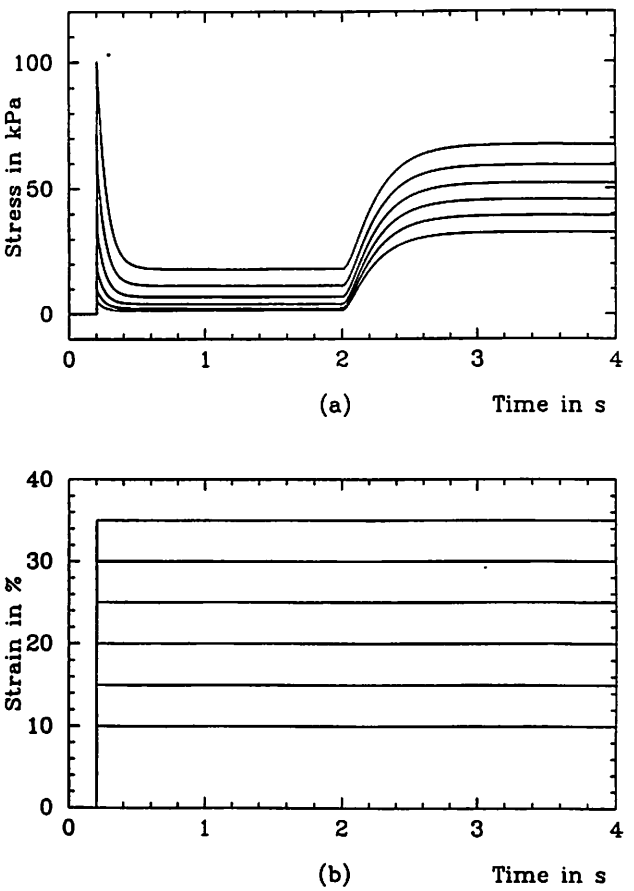


Figure 8. Stress relaxation and muscle activation in a muscle model; (a) the stress response after elongation at 0.1 s and activation at 2.0 s, (b) the step stimulus at six levels of strain.

Table 1.
Parameters for the Biomechanical Model

CT Muscle	Symbol	Value	Unit
Activation time	t_a	.03	s
Series contraction time	t_s	.04+.08ε	s
Parallel contraction time	t_p	.10+.80ε	s
Reflex loop time	t_n	.07	s
Resting length	L_{OCT}	.015	m
Maximum active stress	σ_{am}	100-420 (ε-.4)²	kPa
Young's modulus	E	20-120ε+600ε²	kPa
Cross-sectional area	A	.000078	m²
TA Muscle (Vocal Fold)			
Activation time	t_a	.03	s
Series contraction time	t_s	.04+.08ε	s
Parallel contraction time	t_p	.10+.80ε	s
Reflex loop time	t_n	.07	s
Resting membranous length	L_o	.016	m
Maximum active stress	σ_{am}	100-87.5 (ε-.4)²	kPa
Young's modulus	E	20-120ε+600ε²	kPa
Cross-sectional area	A	.000031	m²
Cricothyroid Joint Dynamics			
Rotational inertia	I_r	.000001	N s²
Translational mass	M_t	.01	kg
Rotational time constant	t_r	.02	s
Translational time constant	t_t	.02	s
TA moment arm	h	.0155	m
CT moment arm	w	.0148	m
Rotational stiffness	k_r	.08185	N-m
Angle between TA and CT	cosϕ	.263	-
Translational stiffness	k_t	500 (1+10⁶$L_o^2\theta^2$)(1+20/πh |ε|)	N/m

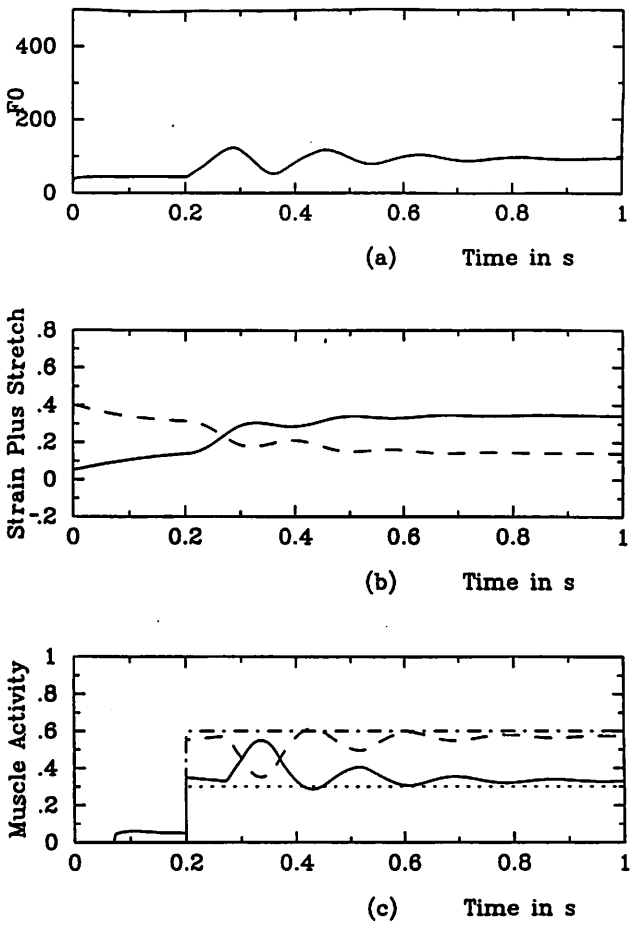


Figure 9. Oscillatory behavior of the combined model of muscle contraction, cricothyroid joint dynamics, and feedback delay in the CT and TA reflex loop. (a) the F_o contour. (b) Strain in the TA muscle (solid) and the CT muscle (dashed). (c) Muscle activities; the solid line is modulated TA activity, the dotted line is unmodulated TA activity, the dashed line is modulated CT activity, and the dot-dashed line is unmodulated CT activity.

latencies in the muscles and CTJ rotation. These combined latencies could produce positive feedback in a negative feedback control loop. A reflex latency of 70 ms, for example, combined with a 30 ms muscle activation time and a 20 ms CTJ rotation time, produced an oscillation on the order of 6 Hz in the model (Fig. 9). The reflex loop was modeled as

$$a = a_o - g\varepsilon(t - t_n) \quad , \qquad (13)$$

where a is the modulated muscle activation after feedback has been applied, a_o is the unmodulated activation, g is a loop gain (chosen to be 0.1 in Fig. 9), ε is the vocal fold strain, and t_n is the neural reflex delay time (70 ms as listed in Table 1).

To explain Fig. 9 further, the calculated F_o variation is shown in part (a). This F_o calculation was based on a model described earlier [10]. In part (b) is shown the variation in vocal fold strain (solid line) and the corresponding variation in the strain of the CT muscle (dashed line). Finally, in part (c) we see the modulated and unmodulated (tonic) muscle activities. For the TA muscle, the chosen unmodulated activity (dotted line) was 30% (0.3) and for the CT muscle it was 60% (0.6). These activities were switched on at t = 0.2 s. Note that the modulated activities oscillate around the tonic values and finally asymptotic toward them when oscillation damps out. Self-sustained oscillation was also achieved with higher loop gain, but we don't know how realistic any of these gains are.

Conclusion

This investigation is incomplete at this point. Several of the subsystems of the vibrato and tremor model have been quantified, but other subsystems are just beginning to

be explored. Hence, no final conclusion can be drawn about the explicit mechanism that produces vibrato or tremor in the larynx. A long reflex latency remains a prime candidate, however, because it does seem to produce the type of resonance measured as a human subject. We expect that better answers will be obtained once the entire system has been quantified adequately. In particular, we need to know more about the mechanical response times of the muscles and the loop gain of the late reflex [9].

Acknowledgements

This work was supported by Grant No. P60 DC00976 from the National Institute on Deafness and Other Communication Disorders.

References

1. G. Fant *Acoustic Theory of Speech Production*, (The Hague: Mouton & Company 1960)

2. L. Ramig and T. Shipp "Comparative measures of vocal tremor and vocal vibrato" Journal of Voice 1(2), 163-167 (1987)

3. H. Herzel, D. Berry, I.R. Titze and M. Saleh "Analysis of vocal disorders with methods from nonlinear dynamics" Journal of Speech and Hearing Research 37(5), 1001-1007 (1994)

4. I.R. Titze, N.P. Solomon, E.S. Luschei and M. Hirano "Interference between normal vibrato and artificial stimulation of laryngeal muscles at near vibrato rates" Journal of Voice 8(3), 215-223 (1994)

5. R.B. Stein and M.N. Oguztöreli "Does the velocity sensitivity of muscle spindles stabilize the stretch reflex?" Biological Cybernetics, 219-228 (1976)

6. E. Vilkman, R. Pitkänen and H. Suominen "Observations on the structure and the biomechanics of the cricothyroid articulation" Acta Otolaryngology (Stockholm) 103, 117-126 (1987)

7. F. Alipour-Haghighi, I.R. Titze and A. Perlman "Tetanic contraction in vocal fold muscle" Journal of Speech and Hearing Research 32, 226-231 (1989)

8. F. Alipour-Haghighi, A. Perlman and I. Titze "Tetanic response of the cricothyroid muscle" Annals of Otology, Rhinology & Laryngology 100(8), 626-631 (1991)

9. C. Ludlow, F. VanPelt and J. Koda "Characteristics of late responses to superior laryngeal nerve stimulation in humans" Annals of Otology, Rhinology and Laryngology 101(2), 127-134 (1992)

10. I.R. Titze, E.S. Luschei and M. Hirano "The role of the thyroarytenoid muscle in regulation of fundamental frequency" Journal of Voice, 3(3), 213-224 (1989)

NCVS Status and Progress Report - 8
July 1995, 33-41

Mechanisms of Jitter-Induced Shimmer in a Driven Model of Vocal Fold Vibration

Darrell Wong, Ph.D.
Robert Lange, Ph.D.
Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts
Ingo Titze, Ph.D.
Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts
Department of Speech Pathology and Audiology, The University of Iowa
Chwen Geng Guo, M.S.
Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts

Abstract

The presence of shimmer (changes in the cycle-to-cycle amplitude) in frequency modulated synthetic vowels (used to simulate jitter phenomena) has often been attributed to time-aliasing. Time aliasing occurs when adding aperiodic overlapping impulse responses. The resulting cycle becomes distorted relative to previous cycles, resulting in measurable shimmer. As a consequence, whenever jitter is present, any measurement of shimmer is the sum of this jitter-induced shimmer plus the inherent amplitude modulation from the vocal fold oscillations. In this study mechanisms of jitter-induced shimmer other than time-aliasing are considered. The behaviour of a time domain simulation model that generates vocal fold motion and the pressure and flow signals generated in the vocal tract was observed. The simulations used a driven model of vocal fold tissue displacement, permitting the characterization of sources of shimmer induced by frequency modulating the displacement. It was found that shimmer appears as a result of slope changes in the glottal area and glottal flow waves, and from direct amplitude modulation of the maximum tissue displacement. These jitter-induced amplitude perturbations then propagate through the vocal tract. In this paper, we discuss these mechanisms, and also the effect of wave propagation speed. A self-oscillating mass-spring model is also simulated for comparison.

Introduction

Measurements of jitter and shimmer on voice signals are obtained for the purpose of discerning perturbations in the oscillatory behaviour of the vocal folds. Jitter is defined as the average cycle-to-cycle change in the fundamental period length, and shimmer is the average cycle-to-cycle change in amplitude. Although the fundamental period length is well defined, there is no generally accepted definition of amplitude for a complex signal. Traditional definitions rely on specific cyclic events in time, such as the largest positive peak-to largest negative peak, or the largest negative peak value. Other definitions include the root-mean-squared (RMS) value of each cycle (given that the fundamental period markers have been correctly placed) used by Kempster and Kistler (1984) and Hillenbrand (1987), or the gain factor defined by Milenkovic (1987).

All of these amplitude definition and extraction algorithms work well with amplitude modulated (AM) voicing signals, reporting linear increases in shimmer as the extent of the AM is increased. Their behaviour under frequency modulation (FM), however, is nonlinear. This nonlinearity is often attributed to time-aliasing (Qi, 1994); (Oppenheim, 1989). The phenomenon has been observed by both Hillenbrand and Milenkovic and found to vary with vowel and F_0. The concept of time-aliasing stems from the impulse driven nature of the source-filter model of speech. The vocal tract is modeled as a linear filter excited by an impulse train. From cycle to cycle, preceding vocal tract impulse responses overlap and add to the current cycle.

When the train is aperiodic, the changing phase relationships between consecutive impulse responses cause the signal to become distorted relative to previous cycles, resulting in measurable shimmer. Since all of the amplitude definitions defined above extract measurements at a stage where the time aliasing has already occurred, they all suffer from its effects.

As a consequence, whenever jitter is present, any measurement of shimmer cannot be solely attributed to amplitude modulation of the vocal fold oscillations. It would thus be useful to devise techniques that can ignore or overcome aperiodic time aliasing as a shimmer source, or else identify and analyze other signals that more accurately reflect the fold behaviour.

In this study the second alternative is pursued. We examined the behaviour of a time domain simulation model that generates signals describing vocal fold motion and the resulting pressure and flow signals generated throughout the vocal tract. This enabled us to characterize possible sources of shimmer induced by the FM modulation of a driven model of vocal fold tissue displacement. It was found that vocal tract time aliasing is only one of a number of sources of shimmer originating from different mechanisms and affecting various signals. Shimmer also appears to be a result of the nonlinear interactions between the transglottal acoustic pressure, mucosal wave velocity, and tissue displacement. These jitter-induced amplitude perturbations then propagate to the supraglottal pressure P_{in}, the subglottal pressure P_s, the minimum glottal area A_g, the glottal flow U and its time derivative (dU/dt in this paper), and, finally, the output pressure from the vocal tract P_0. In this paper, we study the possible mechanisms for these perturbations.

Method and Model

An interactive computer simulation of the vocal fold and vocal tract system has been used to model the behaviour of the folds under conditions of FM subharmonic modulation of the vocal fold tissue displacement. Subharmonic modulation of order 1/2 was chosen so that any changes in the plots could be observed by inspection. An FM extent level of 30% was chosen to exaggerate the effects, although it is acknowledged that this is likely to be much greater than values typically found in human phonation. The model, SPEAK-model 2, was developed at the University of Iowa by one of the authors. It incorporates source-tract interaction, but does not incorporate self oscillation. Instead, there is direct control of tissue displacement by a mathematically described driving function. Empirical relationships previously described in the literature involving F_0, vocal fold length and thickness, mucosal wave velocity, lung pressure, and transglottal pressure are used to define the behaviour of the system. The effects of modulation are then demonstrated as mechanisms supported by these equations. Figures from the simulations are used to illustrate the phenomena.

Modeling Equations

An early description of the model was given by Titze (1984) and a brief outline is given here. Consider the top view of the vocal folds (Figure 1). The vocal processes are at $x = +/-\zeta_0$, at $y = 0$. The lowest mode of vibration occurs as a half-sinusoid in the y-direction. When ζ_0 is small, complete closure can occur, while for large values, a glottal chink causing flow leakage occurs (the parameter h in Figure 1 indicates the height of the chink).

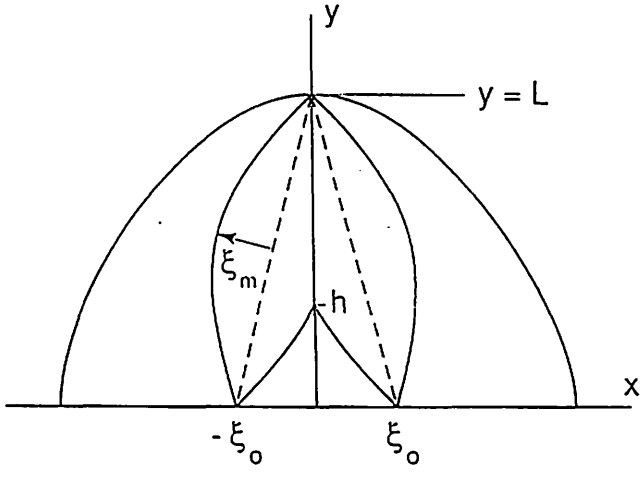


Figure 1. Top view of vocal fold model. From (Titze 1994).

For the simple case without jitter, the edges of the vocal folds oscillate sinusoidally in time with an angular frequency $\omega = 2\pi F_0$. The glottal width at any point on the y axis (as the folds move outward) is defined as

$$g(y,t) = 2[\zeta_0(1-y/L) + \zeta_m \sin(\omega t)\sin(\pi y/L)] \qquad (1)$$

where L is the vocal fold length, and ζ_m is the maximum vocal fold displacement (at $y = L/2$). The first term in (1) is the prephonatory (static) glottal width.

To make the model respond in a manner similar to the human folds, empirical relationships extracted from the literature have been utilized. For example, it is known that ζ_m is dependent on lung pressure and F_0. The following rule is adopted from Titze (1988):

$$\zeta_m = 17.4 P_L^{0.5} F_0^{-1.6} \qquad (2)$$

where P_L is the lung pressure.

Figure 2 shows a three dimensional view of the folds. A z-axis has been added to describe the motion of the mucosal wave as it travels from the bottom to the top of the

folds. If the vertical dimension is sliced into layers, each layer k can have its own value of initial condition ζ_{0k} and maximum displacement ζ_{mk}. The propagation of the mucosal surface wave can then be obtained by replacing sin(ωt) in equation (1) with sin(ω(t-z/c)), where z is the vertical point under consideration and c is the mucosal wave velocity. If the vertical axis is sliced into N layers, the phase between layers is assumed to vary linearly between layers from the bottom (k=1) to the top (k=N) of the z-axis:

$$\phi_k = \omega(t-z/c) = \omega(t-kT'/(Nc)) \quad (3)$$

where ϕ_k replaces ωt in equation (1) and T' is the thickness from bottom to top.

Titze, Jiang and Hsiao (1993) conducted an experiment to measure c. They examined the motion of two sutured points on a canine hemi-larynx placed in the vertical dimension. The time required for the displacement of the superior suture to 'catch up' to the displacement of the inferior suture was measured by observing the time taken to pass a certain point (i.e., the arrival time with respect to a coordinate along the transverse dimension). This time was interpreted as an indication of the wave propagation speed. In the results section of the paper, the authors discuss the fact that the inter-suture distance (and the overall thickness of the folds) may vary during vibration as a function of F_0.

We present an equation which takes this thickness variation into account:

$$c/T' = \underline{a}F_0/T_0 \quad (4)$$

where T_0 is the nominal thickness at rest, T' is the thickness during vibration for a given F_0, and $\underline{a}$ is a constant parameter of value 0.01 meters (Titze, unpublished data).

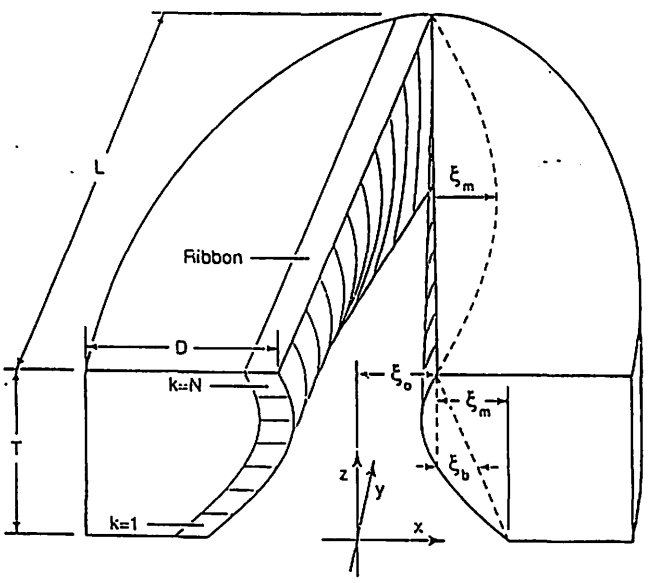


Figure 2. Three dimensional view of vocal folds. From (Titze 1994).

From the above driving equations, the glottal width g at any point (y,k) may be obtained:

$$g(y,k,t) = 2[\zeta_{0k}(1-y/L) + \zeta_{mk}\sin(\phi_k)\sin(\pi y/L)] \quad (5)$$

The glottal area between the symmetric folds at any layer k can be calculated by integration, and the minimum glottal area A_g, observable by a photoglottogram, can be estimated by finding the minimum area layer at each point in time.

The glottal flow is determined by the subglottal pressures P_s, and the supraglottal pressure P_{in}, the glottal area A_g, the density ρ, and an empirically determined pressure coefficient k_t:

$$P_s - P_{in} = k_t\rho|U|U/A_g^2 \quad (6)$$

An /a/ vocal tract shape is modeled using wave reflection equations (Liljencrants, 1985; Story, 1995).

Frequency Modulation

If F_0 is sinusoidally modulated, then the instantaneous fundamental frequency can be written as:

$$F_0' = F_0(1 + E\cos(2\pi F_m t)) \quad (7)$$

where F_0 is the unmodulated fundamental frequency. In equation (1), g(y,t) now becomes frequency modulated at a modulation frequency F_m and an extent E, proportional to F_m/F_0. Figure 3 illustrates the basic shape of the modulated displacement when F_0 = 125 Hz, F_m = 62.5 Hz, and E = 0.2. The effects of Equation 7 propagate through the model in the following ways:

Modulation of the Mucosal Wave Velocity

The mucosal wave velocity c is modulated because it varies with F_0' according to Equation 4. This modulation effect then propagates into Equations 3 (where the parameter ù becomes modulated) and 5 (where the parameter ϕ_k becomes modulated). Consider the signal in Figure 3. It might represent the oscillation of the bottom layer (k=1) of the folds. If a constant time delayed version of this signal is used to represent the top layer (k=N) of the vocal folds,

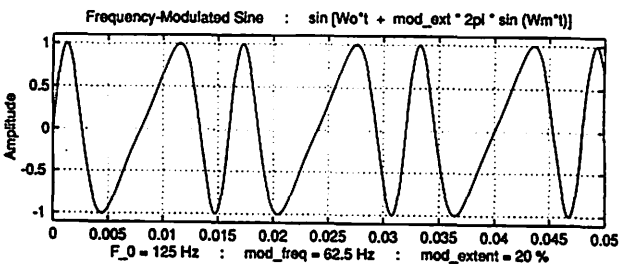


Figure 3. Subharmonic Frequency Modulated Sinusoid. F_0 = 125 Hz, F_m = 62.5 Hz, E = 20%

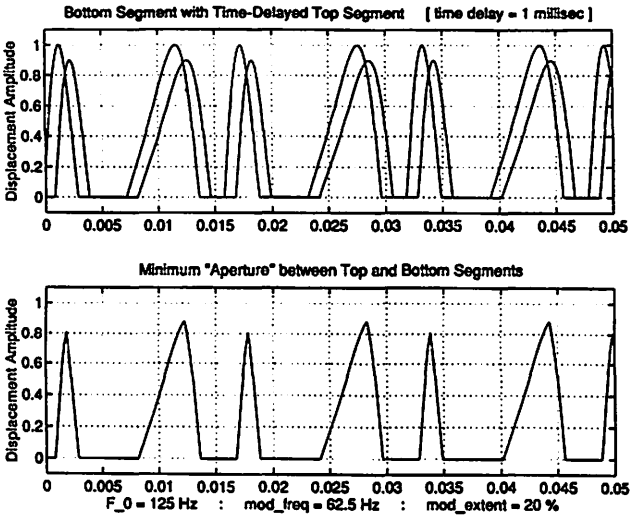


Figure 4. Time delayed FM subharmonic (1/2) modulated sine waves and the minimum aperture wave result.

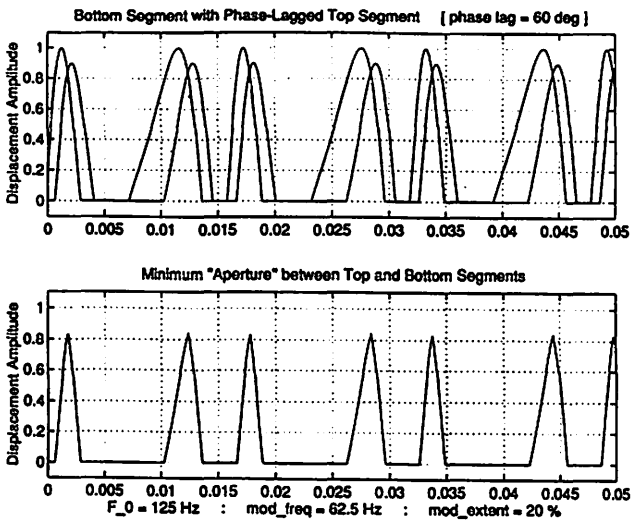


Figure 5. Phase delayed FM subharmonic (1/2) modulated sine waves and the minimum aperture wave result.

amplitude variations will occur in the minimum glottal area wave A_g. These variations in the minimum glottal area are depicted in Figure 4 (described as minimum aperture). Such peak-to-peak variations in the minimum aperture appear regardless of the relative amplitudes of the waves. However, consider the top of the folds to be oscillating with a constant phase delay relationship. The minimum glottal area exhibits no amplitude variations here (Figure 5). It is shown in Appendix 1 that the minimum glottal area in Figure 5 will always be free of amplitude variations, regardless of the relative amplitudes of the top and bottom displacements provided that a constant phase lag is maintained between the layers and that each of the contributing displacements exhibit no amplitude variations independently.

In our vocal fold model, if c is proportional to the 'nominal' value F_0, it can be shown analytically that a constant time delay mucosal wave will be observed (see Appendix 2). However, if c is instead proportional to the instantaneous value of fundamental frequency F_0', it is also shown in Appendix 2 that a constant phase delay relationship will be produced. Studies on the mucosal wave e.g. (Titze 1989) report a constant phase delay, although an FM situation is not considered. It remains to be demonstrated which type of delay - constant time, constant phase, or neither - actually occurs for FM modulations of the vocal folds.

Figure 6 demonstrates what happens in the driven vocal fold model when the nominal average value of F_0 is used for both the maximum vocal fold displacement and the speed of mucosal wave propagation. The displacement signal x (the glottal width g(y,k,t) for y = L/2 and k = 21 [the top layer]) exhibits FM behaviour, but no peak amplitude perturbation. The glottal area waveforms (nonminimum) for three layers on the z axis (bottom (k=1), middle (k=10), and top (k=21)) show a similar result to Figure 4. The FM subharmonic in x produces amplitude perturbations in minimum glottal area A_g, the flow U, the flow derivative dU/dt and on into other signals in the vocal tract. The results indicate shimmer in A_g due to a constant time delay, as predicted earlier. Note that the contact area CA does not exhibit any amplitude perturbations, and that the x and A_{g1},

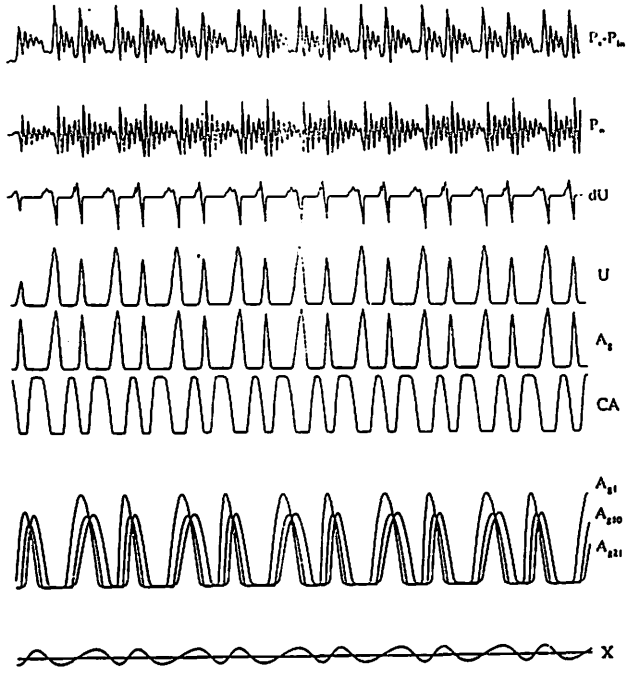


Figure 6. SPEAK generated subharmonic (1/2) array of signals. Assumes F_0 is constant for Equations 2 and 4, causing a constant time delay between displacements.

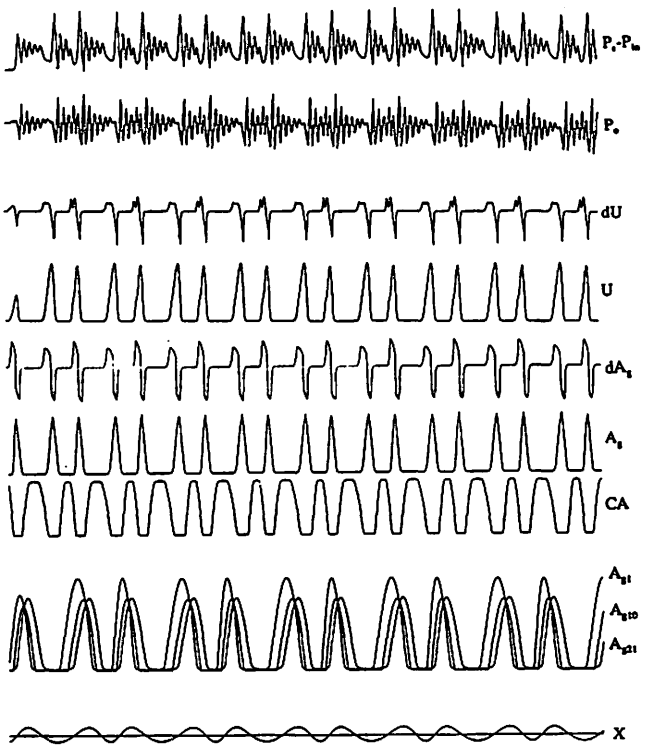


Figure 7. SPEAK generated subharmonic (1/2) array of signals. Assumes F_0 is constant for Equation 2, and varies for Equation 4, causing a constant phase delay between displacements.

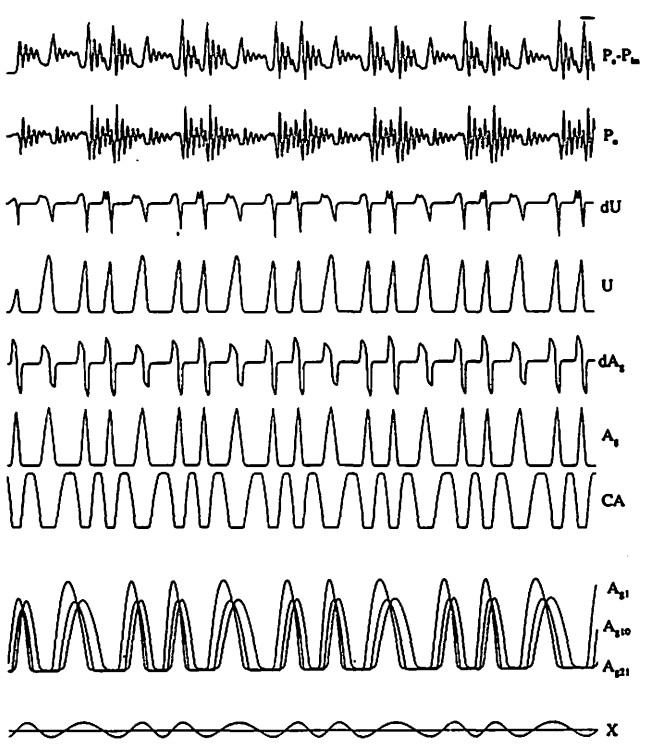


Figure 8. SPEAK generated subharmonic (1/3) array of signals. Assumes F_0 is constant for Equation 2 and varies for Equation 4, causing a constant phase delay between displacements, but uses 1/3 subharmonic to generate displacement slope changes.

A_{g10}, and A_{g21} plots are on an expanded time scale relative to the other plots, so that the effect of overlapping A_g waves can be seen (this applies to Figures 6 through 10).

Figure 7 demonstrates the situation when the instantaneous F_0' is used for the mucosal wave speed. A constant vibrational amplitude is assumed (time varying ζ_m will be discussed later). The three glottal area waves induce no visible minimum aperture amplitude variation in A_g or CA, as predicted by the previous discussion. The glottal flow U, however, does exhibit amplitude variations, which then propagates into dU/dt and on into the vocal tract. The cause of this variation is discussed next.

Transglottal Pressure, Glottal Area, and Glottal Flow Signal Slopes

The glottal flow U in Figure 7 demonstrates amplitude perturbation even though A_g does not. This occurs because of a combination of two effects. First, from Equation 6, U is related to A_g and the transglottal pressure P_s-P_{in}. P_{in} is affected by vocal tract inertance, causing the flow to skew to the right (i.e. the peak in U is delayed relative to the peak in A_g). Second, while the peaks of A_g are the same height, the negative closing slope varies due to the FM modulation. This is illustrated in dA_g in Figure 7 (although the negative peaks in dA_g vary only slightly in amplitude, it is enough to modulate the peak flow). As a result, the value of A_g at the instant of peak flow will vary, causing changes in U via Equation 6.

This phenomena can be seen in Figure 8, in which the model is modulated with a 1/3 FM subharmonic. Again, A_g exhibits no amplitude perturbation, while U, dU/dt and subsequent propagated signals exhibit amplitude changes. The slopes of A_g (or equivalently, the negative peaks in dA_g) clearly influence U, which in turn influences dU/dt. The negative slopes in both U and A_g appear to be the determinants of peak magnitudes in the vocal tract pressure signals.

Modulation of ζ_m

If ζ_m were constant, the glottal width equation (1) would vary sinusoidally for constant F_0. When a time varying relationship between ζ_m and F_0' is assumed for Equation 2 (again assuming that these equations apply to dynamically modulated conditions), direct amplitude modulation of the maximum tissue displacement occurs. Figure 9 assumes a constant phase delay (F_0' used in Equation 4) and dynamically varying ζ_m. The displacement x shows both amplitude and frequency modulation. As a result, amplitude perturbation appears in all other signals (except for CA) in Figure 9.

Discussion

This study has identified several sources of shimmer. One source is the modulation of the maximum amplitude of vibration (ζ_m), which directly influences the amplitude of the glottal areas. A second source is the slope of the

minimum glottal area (A_g), which determines the peak in the flow wave (in conjunction with the inertive load of the vocal tract). It should be noted that it is the slope of the flow wave, not the peak, which is closely tied to the excitation of the vocal tract pressure wave (P_0). Often it is this peak in P_0 that is marked and measured in voice waveform analysis.

It is necessary at this point to consider whether c and ζ_m are likely to vary instantaneously with F_0' or, as an alternative, they are likely to be proportional to the nominal F_0 instead. In the driven model, c and ζ_m are empirically related to F_0, and therefore, F_0'. In self oscillating models using multi-mass-springs or finite elements, the relationship of c and ζ_m to F_0 is not explicit, but occurs as a natural consequence of the tissue stiffness parameters.

Two alternative scenarios to the proposed mechanisms remain to be considered: (1) the mechanisms of shimmer when c and ζ_m do not vary with F_0' in the driven model, and (2) the mechanisms of shimmer in a self oscillating model.

Consider the first case. If c is proportional to the nominal F_0, a constant time delay system between the top and bottom of the folds results, producing shimmer in the A_g peaks (demonstrated in Figure 4). The amplitude of vibration ζ_m, however, will be constant, and will thus disappear as a source of shimmer. Therefore, regardless of whether F_0 or F_0' is used in the empirical relationships, some form of jitter-induced shimmer occurs. In addition, the shimmer in the flow and flow derivative due to changing A_g slopes and inductive vocal tract loading remains.

For a self oscillating multiple-mass-spring model, F_0 is determined approximately by the equation $F_0 = \sqrt{(K/m)}/2\pi$ where K is an approximation to the overall tissue stiffness. K is often approximated as a nonlinear cubic stiffness i.e. the force contribution from K is proportional to the cube of the displacement. In a human vocal fold, changes in the stiffness of the muscle can occur due to dynamic muscle tensioning, where the cubic stiffness relationship increases the force contribution greatly as the tissue is displaced. The mass-spring model developed by Wong (1991) exhibited 1/2 order subharmonics when the stiffness was decreased asymmetrically. This was likely due to the nonlinear cubic response of the spring (Minorsky, 1962), or to the left/right vocal fold asymmetry (Herzel, 1995).

We have also used the three mass body-cover model developed by Story and Titze (1994), and simulated a case of bilateral vocal fold paralysis. This was done by decreasing the stiffnesses symmetrically by a factor of 10 from the normal case. The results appear in Figure 10. An amplitude and F_0 subharmonic can be observed. The first plot represents the speech output P_0, and the second plot shows the upper and lower cover displacements (the lower displacement leads the upper). Note the changing peak amplitudes from cycle to cycle in the displacements. This is analogous to the ζ_m modulation in the driven model. A numerical time/phase delay analysis of the displacement signals (Lange et al., 1995) indicates that the relationship is neither purely time delayed nor purely phase lagged. This would suggest that c is neither a constant nor a strict linear function of F_0. Although A_g is not plotted, it can be approximated as the minimum between the two displacements

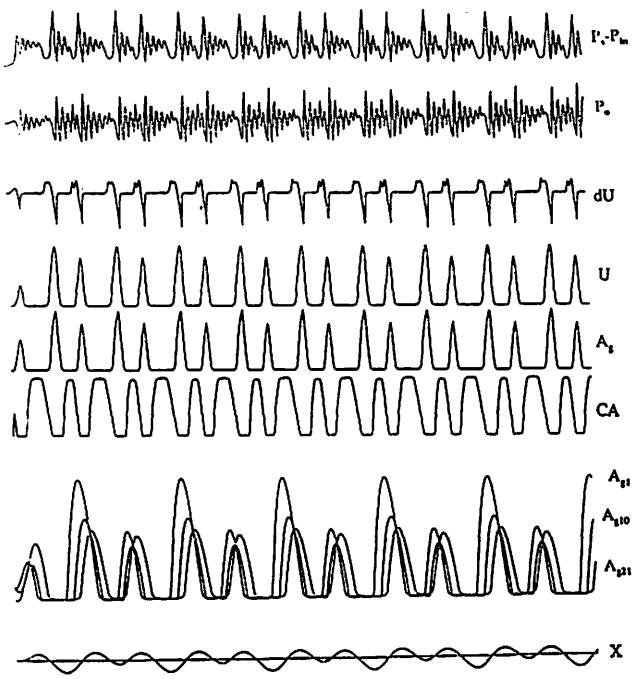


Figure 9. SPEAK generated subharmonic (1/2) array of signals. Assumes F_0' (time varying) is used for Equations 2 and 4, causing a constant phase delay between displacements, and amplitude modulation of the displacement.

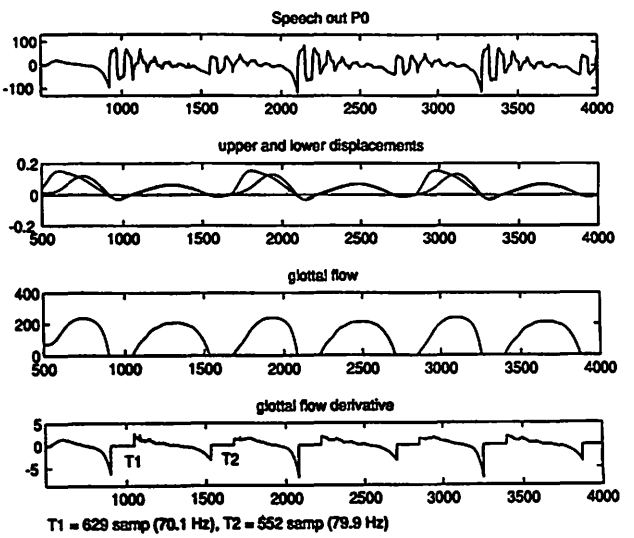


Figure 10. Results from the 3 mass body-cover self-oscillating model of the vocal folds (Story, 1995). Shown are Speech output P_0, displacements, glottal flow, and flow derivative.

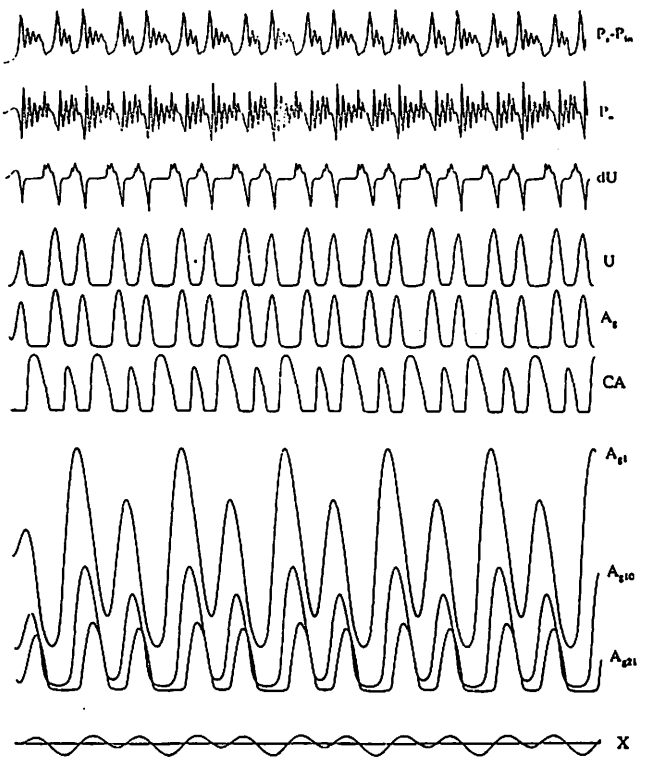


Figure 11. SPEAK generated subharmonic (1/2) array of signals. Assumes F_0' (time varying) is used for Equations 2, and 4, causing a constant phase delay between displacements, and amplitude modulation of the displacement. Lower folds do not close. CA exhibits amplitude shimmer.

above the zero axis. From this approximation and the glottal flow wave, one can identify the A_g peaks and slopes as sources of shimmer. The flow slopes also become a source of shimmer in the flow derivative signal.

Returning to the SPEAK simulation, it is interesting to note that the contact area CA is amplitude insensitive to FM. CA can be measured with the electroglottograph, and it would be useful to know whether the CA demonstrates shimmer or not. It is a measure of the electrical conductivity measured from one side of the folds to the other. In all the examples that have been given, complete closure was achieved (the z layers were preconfigured so that the initial ζ_{0k} were close together, causing CA to reach a maximum for all cycles). Figure 11 illustrates the case where this assumption is relaxed, resulting in incomplete closure for some of the layers in the z-axis. As a result, CA exhibits amplitude perturbation. It should thus be noted that the ability of CA to demonstrate glottal displacement amplitude perturbation is limited due to 'saturation', because it is inherently a measure of behaviour during collision rather than glottal opening.

Since the electroglottograph does not directly reflect maximum tissue displacement, it may be an unreliable indicator of shimmer, especially if complete closure of the glottis is maintained.

Summary

The interactions among several voice production variables have been qualitatively described for a driven model of vocal fold displacement. The mechanisms identified here suggest that shimmer in the glottal flow U occurs from slope changes in A_g, while direct modulation of the peak tissue displacement ζ_m will affect A_g.

The observations made here suggest further study. The simple model we have studied drives the tissue displacement explicitly using predetermined fundamental frequency behaviour. While the mechanism due to slope changes in A_g is likely to be common to all classes of vocal fold models, the empirical equations relating ζ_m and c to F_0 are not used in self-oscillating models, since F_0 is not directly controlled. An examination of the three mass body-cover self-oscillating model has been attempted, in which the stiffness of both folds was decreased. The results suggest that our observations regarding ζ_m and A_g shimmer in the driven model are confirmed by the self-oscillating model.

Acknowledgment

This work has been funded by the National Institutes of Health, Grant #DC 00387-08.

Appendix 1

Two sinusoidal functions which differ by a constant phase lag will intersect each other at a constant value (a constant "height" $\underline{y} = \underline{y}_0$).

Define:

$\underline{f}_1(\alpha)$	$\equiv$	$\underline{a}\sin(\alpha)$
$\underline{f}_2(\alpha)$	$\equiv$	$\underline{b}\sin(\alpha+\phi)$
α		is an arbitrary phase angle.
$(\underline{a},\underline{b})$		are the coefficients on the sinusoid which yield arbitrary magnitudes in $\underline{f}_1$ and $\underline{f}_2$.
ϕ		is the constant phase lag between $\underline{f}_1$ and $\underline{f}_2$.

α may be a function of $[\omega\underline{t}]$; wherein, ω is a constant frequency of arbitrary value. α may also be a function of $\omega(\underline{t})$; wherein, the sinusoidal frequency is modulated continuously over time. In other words, $\underline{f}_1$ and $\underline{f}_2$ are restricted to being constant maximum-amplitude oscillations, but they may exhibit jitter.

Let $(\alpha = \alpha_0)$ be a particular phase angle where $\underline{f}_1$ and $\underline{f}_2$ intersect.

Then

$$\underline{f}_1(\alpha_0) = \underline{f}_2(\alpha_0) = \underline{y}_0$$
$$\underline{a}\sin(\alpha_0) = \underline{b}\sin(\alpha_0+\phi) = \underline{y}_0$$
$$\underline{a}\sin(\alpha_0) = \underline{b}\sin(\alpha_0)\cos(\phi) + \underline{b}\cos(\alpha_0)\sin(\phi)$$

Thus

$$\sin(\alpha_0)\,[\underline{a} - \underline{b}\cos(\phi)] = \underline{b}\cos(\alpha_0)\sin(\phi)$$

and

$$\sin(\alpha_0) / \cos(\alpha_0) = \underline{b}\sin(\phi) / [\underline{a} - \underline{b}\cos(\phi)]$$

ϕ, $\sin(\phi)$, $\cos(\phi)$, $\underline{a}$, and $\underline{b}$ are all constants. Therefore, define:

$$\underline{k} \equiv \underline{b}\sin(\phi) / [\underline{a} - \underline{b}\cos(\phi)]$$
$$\tan(\alpha_0) = \underline{k}$$
$$\alpha_0 = \tan^{-1}(\underline{k})$$

This means that $\underline{f}_1$ and $\underline{f}_2$ intersect at the phase angle ($\alpha = \alpha_0$) and the value ($\underline{y} = \underline{y}_0$):

$$\underline{f}_1(\alpha_0) = \underline{f}_2(\alpha_0) = \underline{y}_0 = \underline{a}\sin[\tan^{-1}(\underline{k})]$$

But

$$\sin(\alpha) = \sin(\alpha + \underline{n}2\pi)$$

where

$$\underline{n} = 0, 1, 2, \ldots$$

Therefore,

$$\underline{f}_1(\alpha_0) = \underline{a}\sin(\alpha_0) = \underline{a}\sin(\alpha_0 + \underline{n}2\pi) = \underline{f}_1(\alpha_0 + \underline{n}2\pi)$$
$$\underline{f}_2(\alpha_0) = \underline{b}\sin(\alpha_0 + \phi) = \underline{b}\sin(\alpha_0 + \underline{n}2\pi + \phi) = \underline{f}_2(\alpha_0 + \underline{n}2\pi)$$

Thus

$$\underline{f}_1(\alpha_0 + \underline{n}2\pi) = \underline{y}_0 = \underline{f}_2(\alpha_0 + \underline{n}2\pi)$$

In other words, $\underline{f}_1$ and $\underline{f}_2$ intersect at many additional points. Each point of intersection is separated by a phase angle ($\pm 2\pi$). Note that the value of the intersection (its height) is always the same:

$$\underline{y}_0 = \underline{a}\sin[\tan^{-1}(\underline{k})]$$

And the phase angle of coincidence is governed by:

$$\alpha_0 = \tan^{-1}(\underline{k})$$

Furthermore,

$$\sin(\alpha + \pi) = -\sin(\alpha)$$

So that

$$\underline{f}_1(\alpha_0 + \pi) = -\underline{a}\sin(\alpha_0) = -\underline{y}_0$$
$$\underline{f}_2(\alpha_0 + \pi) = -\underline{b}\sin(\alpha_0 + \phi) = -\underline{y}_0$$

This means that $\underline{f}_1$ and $\underline{f}_2$ intersect again at an another set of points whose value is constant:

$$\underline{f}_1(\alpha_0 + \pi + \underline{n}2\pi) = -\underline{y}_0 = \underline{f}_2(\alpha_0 + \pi + \underline{n}2\pi)$$

If the set of points intersecting at the value $\underline{y}_0$ appear in the upper half-plane, then their counterparts (intersecting at the value $-\underline{y}_0$) appear in the lower half-plane.

Appendix 2

Consider the tissue displacement equations ζ_t and ζ_b for the top and bottom of the fold:

$$\zeta_b = \zeta_{0b} + \sin(\omega_0' t) \qquad \text{A2.1}$$
$$\zeta_t = \zeta_{0t} + \sin(\omega_0'(t - T'/c)) \qquad \text{A2.2}$$

where ζ_{0b} and ζ_{0t} are the initial displacements, ω_0' is the fundamental frequency (possibly time varying), t is the time variable, T' is the variable vertical thickness, of the fold and c is the velocity of the traveling wave.

If F_0 is assumed constant, $c = \underline{a}F_0 T'/T_0$, then a constant time delay equation replaces A2.2:

$$\zeta_t = \zeta_{0t} + \sin(\omega_0'(t - T_0/\underline{a}F_0)) \qquad \text{A2.3}$$

On the other hand, if F_0 (and therefore c) is assumed time varying, then $c = \underline{a}F_0' T'/T_0$ or $c = \underline{a}\omega_0' T'/2\pi T_0$, and A2.2 becomes a constant phase lag equation:

$$\zeta_t = \zeta_{0t} + \sin(\omega_0' t - T'\omega_0' 2\pi T_0/(T'\underline{a}\omega_0'))$$
$$= \zeta_{0t} + \sin(\omega_0' t - 2\pi T_0/\underline{a}) \qquad \text{A2.4}$$

Bibliography

Hillenbrand, J. (1987). A methodological study of perturbation and additive noise in synthetically generated signals. Journal of Speech and Hearing Research, 30 (4), 448-461.

Kempster, G.B., & Kistler, D.J. (1984). Perceptual dimensions of dysphonic voices. Journal of the Acoustical Society of America, 75 (Suppl. 1) S8 (A).

Lange, R., Wong, D., Long, R., & Titze, I.R. (1995, Oct.). Vocal fold movies and voice transformations, to be presented at the 1995 Matlab Conference, Boston.

Liljencrants, J. (1985). Dynamic line analogs for speech synthesis. Quarterly Progress and Status Report, STL-QPSR, 1/1985, 1-14. Speech Transmission Laboratory, Royal Institute of Technology (KTH), Stockholm, Sweden.

Milenkovic, P. (1987). Least mean square measures of voice perturbation. Journal of Speech and Hearing Research, 30 (4), 529-538.

Minorsky, N. (1962). Nonlinear Oscillations, Van Nostrand, Princeton, New Jersey.

Oppenheim, A., & Schafer, R. (1989). Discrete-Time Signal Processing. Prentice-Hall, Englewood Cliffs, New Jersey.

Qi, Y.Y., Weinberg, B., Bi, N., & Hess, W.J. (1984, Jan.). Minimizing the effect of period determination on the computation of amplitude perturbation in voice. NCVS Acoustic Voice Analysis Workshop Proceedings, Denver, Colorado, Jan. 17-18, 1994.

Story, B. (1995). Physiologically-Based Speech Simulation Using an Enhanced Wave-Reflection Model of the Vocal Tract. [Unpublished Ph.D. Thesis]. The University of Iowa, Iowa City, IA.

Story, B. & Titze, I. (1995). Voice simulation with a body cover model of the vocal folds. Journal of the Acoustical Society of America, 97(2), 1249-1260.

Titze, I. (1984). Parameterization of the glottal area, glottal flow, and vocal fold contact area. Journal of the Acoustical Society of America, 75(2), 570-580.

Titze, I.R. (1988). The physics of small amplitude oscillation of the vocal folds. Journal of the Acoustical Society of America, 83 (4), 1536-1552.

Titze, I.R., Jiang, J., & Hsiao, T.Y. (1993). Measurement of mucosal wave propagation and vertical phase difference in vocal fold vibration. Annals of Otology, Rhinology and Laryngology, 102, 58-63.

Titze, I.R. (1989). On the relation between the subglottal pressure and fundamental frequency in phonation. Journal of the Acoustical Society of America, 85 (2), 901-906.

Wong, D., Ito, M.R., Cox, N.B., & Titze, I.R. (1991). Observation of perturbation in a lumped-element model of the vocal folds with application to some pathological cases. Journal of the Acoustical Society of America, 89, 383-394.

NCVS Status and Progress Report - 8
July 1995, 43-48

Immunocytochemical Study of Proteoglycans in Vocal Folds

Agnieszka Pawlak, M.D.
Division of Otolaryngology, University of Utah
Thomas Hammond
Elizabeth Hammond, M.D.
Department of Pathology, LDS Hospital, Salt Lake City, Utah
Steven Gray, M.D.
Division of Otolaryngology, University of Utah

Abstract

We evaluated the proteoglycan composition of normal vocal folds using immunocytochemical techniques.. Frozen sections of 14 normal cadaveric vocal folds were obtained within 12 hours of death and sectioned immediately. Vocal fold sections were stained with antibodies against keratan sulfate, chondroitin sulfate, heparan sulfate proteoglycan (HSPG), decorin and the hyaluronate receptor. We found that the lamina propria has diffuse staining of fibrillar components with keratan sulfate and decorin. Intense staining was observed in the vocal ligament area with keratan sulfate. HSPG was localized to the basement membrane zone. Chondroitin sulfate, HSPG and hyaluronate receptor were detected in the cytoplasm of interstitial cells with immunocytochemical characteristics of macrophages.

Keratan sulfate distribution suggests that fibromodulin may be significant in normal vocal folds. Production of HSPG and probably versican occurs in macrophages and fibroblasts in the lamina propria.

Introduction

Benign laryngeal lesions such as polyps, nodules, scars and polypoid corditis may represent a disturbance in the balance of the extracellular matrix constituents of the vocal folds. To evaluate this hypothesis, it is first necessary to characterize the components contributing to the normal extracellular matrix in the lamina propria.

The vocal fold lamina propria consists of two important tissue components: cellular and extracellular matrix. As the cellular component is sparse, it appears that the extracellular matrix has the determining biochemical role.[1] The extracellular matrix is composed of two families of macromolecules: collagen/elastins which provide the fibrous scaffolding of the lamina propria, and the substance of the interstitium or "filling" between the fibrous scaffolding - the proteoglycans and structural glycoproteins.[2]

Proteoglycans are ubiquitous substances that have been implicated in a wide variety of processes, such as cell adhesion and migration, matrix assembly, growth factor sequestration and regulation, immune function, binding of plasma proteins and control of thrombogenesis.[3-5]

Proteoglycans are composed of glycosaminoglycan chains covalently attached via a linkage protein to a protein core.[4] Glycosaminoglycans are linear polymers and include keratan sulfate, chondroitin sulfate, dermatan sulfate, heparan sulfate and hyaluronic acid. Hyaluronic acid differs from other glycosaminoglycans in that it is not covalently attached to proteins.[6]

Three subgroups within the extracellular matrix proteoglycan family have been identified: (1) small, interstitial matrix proteoglycans -- decorin, biglycan and fibromodulin; (2) large, aggregating chondroitin sulfate proteoglycans -- aggrecan and versican; and (3) heparan sulfate proteoglycans. This classification is based on the similarities of the protein cores and glycosaminoglycan chains.[6-7] Each of the proteoglycans has characteristic features in relation to glycosaminoglycan chains, structure, function and localization in other tissues, as summarized in Table 1 (following page).

Proteoglycans have not received as much attention as the other components of the lamina propria of vocal folds. Understanding the normal distribution and interactions of these macromolecules may be a guide to discovering the

Table 1.
Characteristics of Proteoglycans

Proteoglycan	Glycosaminoglycan	Size	Function	Localization
Aggrecan	CS* and KS†	2,500 kDa	• binds hyaluyronic acid[23] • provides a strongly hydrated space filling gel[23] • modulates the proliferative and metabolic activities of chondrocytes and fibroblasts[11] • plays a structural role in the development of cartilaginous structures[13]	• **not found in LP** or VF††** • ECM‡ of cartilage, tendon, sclera, aorta and bone[24-25]
Versican	CS	>1000 kDa	• binds hyaluronic acid[14] • ability to fill space, bind and organize water molecules[6] • role in hydration of cartilage[6]	• **found in macrophages and fibroblasts in LP and VF** • expressed by fibroblast cells • isolated from various cell cultures and tissue sources, such as intima of aorta[26]
Biglycan	CS/DS‡ (2 chains)	50-200 kDa	• binds transforming growth factor-β (TGF-β) in vitro -- feedback mechanism regulating the synthesis of biglycan[7] • absorbs to hydroxylapatite crystals after being shed from the osteoblasts	• **has not been searched for in the VF** • CS: fetal/young bone • DS: articular cartilage, skin, tendon and sclera[13,27] • specialized cell types in developing human tissues including bone, cartilage, blood vessel endothelial cells, skeletal myofibrils, renal tubular epithelia, and differentiating keratinocytes[3,28-29]
Decorin	CS/DS (1 chain)	50-200 kDa	• binds to collagen types I and II in *in vitro* assays, resulting in a delayed fibril formation and the formation of thinner fibrils[11,13] • binds to fibronectin and transforming growth factor-β (TGF-β)[11]	• **found in ECM (fibrillar component) of superficial LP of VF** • CS: developing bone • DS: cartilage, skin, tendon and sclera[30]
Fibromodulin	KS	59 kDa	• binds to collagen types I and II in *in vitro* assays, resulting in a delayed fibril formation and the formation of thinner fibrils[11-13]	• **found in ECM (fibrillar component) of intermediate and deep layers of LP and VF** • **found in vocal ligament area** • cartilage, tendon, skin, sclera and cornea[12,24]
Heparan Sulfate Proteoglycan	HS§	55,000 to > 400,000 kDa	• binds to fibronectin, collagen IV and laminin[16] • integral component of basement membranes • may play a role in tissue morphogenesis, growth control and differentiation[10]	• **found in BMZ¶ of VF** • **found in macrophages and fibroblasts of LP of VF** • ECM of cultured epithelial and mesenchymal cells (i.e. human lung fibroblasts)[10]
	Hyaluronic Acid	100-5,000 kDa	• contributes to structure of ECM via interaction with aggrecan and versican • creates a highly viscous environment that plays a role in molecular exclusion, flow resistance and tissue osmosis • influences cell adhesion and migration[21]	• **found in macrophages and fibroblasts of LP of VF** • ECM of synovial fluid, umbilical cord, dermis, subcutaneous tissue, and cartilage[15]

*CS-chondroitin sulfate
†KS-keratin sulfate
‡DS-dermatan sulfate
§HS-heparan sulfate
‡ECM-extracellular matrix
¶BMZ-basement membrane zone
**LP-lamina propria
††VF-vocal folds

function and interaction of components of the extracellular matrix and eventual correlation with various disease states of the vocal folds. We therefore undertook this descriptive study of normal vocal cord proteoglycans.

Materials and Methods

Tissues Studied

Frozen sections of 14 normal cadaveric vocal folds (no known laryngeal injury, no history of laryngeal intubation) were obtained within 12 hours of death and sectioned immediately. Sections were cut using an Instrumedics attachment to a Tissue Tek Cryostat.

Antibodies

Commercially available antibodies and their dilutions used in this study were as follows:

·Anti-keratan sulfate (Clone 5D4) (ICN Biochemicals) 1:500

·Anti-proteoglycan DI-6S (Clone 3B3) (ICN Biochemicals) 1:50 (epitopes recognized: 0-, 4-, and 6-chondroitin sulfate)

·Anti-Heparan Sulfate Proteoglycan (Chemicon) 1:20

·Anti-hyaluronate receptor (Life Technologies, Inc.) 1:100

·Anti-decorin (PG40) (Life Technologies, Inc.) 1:50

·Anti-CD 68 (DAKO) 1:100 (stains cells of macrophage lineage)

Immunocytochemical Method

For the immunocytochemical studies, frozen sections of normal vocal fold were stained with each antibody and compared with the staining of normal skin frozen sections. Antibodies directed against specific proteoglycans and glycosaminoglycans were detected in tissues using the avidin-biotin immunoperoxidase method and diaminobenzidine as a substrate.[8] In the case of decorin, an indirect immunofluorescence method was used with a fluorescein conjugated anti-mouse IgG as the secondary reagent.[9] Controls of skin were used with each staining process.

Interpretation of Results

The intensity and distribution of staining of the peroxidase and immunofluorescence localized antibodies were evaluated by one observer who examined vocal fold slides and compared them to control skin slides stained similarly. The intensity of staining in vocal folds was graded as 0+ to 3+ based on its relationship to the control. If the staining of the specimen was as intense as that found in the control, the grade was considered 3+. Weak staining was considered as 0 to 1+. Intermediate grades of staining were considered to be 2+.

Results

Keratan Sulfate and Decorin

We have found that normal vocal fold lamina propria has staining of fibrillar components with keratan sulfate and decorin (intensity of 2+). Decorin was more prominent in the superficial rather than deeper layers (Photo 1, see center plate). It is also of interest, that a very intense staining of the vocal ligament area occurred with keratan sulfate (intensity 3+) (Photo 2, see center plate).

Chondroitin Sulfate and Hyaluronate Receptor

Chondroitin sulfate and hyaluronate receptor staining were detected only in the cytoplasm of cells in the lamina propria (intensity 2+ for both) (Photo 3, see center plate). Many cells which displayed cytoplasmic staining had immunocytochemical characteristics of macrophages. Serial sections stained with antibody directed against CD68 illustrated that many cells (up to 25%) were positive for CD68 and proteoglycans. Remaining cells not staining with CD68 were spindle-shaped and were presumed to be fibroblasts.

Heparan Sulfate Proteoglycan

Heparan sulfate proteoglycan, an integral component of basement membrane in probably all mammalian tissues,[10] was localized to the basement membrane zone of the vocal folds and blood vessels (intensity 2+). In addition, a cytoplasmic localization in macrophages and spindle cells (presumed to be fibroblasts) in the lamina propria was also observed (intensity 2+).

In summary, the fibrillar component of the lamina propria stained with keratan sulfate and decorin, vocal fold ligament stained with keratan sulfate, basement membrane zone stained with heparan sulfate proteoglycan, and cytoplasm of many macrophages and perhaps fibroblasts stained with heparan sulfate proteoglycan, chondroitin sulfate and hyaluronate receptor (summary results in Table 1).

Discussion

The results of this work indicate that diverse proteoglycans are found in the lamina propria and may play an important biologic role, as demonstrated for proteoglycans present in other tissues.

The intense staining of the fibrillar components of the lamina propria and specifically of vocal fold ligament with keratan sulfate suggests an important role for the proteoglycan fibromodulin. The staining pattern for this proteoglycan outlines the vocal ligament area in a way not previously fully appreciated by the authors. Previous descriptions of the vocal ligament have been based on histologic stains of elastin and collagen. The delineation between the superficial tissues and the deeper structures such as the vocal ligament is often not clear histologically. The keratan sulfate stain demarcates an area which appears to correlate

with the vocal ligament. Fibromodulin, with its keratan sulfate glycosaminoglycan attachments, has been found to bind collagen types I and II in other tissues. Collagen I and II also have been localized to lamina propria and are known to be major components of the vocal ligament. The presence of fibromodulin in the extracellular matrix results in delayed collagen fibril formation leading to formation of thinner fibrils and inhibition of collagen matrix contraction.[11-13]

Decorin, a proteoglycan present mainly in the superficial layer of the lamina propria, has functions in common with its family member fibromodulin. It binds to types I and II collagen fibrils and changes the kinetics of collagen fibril formation. This binding is thought to regulate the size and morphology of collagen fibrils. In vitro, the presence of decorin has resulted in thinner and smaller collagen fibrils.[11,13] It is curious that fibromodulin and decorin are closely related proteoglycans, both in function and structure, yet decorin is found predominantly in the superficial layer of vocal fold lamina propria and fibromodulin is intensely present in the intermediate and deep layers of the vocal fold lamina propria. As these findings are correlated with our knowledge of collagen fiber population density in the various lamina propria layers, a biologic and perhaps therapeutic role may be established for these proteoglycans.

Aggrecan does not appear to have an important role in vocal folds, as indicated by our results. Since aggrecan has both chondroitin and keratan sulfate glycosaminoglycans, it would have been expected that the pattern of staining of these glycosaminoglycans would overlap if aggrecan were present. We did not find this congruency of staining. By contrast, versican appears to play a role in the lamina propria of the vocal fold. Its presence is suggested by the finding of abundant cytoplasmic staining of macrophages with antibody against the hyaluronate receptor. Versican is involved in interaction with hyaluronic acid. Versican's structure includes a hyaluronic acid binding domain and consequently its activity is thought to be related to hyaluronic acid in the extracellular matrix.[14] Hyaluronic acid, which can have voluminous configurations with large, repeating networks of molecular structure, is important in tissue viscosity, osmosis and flow resistance.[15] These properties may make this proteoglycan very important in tissue oscillation.

Heparan sulfate proteoglycan is found in basement membrane zones and also binds matrix components such as collagen IV, laminin and the structural glycoprotein fibronectin.[16] Fibronectin is present in normal vocal fold lamina propria although it has been implicated in benign laryngeal disease.[17] Thus, the importance of interaction of proteoglycans with fibronectin should not be overlooked and needs further research.

The significance of heparan sulfate proteoglycan can be further deduced by examining tissues undergoing fibrosis and scars. The presence of heparan sulfate at these sites during collagen polymerization may result in local changes in the collagen matrices that affect the ability of fibroblasts to remodel the matrix. The discontinuous packing of collagen fibrils as promoted by heparan sulfate, results in abnormal arrangements of collagen fibrils and decreased tensile strength of scar tissue.[18] The contribution of heparan sulfate proteoglycan must be considered in vocal fold injury.

Chondroitin sulfate, heparan sulfate proteoglycan and hyaluronate receptor exhibit a common feature of staining the cytoplasm of macrophages and fibroblasts in the lamina propria, as indicated in our research. A study comparing glycosaminoglycan synthesis by human fibroblasts from normal skin, normal scar and hypertrophic scar suggests the influence of these specific proteoglycans/glycosaminoglycans on organization of collagen following vocal fold injury. Hypertrophic scar tissue contains increased amounts of chondroitin sulfate, dermatan sulfate, heparan sulfate and hyaluronic acid. In addition, the collagen from hypertrophic scar has an altered 3-dimensional arrangement compared to that found in normal skin and normal scar. It appears that increased amounts of glycosaminoglycans protect the collagen from normal remodeling by collagenase, leading to the collagen accumulation and excessive scar formation.[19] Therefore, increased synthesis of proteoglycans/glycosaminoglycans by fibroblasts secondary to a vocal fold injury, may result in scar formation (as evidenced by increase in collagen fibers). This subsequently could lead to dampening of the mucosal wave and result in decreased vocal fold vibration. Assessment of those proteins in benign laryngeal disease is needed.

The cytoplasmic localization of glycosaminoglycans in lamina propria of normal vocal folds suggests a role for these macromolecules in responses to inflammatory stimuli. It has been shown that glycosaminoglycan synthesis is altered in the presence of the cytokine IL-1, a central mediator of inflammation. IL-1 causes fibroblasts to produce less versican and more hyaluronic acid (see table I, biological properties of both).[20-21] Such changes could have pronounced effects on vocal fold lamina propria properties. By contrast, macrophage chondroitin sulfate production has shown to be increased by exposure to inflammatory stimuli such as bacterial polysaccharide.[22] Since macrophages can produce IL-1 in response to inflammation, they can directly influence fibroblast production of these glycosaminoglycans as well.

In summary, using monoclonal antibodies, we have demonstrated that decorin and probably fibromodulin have a lamina propria layer specificity. Other proteoglycans are also present in the lamina propria. Alteration in the concentration or production of the proteoglycans could have a profound effect on the physical properties of the extracellular matrix. Our observations establish a baseline for the

normal distribution of proteoglycans. Further study will allow precise quantitative analysis of macromolecules, as well as elucidation of specific vocal fold disease states that exhibit distinctive changes in the proteoglycan/glycosaminoglycan content and distribution in the lamina propria.

Acknowldgements

Work supported by Public Health Service Grant 5 K08 DC00036-06.

References

1. Gray SD, Titze I, Lusk RP. Electron microscopy of hyperphonated vocal cords. J Voice 1987;1:109-15.

2. Labat-Robert J, Bihari-Varga M, Robert L. Extracellular matrix. FEBS Lett 1990;268:386-93.

3. 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: Academic Press, 1993:33-82.

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

5. McCarthy KJ, Accavitti MA, Couchman JR. Immunological characterization of a basement membrane-specific chondroitin sulfate proteoglycan. J Cell Biol 1989;109:3187-98.

6. Lander AD. Proteoglycans. In: Kreis T, Vale R, eds. Guidebook to the extracellular matrix and adhesion proteins. New York: Oxford University Press, 1993:12-16.

7. Nietfeld JJ. Cytokines and proteoglycans. Experientia 1993;49:456-69.

8. Hammond E, Griffin J, Odell WD. A chorionic gonadotropin secreting human pituitary cell. J Clinical Endocrinol and Metab 1991;71:747-54.

9. Hammond EH, Hansen JK, Spencer LS, Jensen A, Yowell RL. Immunofluorescence of endomyocardial biopsy specimens: methods and interpretation. J Heart Lung Transplant 1993;12:S113-24.

10. Heremans A, Van der Schueren B, De Cock B, et al. Matrix-associated heparan sulfate proteoglycan: core protein-specific monoclonal antibodies decorate the pericellular matrix of connective tissue cells and the stromal side of basement membranes. J Cell Biol 1989;109:3199-3211.

11. Hardingham TE, Fosang AJ. Proteoglycans: many forms and many functions. FASEB J 1992;6:861-70.

12. Oldberg A. Fibromodulin. In: Kreis T, Vale R, eds. Guidebook to the extracellular matrix and adhesion proteins. New York: Oxford University Press, 1993:55-56.

13. Goetnick PF. Proteoglycans in development. In: Kimmel CB, eds. Current topics in developmental biology, Vol 25. New York: Academic Press, 1991:111-31.

14. Zimmermann DR. Versican. In: Kreis T, Vale R, eds. Guidebook to the extracellular matrix and adhesion proteins. New York: Oxford University Press, 1993:100-101.

15. Toole BP. Hyaluronan and hyaluronan binding proteins (hyaladherins). In: Kreis T, Vale R, eds. Guidebook to the extracellular matrix and adhesion proteins. New York: Oxford University Press, 1993:64-65.

16. Heremans A, De Cock B, Cassiman JJ, Van den Berghe H, David G. The core protein of the matrix associated heparan sulfate proteoglycan binds to fibronectin. J Biol Chem 1990;265:8716-24.

17. Gray SD, Hammond E, Hanson DF. Benign pathologic responses of the larynx. Ann Otol Rhinol Laryngol 1995;104:13-18.

18. Guidry C, Grinnell F. Heparin modulates the organization of hydrated collagen gels and inhibits gel contraction of fibroblasts. J Cell Biol 1987;104:1097-103.

19. Savage K, Swann DA. A comparison of glycosaminoglycan synthesis by human fibroblasts from normal skin, normal scar, and hypertrophic scar. J Invest Dermatol 1985;84:521-26.

20. Qwarnstrom EE, Järveläinen HT, Kinsella MG, et al. Interleukin-1 beta regulation of fibroblast proteoglycan synthesis involves a decrease in versican steady-state mRNA levels. Biochem J 1993;294(pt 2);613-20.

21. Laurent TC, Fraser JRE. The properties and turnover of hyaluronan. In: Ciba Foundation Symposium 124: Functions of the proteoglycans. Great Britain: John Wiley and Sons, 1986:9-29.

22. Uhlin-Hansen L, Wik T, Kjellen L, Berg E, Forsdahl F, Kolset SO. Proteoglycan metabolism in normal and inflammatory human macrophages. Blood 1993;82:2880-9.

23. Doege KJ. Aggrecan. In: Kreis T, Vale R, eds. Guidebook to the extracellular matrix and adhesion proteins. New York: Oxford University Press, 1993:17-18.

24. Heinegård D, Franze A, Hedbom E, Sommarin Y. Common structures of the core proteins of interstitial proteoglycans. In: Ciba Foundation Symposium 124: Functions of the proteoglycans. Great Britain: John Wiley and Sons, 1986:69-88.

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

26. Oldberg A, Antonsson P, Hedbom E, Heinegård D. Structure and function of extracellular matrix proteoglycans. Biochem Soc Trans 1990;18:789-92.

27. Fisher LW. Biglycan. In: Kreis T, Vale R, eds. Guidebook to the extracellular matrix and adhesion proteins. New York: Oxford University Press, 1993:21-22.

28. Bianco P, Fisher LW, Young MF, Termine JD, Robey, PG. Expression and localization of the two small proteoglycans biglycan and decorin in developing human skeletal and non-skeletal tissues. J Histochem Cytochem 1990;38:1549-63.

29. Järveläinen HT, Kinsella MG, Wight TN, Sandell LJ. Differential expression of small chondroitin/dermatan sulfate proteoglycans, PG-I/biglycan and PG-II/decorin, by vascular smooth muscle and endothelial cells in culture. J Biol Chem 1991;266:23274-81.

30. Rosenberg LC, Choi HU, Tang LH, et al. Isolation of dermatan sulfate proteoglycans from mature bovine articular cartilages. J Biol Chem 1985;260:6304-13.

NCVS Status and Progress Report - 8
July 1995, 49-54

The Intermediate Layer: A Morphologic Study of the Elastin and Hyaluronic Acid Constituents of Normal Human Vocal Folds

Thomas Hammond
Department of Pathology, University of Utah
LDS Hospital, Salt Lake City UT
Ruixia Zhou, M.S.
Department of Medical Informatics, University of Utah
Elizabeth Hammond, M.D.
Department of Pathology, University of Utah
LDS Hospital, Salt Lake City UT
Agnieska Pawlak, M.D.
Steven Gray, M.D.
Department of Otolaryngology, University of Utah

Abstract

The lamina propria of vocal folds are important in voice production. We undertook an evaluation of the morphologic features of elastin and hyaluronic acid, two important constituents of the lamina propria. Thirty normal human vocal folds were obtained from patients dying of traumatic causes without vocal fold injury. These tissues were immediately prepared for histologic and ultrastructural examination by standard methods. For specific study of the ultrastructure of the layers of the lamina propria, 6 vocal folds were divided horizontally through the mid plane of the lamina propria.

We found that the elastin composition of the vocal folds is variable, the largest amount being seen in the mid portion on elastin-van Gieson (EVG) staining and ultrastructural evaluation.. The superficial layer of the lamina propria contains fewer large elastin fibers. In this region, we found that elastin was predominantly composed of elaunin and oxytalan which stain poorly with EVG. Using computer assisted image analysis, we quantified the differences in elastin composition between the layers. The amount of elastin varied between men and women and these differences could not be accurately measured by the methods employed. Hyaluronic acid was abundant especially in the mid portion of the lamina propria and was significantly more abundant in men than women on quantification. The significance of these observations in normal vocal folds is discussed.

Introduction

Although it is known that the composition the lamina propria of the normal vocal fold plays an important role in the in the biomechanics of sound production, few detailed morphologic studies have described the characteristics of this vocal fold region in normal human subjects. It is imperative to determine the normal homeostasis in the lamina propria in order to evaluate conditions in which that balance has been disrupted. Knowing what physiological components have been altered may permit the design of better restorative measures.

Previous work has detailed that the lamina propria in normal human vocal fold which is comprised of cells, mostly fibroblasts, and matrix substances secreted by these cells including glycosaminoglycans (GAG), proteoglycans, and fibrous proteins including collagen and elastin (1). Capillaries, nerves and macrophages are also found here.

In this report, we describe our morphologic studies concerning two important matrix constituents, elastin and hyaluronic acid. Elastin plays a large role in the actual

vibration of the vocal cord and is found in three distinct forms in the lamina propria: oxytalan, which is composed of microfibrils 10-12nm in diameter, elaunin, which is composed of microfibrils and an elastic amorphous component, and mature elastic fibers which are characterized by a large amorphous component surrounded by small microfibrils. Hyaluronic acid is one of the major GAGs in biological systems. GAGs are hydrophilic molecules that are relatively inflexible which gives them their "space filling" characteristic through gel formation. They also are osmotically active because their negative charge attracts Na+ ions and therefore water. This osmotic activity is what the extracellular matrix uses to resist compressive forces (2). The morphologic studies employed a computer imaging system which enabled us to obtain quantitative data on the structure of normal vocal fold matrix, which has not previously been available (3).

Materials and Methods

Tissue Preparation

Larynges were obtained from the Medical Examiner from individuals dying of traumatic causes in which the vocal folds were known not to have undergone injury or instrumentation. Patient age ranged from 20-60 years old. All vocal folds were obtained within 24 hours of death. Forty vocal folds were prepared and processed for permanent sections.

Larynges were examined grossly for evidence of injury. Vocal folds were removed and cut vertically, perpendicular to the free edge of the vocal cord. One 2 mm mid section was frozen in OCT embedding compound in a Tissue tek cryostat equipped with an Instrumedics modification. (Instrumedics Corp, Hackensack, NJ). A facing midsection was fixed in Carson's Fixative (4) and embedded in paraffin using routine histologic methods. An adjacent section was fixed in Carson's Fixative and processed for electron microscopy.

Electron Microscopy

Thirty vocal cords were examined ultrastructurally. In most cases, thin sections of matrix adjacent to the epithelium were evaluated. To compare the ultrastructure of elastin in the superficial and deep layers 6 vocal cords were specially cut into two sections each. The full thickness section of the vocal fold was bisected horizontal (parallel) to the epithelial surface, using a dissecting scope to visualize the epithelial surface and muscularis. This divided each specimen into superficial and deep portions of the lamina propria. One mm thick blocks of each were processed routinely for electron microscopy using osmium tetroxide post fixation, acetone dehydration and embedment in EMBED (EM Sciences; Fort Washington, PA). Ultrathin sections were prepared using a NOVA ultratome. Sections were stained with Uranyl acetate and lead citrate and examined in a JEOL 100S transmission electron microscope. For staining of elastin fiber, a 0.1% tannic acid solution was used. Unstained sections were incubated in tannic acid for 10 minutes at 70°C. This staining causes elastic fibers to stain black (5). Six vocal folds were examined using this procedure.

Histochemical Evaluation of Histologic Sections

Four micron-thick sections of normal vocal fold were stained to detect their elastic tissue content using Verhoeff's elastic tissue stain (EVG) (6). This stain uses ferric chloride and iodine to detect elastin and acid fuchsin as a counterstain. Forty vocal folds were examined with this stain.

Hyaluronic acid was stained using a Muller-Mowry Colloidal Iron stain (7). The hyaluronic acid component was evaluated by incubating the tissue sections for 2 hours at 37 C with (AMPc) and without (AMPs) treatment by testicular hyaluronidase. The hyaluronidase enzyme removes hyaluronic acid. Masson's Trichrome (TRIC) (8) stain was used to assess the collagen content. Twenty cases were examined with this stain.

Image Analysis

The slides were analyzed using an image analysis system specially configured for histological and cytological studies. This system is run on a 486 PC compatible computer with a 70 megabyte hard disk. A Microimager 1400 high resolution (1280 X 1024 X 10 bit) image scanner (XILLIX Tech Corp., Vancouver, B.C. Canada) is mounted on an Olympus BH2 microscope with a DPlanApo 20UV 20X Lens and BHT 5X Photo eye piece(Olympus Corp., Lake Success, New York, U.S.A.) for the image acquisition. The illumination was carefully controlled via a DC Lamp Power Supply (Olympus Corp, Lake Success, New York, U.S.A.). Optical interference filters which restrict the wavelengths of light bombarding the specimen are chosen to optimize the measurement of the desired substance. The image was viewed on a high resolution color monitor to be sure that the proper image was acquired. Optimas Image Analysis Software (Bioscan Corp, Edmond, WA) developed from macros was used to analyze the images. The distribution and the intensity of the substance of interest are measured and recorded by the software.

For the EVG stain, a single interference centralized at a wavelength of 630nm with 10nm bandwidth was used which deemphasizes the pink in the stain and accentuates the black staining of the elastin fibers. The black fibrillar material (elastin) was measured and a value for the extensiveness (positive staining area per unit matrix area) and intensity (integrated amount of stain per unit matrix area) (3) within a given region was recorded (magnification was 100X). Initially, random samplings with the scanner were made. It

was subsequently found to be more effective to acquire images of the entire thickness of the lamina propria thereby controlling for zonal distribution of the elastin fibers. The number of images required to cover the matrix varied with each case depending on the thickness. The range of images was between 7 and 13. We could not use these measurements to assess the actual width of the lamina propria because the thickness depended on the angle of cut. A vocal cord with a perpendicular cut would appear to have a much thicker lamina propria than the same vocal cord cut at an angle. Relative distances (percentage of the lamina propria crossed) can be used because they do not change. 25 vocal folds were examined by this process.

We used the AMPc and AMPs stains to measure the amount of hyaluronic acid present in a given region in the vocal fold. Two identical serial slides were stained. One with AMP and hyaluronidase (AMPc) and one with AMP alone (AMPs). The amount of hyaluronic acid was measured by imaging the two slides and calculating the difference in extent between them (the slide treated with hyaluronidase would have less staining proportional to the amount of hyaluronic acid present). No filter was used with the transmitted light because only one color of stain was recorded (pink) unlike the EVG stain where there were two (black and pink).

Results

Elastin Composition of the Lamina Propria

Histologically, the elastin could best be observed in an EVG stain. The large elastin fibers were most numerous in the intermediate layer of the lamina propria. However, abundant, tiny fibrils were also observed in the superficial portion (photo 4; see center plate). By electron microscopy, we found that the type of elastin present in the superficial layers was distinct from mature elastin fibers. The elastin in the superficial layer was either in the fibrillar form (oxytalan), or consisted of small amounts of amorphous material mixed with the fibrillar component (elaunin) (photo 5 and 6; see center plate). Mature elastic fibers were mostly found in the mid portion and deeper layers. In these regions large amounts of elastin fibers were found (photo 7; see center plate). It is important to note that in comparing the electron micrographs and the elastin found therein with the sections stained with EVG, it appears that the immature fibrillar elastin does not stain nearly as well as mature elastic fibers. The data is tabulated in Table 1.

Quantitative Elastic Layering in the Lamina Propria

Using image analysis on EVG stained slides of normal vocal fold,, we found that the vocal fold was separable into three distinct layers. Although these layers differed in thickness with each patient, the presence of the layers in every patient was consistent with this model (photo 8; see center plate). By EVG staining, the superficial layer of the lamina propria appears to contain little elastin staining. The mid portion shows increased amounts of elastin. In the deepest layer, the amount of elastin decreases. This is consistent with the literature (1).

Table 1.
Hyaluronic Acid Content in Males and Females

Male Hyaluronic Acid Content	Female Hyaluronic Acid Content
26	32
14	3
44	2
14	5
31	4
29	7
31	1
35	23
25	
24	
Average = 27.3	Average = 9.6

This table displays the relative amount of hyaluronic acid present in males and females using an image analysis system. Each value represents an individual patient.

Comparisons were done with respect to gender and the EVG stain. While the layering remained constant, the amount of elastin present varied in each case. There was enough variance to override any conclusions with respect to gender.

Quantitative Hyaluronic Acid Composition of the Lamina Propria

Table 1 shows the distribution and quantification of hyaluronic acid in the lamina propria of men and women. We observed that men had a threefold increase in the amount of hyaluronic acid present over females. This difference was highly significant ($p<<.01$), using the Wilcoxon Signed Rank Test which tests the hypothesis that the data in a given data set are randomly distributed and that there are no underlying groups.

Discussion

In this study we have described and illustrated evidence indicating that layers of the lamina propria of the normal human vocal fold are morphologically distinct. These observations are important because much of the mechanics of the vocal fold is dependant on the size and composition of the lamina propria (9). Changes in the lamina propria affect the performance and elasticity of the vocal fold which can cause dysynchronous vibrations. For

example, a stiffer, thinner lamina propria will cause the vocal fold to oscillate faster and have a lower amplitude(9).

The individual layers of the vocal fold have been studied using the EVG stain (1). In this report, it was found that in the superficial layer, elastin is scanty, the mid and mid-deep layers contain a significant amount of elastin, and the layer proximal to the vocalis muscle is somewhat less in elastic composition. Using refined observation image analysis techniques, we have established that the lamina propria is divisible into three definable layers which have been quantified. By using electron microscopy, we have shown that the superficial layer does contain substantial quantities of elastin of types which stain poorly with EVG. These types of elastin are elaunin and oxytalan (photo 5 and 6; see center plate) which are immature forms of elastic fibers (5).

Quantitative studies done on the difference in elastic composition between men and women proved inconclusive. There was a wide range of extensiveness of elastic fibers with each individual. Most individuals regardless of sex showed increased elastin in the midportion (1) but the amount of elastic fibers within each layer was highly variable. We intend to do similar quantitative studies with different elastic stains (aldehyde fuschin, orcein, resorcin fuschin) which stain immature forms of elastin as well (10)

Elastic fibers are composed of two types of proteins: an amorphous hydrophobic protein and a fibrillar hydrophilic protein. The elastic properties of an elastic fiber are believed to be directly related to the amount of amorphous elastin present (11). Complete elastic fibers are capable of being stretched to roughly two times their length without loss of resiliency (12). The development of an elastic fiber occurs in stages beginning with the secretion of the fibrillar protein. This protein is then modified with an amorphous component and then compiled into a larger mature elastic fiber (10). This amorphous modification is a result of the cross-linking of elastin in previously deposited microfibrils and is the first step of elastogenesis (13). Oxytalan and elaunin are believed to be interrupted states of elastic fiber maturation differing only in the amount of cross-linking present (5, 13). We conclude that the functions of the differing types and stages of elastin are important in regulating the mechanics of the superficial layer of the human vocal fold. Therefore the location, type, and quantity of elastin in a given vocal fold help determine how that vocal fold will vibrate and perform.

Oxytalan is a grouping of microfibrils 10-12nm in diameter with no amorphous component. The core of each fibril is electrolucent which gives them a hollow appearance (11). Oxytalan consists mostly of hydrophilic (polar) amino acids, contains less glycin than elastin and has no hydroxyproline, desmosin or isodesmosine. The thicken by lateral addition of each microfibril and at some point are coated with fibronectin and amyloid P components (10,14). It does not elongate under mechanical stress because of its paucity of amorphous elastic material and therefore is more abundant in tissues where stress is common (superficial layer of tendons, cartilage) (11). This may be why it is more abundant in the superficial layer of the lamina propria. It was previously thought that oxytalan was completely fibrillar having no elastic component, but recent studies have shown that oxytalan does indeed exhibit elastic properties (13) It has also been hypothesized that oxytalan is partially responsible for the folding of collagen in tendons (15). Oxytalan gets its name because it can be stained only in the oxidized state (10).

Elaunin is an intermediary between the completely fibrillar oxytalan and the large amorphous and fibrillar elastic fibers (hence its name, elaunin). Elaunin fibers are smaller and shorter than complete elastic fibers (10) but do contain some amorphous component. This amorphous component is hydrophobic and rich in valine. It is lacking in hydroxyproline and has no hydroxylysine. It is similar to collagen in that about one third of the amino acids are glycine and about 11 percent are proline (10). Little is known about the specific function of these intermediaries. They are known to be an interruption in elastogenesis (5) and are thought to show elastic properties of both oxytalan and complete elastic fibers to some degree (11).

The hyaluronic acid composition of the lamina propria was quantified in the vocal folds of men and women. Based on our data, we conclude that the lamina propria of vocal folds of human males and females are morphologically distinct. In past studies, it has been shown that the lamina propria in males is usually thicker than that in females. The thickness of the lamina propria is important because it has been shown that there is a direct relation of thickness to pliability of the fold (16). We suggest that this gender specific difference is due in part to the amount of hyaluronic acid present. Hyaluronic acid is a negatively charged, osmotically active macromolecule which attracts sodium ions when present (2). An abundance of hyaluronic acid would therefore increase the amount of sodium ions present and thus create an osmotic potential. The extra hydration of the vocal cord would therefore make it larger and distinct from that of a vocal fold with less hyaluronic acid. It has been shown that the swelling involved in injury is related to a buildup and secretion of hyaluronic acid. Hyaluronic acid has been proven to be an osmotic regulator in other systems(17).

One other possible application of the difference in hyaluronic acid in men and women deals with the subject of vocal fold nodes. Nodes are generally found in adult females and are relatively sparse in adult males. It is thought that the reason behind this is that the adult female has a higher fundamental frequency of vibration than the adult male. This high frequency leads to more shock and more numerous collisions. Another possible explanation of why men don't develop nodes could be the difference in hyalu-

ronic acid content. The heightened amounts of hyaluronic acid would give them a more shock absorbing character through gel formation.

Hyaluronic acid has also been shown to be responsive to growth factors (18). It was suggested that hyaluronic acid transports the growth factors to the individual cells in the extracellular matrix thereby stimulating mitosis and the cell cycle. Many of these same growth factors had no effect when associated with other GAGs. This is believed to occur because most other GAGs are sulfated whereas hyaluronic acid is not. This responsiveness to growth factors would support the hypothesis that the size of the lamina propria is directly related to the amount of hyaluronic acid present.

Hyaluronic acid is also involved in the support of collagen which suggests another biomechanical significance (19). It also 'protects' proteoglycans from degradation (20) and is involved in tissue proliferation, regeneration, and repair including fibrogenesis and clotting (19). The metabolism and catabolism of hyaluronic acid has been shown to occur in matrix-forming cells such as fibroblasts (which are found in the lamina propria of the human vocal fold) (17). It is believed to be synthesized as a large molecule that is gradually depolymerized. It usually binds with aggrecan upon secretion to form new proteoglycans (17). In a previous study, we found that the GAG precursors to aggrecan (chondroitin sulfate and keratan sulfate) were present in the cells of the lamina propria of human vocal folds (21). Hyaluronic acid is also one of the major contributor to a tissue's resistance to compression forces, likely an important factor in voice mechanics.

In summary, the differences in the extracellular matrix are presumably important in the mechanics and function of the human voice. Further research will determine if differences in the makeup of this region lead to differences in vocal performance. We hope that a better understanding of the exact composition will help with future clinical and preventative medicine.

Acknowledgements

We would like to acknowledge the diligent assistance of Janet Hansen, who aided in many of the staining techniques and tissue processing. We would also like to thank Susan Horn for her assistance in data analysis.

References

1. Hirano M. Structure of the vocal fold in normal and disease states anatomical and physical studies. In: Ludlow CL, Hard MO, ed. *Proceedings of the Conference on the Assessment of Vocal Pathology.* Rockville: The American Speech-Language-Hearing Association, 1981:11-30

2. Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD. Cell Junctions, Cell Adhesion and the Extracellular Matrix. *Molecular Biology of the Cell.* 3rd ed. New York and London: Garland Publishing, Inc. 1994:971-95

3. Zhou R, Parker DL, Hammond EH. Quantitative peroxidase-antiperoxidase complex substrate mass determination in tissue sections by a dual wavelength method. *Analyt Quant Cytol Histol* 1992;14(2):73-80

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

5. Porto LC, Chevallier M, Peyrol S, Guerret S, Grimaud JA. Elastin in human, baboon, and mouse liver: an immunohistochemical and immunoelectron microscopic study. *Anat Rec* 1990:228(4):392-404

6. Carson FL. *Histotechnology. A Self-Instructional Text.* Chicago: ASCP Press. 1990: 140-1

7. Carson FL. *Histotechnology. A Self-Instructional Text.* Chicago: ASCP Press. 1990: 127-9

8. Carson FL. *Histotechnology. A Self-Instructional Text.* Chicago: ASCP Press. 1990: 142-3

9. Kahane JC. Connective tissue changes in the larynx and their effects on voice. *J of Voice* 1987;1(1):27-30

10. Ghadially FN. Extracellular matrix (extracellular components). *Ultrastructural Pathology of the Cell and Matrix.* 3rd ed. London, Boston, Singapore, Sydney, Toronto, and Wellington: Butterworths. 1988:1252-59

11. Ferreira Jr. JMC, Caldini EG, Montes GS. Distribution of elastic system fibers in the peripheral nerves of mammals. *Acta Anat* 1987;130:168-73

12. Gray SD, Hirano M, Sato K. Molecular and cellular structure of vocal fold tissue. In: Titze, IR, ed. *Vocal Fold Physiology.* San Diego: Singular Publishing Group, Inc. 1993:1-3519.

13. Schwartz E, Fleishmajer R. Association of elastin with oxytalan fibers of the dermis and with extracellular microfibrils of cultured skin fibroblasts. *J Hist Cyt* 1986;34(8):1063-8

14. Kostovic-Knezevic L, Bradamante Z, Svajger A. On the ultrastructure of the developing elastic cartilage in the rat external ear. *Anat Embryol.* 1986; 173(3):385-91

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;139:341-348

16. Yumoto E, Katota Y, Kurokawa H. Infraglottic aspect of canine vocal fold vibration: effect of increase of mean airflow rate and lengthening of vocal fold. *J of Voice* 1993;7(4):311-8

17. Ng CK, Handley CJ, Preston BN, Robinson HC. The extracellular processing and catabolism of hyaluronan in cultured adult articular cartilage explants. *Arch Biochem Biophys* 1992;298(1):70-9

18. Roy F, DeBlois C, Doillon CJ. Extracellular matrix analogs as carriers for growth factors: in vitro fibroblast behavior. *J Biomed Mater Res* 1993;27(3):389-97.

19. Meyer LJ, Sern R. Age-dependent changes of hyaluronan in human skin. *J Invest Dermatol* 1994;102(3):385-9.

20. Bolis S, Handley CJ, Comper WD. Passive loss of proteoglycan from articular cartilage explants. *Biochim Biophys Acta* 1989;993(2-3):157-67

21. Pawlak A, Hammond TH, Hammond EH, Gray SD. Immunocytochemical study of proteoglycans in vocal folds. *Ann Otol Rhinol Laryngol.* 1995, in press.

NCVS Status and Progress Report - 8
July 1995, 55-64

Midbrain Regions for Vocalization Identified by Electrical Stimulation in Anesthetized Dogs

Kang Liu, M.D.
Nancy Pearl Solomon, Ph.D.
Erich Luschei, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa

Abstract

Vocalization and other motor responses were elicited by electrical stimulation of the midbrain in 9 dogs anesthetized with pentobarbital. Of a total 236 stimulus sites at which vocalization was elicited, more than 75% were located just medial and dorsal to the medial lemniscus in a region of the midbrain ventral to the superior colliculus. Within this region, several types of vocalization (howl, growl, whine, and bark) were elicited, with some occurring in the absence of other motor responses. Brain tissue in this vocalization region is presumed to consist of descending efferent fibers connecting to or descending from the midbrain periaqueductal grey (PAG) and the adjacent tegmentum. The relation of stimulus sites at which vocalization was elicited to sites producing other motor responses is described.

Introduction

Vocalization has been elicited by electrical or chemical stimulation of the midbrain in monkeys [22,24,30,32], cats [4,9,10,12,25,32], rats [48], guinea pigs [34], bats [39,44], and dogs [20,43] under either anesthetized, decerebrate, or awake conditions. Such stimulation has been demonstrated or presumed to excite the midbrain periaqueductal grey (PAG) and the adjacent tissue of the tegmentum. At least in the monkey and cat, it has been confirmed that these areas of the midbrain are critical for the integrated control of vocalization. This conclusion is based on the findings that (1) species-specific vocalization could be elicited only by electrical stimulation to regions rostral to or in the midbrain [21,22,36]; (2) lesions in the midbrain PAG and adjacent tegmental area abolished vocalizations in the monkey [23], reduced or eliminated vocalization in the cat [1,26,37,41] and in the dog [40], whereas, lesions in the diencephalon and forebrain did not affect the acoustic structure of the vocalizations tested [27]; (3) chemical stimuli, which excite only neuron cell bodies, have been found to elicit vocalization when applied in the PAG rather than other regions in monkeys [23] and cats [4]; (4) tract-tracing studies have revealed connections from the PAG to motoneurons of laryngeal muscles [11,22] and from the PAG to nucleus retroambiguus (NRA) which in turn projects to the motoneurons of laryngeal muscles [19]; and (5) Activity of the PAG neurons has been observed to correlate with the activity of laryngeal muscles during spontaneous vocalization in the monkey [30].

Based on the previous literature and our own experiences, we have developed a technique for electrical stimulation of the midbrain for evoking vocalization. It was successfully applied in anesthetized cats for investigating the relations between stimulus parameters (intensity, duration, and frequency) and characteristics (pattern, timing, and amplitude) of laryngeal and respiratory muscle activity for vocalization [42]. The dog is, however, the preferred species for our experiments because of the relatively large size of the canine larynx, and the morphological similarity between the canine and human larynges [16,28]. In addition, the canine larynx has been studied extensively regarding laryngeal anatomy, histology, physiology, and biomechanics [2,3,7,8,15,17,18,35,46].

We have used this technique of midbrain electrical stimulation in the dog to study the influence of tracheal pressure, manipulated experimentally, on the fundamental frequency of vocalization [20], and to describe different types of elicited vocalization [43]. To make the technique generally useful to other investigations, we feel it would be instructive to describe the method in detail, particularly to document the regions of the midbrain that are most responsive to electrical stimulation.

When conducting previous experiments, we searched for vocalization sites in the midbrain by advancing the stimulating electrode along several tracks through the midbrain, guided by previous reports for the monkey [22,23,32] and cat [12,25,32], and using an atlas based on the beagle brain [13]. Motor responses other than vocalization usually were observed before vocalizations. If particular responses could be observed consistently when the stimulation site was, for example, too dorsal, anterior, or lateral to sites producing vocalization, then the search for vocalization sites in subsequent experiments would be more efficient. Thus, the purpose of this study was to develop a description of functional relations between various motor responses to electrical stimulation of the canine midbrain. The specific aim was to develop a "behavioral-functional map" to define the boundaries of a midbrain region for the elicitation of vocalization. Preferably, such a region would include sites at which natural-sounding vocalization were elicited reliably without other non-vocal motor responses.

Method

Preparation of Experimental Animals

Nine adult mongrel hound-like dogs, 18-22 kg, were sedated initially by intramuscular injection of ketamine (7.5 ml/Kg) and xylazine (2.5 mg/kg) mixture. A catheter was then placed in the femoral vein for administering physiologic saline and anesthesia. A surgical level of anesthesia (muscle relaxation and the absence of corneal and deep pain reflexes) was induced and maintained by pentobarbital (25 mg/kg) throughout the experiment. Body temperature was maintained at 36-38° C.

The head of each dog was fixed in a stereotaxic frame (model 1430, Kopf Instrument, Tajunga, CA) using the Kopf dog adapter. This stereotaxic frame may be rotated around it's front mounting, thus allowing the head and body to be turned to the supine position. This made it possible to access the larynx after the head had been placed in the stereotaxic frame. To allow rotation of the head with the electrode carrier in place, a slot was cut in the table top below the stereotaxic frame.

The larynx and trachea were exposed by a midline incision extending from the hyoid bone to the fifth tracheal ring. A silastic rubber tube (2.5 mm inside diameter) was inserted into the trachea between the second and the third rings, sutured tightly in place, and connected to a pressure transducer (Microswitch 143PC03G). A custom-designed brace for providing a clear view of the larynx was inserted into the mouth to open the jaw, anchor the tongue, and fix the epiglottis. The animal then was rotated to a prone position, the scalp incised along the midline, and the temporalis muscle retracted laterally to expose the dorsal surface of the skull. The cerebral cortex of one cerebral hemisphere was exposed by drilling a hole in the skull and carefully cutting and retracting the dura mater. A monopolar microelectrode held in a stereotaxic carrier was inserted into the brain through the cortex to deliver the electrical stimulus. The microelectrode was made from an epoxy-insulated tungsten, 0.2 mm in diameter, which had electrolytically etched to a very fine point. The tip of the electrode was uninsulated for 0.1 mm.

A stereotaxic coordinate system with the origin (earbar zero) referenced to the midpoint between two earbars inserted in external ear canals and fixed to the stereotaxic frame was used to define the position of the tip of the electrode. Every placement in the brain was described by a group of 3 coordinates: the distances (in millimeters) anterior (A), lateral (L), and dorsal (D) to earbar zero. The anatomical locations of the electrode tip in the midbrain were estimated according to the atlas of the canine (beagle) brain [13]. Electrical stimuli consisted of trains of 0.2 ms rectangular constant-current pulses, with electrode as cathode, delivered at a rate of 200 Hz. The train of stimuli was usually 2 seconds, but was sometimes extended up to 4 seconds. Current level varied, but generally was between 0.2 and 1.5 mA.

In our previous experiments, subdural hemorrhage occurred in some dogs, and blood was observed in the tracks created by the passage of the electrode through the midbrain. The bleeding probably resulted from damage to the vascular plexus covering the tectum of the dog's brain. This problem was overcome in the present experiment by inserting the electrode at a 14° angle lateral to the sagittal plane. With this approach, the electrode avoided the vascular plexus when it was advanced to most places in the midbrain. In data analysis, the lateral (L) and dorsal (D) dimensions of the stereotaxic coordinate system were corrected for the deviations resulting from the 14° angle of electrode insertion.

Experimental Protocol

All 9 dogs received unilateral stimulation of the midbrain, seven of them in the left midbrain and two in the right side. The exploration was done with the animal in the prone position and commenced with a track located at A8, L6 (L5 in Dog 3), then extended track by track with 2 mm (1 mm in Dog 3) intervals between tracks. In each track, the electrode was advanced ventrally from D20 to D5 (D3 in some tracks) in 1 mm steps. At each of these sites, electrical stimuli were delivered at the end of the inspiratory phase of a respiratory cycle. The delivered current typically started at 1.5 mA and was decreased by 0.1 mA steps, if a response occurred, until no response was elicited. At some sites, especially those in the most posterior and ventral locations, the elicited movements of the head and body were so strong that the current had to be limited to relatively low levels (e.g., below 0.5 mA). In every dog, the explored areas in the midbrain enveloped at least a clear vocalization region. That is, once finding a site that evoked vocalization, the electrode was advanced in every direction until the vocalization no longer occurred with stimulation.

After the final stimulation, the electrode was returned to the site at which the "best" vocalization (a clear loud voice with little or no other motor responses) had been elicited. The previously-observed responses were always repeated, although in some cases the electrode had to be moved up or down 0.5 to 1 mm presumably due to tissue shifting. A lesion was made at this best-vocalization site by delivering a 1 mA DC current for 20 seconds. Then, the dog was deeply anesthetized and perfused with saline followed by 10% formalin solution. The brain was postfixed in 10% formalin for at least one week. The midbrain was later sectioned to examine the anatomical locations of the electrode tracks and the lesion.

Data Recording

Evoked vocalization in the present experiment was defined as a laryngeally generated sound during or immediately following the stimulation. Even though different types of vocalizations (howls, growls, whines, and rarely barks) were elicited in these dogs, all vocalizations were treated the same in the analysis to identify the midbrain's vocalization region. Vocalizations were recognized by at least two experimenters and were transduced by a microphone fixed on a maxillary canine tooth.

Signals representing the period and intensity of the stimulation, tracheal pressure, and sound were recorded on an FM tape recorder (HP 3968A, band pass DC-2.5 kHz). These signals were displayed simultaneously on a laboratory computer using data acquisition software (DATAQ Instruments Inc., Akron, OH). After the experiment, these signals were low-pass filtered at 500 Hz and then digitized at a sampling rate of 1 kHz per channel. The movement of the larynx was visualized by a video camera focused on the vocal folds through the mouth and displayed on a color monitor. The video image of laryngeal movements occasionally was recorded on videotape. In addition, all observed responses (vocalization, vocal fold movement, movements of body parts, and respiratory activity) and stimulation conditions were recorded in a notebook.

Results

The number and ranges of anatomical locations of stimulation sites for each dog are listed in Table 1. The number of stimulus sites varied across dogs, ranging from 195 sites in Dog 9 to 615 in Dog 3. The number of the sites at which vocalization was elicited was similar across animals (18 to 29) with the exception of Dog 3 (54). Recall that tracks were separated by 1 mm for Dog 3 rather than 2 mm as for the other dogs. It follows, therefore, that the number of sites for Dog 3 is about 2-3 times of those for other dogs. The ranges of anatomical locations of sites are expressed in millimeter distances anterior (A), lateral (L), and dorsal (d) to the reference point, earbar zero. Of the total number of stimulation sites for all 9 dogs (3120), motor responses of any type were observed at 40% (1247), and vocalization was evoked at 8% (236).

Table 1.
The numbers (in parentheses) and ranges of the sites at which vocalization was elicited, motor responses of any type was induced, and stimulation was applied for each dog. Ranges are expressed as millimeter distances anterior (A), lateral (L), and dorsal (D) to earbar zero.

	Vocalization	Any Response	Stimulation
Dog 1	(29) A: 6-12, L: 2-6, D: 5-12	(189) A: 2-14, L: 2-8, D: 5-16	(465) A: 0-16, L: 0-10, D: 3-25
Dog 2	(28) A: 4-12, L: 2-8, D: 6-12	(176) A: 0-1, L: 2-8, D: 4-17	(510) A: 0-16, L: 0-10, D: 3-25
Dog 3	(54) A: 5-11, L: 2-7, D: 4-15	(305) A: 2-15, L: 1-8, D: 3-18	(615) A: 0-16, L: 0-10, D: 3-25
Dog 4	(28) A: 4-12, L: 2-8, D: 8-14	(125) A: 2-14, L: 2-8, D: 5-17	(333) A: 0-16, L: 0-10, D: 4-20
Dog 5	(20) A: 2-10, L: 2-6, D: 5-13	(97) A: 0-12, L: 2-6, D: 5-18	(241) A: 0-14, L: 0-8, D: 4-20
Dog 6	(19) A: 6-10, L: 4-6, D: 7-12	(101) A: 0-14, L: 2-6, D: 5-16	(245) A: 0-14, L: 0-8, D: 5-20
Dog 7	(18) A: 6-10, L: 4-6, D: 7-11	(87) A: 0-14, L: 2-8, D: 4-16	(255) A: 0-14, L: 2-8, D: 4-20
Dog 8	(19) A: 4-10, L: 2-6, D: 7-12	(74) A: 2-14, L: 2-8, D: 5-17	(264) A: 2-14, L: 0-8, D: 5-20
Dog 9	(21) A: 4-10, L: 2-6, D: 6-12	(93) A: 2-12, L: 2-8, D: 5-15	(195) A: 0-14, L: 0-8, D: 5-20

General Responses

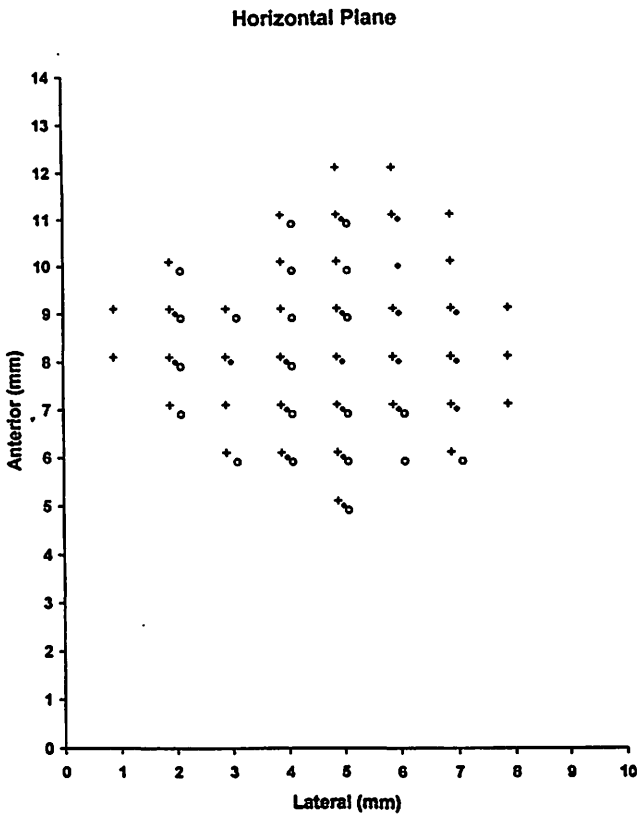
In general, the observed motor responses of body parts to electrical stimulation to the midbrain included movements of muzzle, eyelid (raising and lowering), eyeball (rotating, elevating, and depressing), ear (flattening against the side of the head, pricking up and rotating), jaw (opening and closing), neck, shoulder, leg, trunk, and tail. In addition, respiratory and laryngeal responses included increased depth or rate of respiration, inspiration followed by prolonged expiration, adduction, abduction, and lengthening of the vocal folds, and vocalization. Finally, what appeared to be "retching" was evoked from stimulation delivered close to earbar zero of the stereotaxic coordinate system. This response consisted of alternating opening and closing movements of the glottis and upper esophageal sphincter, accompanied by attempted opening of the mouth.

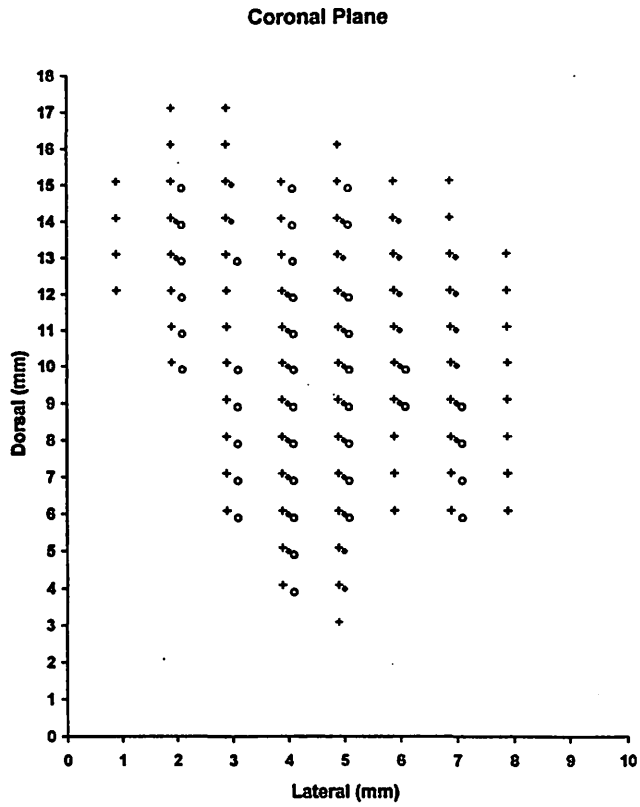
Stimulation sometimes elicited multiple responses, especially at high current levels. Usually, muzzle, ear, and eyelid responses were elicited together and therefore were grouped as "facial movement" for data analysis. For the same reason, shoulder, leg, and trunk responses were grouped as "body movement." The majority of facial-movement responses were contralateral to the stimulation, but they sometimes occurred bilaterally at relatively high current levels. Body-movement responses usually occurred bilaterally.

These two groups of motor responses, especially facial movements, were induced very frequently at or close to the sites at which vocalization was evoked. In addition, jaw closure was induced frequently when the stimulus was delivered close to vocalization sites. Respiratory responses occurred with each evoked vocalization and also occurred without vocalization when stimulation was close to vocalization sites. The relation of sites producing these responses to sites producing vocalization is described later. Adduction of the vocal folds was observed during evoked vocalization. Sometimes, vocal fold adduction was induced without vocalization and without an observable respiratory response. Lengthening of the vocal folds was observed during high-pitched vocalizations.

Vocalization

Locations of stimulation sites at which vocalization was elicited in Dog 3 are illustrated in Fig. 1. The three graphs are representations of the left side of the midbrain, showing horizontal, coronal, and sagittal planes. Sites producing facial movement and/or body movement area also illustrated in this figure. For this dog, vocalization sites were





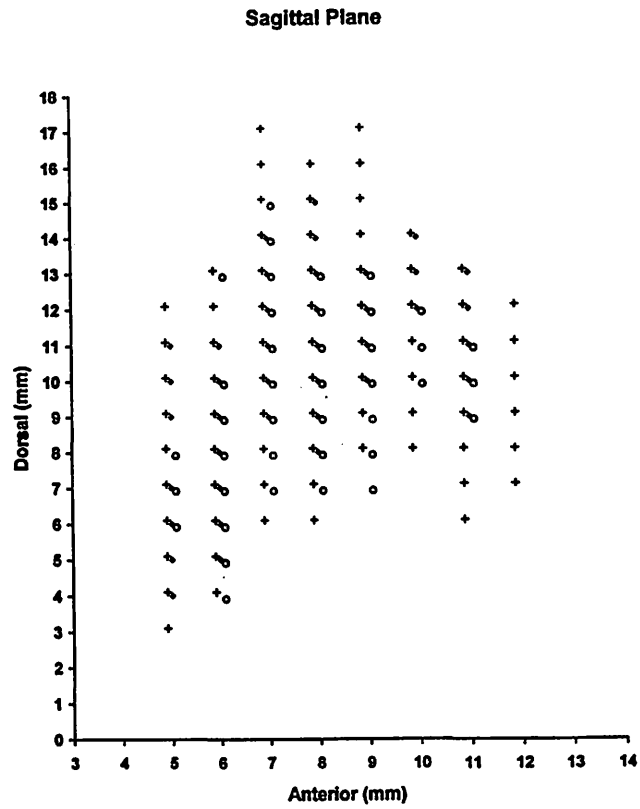


Figure 1 (above and right). Locations of stimulation sites at which vocalization (filled diamond) was elicited, in horizontal, coronal, and sagittal planes of the stereotaxic system in the midbrain of Dog 3. Induced facial (+) and body (open circle) movements also are illustrated at these sites. The origin represents earbar zero. Numbers in each dimension reflect distances in millimeters from earbar zero.

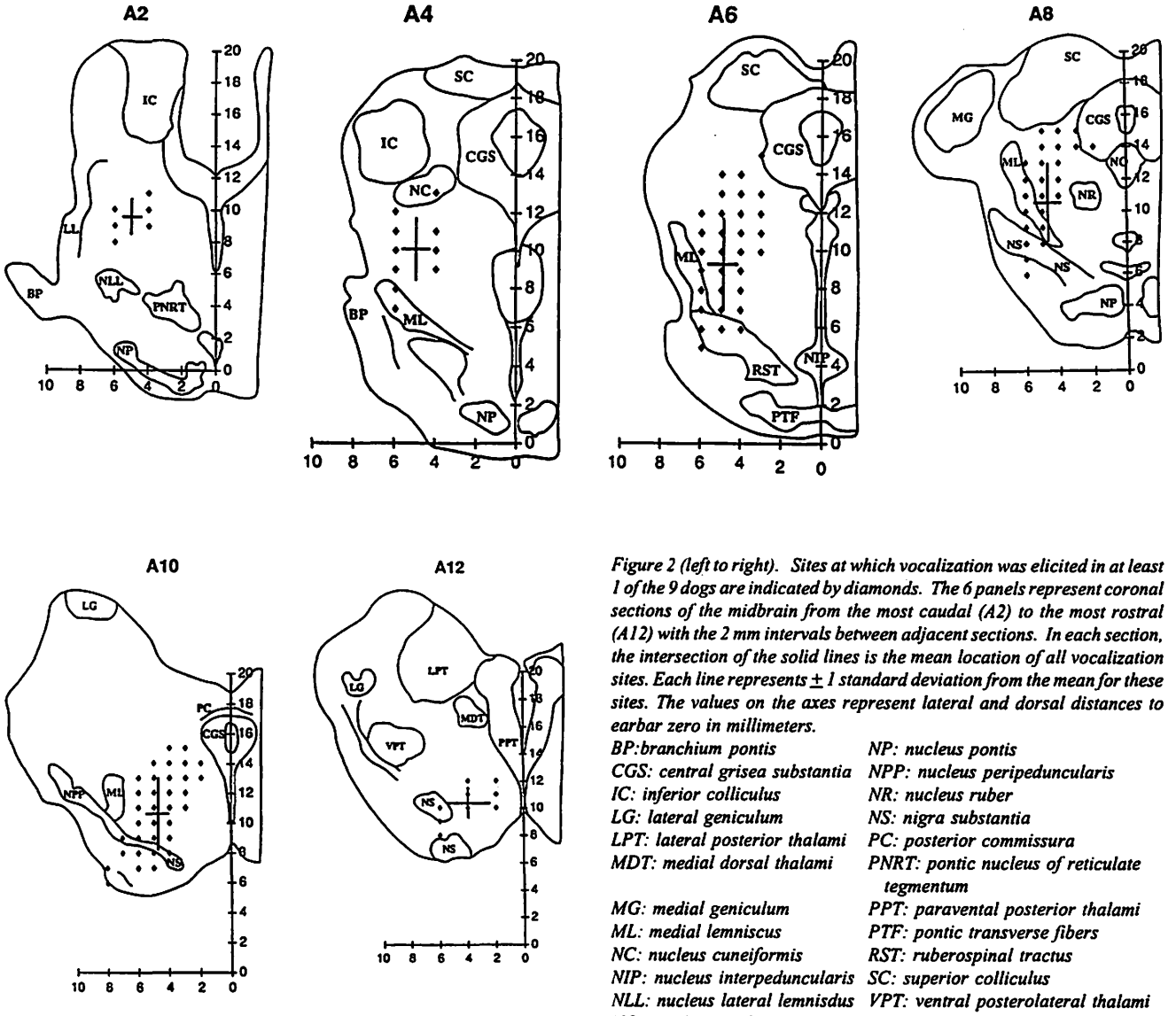


Figure 2 (left to right). Sites at which vocalization was elicited in at least 1 of the 9 dogs are indicated by diamonds. The 6 panels represent coronal sections of the midbrain from the most caudal (A2) to the most rostral (A12) with the 2 mm intervals between adjacent sections. In each section, the intersection of the solid lines is the mean location of all vocalization sites. Each line represents ± 1 standard deviation from the mean for these sites. The values on the axes represent lateral and dorsal distances to earbar zero in millimeters.

BP:branchium pontis
CGS: central grisea substantia
IC: inferior colliculus
LG: lateral geniculum
LPT: lateral posterior thalami
MDT: medial dorsal thalami
MG: medial geniculum
ML: medial lemniscus
NC: nucleus cuneiformis
NIP: nucleus interpeduncularis
NLL: nucleus lateral lemnisdus
NO: nucleus oculomotoris
NP: nucleus pontis
NPP: nucleus peripeduncularis
NR: nucleus ruber
NS: nigra substantia
PC: posterior commissura
PNRT: pontic nucleus of reticulate tegmentum
PPT: paravental posterior thalami
PTF: pontic transverse fibers
RST: ruberospinal tractus
SC: superior colliculus
VPT: ventral posterolateral thalami

located in a region defined by A5-11, L2-7, and D4-15. Facial movements were elicited at all vocalization sites except those in the track A10, L6, as well at most surrounding sites (posterior sites were the exception). Body-movement responses occurred at more than half of the vocalization sites. Compared with the sites for facial responses, body movements were elicited at relatively posterior and ventral sites in or around vocalization region. Vocalization without other responses was only elicited at 4 sites, all within one track: A10, L6 (see horizontal plane for location of the track) from D10 to D13. In the most medial tracks (L2 and L3), vocalization was elicited only at 4 sites (see Coronal plane). According to the atlas [13], these sites were presumably in the PAG.

Vocalization sites from all 9 dogs are shown in Fig. 2. These points are overlaid on an illustration, traced from Dua-Sharma et al. [13], of the presumed anatomical location within the midbrain. The six panels of this figure represent a sequence of coronal sections from caudal to rostral. Each diamond represents a site at which vocalization was elicited in at least one dog. The midpoint of the solid lines in each panel is the mean location, in lateral and dorsal dimensions, of all sites represented in this section. The standard deviation of these sites is represented by the length of the solid lines. As this figure illustrates, vocalization was elicited at sites from A2-12, L2-8, and D4-15. From the most posterior to anterior locations, the centroid of these sites changes dorsally from D7 to D11, but changes little in the medial-lateral dimension. It appears that locations of these sites are relatively stable in the medial-lateral dimension, slightly variable in the anterior-posterior dimension, and quite variable in the dorsal-ventral dimension.

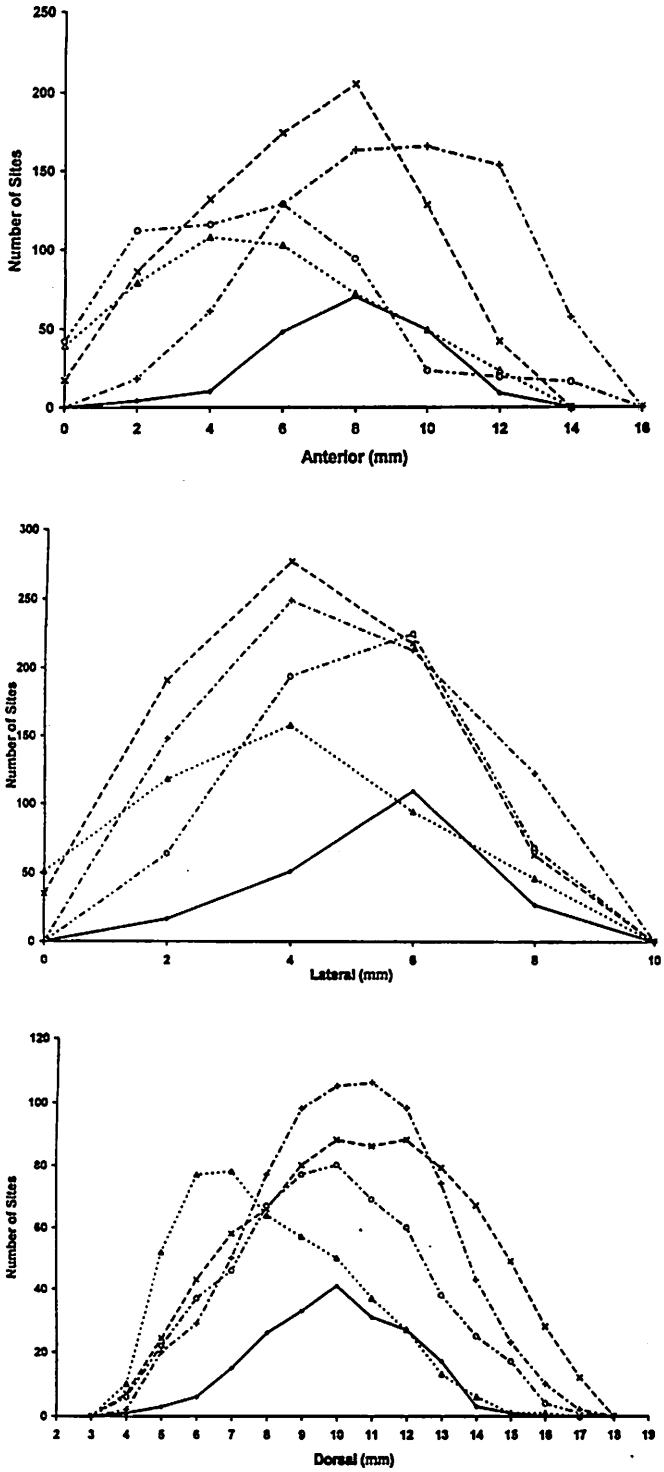


Figure 3. Frequency histograms for the total number of sites, except those having odd-numbered anterior or lateral coordinates for Dog 3, at which vocalization (filled diamond), facial (+) movement, body (open circle) movement, jaw closure (open triangle), and respiratory responses (x) were elicited in all 9 dogs. The three-dimensional locations of these sites are illustrated separately in the three graphs for anterior, lateral, and dorsal sections relative to earbar zero. Solid and dashed lines between adjacent symbols are not meant to indicate continuous data, but are included for clarity of presentation.

To illustrate the frequency of sites at which vocalization was elicited, frequency histograms were constructed for each of the three coordinates in the stereotaxic system (Fig. 3). The frequencies of the sites at which facial movement, body movement, respiratory response, and jaw closure were induced are illustrated as well to show their distributions in and around the vocalization region. The histograms represent the data from all 9 dogs. Only the odd-numbered tracks in the anterior and lateral dimensions were used for Dog 3, however, to make the data comparable with those from the other dogs, all of whom had 2 mm intervals between tracks. For clarity of presentation, frequency counts are represented by symbols connected by lines which do not represent values between samples rather than the more typical bar graph.

Of the 204 vocalization sites included in Figure 3, 78% (n=160) were located in a region defined by the coordinates A6-10, L4-6, and D7-13. In Fig. 2, this modal region for elicited vocalization corresponds to the general areas encompassed by the error bars in sections A6, A8, A10. Generally speaking, this region is just medial and dorsal to the medial lemniscus at the level of the superior colliculus. Distributions of other motor responses (facial movement, body movement, respiratory change, and jaw closure) also are documented in Fig. 3.

Most of the elicited vocalizations could be described as howls (medium-high pitch, smooth), growls (low-medial pitch, rough), whines (very high pitch), and barks (medium-high pitch and short). Additional vocalizations were too weak, rough, or unclear to include them with confidence into any of these categories. Among all types of vocalizations, the howl was the most common in each dog (54-67%). Growls and whines constituted 28-35% and 11-19% of the responses, respectively. Barks were only elicited at a few sites in two dogs. These frequency counts are similar to those reported in another study in which we examined acoustic and aerodynamic characteristics of evoked vocalizations in the dog [42].

A single type of vocalization usually was elicited at each stimulation site. The vocalizations generally sounded loud and clear at high current levels, then became weak and/or rough with the reduction of current, and disappeared at sub-threshold current levels. However, at some sites the type of vocalization changed with stimulus intensity (current). Usually, the alteration was from a howl to a growl as current was reduced. Changes from a whine to a rough or unclear sound also were observed. The vocalization never changed from a low- to high-pitched production with a reduction in current.

Low-threshold vocalization sites were relatively ventral and medial within the vocalization region. High-threshold sites were relatively dorsal, lateral, anterior, or posterior in the region. At some low-threshold sites, natural-sounding vocalizations were elicited without facial and body responses even at high current levels (e.g. 1.5 mA).

Regardless of the type of vocalization, the threshold varied from 0.03 to 0.7 mA for all vocalization sites.

Discussion

Vocalization was elicited by electrical stimulation of the midbrain in 9 dogs anesthetized with pentobarbital. Although the total area for the elicitation of vocalization was found to occupy much of the midbrain, the vocalization was elicited most frequently in a region just dorsal and medial to the medial lemniscus and ventral to the inferior and superior colliculi (A6-10, L4-6, and D8-13). According to the atlas of the canine brain [13], this vocalization region is lateral to but not in the midbrain PAG. Although this atlas is based on the beagle brain, it is likely to be accurate for most canine species, especially for structures close to earbar zero.

This midbrain region identified for electrically evoked vocalization in the dog is comparable to those reported for other animals. In monkeys, species-specific calls were elicited when electrical stimulation was applied at lateral and ventral to rather than in the central nuclei of the brainstem [32,22]. In cats, sites for electrically evoked vocalization without any autonomic and other motor responses were located near the medial lemniscus [12,25].

In monkeys and cats, it has been presumed that the vocalization region beyond the PAG, defined by electrical stimulation, consists of fibers of passage because this region was rarely sensitive to chemical stimulation [4,24,38]. Based on the fact that these brain structures in the monkey and the cat are similar to those in the dog, it is possible that the canine midbrain region for electrical evoked vocalization also is composed of axons.

The course of the axons surrounding the PAG, which apparently synapse with motoneurons of vocalization-related muscles, is unclear. Our finding that most of the vocalizations evoked by electrical stimulation to the midbrain were accompanied by facial movement suggests that these axons may be involved in the central control of vocalization for emotional expression. In the monkey and cat, emotive vocalizations were induced by exciting the lateral group of caudal PAG neurons and the adjacent lateral tegmental tissues in the midbrain [5,6,23,29]. If similar structures and pathways exist in the dog, it is reasonable to assume that this group of axons is a PAG-connected or PAG-projected descending pathway for emotive vocalization.

In the monkey, species-specific calls (cackle and growl) were elicited by electrical stimulation of a number of nuclei in the diencephalon and forebrain, including the hypothalamus, midline thalamus, amygdala, stria terminalis, preoptic region, septum, rostral hippocampus, posteromedial orbitalcortex, cingulate gyrus, and rostroventral temporal cortex [22]. In the cat, call sites in the area of the medial lemniscus were found, by tracing studies, to connect to the red nucleus and contralateral oculomotor nucleus [12]. Any of these nuclei may send descending projections to the PAG. Although the axons presumably stimulated in our study may come from higher structure, it is certainly possible that they comprise descending fibers from the PAG to premotoneurons and/or motoneurons in the pons, medulla, and spinal cord. Holstege [19] found in the cat that nuclei in the lateral part of the caudal PAG and in the adjacent lateral tegmentum project bilaterally to the nucleus retroambiguus (NRA). Cells in the NRA then project to motoneurons of vocalization-related muscles. In addition, direct connections from the PAG to nucleus ambiguus, which contains laryngeal motoneurons, have been reported in the monkey [23,33].

Vocalization evoked by electrical stimulation of the midbrain has been reported to vary in type, depending on stimulus position and/or intensity in the monkey [22,32], cat [12], and dog [43]. We again observed variations in vocalization type in the dog. Indeed, our finding that multiple types of vocalization could be elicited at one stimulation site supports the hypothesis that there is a common pathway for vocalization in the area of the medial lemniscus [12,25].

Motor responses other than vocalization elicited by electrical stimulation of the midbrain have been reported to include opening and closing of the mouth, elevation of the tongue, dilatation of the nostrils, retraction of an angle of the mouth, wrinkling of the forehead, retraction and flattening of the ear in cats and monkeys [32], stretch and contraction of legs, movement of the body in monkeys [22], and respiratory effects in cats and monkeys [22,32]. All of these responses were observed in the present experiment except movement of the tongue. The tongue was fixed by a brace, however, so we can not exclude the possibility that these stimuli produced activation of tongue muscles. In addition, movements of the eyelid and eyeball were observed. We described 4 groups of responses (facial movement, body movement, jaw closure, and respiration) induced frequently in our experiment and displayed their distributions in Fig. 3. The relationships of stimulation sites producing these responses to the vocalization region are illustrated in Fig. 4 (following page), in comparison to the center of the vocalization region.

The purpose of our study was to guide the search for sites in the midbrain at which vocalization is most likely to be elicited by electrical stimulation. It should be understood that region identified for vocalization and its positional relationship to other motor responses are based on frequency counts of observations from a group of mongrel dogs. Differences may exist between canine subspecies. In addition, we observed minor differences between animal subjects. These differences could be related to individual variations in midbrain anatomy, variations in skull configuration, which could affect the stereotaxic coordinates of the brain, and slight variations in the calibration of the stereotaxic coordinate system. Nevertheless, our results provide, objective evidence for making a good guess, based on stereotaxic procedures, for locating vocalization sites in a

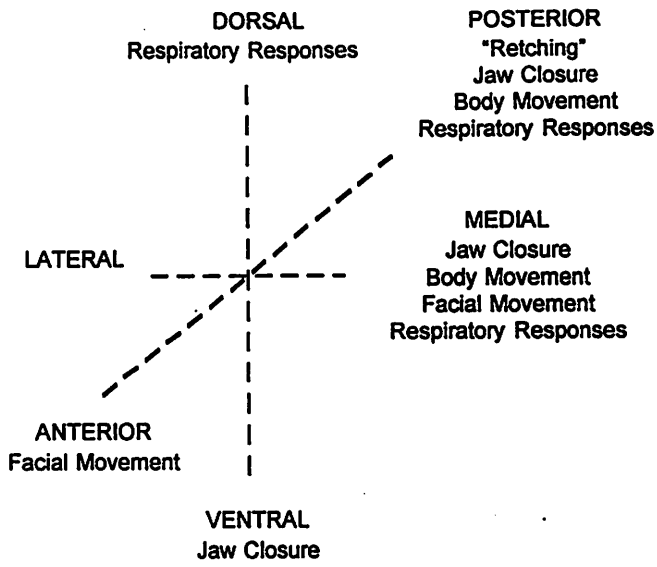


Figure 4. A schematic representation of positional relationships of non-vocal motor responses to the vocalization region in the canine midbrain. The midpoint of lines represents the center of the vocalization region.

non-pedigreed dog. We have used this information in subsequent experiments, and have found that locating a midbrain site for eliciting vocalization has been substantially more efficient than in the past.

In conclusion, a region for elicitation of vocalization by electrical stimulation was identified in the canine midbrain. This region appears to be located just dorsal and medial to the medial lemniscus at the level of the superior colliculus. Within this region, natural-sounding vocalizations could be elicited without other motor responses, and different types of vocalizations were evoked at various stimulation sites or with different stimulus intensities. The vocalization region apparently is composed of axons from a motor pathway either projecting to or descending from the midbrain PAG. Finally, the functional relationships that were described between stimulation sites resulting in vocalization or other motor responses serve as a "map" that can be used for locating the midbrain vocalization region efficiently. These results should aid in the implementation of future investigations of midbrain control of vocalization.

References

[1] Adametz, J., and O'Leary, J. L., Experimental mutism resulting from periaqueductal lesions in cats, *Neurology (Minneap)*, 9(1959):636-642.

[2] Alipour-Haghighi, F., Titze, I. R., and Perlman, A. L., Tetanic contraction in vocal fold muscle, *J. Speech Hearing Res.*, 32(1989):226-231.

[3] Alipour-Haghighi, F., Titze, I. R., and Durham, P., Twitch response in the canine vocalis muscle, *J. Speech Hearing Res.*, 30(1987)290-294.

[4] Bandler, R.J., Induction of 'rage' following microinjections of glutamate into midbrain but not hypothalamus of cats, *Neurosci. Lett.*, 30(1982):183-188.

[5] Bandler, R. J., and Carrive, P., Integrated defence reaction elicited excitatory amino acid microinjection in the midbrain periaqueductal region if the cat, *Brain Res.*, 439(1988):95-106.

[6] Bandler, R. B., and Depaulis, A., Midbrain periaqueductal gray control of defensive behavior in the cat and the rat. In: A. Depaulis and R. B. Bandler, (Eds.), *The Midbrain Periaqueductal Gray Matter*, Plenum Press, New York, 1991.

[7] Berke, G. S., Moore, D. M., Hantke, D. R., Hanson, D. M., Gerratt, B. R., and Burstein, F., Laryngeal modeling: Theoretical, in vitro, in vivo, *Laryngoscope*, 97(1987)871-881.

[8] Berke, G. S., and Gerratt, B. R., Laryngeal biomechanics: An overview of mucosal wave mechanics. *J. of Voice* 7(1993):123-128.

[9] Davis, P. J., Zhang, S. P., and Bandler, R., Pulmonary and upper airway afferent influences on the motor pattern of vocalization evoked by excitation of the midbrain periaqueductal gray of the cat, *Brain Res.*, 00(1993):1-20.

[10] de Molina, A., F., and Hunsperger, R. W., Central representation of affective in the forebrain and brainstem: electrical stimulation of amygdala, stria terminalis adjacent structures, *J. Physiol.*, 145(1959):251.

[11] de Lanerolle, N. C., and Lang, F. F., Functional neural pathways for vocalization in the domestic cat. In: J. D. Vewman, (Ed.), *The Physiological Control of Mammalian Vocalization*, New York, Plenum Press, 1988.

[12] de Lanerolle, N. C., Apontine call site in the domestic cat: Behavior and neural pathways, *J. Neurosci.*, 37(1990)201-214.

[13] Dua-Sharma, S. D., Sharma, K. N., and Jacobs, H. L. (Eds.), *The Canine Brain in Stereotaxic Coordinates*, The MIT press, London, 1970.

[14] Gacek, K. E., Localization of laryngeal motoneurons in the kitten, *Laryngoscope*, 85(1975)1841-1860.

[15] Green, D. C., Berke, G. S., and Ward, P. H., Vocal fold medialization by surgical augmentation versus arytenoid adduction in the in vivo canine model. *Ann. Otol. Rhinol. Laryngol.*, 100(1991):280-287.

[16] Hast, M. H., The primate larynx, *Acta. Oto-laryngologica,* 6(1969)84-92.

[17] Hirano, M., Morphological structure of the vocal cords as a vibrator and its variations, *Folia. Phoniat.*, 26(1974):89:94.

[18] Hirano, M., Structure and vibratory behavior of the vocal folds, In: M. Sawashima and F. S. Cooper, (Eds.), *Dynamic Aspects of Speech Production*, University of Tokyo Press, 1977.

[19] Holstege, G., Anatomical study of the final common pathway for vocalization in the cat,*J. Comparative Neurology*, 284(1989):242-252.

[20] Hsiao, T. Y., Solomon, N.P., Luschei, E.S., Titze, I.R., Liu, K., Fu, T. C., and Hsu, M. M., The effect of subglottic pressure on fundamental frequency of the canine larynx with active muscle tensions, *Annals of ORL. Rhinol. & Laryngol.*, 103(1994)817:821.

[21] Hunsperger, R. W., and Bucher, V., Affective behavior produced by electrical stimulation in the forebrain and brainstem of the cat, *Progress in Brain Res.*, 27(1967):103.

[22] Jürgens, U., and Ploog, D., Cerebral representation of vocalization in the squirrel monkey, *Exp. Brain Res.*, 10(1970):532-554.

[23] Jürgens, U., and Pratt, R., Role of periaqueductal gray in vocal expression of emotion, *Brain Res.*, 167(1979):367-378.

[24] Jürgens, U., and Richter, K., Glutamate-induced vocalization in the squirrel monkey, *Brian Res.*, 373(1986):349-358.

[25] Kanai, T., and Wang, S. C., Localization of central vocalization mechanisms in the brainstem of the cat, *Exp. Neurol.*, 6(1962):426.

[26] Kelley, A. H., Beaton, L. E., and Magoun, H. W., A midbrain mechanism for facio-vocal activity,*J. Neurophys.*, 9(1946):181.

[27] Kirzinger, A., and Jürgens, U., The effects of brainstem lesions on vocalization in the squirrel monkey, *Brain Res.*, 358(1985):150-162.

[28] Kurita, S., Nagata, K., and Hirano, M., A comparative study of the laryngeal structure of the vocal fold, In: D. M., Bless and C. H. Abbs, (eds.), *Vocal Fold Physiology*, San Diego, College-Hill Press, pp 3-21.

[29] Larson, C. R., The midbrain periaqueductal gray: a brainstem structure involved in vocalization,*J. Speech Hear Res.*, 28(1985):241-249.

[30] Larson, C. R., and Kistler, M. K., The relationship between periaqueductal gray neurons to vocalization and laryngeal EMG in the behaving monkey, *Exp. Brain Res.*, 63(1986):596-606.

[31] Larson, C. R., and Kempster, G. B., Voice fundamental frequency changes following discharge of laryngeal motor units, In: I. R. Titze, and R. C. Scherer, (eds.), *Vocal Fold Physiology: Biomechanics, Acoustics And Phonatory Control*, Denver, The Denver Center for the Performing Arts, 1983:91-104.

[32] Magoun, H. W., Atlas, D., Ingersoll, E. H., and Ranson, S. W., Associated facial, vocal and respiratory components of emotional expression: an experimental study. *J. Neurophysiol.*, 17(1937):241-255.

[33] Martin, J. R., Motivated behaviors elicited from hypothalamus, midbrain, and pons of guinea pig (Cavia porcellus), *J. Comp. Physiol.*, 90(1976):1011-1034

[34] Moore, D. M., and Berke, G. S., The effect of laryngeal nerve stimulation on phonation: A glottographic study using an in vivo canine model,*J. Acoust. Soc. Am.*, 83(1988):705-715.

[35] Nakao, H., Yoshida, M. M., and Sasaki, T., Midbrain central gray and switch-off behavior in cats,*Jap. J. Physiol.*, 18(1968):462.

[36] Randall, W. L., The behavior of cats (felis catus) with lesions in caudal midbrain region,*Behavior*, 36(1964):107-139.

[37] Richter, K. and Jürgens, U., A comparative study on the elicitablity of vocalization by electrical brain stimulation, glutamate, aspartate and quisqualate in the squirrel monkey, *Neuroscie. Letters*, 66(1986)239-244.

[38] Schuller, G., and Radtke-Schuller, S., Neural control of vocalization in bats: mapping of brainstem areas with electrical stimulation eliciting species-specific echolocation calls in the rufous horseshoe bat, *Exp. Brain Res.*, 79(1990):192-206.

[39] Skultety, F. M., Experimental mutism in dogs, *Arch. Neurol.* (chic.), 6(1962):235-241.

[40] Skultety, F. M., The behavioral effects of destructive lesions of the periaqueductal gray matter in adult cats, *J. Comp. Neurol.*, 110(1958):337-365.

[41] Solomon, N. P., Grayhack, J. P., Liu, K., and Luschei, E. S., Patterns of laryngeal and respiratory activation during elicited phonation, *ASHA [Abstract]*, 34(1992):175.

[42] Solomon, N. P., Luschei, S. E., and Liu, K., Fundamental frequency and tracheal pressure during three types of vocalizations elicited from anesthetized dogs, *J. Voice*, 1995, manuscript in press.

[43] Suga, N., Schlegel, P., Shimozawa, T., and Simmons, J., Orientation sounds evoked from echolocating bats by electrical stimulation of the brain, *J. Acoust. Soc. Am.*, 54(1973):793-797.

[44] Titze, I. R., On the relation between subglottal pressure and fundamental frequency in phonation. *J. Acoust. Soc. Am.*, 85(1989)901-906.

[45] Titze, I. R., Jiang, J., and Drucker, D. G., Premilinaries to the body-cover theory of pitch control, *J. Voice*, 4(1988)314-319.

[46] Van den Berg, J., Subglottal pressure and vibrations of the vocal folds, *Folia. Phonitr.* (Basel), 9(1959):65-71.

[47] Yajima, Y., and Hayashi, Y., Ambiguus motoneurons discharging synchronously with ultrasonic vocalization in rats, *Exp. Brain Res.*, 50(1983):359-366.

NCVS Status and Progress Report - 8
July 1995, 65-73

Sulcus Vocalis: A Rational Analytical Approach to Diagnosis and Management

Charles Ford, M.D.
Katsuhide Inagi, M.D.
Division of Otolaryngology, Department of Surgery, The University of Wisconsin
Diane Bless, Ph.D.
Division of Otolaryngology, Department of Surgery, The University of Wisconsin
Department of Communicative Disorders, The University of Wisconsin
Aliaa Khidr, M.D.
Department of Communicative Disorders, The University of Wisconsin
Kennedy Gilchrist, M.D.
Department of Pathology, The University of Wisconsin

Abstract

The term sulcus vocalis has been applied to a spectrum of pathology ranging from minor vocal fold indentations to destructive lesions causing severe dysphonia. To clarify the pathophysiology and to develop a more rational approach to treatment we report a series of sulcus patients including 20 surgical cases. Clinical and histopathological analysis produced a clinically useful classification: Type I is a physiological variant accentuated by atrophy but with intact lamina propria; Types II (sulcus vergeture) and III (sulcus vocalis) are characterized by severe dysphonia, loss of vibratory activity, and destruction of the functional superficial lamina propria. These latter cases respond favorably to microsurgery designed to remove destroyed tissue, release scar contracture, and promote mucosal re-draping by regional undermining. Further study of the extracellular matrix of the superficial lamina propria (Reinke's space) might indicate a common pathway in the pathogenesis of sulcus deformities and other related benign vocal fold lesions.

Introduction

Progress in assessment and successful management of sulcus vocalis has been hampered by inconsistent terminology and a lack of agreement on the basic pathology. *Sulcus* is a generic anatomic term referring to an indentation, groove, or furrow. The earliest mention of *sulcus vocalis* was probably in 1892 by the anatomist Giacomini[1], who described a sulcus deformity of the vocal fold. Bouchayer[2] distinguished a focal invagination of epithelium attaching to the vocal ligament (sulcus vocalis), from *sulcus vergeture*, which he described as a linear depression along the medial margin of the vocal fold. Lindestad and Hertegard[3] chose not to distinguish these, lumping them in the "sulcus" group; both groups exhibited *spindle-shaped glottal insufficiency* (SGI). Since the clinical appearance of many forms of SGI can be similar, patients with sulcus vocalis, furrows and bowing, presbylaryngis, paralysis, and connective tissue and/or thyroarytenoid muscle atrophy are often considered together. Adhering to rigorous anatomical criteria, Nakayama[4] identified shallow sulci in 4 of 20 larynges incidentally removed from random autopsy specimens.

There is agreement on the general clinical presentation of sulcus vocalis. Hirano et al.[5], Lindestad et al.[6], and Bouchayer and Cornut[7] have described the presenting symptoms of hoarseness, vocal fatigue, aesthenic voice quality, and effortful voice projection. In a recent comprehensive functional assessment of 126 sulcus vocalis patients, Hirano et al.[5] described perceptual ratings, acoustic and aerody-

namic functions, and analysis of videostroboscopy in this population. Most authors agree that laryngoscopic findings include spindle-shaped bowed vocal folds that fail to approximate, medial furrows seen during inhalation, impaired vibration, and frequent supraventricular hyperfunction.

Discussions on pathogenesis and treatment are less consistent and reflect a lack of agreement on terminology. Arnold[8] reviewed comparative anatomical studies supporting a congenital origin for sulci. Bouchayer[9] proposed that sulcus vocalis is congenital, a result of faulty development of the fourth and sixth branchial arches, and probably resulting from ruptured epidermoid cysts. He documented onset of dysphonia during childhood in 55% of sulcus patients, frequent associated cysts, failure of lesions to return after adequate excision, and descriptions of familial occurrence. Evidence for an acquired etiology dates back to 1928 when Van Caneghen[10] noted onset of sulcus vocalis in adult patients with potentially provocative traumatic or infectious factors (such as tuberculosis) evident. Also, the recent review of whole-mount laryngeal cancer specimens from the University of Wisconsin[4] identified adjacent sulcus deformities in 48% of specimens, supporting argument for an acquired pathogenesis.

Numerous reports of treatment results have failed to agree on suitable management strategies. This is partly due to lack of consensus about the definition of sulcus vocalis and a tendency to address only one component of the pathology. Voice therapy as a single modality has had limited success in the management of sulcus vocalis, but it is considered an essential adjunctive measure[11]. Injections of bioimplants[2], Teflon®[13], and steroids[7], as well as medialization thyroplasties[14], have been helpful but are inadequate without additional measures. Incision with undermining has been done in an attempt to reconstitute a looser sub-epithelial plane[7] and this has been combined with vertical incisions (slicing) to break up linear tension lines[15]. Direct excision by microdissection[7] and with laser ablation (Marc Remarcle, M.D., Yvoir, Belgium, 1989, personal communication and review of videotapes) have been used successfully in some sulcus vocalis cases.

In clinical practice we have observed many patients with vocal fold configurations that resemble sulcus deformities. The spectrum of pathology ranges from physiological vibratory contours revealing medial vocal fold furrows to vocal fold destructive lesions causing severe dysphonia. In an effort to clarify the pathophysiology and develop a more rational approach to treatment, we focus here on a review of 20 patients with surgically-treated pathological sulcus lesions. Clinical, surgical, and histopathological observations are described and these lesions contrasted to those in other patients exhibiting physiological sulcus configurations. In describing the clinical features and histopathology, we explore common denominators that might be useful as a guide in treatment selection.

Materials and Methods

Observation of a sulcus configuration to the vocal fold is evident in many patients evaluated in our voice clinic. To compare physiological and pathological cases of sulcus, we looked at two groups of subjects. We examined 116 volunteers who had no complaints, but who responded to an invitation to have their voice evaluated at no charge. All 116 presented with normal-to-mildly-impaired vibratory activity of the vocal fold and were considered as a group to have *physiological sulcus*. Of these, nine exhibited a physiological sulcus configuration on videostroboscopy. All 116 patients served as controls, and underwent no treatment.

The focus of this study was 20 patients with sulcus deformity and marked vocal fold stiffness indicating loss of a functional superficial lamina propria (Table 1). The vocal fold pathology in this *pathological sulcus* group was addressed by surgical intervention. Patients were nine males and 11 females, ages 10 to 68 years (average 32.0, SD 13.1). The most common presenting complaint was prolonged hoarseness, with age of onset equally divided between childhood and adulthood. Each case was reviewed to identify possible environmental, focal, or systemic medical factors that might have caused the condition. Most patients had a history of laryngitis or vocal abuse; allergy and prior intubation were also common; and 25% of patients had a history of gastroesophageal reflux. Alcohol intake and smoking were not particularly common in this group. Two patients had connective tissue diseases (arthritis and systemic lupus erythematosus). A family history of chronic hoarseness was present in four cases.

All patients underwent preliminary voice evaluations including videostroboscopy and aerodynamic and acoustic evaluations as previously described[16]. In addition to voice therapy, interventions included injections of collagen and steroids, medialization thyroplasty, incision with undermining, undermining with vertical slicing, and excision. Excised specimens were fixed and processed with hematoxylin and eosin stains. The most recent follow-up assessment was used for comparison with pre-treatment studies.

Treatment Methods

Voice Therapy

All patients exhibited glottic insufficiency which perhaps predisposed them to compensatory hyperfunctional activity. The goals of therapy were to encourage vocal hygiene, enhance breath support, and avoid hyperfunctional activity. Voice therapy was considered an important perioperative adjunctive measure.

Injections

Intracordal injections were performed using the oro-tracheal injector (Xomed-Treace, Jacksonville, Florida).

Table 1.
Demographic Summary and Possible Etiologic Factors in Pathological Sulcus Group

Operative Patient	Age	Sex	Chief Complaint	Onset (Years Ago)	Possible Causes								
					Infection/ Laryngitis	Vocal Abuse	Vocal Fold Surgery	Intubation	GERD	Smoke	Daily ETOH	Family History	Other Disease/Allergy
1	10	M	Hoarseness	3		+						+	
2	26	M	Hoarseness	22	+		+	+	+				
3	27	M	Hoarseness	12				+			+		
4	31	M	Hoarseness	20		+							+
5	32	M	Hoarseness	6		+			+			+	+
6	33	M	Chronic Hoarseness	>20									+
7	36	M	Hoarseness	>30		+			+	+	+	+	
8	38	M	Abnormal Voice	>30	+								
9	54	M	Breathy Dysphonia	4				+		+		+	
10	13	F	Abnormal Voice	13									
11	21	F	Hoarseness	1	+								
12	23	F	Recurrent Raspy Voice	5	+		+	+			+		
13	23	F	Hoarseness	3	+	+				+			+
14	27	F	Hoarseness	3	+	+	+	+	+				+
15	29	F	Recurrenct Hoarseness	11		+	+	+					+
16	32	F	Hoarseness, Raspy Voice	3	+	+	+	+	+				
17	35	F	Hoarseness	3		+							
18	39	F	Voice Loss	3	+								+
19	43	F	Recurrent Hoarseness	1		+	+	+					+
20	68	F	Hoarseness	>20		+	+	+			+		+

Saline or 1% lidocaine with 1:10,000 epinephrine were injected to facilitate microdissection. Therapeutic injections of triamcinolone acetonide (10 mg/ml) was also performed intraoperatively in some cases. In an effort to correct bowing and to address scar tissue softening, cross-linked collagen (Zyplast®, Collagen Corporation, Palo Alto, California) and, recently, autologous collagen (Autologen®, Autogenesis, Acton, Massachusetts) were injected into the most superficial layers accessible as previously described[16]. When injections failed to distribute appropriately, microincisions were placed to create a plane between the vocal ligament and covering epithelium. Neither Teflon® nor fat injections were used to treat sulcus patients.

Incision and Undermining

Incisions were placed just superior to the edge of linear sulci and the epithelium sharply dissected from the vocal ligament. Additional undermining for a distance of 2 mm was done inferior to the sulcus to achieve adequate freeing of the mucosa. Neither suturing nor gluing was used with this technique.

Incision, Undermining, and Sectioning *(Slicing)*

This technique, described by Pontes and Behleu[15], is an extension of the incision and undermining approach just described. When the inferior flap has been undermined, a series of four of five incisions are placed in a vertical direction producing three or four mucosal segments (Photo 9A-H; see center plate). An important detail, illustrated but not discussed in the published description, is that each incision should be a different length (Paula Pontes, M.D., Sao Paulo, Brazil, personal communication, 1993); this impedes subsequent linear contracture. The superior surface mucosa is also undermined about 2 mm to help in redraping. The immediate postoperative appearance reveals denuded gaps along the medial edge that heal remarkably well.

Excision

All excisions were done using microdissection techniques; laser was not employed; hemostasis was accomplished with topical 1:10,000 epinephrine. Incisions were placed to circumscribe the medial pore of the sulcus (Photo 10A-H; see center plate) and an inferior flap developed. The superior approach was then used to dissect the sulcus away from underlying vocal ligament to which it was invariably adherent. Occasionally, sharp dissection into vocal ligament with limited resection of ligamentous structure was necessary. Even with extensive (2-4 mm) undermining, it is usually not possible to achieve good coaptation of the edges. When the residual gap was greater than 2 mm, a single 5-0 chromic catgut suture was placed (Photo 10D, H; see center plate).

Results

Patients with sulcus deformities of the vocal folds included those with apparently intact lamina propria without voice complaints at one end of the spectrum, and those with pathological alteration of lamina propria and grossly disturbed voices at the other. Videostroboscopy was useful in distinguishing these two groups. The physiological sulcus configuration group included patients with presbylaryngis,

Table 2.
Videostroboscopic Findings in Control Subjects with Physiologic Sulcus

Control Patient	Sex	Age	Ampl. R	Ampl. L	Wave R	Wave L	Asym.	Closed PH.
1	F	33	NL	NL	NL	NL	NL	NL
2	M	34	NL	NL	-1	-1	Minimal	NL
3	M	36	-1	NL	-1	NL	Minimal	-1
4	M	51	-1	-1	-1	-1	Minimal	NL
5	M	36	-1	-1	-1	-1	Minimal	-1
6	M	32	NL	NL	NL	NL	Minimal	-1
7	M	39	NL	NL	NL	NL	NL	NL
8	M	31	-1	-1	-1	-1	NL	-2
9	M	35	NL	NL	-1	NL	NL	NL

Ampl. = amplitude of vibration; NL = normal limits; -1 = markedly reduced; -3 = increased
Wave = mucosal wave (same scale)
Closed PH. = closed phase; -1 = predominately open; -3 = predominantly closed

paralysis (Photo 11A; see center plate), and bowed vocal folds, and a control group of volunteers with no specific voice complaints (Photo 11B; see center plate). Normal or minimally altered mucosal wave was characteristic of this latter group of patients. Table 2 shows the videostroboscopic findings in nine patients, of the 116 essentially normal voices, who exhibited physiological sulcus and spindle-shaped glottal areas. In the most subtle cases, the sulcus was seen only on videostroboscopy during phonation but in others was apparent on indirect laryngoscopy during abduction. For purposes of further discussion, all of these patients will be considered as **TYPE I or physiological sulcus.**

Diagnostic Assessment of Pathological Sulcus Patients

Twenty other patients exhibited more advanced changes, with vocal fold stiffness (markedly reduced amplitude of vibration to complete absence of mucosal wave) and anatomical alterations indicating vocal fold pathology. Associated clinical findings included vocal fold edema, capillary ectasia, focal fullness, and vibratory disturbance (Table 3). These patients were considered candidates for surgical exploration and treatment. This pathological sulcus group included two distinct types of lesions that were usually evident on videostroboscopy. One pattern was a linear sulcus along the medial vocal fold edge separating superior and inferior lips of membranous vocal fold by a rigid contracted band; this pattern is consistent with Bouchayer's description of a vergeture[2] and will be designated **TYPE II or sulcus vergeture** (Photo 12A; see center plate). The other pathological sulcus pattern was characterized by severely dysphonic voices in patients who might exhibit minimal abnormality on indirect laryngoscopy but demonstrated focal stiffness on videostroboscopy. Further examination occasionally revealed a medial pit-shaped sulcus that became obvious at microlaryngoscopy (Photo 12B; see center plate). This **TYPE III or sulcus vocalis** is the most specific pathological sulcus designation.

Operative Observations

Micro-surgical intervention was routinely preceded by careful examination with the operating microscope. Inspection of the superior vocal fold surface often provided a clue as to the location and extent of intra-fold pathology. Pathological sulci were invariably associated with prominent mucosal vessels or capillary ectasia on the superior surface of the vocal fold. Unlike the normal anterior-to-posterior oriented vascular pattern, vocal folds with pathological sulcus vocalis typically exhibited vessels branching at 90° angles and running perpendicular to the long axis of the vocal fold (Photo 12A; see center plate).

Additional information found during microlaryngoscopy resulted from palpation and manipulation of the vocal fold. Palpation with a fine curved alligator forceps revealed focal induration either distributed in a linear fashion along the medial vocal fold or associated with a focal invagination or pit. These deep pits could be displayed by grasping the inferior lip to distend the cavity (Photo 12B; see center plate). The curved alligator was also useful to probe the mucosal surface for mucosal bridges and to distend the anterior commissure to display associated microwebs. The associated lesions observed in this group of patients are also tabulated in Table 3.

Histopathology

Specimens from this study were limited to those patients who underwent direct microlaryngoscopy with excision of the sulcus. One excised sulcus vergeture (Type II) demonstrated epithelial thinning and loss of the lamina propria cellular layered structure (Photo 13A; see center plate). The relationship of these pathological changes to the underlying vocal fold structure can be appreciated in the whole-mount preparation (Photo 13B; see center plate). These lesions demonstrated adherence of a thinned epithelium to the intermediate and deep layers of lamina propria but there was no evidence of vocal ligament invasion and there was minimal inflammatory response.

The focal Type III sulcus vocalis lesions were associated with varying degrees of inflammation. In some cases this was limited to a few inflammatory cells and increased capillaries (Photo 14A; see center plate). In one case where the sulcus appeared to arise from a ruptured epidermoid cyst, there was intense inflammation with massive neovascularization and fibrosis extending through much of the vocal ligament (Photo 14B; see center plate). In other

Table 3.
Laryngeal Findings in Operative Patients

Operative Patient	Pre-Operative Findings							Operative Findings		
	Capillary Ectasia	Edema	Bowing	Fullness	Mucosal Wave*	Amplitude*	Glottal Closure**	Cyst	Microweb	Mucosal Bridge
1				+	-1	-1	Incomplete		+	
2			+		-1	-1	Spindle			
3					-2	-2	Incomplete			
4			+		-2	-1	Incomplete			+
5	+	+		+	-2	-2	Posterior	+		multiple +
6			+		-2	-2	Spindle			
7			+		-2	-2	Longitudinal			
8			+		-2	-2	Spindle			
9			+		-2	-2	Incomplete			
10			+		-1	-1	Spindle			
11			+		-2	-2	Spindle	+	+	
12		+	+		-3	-2	Spindle			
13				+	-3	-3	Longitudinal	+		
14	+	+		+	-2	-2	Incomplete	+		
15	+				-3	-3	Posterior	+	+	
16			+		-3	-3	Posterior		+	
17	+		+		-2	-2	Posterior	+		+
18				+	-3	-2	Posterior	+		
19				+	-2	-2	Incomplete	+		
20			+	+	-3	-3	Spindle	+		

* = Mild Decreased = -1; Moderate Decreased = -2; Severe Decreased = -3

** = Incomplete = Incomplete Closure; Spindle = Spindle Shape Gap; Longitudinal = Longitudinal Gap; Posterior = Posterior Gap

instances, the sulcus was associated with intact epidermoid cysts. In contrast to the vergeture lesions, there was a tendency for these Type III sulci and associated cysts to extend into the vocal ligament.

Results of Surgical Intervention

Initially we treated patients conservatively with voice therapy, collagen injections, and medialization thyroplasty. Voice improvements were limited in all of those early cases. Subsequent efforts involved direct surgical approaches to the affected vocal folds with techniques varied depending on the pathology. Subjective improvements were much better in these 20 cases, and videostroboscopic findings confirmed improved glottal closure during phonation, improved vibratory activity, and better symmetry (Table 4, following page). Objective measures confirmed voice improvement in most subjects (Figure 1A-F). This was reflected in a reduction in perturbation measures (jitter and shimmer) and increase in signal-to-noise ratio. Frequency range was generally increased after intervention. Transglottic airflow was somewhat variable but maximum phonation time tended to increase.

Discussion

Making a meaningful diagnosis is essential to a substantive discussion of pathophysiology and pathogenesis, and to develop rational management strategies. It is not helpful to affix the sulcus vocalis label to all patients with vocal fold invaginations and spindle-shaped glottic configurations. The term sulcus literally means a depression or furrow and does not describe a pathological entity. Such depressions on the medial surface of vocal folds can occur in patients without vocal complaints, as in 9 of the 116 volunteers reported here, and might be secondary to atrophic vocal fold changes from aging or paresis. These lesions that we have designated as Type I (physiological sulcus) can be distinguished clinically by noting preservation of vocal fold vibratory activity on videostroboscopy. The presence of mucosal wave activity indicates adequate functional separation of vocal fold body and cover as described by Hirano and Kakita[17]. The histological correlate to such separation is an intact superficial lamina propria. It is important to understand that this population of sulcus patients exists and that it differs from the Type II (vergeture) and III (vocalis) group

Table 4.
Surgical Treatment and Operative Results in Pathological (Types II & III) Sulcus Patients

Operative Patient	Type of Operation*		Type of Sulcus**		Operative Results***					
	R-VF	L-VF	R-VF	L-VF	Glottal Gap	Shape of Vocal Fold	Phase-CL	Amplitude	Mucosal Wave	Asymmetry
1	UND	UND	II	II	-	-	-	-	-	-
2	EXC	LYSIS+COL	II	II	ABS	IMPR	IMPR	IMPR	IMPR	IMPR
3	UND	EXC	II	III	-	-	-	-	-	-
4	EXC	EXC	III	III	-	-	-	-	-	-
5	EXC	EXC	III+CYST	III	ABS	IMPR	NC	DECR	DECR	NC
6	EXC	EXC	III	III	DECR	NC	-	IMPR	IMPR	-
7	-	LYSIS+COL	II	II	DECR	IMPR	IMPR	IMPR	IMPR	IMPR
8	LYSIS+COL	COL	II	II	ABS	IMPR	IMPR	IMPR	IMPR	NC
9	S	UND+S	II	II	DECR	-	-	NC	NC	-
10	UND+S	UND+S	II	II	INCR	IMPR	NC	NC	NC	NC
11	EXC	UND	III	II	ABS	IMPR	NC	IMPR	IMPR	NC
12	UND	EXC	II	II	DECR	IMPR	IMPR	IMPR	IMPR	IMPR
13	EXC	EXC	III	III	DECR	IMPR	IMPR	IMPR	IMPR	IMPR
14	EXC	EXC	III	III	-	-	-	-	-	-
15	EXC	EXC	III	III+CYST	NC	IMPR	NC	IMPR	IMPR	NC
16	UND	THY	II	II	DECR	IMPR	IMPR	IMPR	IMPR	NC
17	EXC	EXC	III	III	ABS	IMPR	-	IMPR	IMPR	IMPR
18	SLICE	UND	II	II	DECR	IMPR	IMPR	IMPR	IMPR	IMPR
19	EXC	EXC	III+CYST	III	ABS	NC	IMPR	IMPR	IMPR	NC
20	UND	SLICE	II	II	ABS	IMPR	IMPR	IMPR	IMPR	NC

* = UND = Undermine; COL = Collagen Injection; EXC = Excision Technique; THY = Thyroplasty Type I; S = Steroid Injection

** = II = Vergeture; III = Vocalis

*** = ABS.= Absent; INCR = Increase DECR = Decreased; NC = No Change; IMPR = Improved

by having essentially intact layers of lamina propria. Considering any of these cases as sulcus vocalis might account for the great variability in reported incidence of sulcus in voice disordered patients.

Appropriate management of Type I sulcus patients varies depending on the underlying pathology. In some instances it might be an incidental finding and no treatment is necessary in such asymptomatic patients. When associated with paresis or loss of muscle bulk, voice therapy and possibly vocal fold medialization techniques might be helpful. As a general rule, it is preferable to avoid surgical disruption of the intact functioning membranous vocal fold tissues in these physiological sulcus cases.

The hallmark of the pathological sulcus (Types II and III) is disturbance of the lamina propria with loss of body-cover separation. This pathological event disrupts normal vocal fold vibration, manifested by stiffness on videostroboscopy and severe dysphonia. The 20 surgical patients reported here were similar in their clinical presentation: They had severe dysphonia, vocal fold stiffness, and incomplete glottic closure. The operative findings and pathology indicated loss of superficial lamina propria and fixation of a thinned epithelium to underlying vocal ligament. In the Type II sulcus vergeture group, the attachment was linear; in the focal Type III sulcus vocalis, an epithelial-lined pit or contiguous cyst extended into or through the vocal ligament evoking an inflammatory reaction.

Unlike the physiological sulcus patients, pathological sulcus patients required direct microsurgical intervention to achieve satisfactory results. Initial attempts to treat these lesions with injections and medialization yielded minimal improvement. Treatment that succeeded removed diseased tissue and achieved improved separation of covering epithelium from vocal ligament and release of contracture. For Type II sulcus vergeture lesions, the best results have occurred in two patients treated with the undermining and segmental slicing technique of Pontes and Behlau[15]. This approach employs the benefits of freeing mucosal attachments and also breaks up the linear contracture that divides the vocal fold. Postoperative videostroboscopy showed remarkable return of mucosal wave and improved glottic closure in these cases. The spectacular results in this small series is insufficient to scientifically endorse such a radical-appearing approach, but it provides a rational approach to correcting the pathology.

The Type III sulcus vocalis is the most demanding because of the interface of lesion with remaining vocal fold structures. These sulci are typically associated with cysts, mucosal bridges, and/or considerable inflammation. Damaged fibrotic tissue and cystic remnants must be removed while preserving as much of the normal adjacent vocal fold

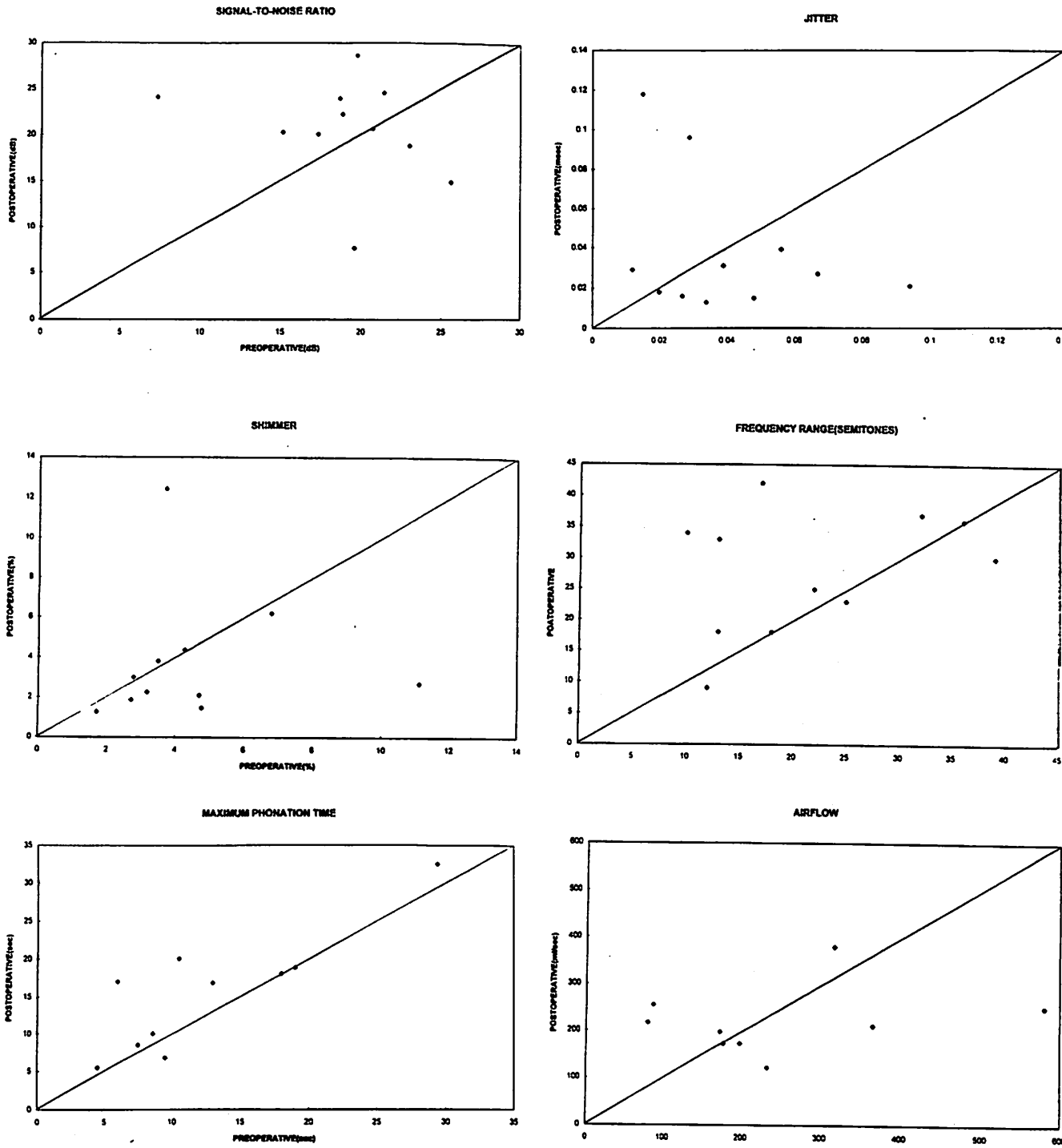


Figure 1. Shows scatter plots of patients who completed pre-operative and post-operative voice assessment measures. Parameters include (A-upper left) signal-to-noise ratio, (B-upper right) jitter, (C-middle left) shimmer, (D-middle right) frequency range, (E-lower left) airflow, and (F-lower right) maximum phonation time. Values that above diagonal indicate post-operative increased values and those below the diagonal show decreased values. Positive results were indicated by increased SNR (A-upper left), frequency range (D-middle right) and phonation time (F-lower right); a decrease in jitter (B-upper right), shimmer (C-middle left), and airflow (E-lower left) were also considered favorable.

tissue as possible. Once the lesion is excised, undermining of the residual mucosa facilitates coaptation and minimizes secondary intention healing.

The proposed classification of sulcus deformities might facilitate treatment decisions and should allow investigators to compare results. The prevalence and pathogenesis of sulcus problems was not specifically addressed in this study. Although physiological sulcus configurations are more obvious and offer a high index of suspicion, the condition is not as common as these reported data might

suggest. It is apparent that 116 volunteers were recruited as control subjects and nine of these had evidence of physiological sulci. This does not indicate an 8% prevalence, for many of these volunteers were concerned enough about their voices to seek a free voice evaluation, and several had minor laryngitis findings.

It remains uncertain as to whether pathological sulci are congenital or acquired. The finding of associated epidermoid cysts in 9 of the 20 operative cases support the hypothesis of Bouchayer[9] that these lesions are the result of ruptured congenital cysts. In one sulcus patient, the medially-ruptured cysts showed extruding epithelial debris that was, on indirect laryngoscopy, suggestive of verrucous carcinoma. In this series, half of the patients reported onset of symptoms during childhood, which would support a possible congenital problem. In contrast, half of the patients reported adult-onset of symptoms and numerous potentially causative factors revealed in the histories of these patients included laryngitis, vocal abuse, allergy, and gastroesophageal reflux. The argument for acquired pathogenesis is also supported by our previous report of histological sulci in laryngeal cancer cases showing nearly one half of the laryngectomy specimens contained adjacent histological sulci[4]. The histology in Photo 14B (see center plate) reveals an intense inflammatory response which was also characteristic of the reported cancer cases. Photo 14A (see center plate) shows a more extracellular response with loss of cellular detail. This type of response is similar to the superficial lamina propria responses that Gray et al.[18] characterized by demonstrating increased fibronectin deposition and basement membrane zone injury. They postulated that such a pattern is a product of repetitive and chronic re-injury of the vocal folds leading to an aberrant healing response. Further study of the macromolecules that provide the filling around the delicate scaffolding of the superficial lamina propria (Reinke's space) might indicate a common pathway in the pathogenesis of sulcus vocalis that might explain how sulcus deformities are acquired from a variety of causes.

In summary, a series of patients with sulcus deformities of the vocal folds was presented. The clinical and pathological findings indicate that some sulci are physiological and can be observed in asymptomatic subjects. When there is damage to the superficial lamina propria, vocal fold vibration is impaired and patients become severely dysphonic. These cases respond favorably to microsurgery designed to remove destroyed tissue, release scar contracture, and promote mucosal re-draping by regional undermining. Classifying patients into Type I (physiological sulcus), Type II (sulcus vergeture), and Type III (sulcus vocalis) should prove useful in understanding the pathogenesis, making treatment decisions, and comparing outcomes.

Acknowledgments

The authors wold like to thank Joan Kozel for the skillful drawings and Celeste Kirk for editorial assistance. This work was supported by the National Center for Voice and Speech Grant No. P60DC00976 from the National Institutes on Deafness and Other Communication Disorders

References

1. Arnold GE. Dysplasia dysphonia. Minor anomalies of the vocal cords causing persistent hoarseness. Laryngoscope 1958;68:148-158.

2. Bouchayer M, Cornut G. Microsurgery for benign lesions of the vocal folds. ENT Journal 1988;67:446-466.

3. Bouchayer M, Cornut G. Instrumental microsurgery of benign lesions of the vocal folds. In: Ford CN, Bless DM, eds. Phonosurgery: Assessment and Surgical Management of Voice Disorders. New York: Raven Press, 1991:143-167.

4. Bouchayer M, Cornut G, Witzig E, Loire R, Roch JB, Bastian RW. Epidermoid cysts, sulci, and mucosal bridges of the true vocal cord: A report of 157 cases. Laryngoscope 1985;95:1087-1094.

5. Cornut G, Bouchayer M. Phonosurgery for singers. J Voice 1989;3:269-276.

6. Ford CN, Bless DM. Selected problems treated by vocal fold injection of collagen. Am J Otolaryngol 1993;14:257-261.

7. Ford CN, Bless DM, Loftus JM. Role of injectable collagen in the treatment of glottic insufficiency: A study of 119 patients. Ann Otol Rhinol Laryngol 1992;101:237-247.

8. Ford CN, Bless DM, Prehn RB. Thyroplasty as primary and adjunctive treatment of glottic insufficiency. J Voice 1992;6:277-285.

9. Giacomini C. Annotazioni sull"anatomia del negro. G Accad Med Torino 1892;40:17-61.

10. Gray SD, Hammond E, Hanson DF. Benign pathologic responses of the larynx. Ann Otol Rhinol Laryngol 1995;104:13-18.

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

12. Hirano M, Yoshida T, Tanaka S, Hibi S. Sulcus vocalis: functional aspects. Ann Otol Rhinol Laryngol 1990;99:679-683.

13. Lee ST, Niimi S. Vocal fold sulcus. J Laryngol Otol 1990;104:876-878.

14. Lindestad PA, Hertegard S. Spindle-shaped glottal insufficiency with and without sulcus vocalis: a retrospective study. Ann Otol Rhinol Laryngol 1994;103:547-553.

15. Lindestad PA, Hertegard S; Hammarberg B. An audioperceptual, videostroboscopic and electromyographic study of spindle-shaped glottal insufficiency with and without sulcus vocalis. Phoniatric Logopedic Progr Report, Karolinska Inst 1994;9:21-32.

16. Nakayama M, Ford CN, Brandenburg JH, Bless DM. Sulcus vocalis in laryngeal cancer: a histopathologic study. Laryngoscope 1994;104:16-24.

17. Pontes P, Behlau M. Treatment of sulcus vocalis: auditory perceptual and acoustic analysis of the slicing mucosa surgical technique. J Voice 1993;7:365-376.

18. Van Caneghem D. L'etiologie de la corde vocale a sillon. Ann des maladies de l'oreille,larynx, nez, et pharynx 1928;47:121-130.

NCVS Status and Progress Report - 8
July 1995, 75-79

Autologous Collagen Vocal Fold Injection: A Preliminary Clinical Study

Charles Ford, M.D.
Paul Staskowski, M.D.
Diane Bless, Ph.D.
Clinical Science Center, The University of Wisconsin

Abstract

This preliminary study reports the first use of injectable autologous collagen for vocal fold augmentation. In our previous studies, cross-linked bovine collagen was shown to be effective in over 150 cases of glottic incompetence, particularly in cases of focal defects or scarred or atrophic vocal folds. Concerns about possible adverse immunologic response to the bovine material has deterred FDA approval and has limited utilization.

We studied eight patients with difficult vocal fold pathology including sulcus vocalis, atrophy, and scarring secondary to trauma and cordectomy. Skin was harvested under local anesthesia, processed into a naturally cross-linked injectable form (Autologen^{R}), and then injected using indirect laryngoscopy in the outpatient clinic. Voice production was evaluated pre-injection and at intervals post-injection using subjective, perceptual, aerodynamic, acoustic, and videostroboscopic assessments.

Results indicate that autologous collagen is comparable to injectable bovine collagen in the management of several difficult glottic insufficiency problems. The major advantage is that, by using the patient's own tissues to constitute the implant, the likelihood of a hypersensitivity response is negligible. Unlike bovine collagen preparations, preparation of Autologen^{R} does not require extensive breakdown of the natural collagen molecule or cleavage of the antigenic telopeptides, so it is anticipated that this material will be better tolerated and more durable over time.

Introduction

Recent advances in phonosurgery have resulted in several new approaches to managing glottic insufficiency. Among the options, injection of collagen preparations is useful to augment vocal folds and to improve glottic function. In addition to meeting minimal requirements of a vocal fold filler, injected collagen closely resembles host tissues (collagen in vocal fold lamina propria) and is an active bioimplant. As with other bioimplants, it promotes ingrowth of host tissues and eventual replacement of injectate with host collagen[1-3]. One of the unique properties of injected collagen is the apparent tissue-softening that occurs when placed in scar tissue[4,5]. These positive biological properties prompted our early investigation of collagen injection in the management of glottic insufficiency.

Clinical studies have demonstrated the efficacy of collagen injection to correct a variety of glottic insufficiency problems[6-8]. It has been found to be uniquely helpful in treating specific vocal fold problems associated with scarring, atrophy, and focal defects[9,10]. The major concern about the use of injectable collagen is the potential immunologic consequences[11]. Although no substantial reactions have been documented, the use of a foreign protein raises the possibility of allergic reactions, airway compromise, anaphylaxis, immunologic destruction of the implant, and autoimmune responses. The materials used for vocal fold injection have included soluble bovine collagen products Zyderm Collagen Implants I and II^{R} (Collagen Corporation, Palo Alto, California), and their more recent cross-linked bovine preparations Zyplast^{R} and Phonagel^{R}.

The emergence of technology to prepare injectable autologous collagen from skin excised during cosmetic surgery (Autogenesis Technologies, Acton, Massachusetts) has allowed surgeons to re-inject collagen to improve facial contouring[2]. The use of the patient's own collagen virtually eliminates the risk of immunologic response to foreign protein. The process also preserves much of the natural cross-linking inherently formed during collagen fibril for-

mation in the body, so there is a greater likelihood of injectate persistence. These features of autologous collagen make it a potentially ideal substance for vocal fold augmentation.

The aim of this study was to demonstrate that autologous collagen could be used in a fashion similar to bovine collagen injections for vocal fold augmentation and that it could be used to treat difficult problems, not ideally suitable for other types of intervention. Preliminary results in eight patients are presented.

Materials and Methods

Patients with glottic insufficiency who were not ideal candidates for other types of treatment were invited to participate in this study. Of the eight patients, ages 29 to 84 years (mean - 47), six had scarred larynges, three associated with sulcus vocalis defects, two had had their vocal folds resected for cancer, and the other had loss of mucosa with secondary intention healing following laryngeal fracture. The remaining two patients had severe atrophy and bowing associated with unilateral paralysis; one of these was an elderly female with presbylaryngis. They all exhibited symptomatic glottic insufficiency with dysphonia as the primary symptom. Each patient was screened for general medical problems that might put him/her at increased risk from a procedure affecting the airway.

Pre-treatment voice assessment was done during initial encounters when the protocol was presented, and repeated prior to the injection; the immediate pre-treatment values were used for efficacy comparison. A comprehensive voice analysis was performed including acoustic analysis, aerodynamic measures, and videostroboscopy as described in our bovine collagen injection studies[6,7]. During each follow-up evaluation, each patient was asked to rate his/her voice on a 7-point scale comparing voice quality and the effort required for phonation to pre-treatment status; maximum decrement was -3, maximum improvement was +3, and 0 represented no appreciable change. In addition to the patient-rated subjective voice assessments, voice recordings were used for perceptual judgments by a panel of three speech pathologists. A master tape was created consisting of scrambled pre-, and post-treatment voice samples presented in random order to judges who were asked to decide which of two voices was better; +1 denoted improvement, -1 deterioration, and 0 no change. Repeats and foils were introduced to assure reliability.

Treatment consisted of two phases. The initial phase consisted of skin excision from the left lower quadrant of the abdomen. An ellipse of skin, approximately 1 x 2 inches, was excised, yielding approximately 4 g of dermis and 2 cm^3 injectable collagen; this was done in the outpatient clinic and incision was closed primarily. Morbidity was minimal. The skin was stored temporarily at -160° C and subsequently shipped with coolant in a styrofoam container (Photo 15; see center plate) to Autogenesis Technologies. The low temperature is necessary to inactivate proteolytic enzymes that would destroy the collagen structure. Once processed, the material (Autologen^R) was returned in a tuberculin syringe (Photo 16; see center plate). The injection procedure was the same as described for bovine collagen[6,13]. The injections were performed under topical anesthesia using 4% lidocaine dripped onto laryngeal mucosal surfaces and the Oro-tracheal InjectorR (Xomed-Treace, Jacksonville, Florida) delivery device used for injection into the vocal fold[14]. As with bovine collagen, Autologen^R is injected superficially into the lamina propria or scar tissue when present. Patients were observed for several hours after injection and monitored carefully for signs of dyspnea, dysphagia or discomfort. Follow-up voice evaluations were arranged at 1 month and 3 months following injection in all eight subjects.

Autologen^R is a sterile suspension of the patient's collagen fibers prepared from his/her own skin. The autologous collagen fibers are suspended in a sterile phosphate buffer at neutral pH. Electron microscopy has shown the actual collagen fibers to be intact with normal banding and periodicity. Unlike the bovine preparations previously used, AutologenR collagen fibers have not been solubilized by enzyme or chemical treatment, and they appear identical to collagen fibers in intact dermal bundles. Autologen^R can be stored at refrigerator temperatures (2-8° C) for up to 6 months. In this study, all patients were injected within 4 weeks of our receipt of the material.

Results

Autologous collagen was successfully injected into the vocal folds of eight patients, who were considered suitable candidates because of difficult glottic insufficiency problems (Table 1). In six of the eight patients, the injection procedure was uneventful. There were no focal or systemic signs of infection, hemorrhage, rejection, or hypersensitivity. The only complications occurred during the injection of two patients (Taiwanese graduate students, close friends with similar sulcus vocalis deformities) who experienced transient anxiety reactions. The reactions were unusual in that they exhibited hypertension for several minutes immediately following injection; they did not have bronchospasm, hives, rash, redness, or any laryngeal changes suggestive of an allergic reaction. One of those who experienced this reaction thought his voice might be slightly worse after injection but this was not corroborated by the perceptual judgment panel. All other patients felt they were improved by three to six months after injection (Table 1).

Subjective patient self-assessment of voice was based on questionnaires at 1 and 3 months post-injection. Each patient was asked to rate his/her voice on a 7-point scale (-3 = worse and +3 best). In response to specific

Table 1.
Descriptive Information on Collagen-Injected Patients

Patient	Age (Yr)	Pathology	Last Follow Up (Mos.)	Subjective[1]	Complications
#1	33	Sulcus vocalis s/p Teflon injection	6	+1	none
#2	29	Sulcus vocalis	3	-1	anxiety react[2]
#3	32	Sulcus vocalis	3	+1	anxiety react.
#4	74	Scarred, s/p cordectomy	6	+1	none
#5	57	Scarred, s/p cordectomy	3	+2	none
#6	31	Scarred, s/p fracture with mucosal loss	3	+1	none
#7	84	Paralysis and advanced presbylaryngis	3	+1	none
#8	38	Paralysis with atrophy and bowing	6	+2	none

1. SUBJECTIVE = Patient overall self-assessment compared to pre-treatment on scale of -3 (worst) to +3 (best), where 0 = no change in voice from pre-injection status.

2. ANXIETY REACTIONS = Transient hyperventilation, tachycardia, headache, and hypertension at the time of injection.

questions regarding voice quality and effort to produce voice, the mean change was favorable for the entire group but more pronounced at 1 month post-injection (Table 2).

Perceptual ratings by independent judges of tape samples revealed a net improvement in five subjects and no appreciable net change in the remaining three patients (+1 = improvement, -1 = deterioration) (Table 3). Assessment of glottic closure during phonatory tasks performed under videostroboscopic examination indicated improved closure in the same five of eight patients (Table 3). Table 3 also summarizes the clinically significant changes in transglottic airflow, maximum phonation time, intensity range, frequency range, and signal-to-noise ratio. The greatest changes were noted in signal-to-noise ratio, where five patients experienced substantial improvements.

Table 2.
Patient Self-Assessment[1] of Voice Based on Questionnaire

Change, Post-Injection:

Voice Quality	N	Mean	STD Dev	STD Error	P(Signed Rank)
1 MO	8	0.8750	0.9910	0.3503	0.0937
3 MO	8	0.9375	0.9425	0.3332	0.0625
Phonatory Ease					
1 MO	8	0.8750	1.1260	0.3980	0.1250
3 MO	8	0.5000	0.0258	0.3273	0.3125

1. Assessment of voice compared to pre-treatment, on scale of -3 to +3, where 0 = no change in voice from pre-injection status. Patients were asked to distinguish the quality of the voice and the ease (relative effort) of phonation. Favorable results were all positive numbers.

Table 3.
Assessment Measures[1] at 3 Months Post-Injection

	Aerodynamic		Acoustic					
PT	AirFlow	MPT	SPL range	Fo range	SNR	LVS closure	Perc	Net
#1	0	0	1	0	1	0	0	+2
#2	1	0	0	1	1	0	0	+3
#3	1	0	1	-1	1	1	1	+4
#4	-1	-1	0	0	1	1	0	0
#5	1	1	0	1	-1	1	1	+4
#6	0	0	0	0	0	0	1	+1
#7	-1	1	0	0	-1	1	1	+1
#8	-1	0	1	-1	1	1	1	+2
All	0	+1	+3	0	+3	+5	+5	--

1. In this sign chart, +1 = clinically significant positive change, 0 = no change, and -1 = negative change. Clinically significant refers to changes greater than 1 standard deviation based on observed variability of normal speakers. MPT = maximum phonation time, SPL = intensity, Fo = frequency, SNR = signal-to-noise ratio, LVS-closure = apparent glottic closure seen on laryngeal videostroboscopy, Perc = perceptual judgments of expert panelists listening to tape recordings, NET = the net positive + negative changes across parameters.

Discussion

Injectable collagen has been demonstrated to be effective in the management of a variety of glottic insufficiency problems. It is particularly useful in cases involving scarring, atrophy, and focal defects where other medialization techniques are not very effective[9,10]. The major limiting factor in employing this treatment in the United States has been the reluctance of the FDA to approve the use of bovine collagen preparations in the larynx where an airway crisis could occur as a result of a hypersensitivity reaction. It should be noted that we have not observed such a response in over 150 injections, and the material is safely used in other countries[15]. Because of the hypersensitivity concerns associated with bovine collagen, autologous collagen derived from the patients' own dermal tissues has significant potential as an alternative. This study is an initial attempt to see if autologous collagen appears useful in the spectrum of patients who might have been helped by bovine collagen injections.

It is important to stress the obvious limitations of the study. This was a preliminary investigation with few patients, all of whom had slightly different pathological conditions and none of whom was ideally suited for other types of treatment. Six of the eight treated patients had scarred larynges, which put them in the sub-group of patients with potentially the most variable and least dramatic responses to bovine collagen[9]. The short length of follow-up did not permit sufficient time for repeat injections or for physiologic softening of scar tissue, which often required months to years after bovine collagen injections. It was anticipated that a favorable response in this study would be

for the patients to demonstrate early voice results comparable to the early results following bovine collagen treatment as previously reported.

Results are reported out to 3 months in this preliminary report. This interval has some validity based on initial clinical studies with bovine collagen[7], where it was shown that voice results tend to be unstable during the first month and then stable after 3 months. Although it is too early to assess the effect on softening scar tissue, it is possible to achieve mechanical correction of focal scarred defects with precision and to improve voices. At 3 months post-injection, all but one patient (#2) in this study was subjectively improved (Table 1). Patients reported greater subjective improvement at 1 month than at 3 months, but two patients confessed that they had difficulty with the question and seemed to compare their voices from one session to another and not to their baseline pre-injection voices. This points out how such questionnaire-generated data might be useful but should not be relied on for valid assessment. Perceptual judgments, by speech pathologists and graduate students listening to random pairs of voice samples, revealed that five of eight patients had improved voices at 3 months post-injection (Table 3). Laryngeal videostroboscopy showed improved closure in the same five ofeight patients (Table 3), although one of these (#4) continued to exhibit extreme stiffness at three months. Objective aerodynamic and acoustic measures were studied but the small number and heterogeneous population (differences in age, sex, pathology, compensatory behaviors, etc.) made statistical analysis irrelevant. As in prior studies, airflow measures are highly variable in this population. In some instances, increased post-treatment airflow measures actually reflected a value closer to the normal range; this could be due to increased effort, loss of compensatory behaviors, or other factors. The results summarized in Table 3 indicate that, as in our previous bovine collagen injection studies, signal-to-noise ratio is probably the best acoustic measure indicator of a positive treatment response. The net change across all parameters was positive in all but one patient (#4) where there was no net change objectively; this patient had been injected for a cordectomy defect and felt subjectively improved.

With the proliferation of treatment options available for managing glottic insufficiency, it is important to recognize the potential of bioimplants as an alternative and adjunct in the management of some difficult problems. Fat has been demonstrated to be effective for management of vocal fold paralysis problems but it undergoes extensive resorption 16, cannot be precisely placed through fine needles, and has not been demonstrated effective in the scarred larynx. Collagen is ideally suited for precise placement through fine 27- to 30-gauge needles and it appears to be effective in augmenting focal defects, correcting tissue atrophy, and softening scar tissue[4,5]. One of the problems with earlier bovine collagen preparations (Zyderm I and II^{R}) was rapidresorption and fear of hypersensitivity reactions. GAX cross-linked collagen (Zyplast^{R}, Phonagel^{R}) was a great improvement; it was more stable, less antigenic, and seemed to last longer.

It is likely that autologous collagen will have even greater persistence because the collagen molecules are preserved in the preparation process. Telopeptides are the non-helical portions of the collagen molecule responsible for antigenicity; in bovine collagen preparations these units must be cleaved to decrease the risk of allergic reaction in human hosts. Intact telopeptides are essential to the quarter-stagger formation of collagen fibers and to formation of intermolecular crosslinks between adjacent collagen molecules. In Autologen^{R}, the *in vivo* architecture of collagen fibers is preserved including inter-molecular enzymatically catalyzed crosslinks. It is likely that these natural collagen fibers will exhibit greater resistance to protease degradation than reconstituted collagen in which telopeptides have been cleaved. The degree of retained cross-linkage can vary as desired and it is not clear at this stage what constitutes the optimal preparation. It does seem likely that such autogenous preparations will ultimately simulate host tissues to a greater degree, exhibit greater host tolerance, and last longer. We are currently investigating the initial and long-term fate of autologous collagen injected into the canine vocal fold.

Conclusion

This preliminary study supports the use of autologous collagen injections for the same spectrum of vocal fold pathology that bovine collagen has been successfully used to manage. The short-term results appear comparable to those with bovine collagen but the variable pathology and small number of patients fail to lend statistically significant efficacy measures. The risk of hypersensitivity reaction is negligible with autologous implants and it is likely that this material will exhibit greater host tolerance and persistence than bovine collagen. As a consequence, autologous collagen deserves further study as an injectable bioimplant for glottic insufficiency.

Acknowledgments

This research is supported by NIH Research Training Grant: NIDCD Grant No. 60-00976.

Bibliography

1. Ford CN: Histologic Studies on the Fate of Soluble Collagen Injected Into Canine Larynges. Laryngoscope, 96:1248-1257, 1986.

2. Ford CN, Martin DW and Warner TF: Injectable Collagen in Laryngeal Rehabilitation. Laryngoscope, 94:513-518, 1984.

3. Okamoto K: A Fundamental Study of Injectable Collagen in Vocal Rehabilitation. J of Otol Tokyo, 90:394-403, 1987.

4. Ford CN and Bless DM: Collagen Injection in the Scarred Vocal Fold. J Voice, 116-118, 1987.

5. Knapp TR, Kaplan EN and Daniels JR: Injectable Collagen for Soft Tissue Augmentation. Plast Reconstr Surg, 60:398-405, 1977.

6. Ford CN and Bless DM: A Preliminary Study of Injectable Collagen in Human Vocal Fold Augmentation. Otolaryngol Head Neck Surg 94:104-112, 1986.

7. Ford CN and Bless DM: Clinical Experience with Injectable Collagen for Vocal Fold Augmentation. Laryngoscope, 96:863-869, 1986.

8. Remacle M, Marbaix E, Hamoir M, Bertrand B and Van den Eeckhaut J: Correction of Glottic Insufficiency by Collagen Injection. Ann Otol Rhinol Laryngol, 99:438-444, 1990.

9. Ford CN, Bless DM and Loftus JM: Role of Injectable Collagen in the Treatment of Glottic Insufficiency: A Study of 119 Patients. Ann Otol Rhinol Laryngol, 101:237-247, 1992.

10. Ford CN and Bless DM: Selected Problems Treated by Vocal Fold Injection of Collagen. Am J Otolaryngol, 14:257-261, 1993.

11. Spiegel JR, Sataloff RT and Gould WJ: The Treatment of Vocal Fold Paralysis With Injectable Collagen: Clinical Concerns. J Voice, 1:119-121, 1987.

12. Swinehart JM: Dermal Pocket Grafting: Implants of Dermis, Fat, and "Autologous Collagen" for Permanent Correction of Cutaneous Depressions. Int J Aesth Restorative Surg, 2:43-52, 1994.

13. Ford CN.: Laryngeal Injection Techniques. In: Phonosurgery: Assessment and Surgical Management of Voice Disorders. Ford CN and Bless DM (Eds.). Raven Press, New York, pp. 123-141, 1991.

14. Ford CN: A Multipurpose Laryngeal Injector Device. Otolarygol Head Neck Surg, 103:135-137, 1990.

15. Remacle M, Marbaix E, Hamoir M, Declaye X and Van den Eeckhaut J: Initial Long-Term Results of Collagen Injection for Vocal and Laryngeal Rehabilitation. Arch of Oto-Rhino-Laryngol, 246:403-406, 1989.

16. Milkus JL, Koufman JA and Kilpatrick SE: Fate of Liposuctioned and Purified Autologous Fat Injections in the Canine Vocal Fold. Laryngoscope, 105:17-22, 1995.

NCVS Status and Progress Report - 8
July 1995, 81-96

Longitudinal Phonatory Characteristics After Botulinum Toxin Type A Injection

Kimberly Fisher, Ph.D.
John W. Keys Speech and Hearing Center, University of Oklahoma Health Sciences Center
Ronald Scherer, Ph.D.
The Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts
Chwen Guo, M.S.
The Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts
Ann Owen, Ph.D.
John W. Keys Speech and Hearing Center, University of Oklahoma Health Sciences Center

Abstract

Following Botulinum Toxin Type A injection, glottal competency of an adductor spasmodic dysphonia patient is thought to vary over a wide range. This study quantifies variability in laryngeal adduction for one such patient over a ten week period. Analyses of kinematic and aerodynamic measures were used to track the voice weekly. The measures included the electroglottographic waveform width (EGGW50), non-dimensional electroglottographic slope quotient (SLQ), glottal flow open quotient (FOQ), dc glottal flow, and non-dimensional glottal flow peak quotient (FPQ). The results suggested that change in degree of glottal adduction over time can be observed even when vocal instability is present within each recording session. Perceptual ratings of vocal quality (breathy to pressed) were related to the laryngeal measures. The coefficient of variation for EGGW50 and the percentage of dichrotic phonations reached minima during sessions with predominantly breathy and hypoadducted phonation. The methods used in this study show potential to aid decisions about dose level and sources of perceptual adductor spasmodic dysphonia symptoms for a given patient.

Adductor spasmodic dysphonia (ADSD) patients who are treated with Botox® can have voices that are highly variable over time. This variability is presumably due to the covarying degree of both spasmodic voice symptoms and Botox®-related vocal fold weakness. For example, ADSD symptoms include hyperadductive vocal fold spasm, intermittently pressed phonation, voice cessations, and strained, effortful phonation (Aronson, 1985; Aronson, Brown, Litin, & Pearson, 1968; Freeman, Cannito, & Finitzo-Hieber, 1985). Injection of Botox® into the thyroarytenoid muscle has been shown to reduce significantly these voice symptoms by chemically inducing a temporary deenervation and vocal fold paresis (Blitzer, Brin, Fahn, & Lovelace, 1988; Brin, Blitzer, Fahn, Gould, & Lovelace, 1989; Brin, Blitzer, Stewart, & Fahn, 1992; Ludlow, 1990; Ludlow, Naunton, Sedory, Schulz, & Hallett, 1988; Miller, Woodson, & Jankovic, 1987; Zwirner, Murry, Swenson, & Woodson, 1991). Potential side effects of Botox® can include temporary breathiness, hoarseness, and sub-clinical aspiration of fluids. Reports also have included the rarer symptoms of breathy aphonia, sore throat, and diplophonia (Brin et al., 1992). Thus, Botox® injection may not result in normal laryngeal function, but rather a reduction or elimination of voice spasms, with the potential temporary replacement of other voice symptoms. Botox® treatment of ADSD also allows a unique opportunity to explore variability and extremes of laryngeal adduction within a single patient.

Gross perceptual or acoustic voice measures are sometimes used to document the voice improvement that occurs following Botox® injection (Ford, Bless, & Patel, 1992; Ludlow et al., 1988, Zwirner, Murry, & Woodson, 1993). Gross acoustic measures such as speech rate, phonatory offset times, intrusive loudness changes, voice cessations,

and pitch breaks have been helpful to characterize patient symptoms (Ludlow, 1990). Analysis of acoustic signals, however, can be difficult for weak voices and may not relate directly to the pathophysiology of voice disruptions (Watson et al., 1991). In some cases, clinician assessment and patient self-assessment (through rating scales and diary entries) are used to compare treatment effects at different dose levels for a given patient (Brin et al., 1992). Fiberoptic laryngoscopy has been useful to describe the laryngeal behaviors of ADSD or to view the folds during toxin injection (Ford, Bless, & Lowrey, 1990). Unfortunately, viewing the vocal folds can be obscured or precluded by laryngeal/pharyngeal hyperconstriction.

Longitudinal physiologic measures of laryngeal function would help to quantify the within-subject phonatory characteristics and variability following Botox® injection. Also, longitudinal measures could ultimately provide the comparison of voice change across repeated injections, and potentially reveal maximized treatment effectiveness for a given patient. Obtaining measures that quantify voice characteristics during the post-injection period is preliminary to identifying, for example, dose levels that result in voice improvement with minimal adverse effects. Ideally, voice measurement should be non-invasive and highly sensitive to subtle variations in vocal fold function and glottal competency.

Investigators studying normal, hypofunctional, and hyperfunctional voices (Childers, 1991; Gauffin & Sundberg, 1980; Hillman, Holmberg, Perkell, Walsh & Vaughan, 1989; Lofqvist, 1982; Rothenberg & Mahshie, 1988; Scherer, 1991, 1991b; Scherer, Sundberg, & Titze, 1989; Sundberg, Titze, & Scherer, 1993; Titze, 1992a; Titze & Sundberg, 1992) suggest that indirect, within-cycle measures of tissue contact area and glottal volume velocity are sensitive to a wide range of adduction and acoustic voicing characteristics. There are no published reports that demonstrate the use of these measures to quantify phonatory characteristics in a highly unstable voice like that of a post-injection ADSD voice.

Temporary changes in adduction associated with the well-documented paretic effect of Botox® injection were expected to alter tissue contact (EGG; Scherer, Druker & Titze, 1988) and glottal volume velocity waveforms. Potentially, the measures could reflect a continum that might extend from pre-injection hyperadduction to post-injection hypoadduction associated with temporary glottal incompetency. Other investigators report that temporary breathiness is common in many patients and is followed by voice improvement, and finally a return of spasmodic, effortful phonation (Brin et al., 1992; Truong, Rontal, Rolnick, Aronson, & Mistura, 1991). Given these previously reported voice changes, we questioned whether the phonatory measures can track adductory variability attributable to change in time, even when phonation may be spasmodic or unstable within a single recording session.

This study quantifies the glottal airflow and kinematic characteristics for one adductor spasmodic dysphonia patient who underwent Botox® injection treatment. Recordings were made at approximately one-week intervals for 10 weeks. The glottal volume velocity signal was obtained by inverse filtering the oral airflow. The electroglottograph (EGG) was used to obtain a measure of glottal adduction. One purpose was to quantify the variability in adduction of a post-injection ADSD voice. A second purpose was to determine whether within-cycle kinematic and aerodynamic measures would correspond to ratings of voice quality. Third, the study sought to describe and quantify any temporal pattern of phonatory change in this patient's post Botox® injection voice.

Method

Subject

A 67 year-old male ADSD patient was selected for study. His voice diagnosis adhered to clinical criteria reported elsewhere (Aronson, 1985; Freeman, Cannito, & Finitzo-Hieber, 1985). The patient, considered to have severe ADSD symptoms, reportedly had suffered from a "strained and jerky" voice for 15 years. His pre-injection speech was characterized by pressed phonatory quality and intermittent phonatory cessation. Transnasal videolaryngoscopic examinations revealed the presence of hyperadductive laryngeal spasms. There were no known neurologic or psychiatric disorders prior to symptom onset, no signs of essential tremor or Meige's syndrome, and no history of phenothiazine use. Familial history for speech or language disorders, tremor, or symptomatic neurologic disease was negative. The patient had received no treatment for ADSD other than speech therapy. The patient reported that improvement with therapy was minimal and deteriorated over time.

Procedure

Two and one half units of Botox® solution were injected into each vocal fold via an EMG needle using a percutaneous approach. Monitoring of muscle activity allowed rough confirmation of the intended injection site. Ten data recording sessions were scheduled over a three-month period. The first recording session was scheduled immediately before injection. The second recording session occurred 1 hour after injection. These sessions hereafter are denoted week 0. The next five sessions occurred weekly (denoted weeks 1 through 5). Measurements resumed at weekly intervals from week 8 through 10. The three week lapse between weeks 5 and 8 was due to the patient's unavailability. Each recording session was scheduled for approximately the same time each day. All recordings were obtained within an acoustically isolated sound suite.

The speech material and recording method were similar to those described in a normative study (Holmberg, Hillman & Perkell, 1988). Briefly, during each session the subject produced a five syllable repetition of /pæ/ under three conditions: normal loudness, soft but not whispered, and loud but not shouting. The vowel portion of the fifth syllable was prolonged for not less than 2 seconds. Before each recording, the subject performed several practice trials. A speech rate of 1.5 syllables per second was modelled for the subject. The syllable string is a mildly loaded linguistic utterance allowing potential sensitivity to phonatory instabilities without eliciting severe symptoms. The subject was free to phonate at a "comfortable" pitch level for each loudness condition. At least three task repetitions were recorded for each loudness condition during each session. Only the normal and loud conditions will be reported here. The soft syllables were recorded to be consistent with the protocol of Holmberg et al. (1988) and in case the other analyses indicated that soft phonation contained valuable information. It was not anticipated, however, that the soft phonation condition would be analyzed due to the lower likelihood of eliciting ADSD symptoms (Izdebski, 1992).

The oral volume velocity (Glottal Enterprises Inverse Filter, Model MSIF-2S; Glottal Enterprises Wide Bandwidth Transducer, PTWL-2), sound pressure level (Bruel and KjAr Sound Level Meter, Model 2209), and electroglottographic (Laryngograph, Laryngograph Ltd. & Kay Elemetrics, Model K6094) waveforms were recorded using an FM converter, 16 bit pulse code modulator and super-VHS recorder combination (Titze, Scherer & Winholtz, 1985). The microphone signal (Bruel and Kjær, 1/2 inch condenser, Type 4155) was recorded directly onto the audio channel of the super-VHS recorder (Mitsubishi, Model HS-U62). Mouth-to-microphone distance was 1 ft. A head support was used to prevent head movement which might result in an air leak around the face mask (Glottal Enterprises Mask, Model MA-1N), or a change of mouth-to-microphone distance and EGG electrode placement.

The airflow and sound level channels were calibrated before each session. Amplifier gain levels were adjusted to allow maximum recording resolution of the subject's phonation output from each transducer. A calibration signal of known amplitude was then recorded through each channel and later digitized to determine the voltage to computer unit ratio. Calibration regression equations also were generated for each transducer. The regression equations and computer unit conversion ratios were used to determine the equivalence between the input units, output voltages, and computer units.

Digitized signals included the oral volume velocity, post-recording inverse filtered glottal volume velocity (using the analog method of the Glottal Enterprises system), the microphone signal, sound level, and the electroglottographic signal. Each signal was sampled at 20.0 kHz by means of a 16 bit A/D converter (Digital Sound Corporation, Model DSC200) for computer analysis (VAX, Model 4200). EGG and inverse filtered flow signals were low pass filtered at 3 kHz (Tektronix Amplifier/Filter, Model AM502). The acoustic signal was band pass filtered with high and low cut-off frequencies of 10 kHz and 60 Hz, respectively (Wavetek Amplifier/Filter Model 432).

Analyses

Signals were analyzed using MULTISIG (Scherer & Guo, 1992), a display and analysis program written especially for multi-channel, glottal measurements. All quantitative single cycle measures were obtained from 10 adjacent cycles that occurred relatively early in the last prolonged vowel within the /pæ/ sequence. Potential signal transitions were avoided by beginning analysis at least 200 msec after the onset of voicing. The mean of each measure over the 10 adjacent cycles was taken as the representative value for that segment. Segments containing dichrotic vibratory patterns were excluded from all measures except that of percent dichrotic phonation. Dichrotic phonation was excluded because it was thought to represent a different mode of phonation, and because it would increase the observed variability in the glottal cycle measures.

The peak-to-peak (AC) flow (FPK) value and flow open quotient (FOQ) (as shown in Figure 1; see following page) were obtained using the inverse filtered glottal volume velocity. FPK was defined for each phonatory cycle as the difference between the maximum of the volume velocity waveform and the left-side waveform baseline, measured in cubic centimeters per second. The flow open quotient (FOQ) was obtained for the same segment by dividing the "open" time (a) by the period (T). The vertical marks on Figure 1 correspond to the selected flow "shut-off" and flow "beginning" times for the purposes of determining the open quotient. Flow "shut-off" was determined by a zero value for the derivative of the glottal volume velocity near the baseline. If this location was in question, the corner formed by the intersection of the decreasing portion of the flow waveform and the relatively flat baseline was chosen. The flow opening point was selected to correspond with the maximum value in the second derivative of the glottal volume velocity waveform. This may be estimated visually as the point of steepest increase of the glottal flow derivative waveform near the location corresponding to glottal opening.[2]

The maximum negative peak value (FDPK) from the first derivative of the glottal volume velocity waveform also was obtained (Figure 1). The ratio of the FPK and FDPK measures (shown at the bottom of Figure 1), divided by $\omega = 2\pi f_0$ is a non-dimensional glottal volume velocity quotient, hereafter referred to as the flow peak quotient (FPQ). The FPQ magnitude increases negatively for steeper spectral tilt as in breathy voice (Gauffin & Sundberg, 1989).

The EGG pulse width at the 50% level of the total wave amplitude (EGGW50, Figure 2) is related to laryngeal adduction (Arends, Pavel, Van Os, & Speth, 1990; cf. Dejonckere & Labacq, 1985; Orlikoff, 1991; Rothenberg & Mahshie, 1988; Scherer, Vail, & Rockwell, in press). The EGGW50 is sensitive to the extent of vocal fold medial compression. The EGGW50 measure was obtained from the EGG segment that corresponded in time to that used for the aerodynamic measures.

Non-dimensional EGG slope quotients, indications of kinematic vocal fold behavior during vocal fold contact, also were obtained (Childers, 1991; cf. Houben, Buekers, & Kingman, 1992; Scherer, Cooper, Alipour-Haghighi, & Titze, 1984). The contact closing slope measure SC1090 (Figure 3) is defined by the normalized rise (b_1/H) divided by the normalized run a_1/T, where b_1 is the EGG signal height corresponding to the segment between 10% and 90% of the peak-to-peak amplitude H, a_1 is the time segment corresponding to b_1, and T is the cycle period (cf. Orlikoff, 1991).[3] The closing slope quotient may vary with the degree of energy in the acoustic signal (Orlikoff, 1991; Titze, 1990). The opening contact slope SO9050 (Figure 3) is defined by the normalized rise b_2/H divided by the normalized run a_2/T; where b_2 is the EGG signal height corresponding to the segment between 90% and 50% of H, and a_2 represents the time for the signal to drop from the 90% to the 50% level.[4] This region of the EGG waveform corresponds to the oscillation of vocal folds during the closed glottis just prior to glottal opening.. The closing and opening contact measures depend upon the speed and manner of vocal fold movement during contact.[5] The ratio of the SD1090 measure to the SO9050 measure is called the slope quotient (SLQ). Interpretation of the slope measures, and the EGG waveshape in general, may be aided by reference to computer modelling (Childers & Lee, 1991; Titze, 1989, 1990).

The percentage of dichrotic cycles, a measure of gross instability within the vowel segment, was determined by visual inspection of the EGG waveform. Any group of cycles with alternately increased and decreased amplitude of greater than an approximated 20% change were considered dichrotic.

Kinematic instability was estimated from the sustained vowel /æ/ beginning 200 msec after the onset of voicing and continuing to the end of the utterance. The selected segment included (but was not limited to) that used for other analyses described previously. To estimate the cycle-to-cycle kinematic variability during each session, the EGGW50 jitter of the longest non-dichrotic segment was determined for each token. The coefficient of variation of EGGW50 (a longer-term variability measure, EGGWCVA) was obtained by the ratio of mean EGGW50 to the standard deviation (Scherer, Gould, Titze, Meyers, & Sataloff, 1988). For these two perturbation measures, the number of consecutive cycles used over the entire set of 58 tokens ranged from 27 to 192, with an average number of 106 cycles (sd=42).

The patient provided three ratings of his voice: 1) how his voice sounded, 2) how his voice felt, and 3) the degree of effort needed to produce voicing (Appendix A). Any report of difficulty in swallowing, which could suggest sub-clinical aspiration, was documented. Finally, the patient scaled how his voice sounded by placing a vertical mark through a continuous line that represented subjective voice ratings from "worst ever" to "best ever."

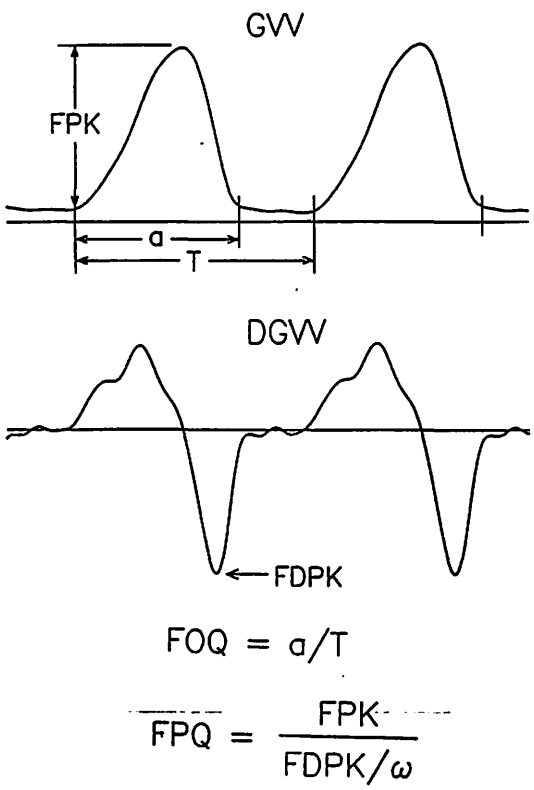


Figure 1. Measurements made from the glottal volume velocity waveform, GVV, and its first derivative, DGVV. Lines intersecting the glottal volume velocity waveform indicate selected glottal closing and opening points for the purpose of obtaining the flow open quotient (FOQ) and flow peak quotient (FPQ).

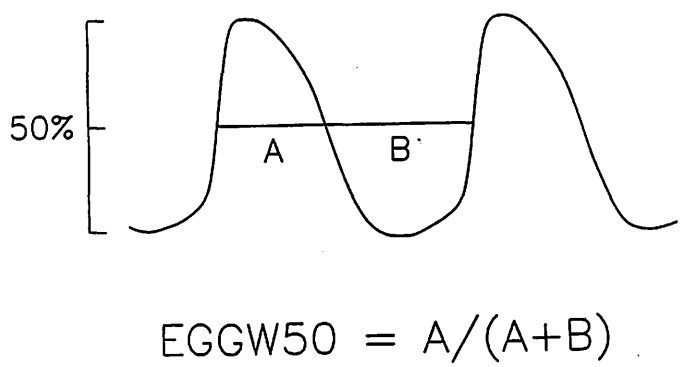
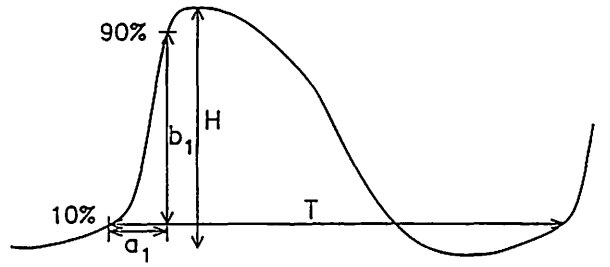
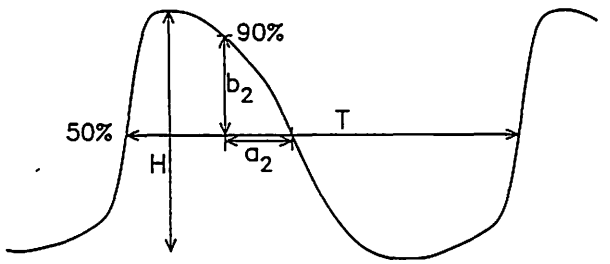


Figure 2. Measurement of the electroglottographic waveform pulse width at the 50% wave amplitude level (EGGW50).

Ten ASHA certified speech language pathologists rated taped samples of the patient's voice on an 11 point equal appearing interval scale that ranged from breathy (-5) to pressed (+5). The judges initially were unfamiliar with the patient's voice. Each of the listening samples was 2 seconds of the normal loudness vowel /æ/ productions from the final prolonged syllable /pæ/ from which aerodynamic and EGG measurements were made. Listening samples were recorded in random order on a reel-to-reel tape recorder (TEAC Tascam 22M) and presented through the headphones of a listening station within an acoustically isolated sound suite. The first eight taped samples were used as training stimuli and reflected the full range of voice quality. These samples were dropped from subsequent analysis. Twenty normal loudness test samples, two from each session, followed. A reliability set of the same 20 tokens also was randomized into the listening tape. This resulted in a total listening tape of 48 samples. Judges were allowed to repeat individual samples as often as necessary in order to choose their rating. The median voice rating was used as the index for an individual sample.



$$SC1090 = \frac{b_1/H}{a_1/T}, \quad T = \text{period}$$



$$SO9050 = \frac{b_2/H}{a_2/T}, \quad T = \text{period}$$

Figure 3. Measurement of the electroglottographic contact closing slope (SC1090) and contact opening slope (SO9050).

Results

Perceptual Ratings

The patient reported that his voice was "smoother" after injection, and that he had some transient problems swallowing soon after injection. This is consistent with reports of other investigators (Brin et al., 1992). By week 9 or 10, the patient reported elevated vocal effort and a return of "choppiness." Figure 4 shows the patient's self scaling of the way his voice sounded. A monotonic trend for self-perceived voice improvement after injection is evident, with the largest between-session change occurring at the first week post injection. It is unclear why the patient continued to indicate steady voice improvement on this scale in spite of his subjective impression that his voice was "worse" by week 10 and his request for a re-injection by week 11 (refer to Appendix A).

Table 1 (next page) shows the median listener ratings of voice quality ranging from "most breathy" to "most pressed." To evaluate intrajudge reliability, the 20 reliability vowel samples were presented twice to the judges. A judge was said to be in agreement with herself/himself if she/he gave the same or adjacent rating for the two occurrences of the 20 reliability tokens. Percentages of intrajudge voice quality rating agreement ranged from 40% to 85% (mean = 63%) and suggests that this task was difficult for some judges. An intraclass correlation coefficient of 0.83, however, suggested adequate interjudge agreement (Ebel, 1951). Note that recordings in week 0 (before and just after injection) as well as week 8 are characterized by median judg-

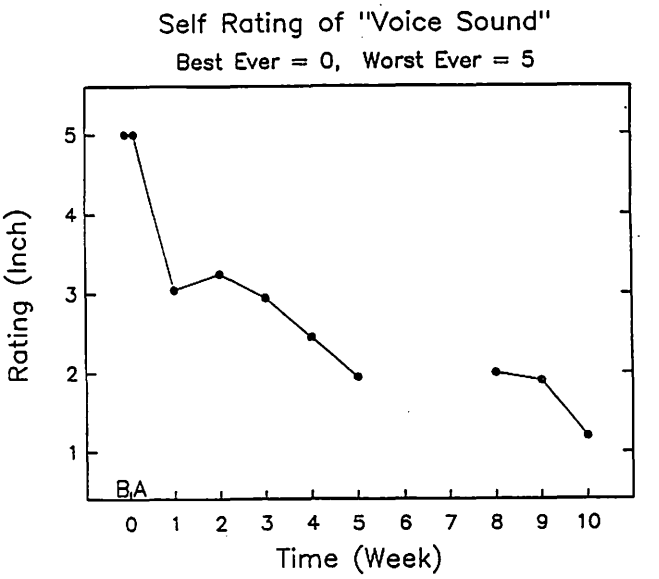


Figure 4. The patient's self scaling of voice for each recording session beginning before and just after injection (week 0), and continuing at weekly intervals until week 10 after injection. No measures were obtained during weeks 6 and 7. The rating scale ranged from "best ever" (0 inches) to "worst ever" (6 inches).

Table 1.
Median voice ratings across ten listeners for the sustained vowel from /pæ/ tokens on each of the 8 recording sessions.

Weeks Post Injection	Median Rating	Weeks Post Injection	Median Rating
0 (before)	2.5	3	-3.0
0 (just after)	2.5	4	1.0
1	-1.5	5	1.0
2	-2.0	8	2.0

Note. Ten speech-language pathologists rated vowel samples on an 11-point equal appearing interval scale that ranged from "most breathy" (-5) to "most pressed" (+5). A zero rating was neutral.

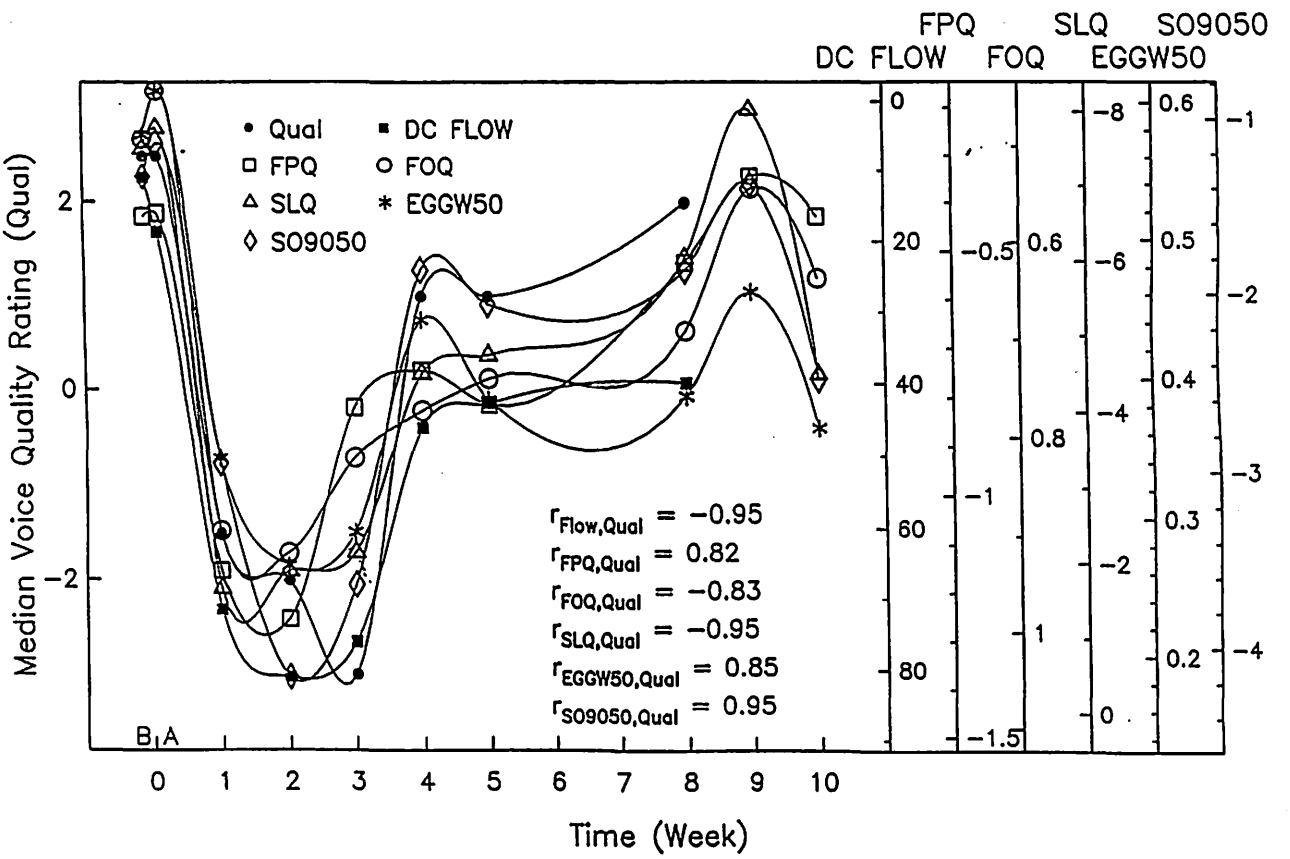


Figure 5. Median voice quality ratings, mean DC flow, flow peak quotient FPQ, glottal flow open quotient FOQ, electroglottographic waveform width EGGW50, opening slope SO9050, and slope quotient SLQ for midvowel tokens of the /pæ/ syllables produced at normal intensity by the ADSD subject during each recording session. The rating scale ranged from -5 (most breathy) to +5 (most pressed). The graph represents a least squares fit of each measure with voice quality.

Table 2.
Proportion of Variance Attributable to Change Over Time (Effect Size)

Electroglottographic Measures		Aerodynamic Measures	
EGGW50	$r^2 = .837$	FOQ	$r^2 = .844$
SC1090	$r^2 = .548$	FPK	$r^2 = .548$
SCSO9050	$r^2 = .824$	FDPK	$r^2 = .459$
SLQ	$r^2 = .748$	FPQ	$r^2 = .768$
EGGW50 Jitter	$r^2 = .501$	DC Flow	$r^2 = .645$
EGGW50 Coef. Var.	$r^2 = .714$		

ments of relatively more "pressed" voice. Judgments of relatively greater "breathiness," however, occurred during weeks 2 and 3 after injection. This is consistent with other reports that temporary breathiness may occur soon after injection, and that the effects of injection wear off over time (Brin et al, 1992, p. 222; Ludlow, Naunton, Sedory, Schulz & Hallett, 1988; Miller, Woodson & Jankovic, 1987).

Variability in Adduction Measures Over Time

Effect size was calculated for each aerodynamic and kinematic variable in order to assess that proportion of variance attributable to change over time. Effect size r^2 values greater than .7 indicate that within session variabilities are small relative to those between measurement sessions. For all but four variables (FPK, FDPK, SC1090, EGGW50 Jitter) effect size calculations exceeded r^2=.7 (Table 2). This does not suggest that phonation was stable during each measurment session. Rather, it suggests that adductory change can be quantified over time even given the within session phonatory instability associated with this patient's ADSD symptoms and/or treatment.

Figure 5 shows the variability over time for the laryngeal measures with moderate to large effect sizes. All kinematic and aerodynamic variables are graphed by least squares best fit to voice quality rating, the latter represented by small filled circles. Greater adduction and pressed voice quality are oriented to the top of the graph. Greater abduction and breathy voice quality are oriented to the bottom of the graph. Note the similar trends in FPQ, FOQ, DC Flow, EGGW50, SLQ, SO9050 and the accompanying perceptual ratings over time. In general, values suggesting minimum adduction occur at weeks 1 through 3, consistent with the judgements of breathy voice quality during this time. Linear regression analyses indicated high correlations (Figure 5) between the median voice quality ratings and the session means for each of the measures. Generally, the variables with the larger effect sizes also were the variables with the stronger correlations to median quality rating.

Variability in adduction measures may be associated with variability in vocal fundamental frequency (f_0) and/or intensity (Figure 6). Across all sessions and loudness levels, vocal f_0 ranged from 92.3 Hz to 132.9 Hz (X=113.3, SD=10.0). The largest between session mean difference in f_0 was 30 Hz (approximately 4.8 semitones between weeks 2 and 9). This difference may show some relation to the

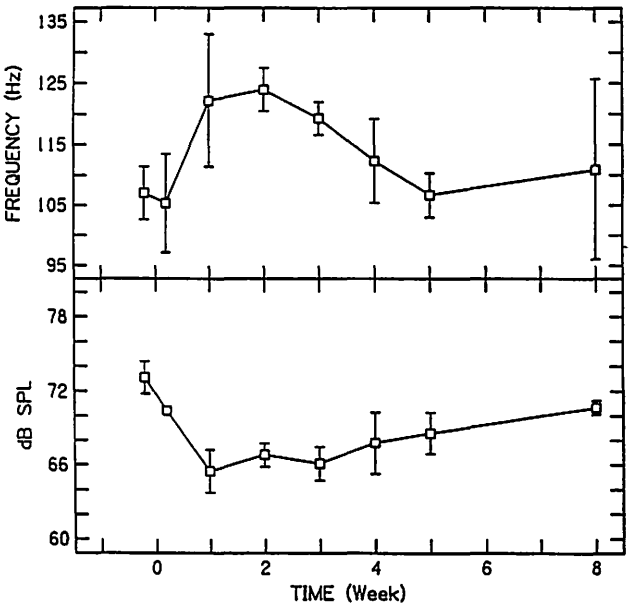


Figure 6. The mean ± 1 STD fundamental frequency and dB SPL for the midvowel portion of each analyzed /pæ/ token produced in normal and loud voice. Recording sessions begin at week 0, before (B) and 1 hour after (A) injection.

Table 3.
Correlation Matrix for Kinematic Variables

	EGGW50	SC1090	SO9050	SLQ	Fo	dB SPL
EGGW50	1.000	.191	* .874	* -.745	-.567	.464
SC1090	.191	1.000	.394	* -.760	-.556	* .608
SO9050	* .874	.394	1.000	* -.830	* -.618	.546
SLQ	* -.745	* -.760	* -.830	1.000	* .717	* -.628
Fo	-.567	-.556	* -.618	* .717	1.000	-.326
dB SPL	.464	* .608	.546	* -.628	-.326	1.000

laryngeal measures because when the vocal folds are in a fully adducted position, vocal fold lengthening can increase abduction slightly (Hirano, Kiyokawa and Kurita, 1988). Across all sessions and loudness levels, vocal intensity ranged from 63.7 dB SPL to 73.9 dB SPL (X=68.2, SD=2.6). The largest between session mean difference in intensity was 4.9 dB (between 1 hour and 1 week post injection). Intensity may show some relation to changes of adduction and glottal aerodynamics because reduced glottal adduction and flow declination rate have been shown to reduce phonatory loudness.

Intercorrelational Analysis

Pearson product moment correlations were computed across all tokens to determine if any of the phonatory measures covaried. Some significant ($r \geq .7$, $p \leq .01$) relations among kinematic variables were observed (Table 3). Likewise, significant relations among aerodynamic variables were present (Table 4). This was expected because different measures derived from the same waveform are not necessarily independent. For example, the maximum negative SO9050 occured at week 2, consistent with the minimum EGGW50 values during this time (Figure 5). The SO9050 and EGGW50 measures had a correlation of r=.874 (Table 3) taken over all tokens. This relation is consistent with the concept that the EGG waveform width is driven by the speed or abruptness of opening.

Only weak to moderate correlations were found among FPK, FDPK, and FOQ (Table 4), even though each of these measures was obtained from the AC portion of the glottal volume velocity waveform. Multiple regression analysis was computed in order to obtain the semi-partial (part) correlation (Howell, 1982) between FOQ and FDPK.

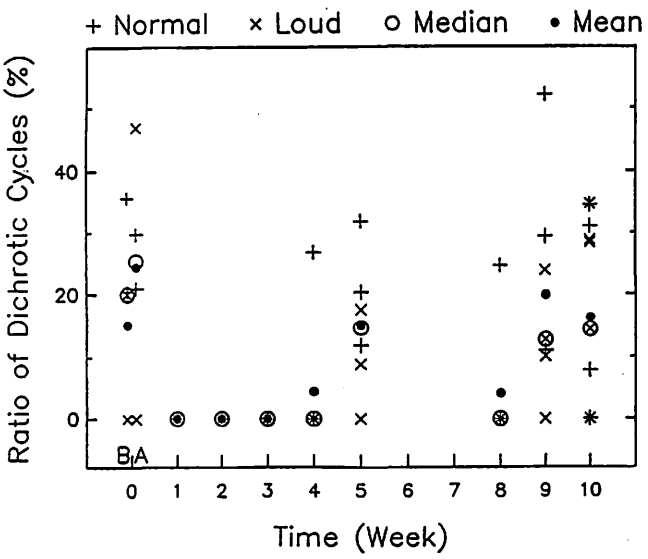


Figure 7. The percentage of dichrotic cycles within the midvowel portion of the /pæ/ syllables produced by the ADSD subject during each recording session.

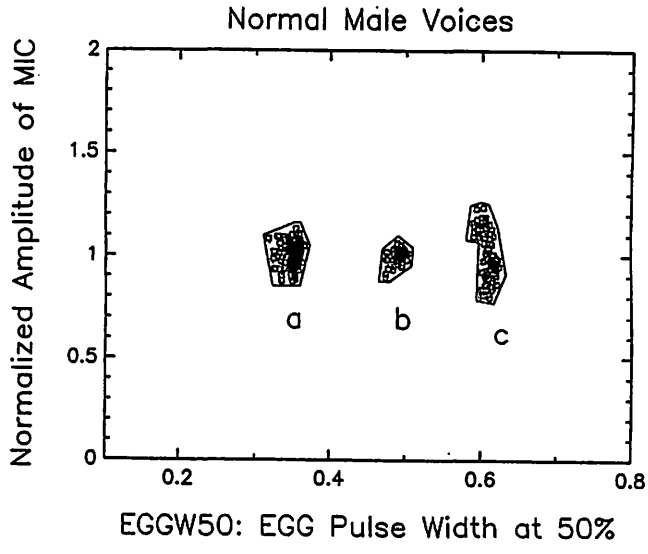


Figure 8. Electroglottographic waveform pulse width (EGGW50) versus the microphone peak-to-peak amplitude value for successive glottal cycles during a single prolonged /pæ/ utterance produced by each of three normal adult male subjects. Subjects (a) and (c) were judged to be slightly breathy and moderately pressed, respectively.

Table 4.
Correlation Matrix for Aerodynamic Variables

	dc flow	FOQ	FPK	FDPK	FPQ	Fo	dB SPL
dc flow	1.000	* .695	.100	.436	* -.675	.477	* -.689
FOQ	* .695	1.000	.108	.553	* -.769	* .622	* -.730
FPK	.100	.108	1.000	* -.597	-.167	.189	-.021
FDPK	.436	.553	* -.597	1.000	* -.616	.153	-.543
FPQ	* -.675	* -.769	-.167	* -.616	1.000	-.516	* .621
Fo	.477	* .622	.189	.153	-.516	1.000	-.326
dB SPL	* -.689	* -.730	-.021	-.543	* .621	-.326	1.000

The part correlation (r=.621) as compared to the simple correlation (r=.553, Table 4) suggested that FPK was supressing the variability in FDPK that was unrelated to FOQ. Thus, FPK and FOQ were relatively independent in their prediction of FDPK. Taken together, FPK and FOQ predicted approximately 74% of the variance (R=.861, p<.0001) in FDPK.

Some significant relations between aerodynamic and kinematic variables were present. For example, SO9050 was negatively correlated with both dc flow (r=-.761, p<.0001), and FOQ (r=-.709, p<.0001). Also, FOQ was correlated significantly with both SLQ and EGGW50 (r=.766, p<.0001, r=-.720, p<.0001). As previously suggested, each of these variables showed moderate correlations with dB SPL or f_0.

Phonatory Instability Measures

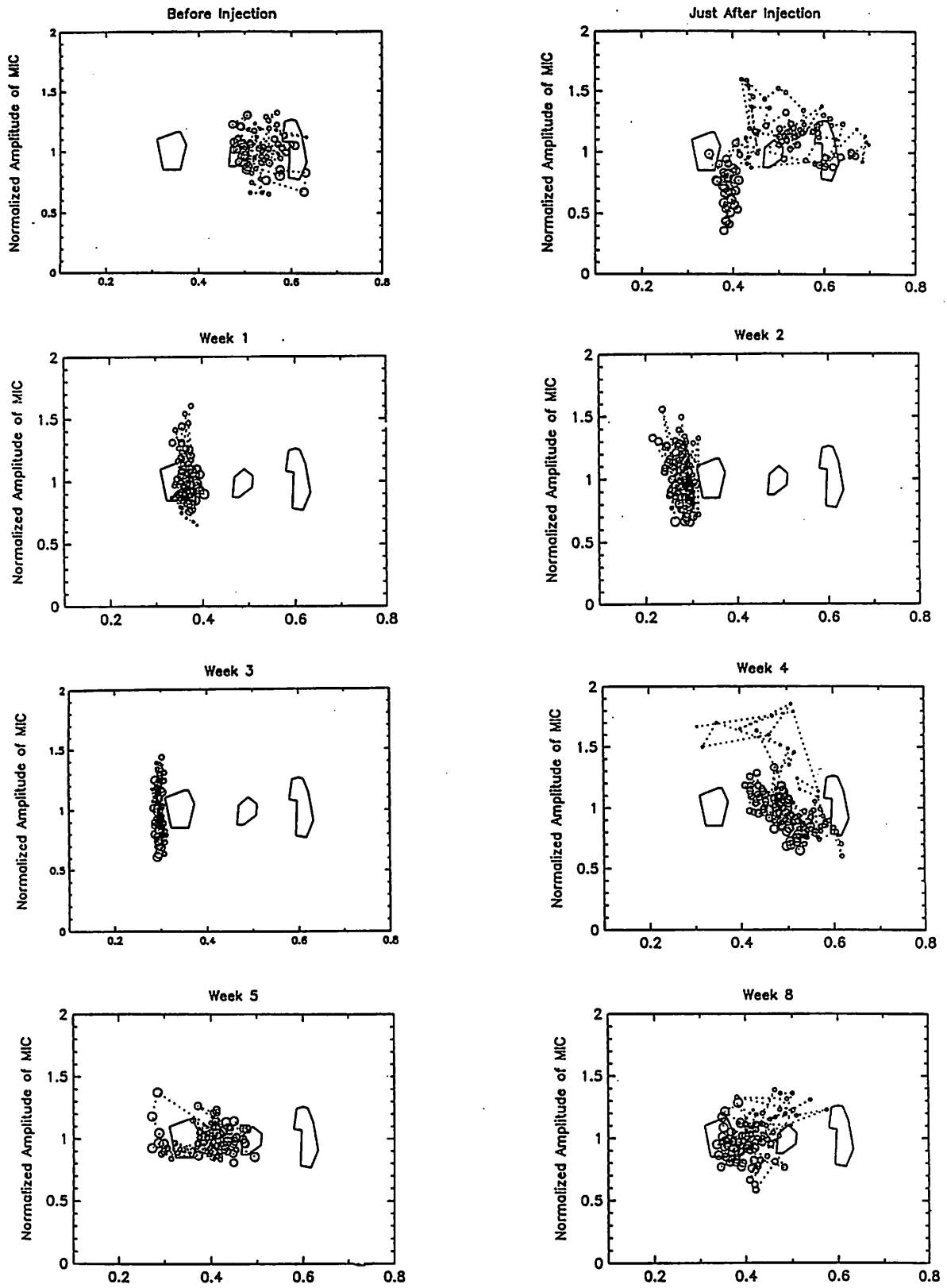
Figure 7 shows the percentage of dichrotic cycles within the final prolonged /æ/ segment from each of the /pæ/ sequences recorded each week. No dichrotic segments were detected during the weeks where voice quality was judged to be relatively more breathy (weeks 1 through 3 after injection). Other weeks show more abundant dichrotic cycles.

Figure 8 shows the adduction measure, EGGW50, versus the microphone signal amplitude for successive glottal cycles during a single prolonged spoken /pæ/ utterance produced by each of three normal adult male subjects (a,b, and c). Each open square represents the EGGW50 and microphone peak-to-peak amplitude value for a single phonatory cycle. Notice that the excursions of adduction and the microphone signal amplitude remain within a relatively small outlined area, even though each subject used a characteristically different range of adduction. Informal judgement of quality by the authors suggested that subject "a" sounded slightly breathy and subject "c" sounded slightly pressed.

Figure 9 shows similar plots for the ADSD subject before injection and at weekly intervals after injection. The displays are taken from the final prolonged /æ/ in the /pæ/ sequence that corresponded to the median EGGW50 value of each session. The token closest to the median value was selected if there was an even number of tokens within a session. Dichrotic segments of phonation are not included because they are thought to reflect a different mode of phonation and would increase the apparent kinematic instability. Each open circle represents the value for an individual cycle. Consecutive cycle values from the beginning of phonation (small circles) to the end of the utterance (progressively larger circles) are connected by a dotted line. Before and one hour after injection (Figures 9a and 9b, respectively; see following page), the excursions of adduction and microphone signal amplitude are relatively large when compared to that of the normal male subjects (shown in the smaller enclosed areas, taken from Fig. 8). Comparable instability in the microphone signal of a spasmodic dysphonia patient has been observed previously (Scherer, Gould, et al., 1988).

Visual inspection of Figures 9c through 9e (relatively more breathy sessions) suggests that the within-syllable EGGW50 instability and variability may reflect levels more simlar to those of the three normal subjects. Large variability (when present) appeared to involve two components. In the short-term, cycle-to-cycle variability (perturbation) in adduction is present. A long-term component appeared to contain small "families" of adjacent cycles that would vary from the center cluster, either in adduction, microphone amplitude, or both (Figure 9b, 9f, 9g). The variability did not appear to be repetitive as might be expected for tremor.

Effect size for the short-term (EGGW50 Jitter, r^2=.501) and long-term (EGGW50CVA, r^2=.714) instability measures suggested that a larger proportion of variation in long-term instability could be attributed to change over time. Figures 10 and 11 show for each week the short-term sequential variability measure (EGGW50 jitter), and the long-term variability measure (EGGW50 CVA), respectively. Both cycle-to-cycle variation and longer term variation of adduction varied over a wide range (by greater than a factor of 6 for each measure).



Figures 9a-9h (a is upper left, b upper right, and so forth). Electroglottographic waveform pulse width (EGGW50) versus microphone peak-to-peak amplitude value for successive glottal cycles during the midvowel portion of single prolonged /pæ/ utterances produced by the ADSD subject. Figures 9a-9h represent the median EGGW50 token for individual recording sessions beginning at week 0 (before and just after injection) through week 8 post-injection.

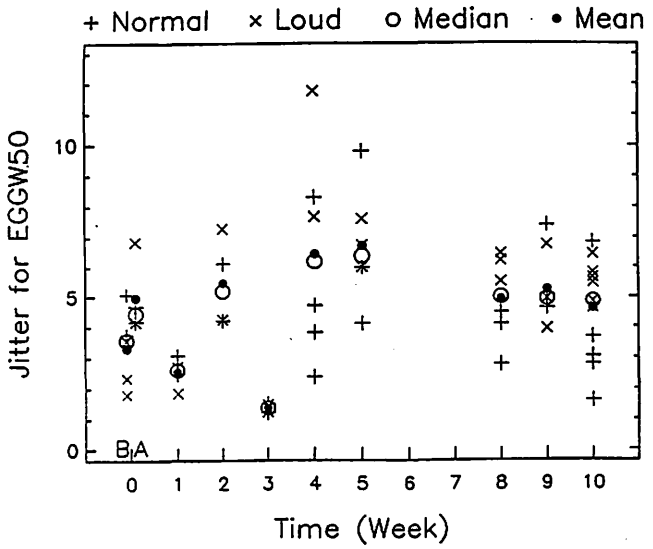


Figure 10. Jitter for the electroglottographic waveform pulse width (EGGW50) during the non-dichrotic midvowel portion of the /pæ/ syllables produced by the ADSD subject during each recording session.

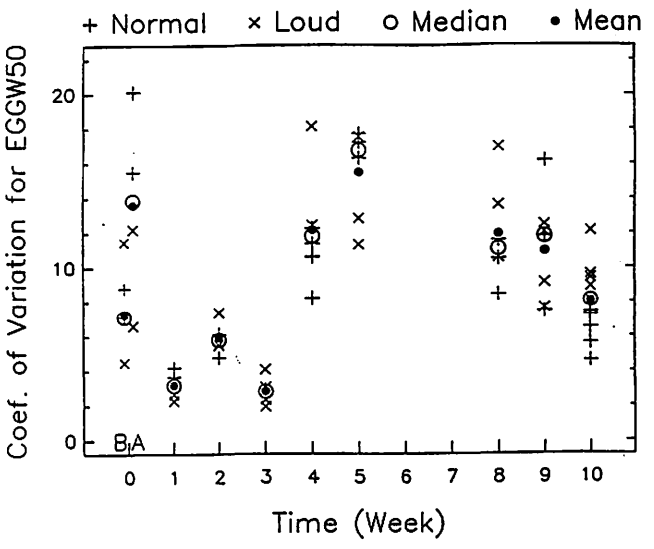


Figure 11. Coefficient of variation for the electroglottographic waveform pulse width (EGGW50) during the non-dichrotic midvowel portion of the /pæ/ syllables produced by the ADSD subject during each recording session.

Discussion

Intra-subject variation in aerodynamic and kinematic phonatory measures was large for this subject, even within each loudness condition. This finding agrees with those of normal voices (Holmberg et al., 1988; Schutte, 1980). Variability was expected to be quite large for this patient due to phonatory instability at the gross spasmodic level, potential Botox®-related changes in laryngeal function, and compensations that may be associated with the onset and recovery from Botox®-related vocal fold weakness. Variation in EGGW50, SLQ, and FPQ of greater than an order of magnitude were observed between adjacent measurement sessions and, in some cases, within a single session. This variability means that a patient may require frequent measurement over a greater number of voiced segments in order to characterize their phonatory function. Non-steady adduction during a phonation limits the ability of averaged laryngeal measures to show large effect size over time. In spite of this limitation, weekly measures characterized trait variability in phonatory behavior that was generally attributable to change over time. Previous investigators, however, have not quantified fine or gross voice symptom variation at such frequent intervals.

Laryngeal measures with moderate and large effect sizes tracked the perceived phonatory quality during the post-injection period. For this patient, measures of EGG waveform symmetry (SLQ and SO9050) and glottal compentency (DC flow) showed strongest relation to the clinical perceptions of voice quality that ranged between breathy and pressed. The time courses of variations in voice quality and laryngeal measures generally were consistent with clinical experiences that suggest hypoadduction after Botox® injection, with eventual reversal toward pre-injection levels of greater adduction. Future application of this method may determine which of the measures, if any, are practical and clinically sensitive to the range of voice qualities presented by other ADSD patients.

Several EGG measures reached maxima or minima during the breathy voice sessions. For example, the tendency for steep contact opening slope (S09050) and small EGGW50 during the breathy weeks also may be explained by less vocal fold contact due to less adduction. During relatively pressed voice sessions, the ratio of closing to opening contact slopes (SLQ, Figure 5) is relatively large negatively due to fast glottal closure and/or a relatively long closure period increasing the time from maximum closure to glottal opening. SLQ was found to approach negative unity during perceptually breathy sessions (weeks 1 through 3) and was due to a relatively more symmetrical EGG waveform during this time. This would be expected for paretic vocal folds that close slowly (or minimally) and open more abruptly. The SLQ also could be more strongly negative as f_0 is increased, a conjecture that is supported by the moderate negative correlation between SLQ and f_0.

Three aerodynamic variables (dc flow, FOQ and FPQ) showed strong relation to perceived vocal quality. An initial expectation was that FPK also would increase with decreased laryngeal adduction and increased peak glottal permittance (Hillman et al., 1989; Sundberg et al. 1993; Rothenberg & Mahshie, 1988). For this patient, however, FPK effect size was small. This might be explained by

hypoadduction and breathiness that could, when extreme, reduce modulation of the air stream. As the hypoadduction subsides, there is increased potential for modulation of the air stream. The possible f_0 increase after injection also would be expected to limit potential FPK increase. Additional information (about respiratory adjustments, subglottic pressure, glottal configuration, and/or supraglottic impedance) would be necessary to give a satisfactory physical explanation.

Quantitative information about adduction might profit medical decision making for this patient. For example, the majority of this patient's FOQ values at weeks 1 through 5 after injection (generally greater than 0.7) were above the expected normal range (Higgins & Saxman, 1991; Holmberg et al., 1988). Furthermore, clinically significant hypoadduction (OQ near 1.0) and minimal AC glottal volume velocity (greater than three standard deviations below normal) after injection might contraindicate larger doses of Botox® due to the potential for aphonia. In order to make such decisions, it will be necessary to establish both baseline and post-injection levels of adduction for the patient. Currently, modification of dose is based largely upon clinical judgment and patient verbal report or anecdotal information from rating scales and patient-kept diaries (Brin et al., 1992).

Clinical impressions and research findings suggest that current methods of perceptual voice rating, and the patient's ability to make certain judgments and keep diaries about their post-injection voice, may be questioned (Ford et al., 1992). In spite of training stimuli and multiple listening opportunities, some listener judges in this study found the rating of breathiness to be a difficult task. This may be due in part to the changes in dynamic spasmodic symptoms that co-occur with gross adductory changes and confound the listening task. Apparent differences in the voice rating scale and objective voice descriptions make interpretation of this patient's self-assessment more difficult. Similar observations have been published elsewhere and may be due to a patient's own perception of vocal effort which can increase as the toxin effect subsides (Zwirner et al., 1993).

Adductor spasmodic dysphonia patients demonstrate gross adductory adjustments that are not normal. Glottal fry and diplophonia have been observed in some ADSD phonation, and inspection of EGGW waveforms revealed a significant number of dichrotic segments. Titze, Baken and Herzel (1993) also observed dichrotic vibratory behavior in an ADSD patient. For the patient reported here, dichrotic cycles were absent during breathy sessions, but were more frequent during sessions with relatively greater vocal fold adduction. This generally was consistent with the clinical expectation that phonatory stability would increase after injection but, over time, reverse toward pre-injection level as the paretic affect subsided. This is not to say that we attribute this change to treatment; rather, the data are not in conflict with previous findings that gross symptoms and acoustic aperiodicity improve after Botox® injections associated with vocal fold paralysis or weakness (Ludlow, 1990).

Several factors might account for the absence of dichrotic cycles during the breathy sessions. Dichrotic behavior and other supraperiodic vibratory behavior have been observed in some computer simulations (Smith, Berke, Gerratt, & Kreiman, 1992; Wong, Mabo, Cox, & Titze, 1991) and in a unilateral vocal fold paralysis patient (Scherer et al., 1988) when vocal fold force, stiffness, or adduction are asymmetrical. This leads to speculation that tension, force or adduction asymmetry might have been reduced during breathy sessions when compared to spasmodic sessions for this patient.

It is difficult to support this hypothesis from a purely aerodynamic or mechanical perspective. Because of possible differences in internalization, leakage, and localization of toxin with the bilateral percutaneous injection procedure, it seems unlikely that each vocal fold would have a similar number or arrangement of motor units that have equally toxin-affected synaptozomic release processes. Thus, vocal fold stiffness or adduction could be asymmetric even after a bilateral injection. Some have speculated, however, that (after injection) altered feedback about pressure, tissue contact, muscle activation, or some other mechanism allows for summation of force that is similar for each fold.

It also seems possible that, during breathy sessions, the amount of hypoadduction exceeded a critical threshold. Even though some tension or adduction asymmetry still might be present, it may not be substantial enough to bring about a supraperiodic mode of vocal fold motion, particularly when abduction is extreme. In normal voices and computer models, bifurcations tend to occur at the borderline of entrainment zones (Isshiki, Tanabe, Ishizaka, & Broad, 1977; Herzel, 1992; Smith et al., 1992). Vocal fold tension and/or adduction for this patient may have been reduced below this borderline zone.

Dichrotic vocal fold vibration in this ADSD patient may arise because of difficulty maintaining stable vocal fold tension and adduction during oscillation. Vibration of the false vocal folds, mechanical interaction between the false and true vocal folds, or aerodynamic interaction of the glottal flow with the ventricular fold boundaries during a pressed mode of phonation are other potential sources of dichrotic vibration. Mechanical or aerodynamic interaction might load one or both folds differently. Resulting horizontal, vertical and/or anterior-posterior vocal fold desynchronization could contribute to dichrotic vibratory modes of phonation (Herzel, 1992).

Some voice irregularities during non-dichrotic segments appeared to be related to sources other than kinematic instability. Peak-to-peak microphone amplitude variability, for example, was not necessarily dependent on the kinematic instability. Figure 9e shows relatively large variability in microphone signal amplitude when compared to the glottal

adduction measure. This could be the result of air turbulence that accompanies perceived breathiness and potentially reduced tissue adduction during this time (Isshiki, Kitajima, Kojima, & Harita, 1978). The possibly confounding effects of aerodynamic and kinematic sources of acoustic perturbation may account in part for apparent non-significant findings regarding acoustic perturbation change after Botox® injection (Ford et al., 1992; Ludlow, 1990; Zwirner et al., 1991). It also has been observed that steady-state phonation (necessary for most perturbation analyses) is less likely to elicit SD symptoms in less severe patients. Note that acoustic perturbation measures were not attempted in this study due to possible acoustic distortion by the flow mask.

Conclusions

This study explored the utility of glottal measures in a highly unstable voice. The voice of this ADSD patient who underwent Botox® treatment provided wide variability and thus posed an ideal challenge for the method. Even with this challenge, the method allowed non-invasive, objective quantification of laryngeal function associated with clinical perceptions. It extended work with a spasmodic dysphonia patient from gross measures of sypmtoms to quantitatification of glottal adduction both in kinematic and aerodynamic terms. The information complimented videostroboscopic evaluation and was obtained even when images of the vocal folds were obscured due to spasmodic laryngeal constriction. Given that phonatory measures for this patient were quite variable, pre-treatment baseline was not studied, and that uninjected or normal controls were not followed over the same time period, the data presented here should not be generalized to patient groups and are not intended to demonstrate Botox® treatment efficacy.

If symptom categories of different classes of spasmodic dysphonic patients are tied to the biomechanics of the larynx, and gross symptoms are the result of compensations in vocal fold adduction and/or stiffness (Watson et al. 1991), kinematic and aerodynamic measures may be helpful to identify the type, pattern and extent of compensation used by patients. Such compensations potentially may vary with individual patients. Several investigations have shown that relations among respiratory or phonatory variables can vary among normal speakers (Stathopoulos & Sapienza, 1993; Higgins, Netsell & Schulte, 1994). The quantitative approach presented here allows objective comparison of phonatory response not only within a patient, but also between patients.

Due to variability between and within recording sessions, regular and frequent measurement was useful in describing the phonatory characteristics of the patient. Weekly measurements were practical for this purpose, although results from animal studies indicate that more frequent measures might detect meaningful change as early as three to four hours post-injection (Prakash, Luschei, Ramos, & Bless, 1994). The contribution of intraindividual variability and change to interindividual variance might account in part for differences among patients' apparent responses to Botox® injection. It will not be possible to determine these time-related variabilities with designs that offer only a single pre- and post-injection comparison for groups of subjects.

The accuracy of the multi-signal approach discussed here relies on the opportunity of the examiner to obtain a valid volume velocity signal and an adequate electroglottographic signal. Tremor, for example, may pose challenges for both signals due to changing resonance structure and larynx movement. Even within this limitation, it appears possible to design treatment efficacy studies that quantify changes in the manner and degree of adduction in kinematic and aerodynamic terms. This information, in addition to that regarding gross symptom reduction, could aid in determining the most effective intervention on an individual patient basis. In the future, multidimensional physiological measures may help to identify sources of perceptual symptom heterogeneity and differential toxin effects among spasmodic dysphonic patients.

Endnotes

1. Botox® is a registered trademark of *Allergan*, 1010 Silverbell Lane, Knightsdale, NC 27545.

2. The portion of the glottal volume velocity waveform that lies above the baseline portion corresponds to the motion of the vocal folds away from their most midline position. The "opening" and "closing" points do not necessarily correspond to moments of full glottal closure or full closure of the membranous glottis, but to minimal flow during the phonatory cycle. On rare occasions it was necessary to estimate subjectively the "closing" point. In a very breathy voice, the glottal volume velocity waveform may have no flat "closed" portion. In this case subjective determination of "opening" and "closing" points is more difficult and may slightly over-estimate the glottal cycle open time (Higgins & Saxman, 1991). Selecting the opening point at the maximum value of the second flow derivative allows for a "closed" segment in almost all waveforms, even in breathiest phonations. Furthermore, it is understood that a flat baseline does not necessarily mean a zero-flow condition, especially if the posterior cartilaginous glottis has a non-zero area when the membranous glottis is completely closed.

3. The SC1090 EGG contact closing slope measure also can be described as the actual closing slope b_1/a_1 normalized by the ratio of the cycle height H to period T. Normalizing b_1 by the height H is important in order to account for variations in height not due to laryngeal differences (e.g., slight differences in electrode placement, EGG amplification, or neck impedance across subjects). Normalizing a_1 by T helps to create a relatively pitch independent

measure. Also, the SC1090 measure is .8 divided by a_1/T, suggesting that it is primarily related to a_1.

4. The rationale for SO9050 is similar to that for SC1090, refer to Endnote 3. SO9050 also can be thought of as the opening slope b_2/a_2 normalized by the ratio H/T. Also, it is 0.4 divided by a_2/T, suggesting that it is primarily related to a_2.

5. Although the terms "closing" and "opening" contact are used, it is understood that neither the 3-dimensional glottal volume nor the exact moment of vocal fold closing and opening can be determined using electroglottography. Also, an increase or decrease of the EGG waveform most likely corresponds to some greater or lesser tissue contact, respectively, between the two vocal folds, but "complete" glottal closure or opening is not necessarily deducable from the waveform.

Acknowledgements

The authors wish to acknowledge Keith Clark, M.D. who provided the Botox® injection treatment, and students Greta Rowe and Jennifer Kurowski who assisted with the listening sessions. We thank reviewers Gail Kempster, Thomas Doherty and Richard Morris for their helpful comments. We also thank Willis Owen and Paul Swank for statistical consultation. Grant P60 DC00976 from the National Institute on Deafness and Other Communication Disorders supported digitizing and analysis of data in Denver. An ASHA Foundation award supported the acquisition of equipment, and contract number HN2-031 from the Oklahoma Center for the Advancement of Science and Technology supported patient recording and data analysis in Oklahoma City.

References

Arends, N., Povel, D., Van Os, E., & Speth, L. (1990). Predicting voice quality of deaf speakers on the basis of glottal characteristics. Journal of Speech and Hearing Research, 33, 116-122.

Aronson, A. E. (1985). Clinical voice disorders. An interdisciplinary approach (2nd ed.). New York: Thieme.

Aronson, A. E., Brown, J. R., Litin, M. E., & Pearson, J.S. (1968). Spastic dysphonia. I. Voice, neurologic and psychiatric aspects. Journal of Speech and Hearing Disorders, 33, 203-218.

Blitzer, A., Brin, M.F., Fahn, S., & Lovelace, R.E. (1988). Localized injections of botulinum toxin for the treatment of focal laryngeal dystonia (spastic dysphonia). Laryngoscope, 98, 193-197.

Brin, M., Blitzer, A., Fahn, S., Gould, W., & Lovelace, R. (1989). Adductor laryngeal dystonia (spastic dysphonia): treatment with local injections of botulinum toxin (Botox). Movement Disorders, 4, 287-296.

Brin, M., Blitzer, A., Stewart, C., & Fahn, S. (1992). Treatment of spasmodic dysphonia (laryngeal dystonia) with local injections of botulinum toxin: review and technical aspects. In Blitzer, A., Brin, M., Sasaki, C., Fahn, S., & Harris, K. (Eds.), Neurologic disorders of the larynx (pp. 214-228). New York: Thieme Medical Publishers.

Childers, D. G. (1991). Assessment of the acoustics of voice production. In Judith A. Cooper (Ed.), Assessment of Speech and voice production: research and clinical applications, NIDCD Monograph, 1, 63-83.

Childers, D. G., & Lee, C. K. (1991). Vocal quality factors: analysis, synthesis and perception. Journal of the Acoustic Society of America, 90, 5, 2394-2410.

Dejonckere, P. H., & Lebacq, J. (1985). Electroglottography and vocal nodules. Folia Phoniatria, 37, 195-200.

Ebel, R. L. (1951). Estimation of the reliability of ratings. Psychometrika, 16, 407-421.

Ford, C. N., Bless, D. M., & Lowrey, J. D. (1990). Indirect laryngoscopic approach for injection of botulinum tosin in spasmodic dysphonia. Otolaryngol Head Neck Surg., 103, 752-758.

Ford, C. N., Bless, D. M., & Patel, N. Y. (1992). Botulinum toxin treatment of spasmodic dysphonia: techniques, indications, efficacy. Journal of Voice, 6, 370-376.

Freeman, F. J., Cannito, M. P., & Finitzo-Hieber, T. (1985). Classification of spasmodic dysphonia by perceptual-acoustic-visual means. In G. Gates (Ed.) Spasmodic dysphonia: The state of the art, 1984. New York: The Voice Foundation.

Gauffin, J., & Sundberg, J. (1980). Spectral correlates of glottal voice source wave form characteristics. Journal of Speech and Hearing Research, 32, 556-65.

Herzel, H. (1992). Bifurcations and chaos in voice signals. In Chaos: The New Idea, T. Kapitaniak (Ed.). Cambridge, MA: University Press.

Higgins, M. B., & Saxman, J. H. (1991). A comparison of selected phonatory behaviors of healthy aged and young adults. Journal of Speech and Hearing Research, 34, 5, 1000-1010.

Hillman, R. E., Holmberg, E. B., Perkell, J. S., Walsh, M., & Vaughan, C. (1989). Objective assessment of vocal hyperfunction: an experimental framework and initialresults. Journal of Speech and Hearing Research, 32, 373-392.

Hirano, M., Kiyokawa, K., & Kurita, S. (1988). Laryngeal muscles and glottis shaping. In O. Fugimura (Ed.), Vocal Physiology: Voice Production, Mechanisms and Functions (pp.49-65). New York: Raven Press, Ltd.

Holmberg, E. B., Hillman, R. E., & Perkell, J. S. (1988). Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal and loud voice. Journal of the Acoustical Society of America, 84 (2), 511-529.

Houben, G. B., Buekers, R., & Kingman, H. (1992). Characterization of the electroglottographic waveform: a primary study to investigate vocal fold functioning. Folia Phoniatrica, 44, 269-281.

Howell, D. C. (1987). Statistical Methods for Psychology. (pp. 487-491). Boston: Duxbury Press.

Isshiki, N. Kitajima, K., Kojima, H., & Harita, Y. (1978). Turbulent noise in dysphonia, Folia Phonatr., 30, 214-244.

Isshiki, N., Tanabe, M., Ishizaka, K., & Broad, D. (1977). Clinical significance of asymmetrical vocal cord tension. Annals of Otology, Rhinology and Laryngology, 86, 58-66.

Izdebski, K. (1992). Symptomatology of adductor spasmodic dysphonia: A physiologic model. Journal of Voice, 6, 306-319.

Lofqvist, A. (1982). Aerodynamic measurements of vocal function. In A. Blitzer, M. Brin, C. Sasaki, S. Fahn, & K. Harris (Eds.), Neurologic Disorders of the Larynx. (pp 98-108). New York: Thieme Medical Publishers, Inc.

Ludlow, C. L. (1990). Treatment of speech and voice disorders with botulinum toxin. Journal of the American Medical Association, 264, 2671-2676.

Ludlow, C., & Connor, N. (1988). Dynamic aspects of phonatory control in spasmodic dysphonia. Journal of Speechand Hearing Research, 55, 2-24.

Ludlow, C.L., Naunton, R.F., Sedory, S.E., Schulz, G.M., & Hallett, M. (1988). Effects of botulinum toxin injections on speech in adductor spasmodic dysphonia. Neurology, 38, 9-16.

Miller, R. H., Woodson, B. E., & Jankovic, J. (1987). Botulinum toxin injection of the vocal fold for spasmodic dysphonia. A preliminary report. Archives of Otology and Head Neck Surgery., 113, 603-605.

Orlikoff, R. F. (1991). Assessment of the dynamics of vocal fold contact from the electroglottogram: data from normal male subjects. Journal of Speech and Hearing Research, 34, 5, 1066-1072.

Prakash, S., Luschei, E. S., Ramos, C. A., & Bless, D.M. (1994). Botox induced changes in canine laryngeal musculature: Electromyographic and videoendoscopic findings. Journal of the American Speech Language Hearing Association [Abstracts], October, 56.

Rothenberg, M., & Mahshie, J. (1988) Monitoring vocal fold adduction through vocal fold contact area. Journal of Speech and Hearing Research, 31, 38-51.

Scherer, R. C. (1991). Physiology of phonation: A review of basic mechanics. In Ford, C.N. & Bless, D. M. (Eds.), Phonosurgery: Assessment of Surgical Management of Voice Disorders, (pp.77-93). Raven Press: New York.

Scherer, R. C. (1991b). Aerodynamic assessment in voice production. In: J. A. Cooper (Ed.), Assessment of Speech and Voice Production: Research and Clinical Applications. NIDCD Monograph, 1-1991, 112-113.

Scherer, R. C., Cooper, D., S., Alipour-Hashishi, F., & Titze, I. R. (1984). Vocal process contact pressure. In V. Lawrence (Ed.), Transcripts of the Twelfth Symposium: Care of the Professionsal Voice(pp. 61-68). New York, New York: The Voice Foundation.

Scherer, R. C., Druker, D. B. & Titze I. R. (1988). Electroglottography and direct measurement of vocal fold contact area. In O. Fujimura (Ed.), Vocal Physiology: Voice Production Mechanisms and Functions. (pp.279-291). New York, New York, Raven Press Ltd.

Scherer, R. C., Gould W. J., Titze, I. R., Meyers, A. D., & Sataloff, R. T. (1988). Preliminary evaluation of selected acoustic and glottographic measures for clinical phonatory function analysis. Journal of Voice, 2, 3, 230-244.

Scherer, R. C., Sundberg, J., & Titze, I.R. (1989). Laryngeal adduction related to characteristics of the flow glottogram. Journal of the. Acoustical Society of America, 85(S1), S129(A).

Scherer, R. C., Vail, V. J., & Rockwell, B. (In Press). Examination of the laryngeal adduction measure EGGW. In F. Bell-Berti & L. J. Rapheal (Eds.), Producing Speech: A Festschrift for Kathering Safford Harris. American Institute of Physics.

Schutte, H.K. (1980). Doctoral dissertation, The efficiency of voice production. University of Groningen.

Smith, M., Berke, G., Gerratt, B., & Kreiman, J. (1992). Laryngeal paralyses: theoretical considerations and effects on laryngeal vibration. Journal of Speech and Hearing Research, 35, 3, 545-554.

Sundberg, J., Titze, I., & Scherer, R. (1993). Phonatory control in male singing: A study of the effects of subglottal pressure, fundamental frequency, and mode of phonation on the voice source. Journal of Voice, 7, 15-29.

Titze, I.R. (1989). A four-parameter model of the glottis and vocal fold contact area. Speech Communication, 8, 191-201.

Titze, I.R. (1990). Interpretation of the electroglottographic signal. Journal of Voice, 4, 1-9.

Titze, I.R. (1992). Acoustic interpretation of the voice range profile (phonetogram). Journal of Speech Hearing Research, 35, 21-24.

Titze, I.R., Baken, R., & Herzel, H. (1993). Evidence of chaos in vocal fold vibration. In I. R. Titze (Ed.), Vocal Fold Physiology: New Frontiers in Basic Science (pp.143-178). San Diego, CA: Singular Publ. Group.

Titze, I.R., Scherer, R.C., & Winholtz, W. (1985). Manual of Instruction for GLIMPES Users. Part I. Recording of Signals. (The Recording and Research Center, The Denver Center for the Performing Arts).

Titze, I.R., & Sundberg, J. (1992). Vocal intensity in speakers and singers. Journal of the Acoustical Society of America, 91, 2936-2946.

Truong, D., Rontal, M., Rolnick, M., Aronson, A., & Mistura, K. (1991). Double-blind controlled study of botulinum toxin in adductor spasmodic dysphonia. Laryngoscope, 101, 630-634.

Watson, B., Schaefer, S., Freeman, R., Dembowski, J., Knodraske, G., & Roark R. (1991). Laryngeal electromyographic activity in adductor and abductor spasmodic dysphonia. Journal of Speech and Hearing Research, 34, 473-482.

Wong, D., Mabo, R., Cox, N., & Titze, I. (1991). Observation of perturbations in a lumped-element model of the vocal folds with application to some pathological cases. Journal of the Acoustical Society of America, 89, 383-394.

Zwirner, P., Murry, T., Swenson, M., & Woodson, G. (1991). Acoustic changes in spasmodic dysphonia after botulinum toxin injection. Journal of Voice, 5, 78-84.

Zwirner, P., Murry, T., & Woodson, G. (1993). Perceptual-acoustic relationships in spasmodic dysphonia. Journal of Voice, 7, 165-171.

Appendix A

Patient's Subjective Voice Descriptions

Before Injection

Week 0 "Voice cracks and at times locks up. I'm not able to speak with a clear voice. Occasionally my voice would have a glottal sound, and sometimes I couldn't speak for a second or two It sounds terrible."

After Injection

Week 0 "Same as before the injection. No discomfort."
Week 1 "A little trouble swallowing water. The voice is much smoother."
Week 2 "No problem swallowing, really. I run out of air quickly when I talk."
Week 3 "Voice feels and sounds pretty good except for a little hoarseness."
Week 4 "Same as last week. Sounds good except for hoarseness."
Week 5 "The voice is improving all the time. It's still hoarse."
Week 8 "Voice not too bad. Hoarseness has gone away some."
Week 9 "Voice feels pretty good. There's some effort to speak, but not bad."
Week 10 "Voice sounds not too good really. It's kind of raspy, not exactly hoarse. It sounds a bit choppy. Worse than last week. It may have been a little choppy then too."
Week 11 "It's a little tight today. I think I'm probably due for another injection. It takes quite a bit, a little, effort to talk today."

NCVS Status and Progress Report - 8
July 1995, 97-111

Electroglottographic Tracking of Phonatory Response to Botox

Kimberly Fisher, Ph.D.
John W. Keys Speech and Hearing Center, The University of Oklahoma Health Sciences Center
Ronald Scherer, Ph.D.
The Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts
Paul Swank Ph.D.
The University of Houston
Cheryl Giddens, M.S.
Dava Patten, B.S.
John W. Keys Speech and Hearing Center, The University of Oklahoma Health Sciences Center

Abstract

Botox® injection into the thyroarytenoid muscle is thought to alter the glottal competency and laryngeal adduction of adductor spasmodic dysphonia (ADSD) patients. These changes have been rated subjectively and inferred from post-injection breathy voice quality, aphonia, midline glottal gap, or subclinical aspiration. Clinical experience suggests that these temporary effects vary in duration and severity among patients. This study utilizes electroglottographic (EGG) measures to model changes over time in laryngeal adduction for five ADSD patients. Three hypotheses were tested: (1) that EGG measures would reveal a trend for reduced laryngeal adduction during the first three weeks post-injection, followed by a reversal toward greater adduction over time; (2) that individuals would differ one from another in their adductory variability over time; and (3) that changes in adduction, if present, would be related to changes in severity of ADSD symptoms. Hierarchical linear modeling was used to examine patient response curves. A trend for change in laryngeal adduction was observed and possibly differed by gender. For the three males, hypoadduction occurred after injection and reversed toward pre-injection level over an 8 week post-injection period. The two female subjectsdemonstrated flat or increased adduction during the early post-injection period. Results also supported the observation that individual subjects varied in duration and magnitude of adductory response to injection. Laryngeal adduction during the prolonged vowels of a syllable string related to ratings of overall symptom severity in two of five patients.

Injection of Botox®[1] into the thyroarytenoid muscle has been shown to reduce adductor spasmodic dysphonia (ADSD) symptoms such as intermittently strained-strangled, effortful phonation (Blitzer, Brin, Fahn, & Lovelace, 1988; Brin, Blitzer, Fahn, Gould, & Lovelace, 1989; Brin, Blitzer, Stewart, & Fahn, 1992; Ludlow, 1990; Ludlow, Naunton, Sedory, Schulz, & Hallett, 1988; Miller, Woodson, & Jankovic, 1987; Zwirner, Murry, Swenson, & Woodson, 1991). Botox® is thought to work at least in part by blocking release of acetylcholine at the neuromuscular junction, thereby preventing muscle contraction (Dolly, Black, Williams, & Melling, 1984; Sellin & Thesleff, 1981; Simpson, 1981, 1992; Thesleff, Molgo, & Lundh, 1983; Valtorta & Arslan, 1993; Tse, Dolly, Hambleton, Wray, & Melling, 1982). This temporary chemical denervation resultsin vocal fold paresis (Brin, et al., 1992; Ludlow, 1990; Ludlow, et al., 1988; Miller, Woodson, & Jankovic, 1987) and, in many cases, hypofunctional laryngeal symptoms (Blitzer & Brin, 1992; Jankovic, Schwartz, & Donavan, 1990; Ludlow, 1990; Maloney & Morrison, 1994).

Hypofunctional symptoms of Botox® injection can include temporary breathiness, aphonia, hoarseness, diplophonia and sub-clinical aspiration of fluids (Brin et al., 1992; Ford, Bless, & Patel, 1992; Ludlow, 1990; Ludlow et al., 1988; Truong et al., 1991). The presence, duration and severity of hypofunctional effects vary to a great extent

[1] Botox is a registered trademark of Allergan, 1010 Silverbell Lane, Knightsdate, NC 27545.

within individual patients (Brin et al., 1992). Researchers hypothesize that this variation may be due to differences in injection protocol, dose level, biological factors, pharmacological factors, and/or individual compensations by patients (Ford, Bless, & Patel, 1992; Ludlow et al., 1988; Ludlow, Bagley, Yin, & Koda, 1992; Truong, Rontal, Rolnick, Aronson, & Mistura, 1991). At present it is not possible to predict which patients may be more susceptible to severe hypofunction. However, unilateral injections may decrease the risks of breathy aphonia and aspiration in some patients (Ludlow et al., 1988; Maloney & Morrison, 1994).

Despite the well-recognized hypofunctional effects, there have been few published quantitative measures of laryngealadduction or glottal competency in patients who are treated with Botox®. Many treatment efficacy studies have addressed perceptual or acoustic voice measurements of gross symptomatic relief from spasmodic voice interruptions and effortful phonation (Ludlow, 1990; Ludlow et al., 1992; Ludlow, Naunton, Fujita, & Sedory 1990; Zwirner, Murray, & Woodson, 1993, 1995). Both dose adjustment and selection of injection protocol, however, are based upon subjective judgment about hypofunctional effects (Blitzer & Brin et al., 1992; Ford et al., 1992; Maloney & Morrison, 1994), as well as symptomatic relief. A noninvasive means to quantify laryngeal adduction, therefore, would be useful to show the efficacy of Botox® treatment and to document objectively any changes in laryngeal adduction associated with treatment. Quantifying adduction longitudinally also could allow modeling of injection response for individual patients, a process that has been attempted only through more invasive techniques in animals (Prakash, Luschei, Ramos, & Bless, 1994).

In a preliminary study, electroglottographic (EGG) measures were shown to track perception of vocal quality in an ADSD patient who underwent Botox injection treatment (Fisher, Scherer, Guo, & Owen , in review). In that study, useful EGG measures were obtained even when spasmodic symptomsobscured viewing of the vocal folds. EGG measures also help to quantify laryngeal adduction in patients with hyper- or hypofunctional voice disorders (Arends, Pavel, Van Os, & Seth, 1990; Childers, Smith & Moore, 1984; cf. Dejonckere & Labacq, 1985; Rothenberg & Mahshie, 1988) as well as in normal subjects who intentionally produce a wide range of glottal adjustments (Scherer, Vail, & Rockwell, in press). EGG measures, though noninvasive, have been used to relate glottal adduction to thyroarytenoid activity (Kempster, Preston, Mack, & Larson, 1987).

This study used EGG measures to quantify laryngeal adduction longitudinally during the Botox® post-injection period for five patients with ADSD. We questioned: (1) whether EGG measures would reveal a trend for reduced laryngeal adduction during the first three weeks post-injection, followed by a reversal toward greater adduction over time; (2) whether individuals would differ one from another in their adductory variability over time; and (3) whether changes in adduction, if present, would be related to changes in perceived severity of ADSD symptoms.

Method

Subjects

Table 1 displays characteristics of the five subjects (mean age= 52 yrs, range= 39 yrs to 68 yrs). The three males and two females presented ADSD sypmtoms that ranged in severity from moderate to profound. Voice diagnosis adhered to clinical criteria reported previously (Aronson, 1985; Freeman, Cannito, & Finitzo-Hieber, 1985) and required the consensus of two speech language pathologists and an otolaryngologist. Transnasal videolaryngoscopic examinations revealed the presence of hyperadductive laryngeal spasms.

The subjects demonstrated no signs of Meige's syndrome and no history of phenothiazine use prior to symptom onset. One female subject, subject F2, presented slight essential head tremor and reported a positive history of familial head and hand tremor. She also had been treated with antidepressant medication for a period of approximately five years prior to symptom onset. There were no known neurologic or psychiatric disorders in the remaining four subjects. The other female subject, F1, reported a sister who displayed similar, but less severe, voice symptoms.

Time from symptom onset for each of the five subjects ranged from 1 to 15 yrs (mean= 5 yrs, sd= 5 yrs). The subjects had received no treatment for ADSD other than speech therapy. All subjects reported lack of benefit from speechtherapy.

Procedure

For each patient, 2.5 units of Botox® solution were injected into each vocal fold via an EMG needle using a percutaneous approach. Monitoring of the muscle activity allowed rough confirmation of the intended injection site.

Table 1.
Characteristics of the three male (M) and two female (F) patients with adductor spasmodic dysphonia.

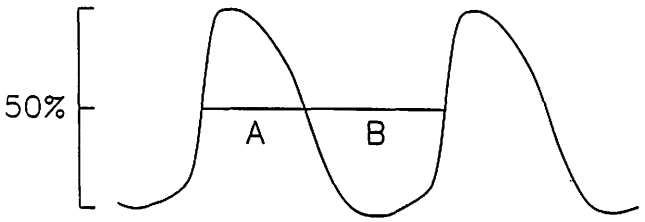
Subject	Years of Age	Years Post Onset	Severity
Males			
M1	68	15	Profound
M2	39	1	Moderate
M3	57	2	Severe
Females			
F1	47	2	Severe
F2	51	3	Moderate
Mean	52	5	

Data recording sessions were scheduled at repeated intervals over an eight week period. The first recording session occurred immediately before injection. The second recording session occurred 1 hour after injection. Each subsequent recording was made at approximately the same time each day, although some patients were not available at precisely weekly or biweekly intervals. All recordings were made within an acoustically isolated sound suite.

The speech materials and recording methods have been described previously (Fisher et al., in review) and were similar to those described in a normative study (Holmberg, Hillman, & Perkell, 1988). Subjects produced a five syllable repetition of /pæ/ under three conditions: normal loudness, soft but not whispered, and loud but not shouting. The vowel portion of the fifth syllable was prolonged for not less than 2 seconds. At least three task repetitions were recorded for each loudness condition during each session. Only the normal andloud conditions will be reported here.[2]

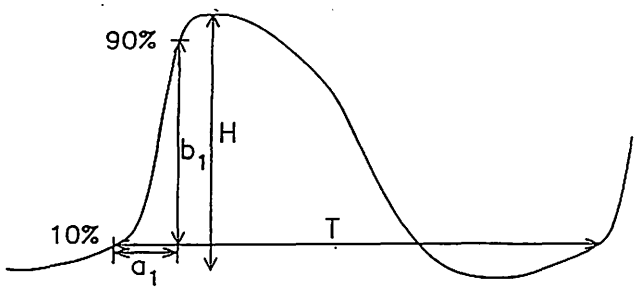
The EGG signal was analyzed using MULTISIG (Scherer & Guo, 1992). All measures were obtained from a segment beginning 200 ms following the onset of voicing in the fifth syllable's prolonged vowel. Analysis ended at least 200 ms before any offset in voicing. This process resulted in analyzed segments of varying length for different subjects (range 6 to 403 consecutive phonatory cycles). The mean of each measure over this segment was taken as the representative value for the phonated segment. Dichrotic phonation was excluded from the analysis because it was thought to represent a different and more variable mode of phonation. All measurements were made blind to treatment status.

Three measures were obtained from the waveforms. First, a duty cycle measure of waveform width at the 50% amplitude level (Figure 1) was made (EGGW50, Scherer, et al., in press; Fisher, et al., in review). A larger EGGW50 value results from a longer apparent closed time of the glottis. Conversely, a relatively narrow waveform, and smaller EGGW50 value, represents less glottal closure, such as expected for breathy voice. Next, the nondimensional opening slope measurement (S09050, Figure 2) was obtained between the 90% and 50% amplitude levels on the downward or glottal opening side. Amore strongly negative slope would be expected for a breathy voice or paretic folds that separate quickly. The nondimensional closing slope (SC1090, Figure 2) was calculated between 10% and 90% of the waveform amplitude level on the upward or glottal closing side. Closure is usually steeper than opening. Steeper contact increase also would be expected for hyperadducted phonation (Orlikoff, 1991). A slope quotient (SLQ) was derived by taking the ratio of SC1090 to SO9050. The slope quotient may be sensitive to waveform symmetry. A more complete discussion of these measures is presented elsewhere (Fisher et al., in review). Interpretation of EGG waveform shape also may be aided by reference to computer modeling (Childers & Lee, 1991; Titze, 1989, 1990).

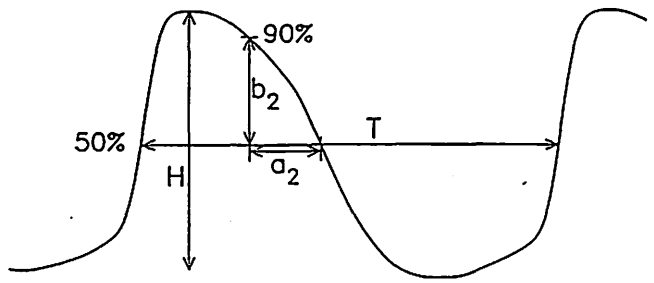


$$EGGW50 = A/(A+B)$$

Figure 1. The electroglottographic waveform width measure (EGGW50).



$$SC1090 = \frac{b_1/H}{a_1/T}, \quad T = \text{period}$$



$$SO9050 = \frac{b_2/H}{a_2/T}, \quad T = \text{period}$$

Figure 2. The closing contact slope measure (top, SC1090), and opening contact slope measure (bottom, SO9050).

[2] The softly phonated syllables were recorded to be consistent with the protocol of Holmberg et al. (1988). It was not anticipated, however, that the soft phonation condition would be analyzed due to the lower likelihood of eliciting ADSD symptoms (Izdebski, 1992).

For each recording session, ratings of ADSD symptom severity were made independently by two speech-language pathologists certified by the American Speech-Language-Hearing Association. Ratings were assigned on an equal appearing interval scale that ranged from 0 (no impairment) to 4 (profound impairment). The clinicians rated each voice after listening to a wide variety of speaking tasks that were recorded on the audio portion of patient data tapes. The tasks included spontaneous conversation, reading of the rainbow passage, production of phrases, laryngealdiadochokinetics, sustained vowels, and the /pæ/ syllable string from which adductory measures were taken (Appendix A). The clinicians were allowed to review the audio tapes as many times as necessary before assigning a severity rating.

Statistical Analyses

A hierarchical linear models (HLM) (Bryk & Raudenbush, 1992) procedure was used to test our first hypothesis that longitudinal EGG measures would reveal a trend for reduced laryngeal adduction during the early post-injection period, followed by a reversal toward greater adduction over time. A between-subjects HLM model was used to determine if individual subject differences in rates of change were a function of chance. For the third hypothesis, within-subject variability in adduction was related to changes in symptom severity using HLM. Gender was included as a control variable in the HLM model in the event that this factor related to the experimental outcomes. In each case, the HLM procedure was used because it: (1) has less stringent assumptions than other modeling procedures, (2) has built-in reliability checks, (3) allows for individual differences, (4) allows for differing times of assessment and numbers of data points for subjects, and (5) provides direct tests of the proposed hypotheses.

Temporal reliability was determined as follows. Ten percent of the waveforms were chosen randomly and reanalyzed. Observed correlations of .99 and small average measurement differences (-.0004, .003, .004 for EGGW50, SO9050, and SC1090, respectively) suggested adequate measurement reliability for the purposes of this study. Ten percent of the audio recordings also were chosen at random and rated a second time by each of the clinicians. An observed correlation of .93 and only two small re-measure differences (of -1 and 1, respectively) in severity rating suggested adequate intra-rater reliability. In all but these two cases, the clinicians' reliability severity rating exactly matched their original rating for the same samples.

Since all five subjects were rated for severity by both judges across at least five sessions, this data was used to estimate inter-rater reliability. Inter-rater reliability was determined using a generalizability procedure (Brennan, 1983). The procedure uses a general linear models approach to estimate the variance components associated with measurement. Judge was included as one facet in the measurement model. In addition, session was included since this was a potential source of consistent variation in the observations. The inter-rater reliability per session was .73. However, a generalizability coefficient of .93 indicated that the judges were relatively consistent across subjects when ratings were averaged over five measurement sessions. The largest sources of relative error were disagreement between judges' ratings of individual subjects' symptom severity from session to session (variance component=.115), followed by smaller inconsistencies in subjects across sessions (variance component = .05). These magnitudes of relative error are comparatively smaller than the variance between subjects, averaged across judges and sessions (.44). Moreover, in all cases, the judges gave the identical or adjacent rating, yielding maximum between-judge differences in severity ratings of 1 and -1, respectively. For all subsequent statistical analyses, the mean of both judges ratings was taken as the estimate of severity for a given patient on a recording day.

Results

Individual Subject Data

Figures 3 and 4 show individual subject data for the variables EGGW50, SO9050, SC1090 and SLQ throughout the post-injection period. Each vertical pair of graphs corresponds to an individual subject. For each graph, time (weeks) is shown on the abscissa. The letter B refers to the measures obtained immediately before Botox® injection. Week 0 refers to the measures obtained 1 hour after injection. Data points and error bars correspond to means and standard deviations, respectively, averaged across both normal and loud attempts. Data were averaged across normal and loud attempts because patients were not able to produce consistently the loudness differences as instructed. Inspection of the data also showed no consistent differences in adduction between loudness conditions. This most likely was due to unstable spasmodic voice symptoms and variable phonatory qualities associated with onset and recovery from a histochemically induced paresis.

Informal visual inspection of these data reveal individual differences in adductory variability both within and between sessions. Further, a wide range of adduction was present for individuals and for the group. For example, variation in EGGW50 from .6 to .1 within individual subjects reflects an extreme range of adduction (Scherer, Vail & Rockwell, in press). Variability differences between sessions ranged over a factor of 5 or 10 for some patients. The most extreme example is SC1090 of patient M2 at 1.5 weeks post-injection (Figure 4). This variability was due to inconsistency in EGG waveform shape near the time of initial vocal fold contact.

The data in Figures 1 and 2 also suggest a natural subgrouping by gender. For the males, minimum values for

the measures occur between weeks 1 and 3; while for the two females, minimum values of adduction occured at pre-injection or at 8 weeks post-injection. Over all subjects the least adduction occurred for subject M3 between weeks 1 and 3. This subject also had the least number of consecutive vibratory cycles during his utterances due to intermittent periods of breathy aphonia.

Modeling of Adductory Change

The HLM procedure was used to determine the best fit models for change of adduction measures over time. The analyses were centered at the time of 3 days post-injection. This time period was chosen because it was well past the expected time of initial paretic effect (3 hours; Prakash et al., 1994), and approximated a time when patients report noticeable effect (Blitzer & Brin, 1992; Jankovic, Schwartz, & Donovan, 1990; Ludlow et al., 1988). It also approximately split the time between 0 and 1 week post-injection, a measurement time interval used in previous studies that have shownefficacious reduction of spasmodic dysphonia symptoms (Ludlow et al., 1988, Truong et al., 1991).

Taking all data into consideration, the best fit was a third degree (four parameter) polynomial. In addition, the best fit for most subject's individual data was a third degree polynomial. This was the maximum degree possible due to a limited number of data points for one subject (M3). The cubic model, therefore, was fit for EGGW50, SC1090, SO9050, and SLQ. The results for each variable will be presented separately. The analysis proceeded in the same way for all variables: first the random coefficients (within subject) model was fit, and then the slopes as outcome (between subjects) model was fit using gender as the explanatory variable.

EGGW50: The four parameter estimates for each of the five subjects with their corresponding group reliability coefficients are given in Table 2. The reliability coefficients represent the proportion of consistent (non-random) variation in the parameters. All reliabilities were above 0.5

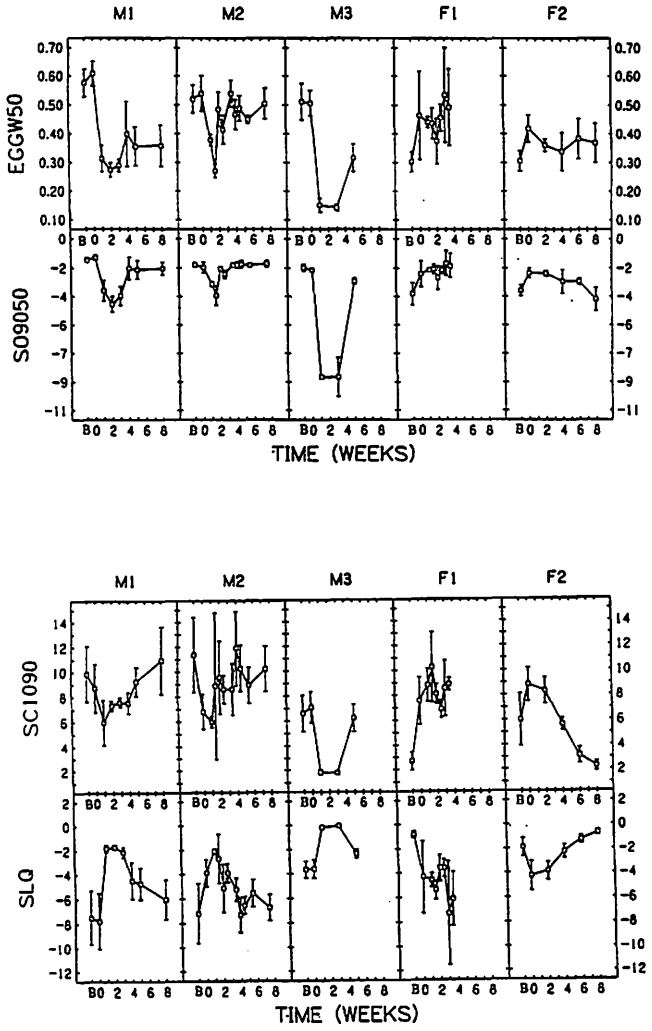


Figure 3(top). The mean ± 1 standard deviation for EGGW50 (top row) and SO9050 (bottom row) for three male (M1-M3) and two female (F1-F2) subjects, individually. For each subject, recording sessions begin at week 0, before (B) injection. Figure 4 (bottom). The mean ± 1 standard deviation for SC1090 (top row) and SLQ (bottom row) for three male (M1-M3) and two female (F1-F2) subjects, individually. For each subject, recording sessions begin at week 0, before (B) injection.

Table 2.
Parameter Estimates and Reliabilities for EGGW50

Subject	Bo	B1	B2	B3
male 1	.39443	-.05427	.04690	-.00594
male 2	.35902	-.15058	.07121	-.00703
male 3	.15085	-.20936	.14903	-.01950
female 1	.43402	-.00508	-.02876	.03463
female 2	.35516	-.00702	.00333	-.00028
reliability	.719	.865	.589	.408

Table 3.
Fixed and Random Effects Terms for EGGW50

effect	fixed coefficient	t	Variance Component	chi-square
Bo	.353270	9.34*	.00558	13.89*
B1	-.066873	-1.46**	.00943	24.81*
B2	.039909	2.17*	.00130	11.99**
B3	-.004310	-2.46*	.00001	8.77**

* p< .05
** p< .10

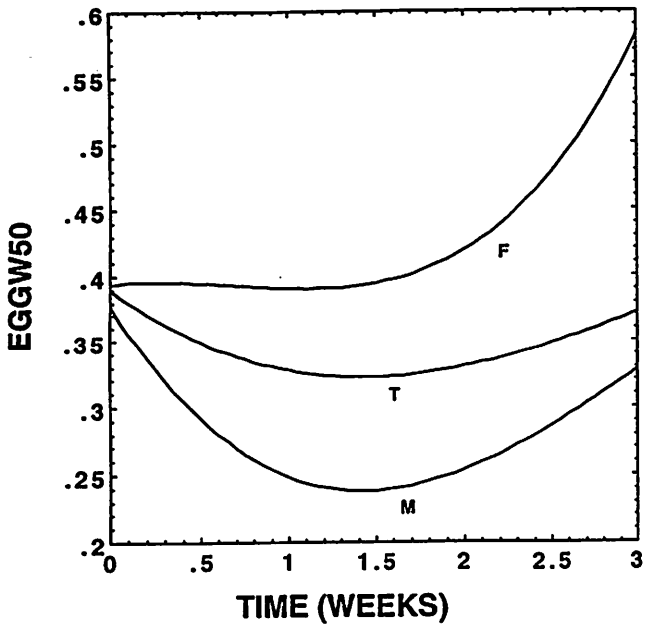


Figure 5. The curves for males (M), females (F), and all subjects combined (T) EGGW50 models during the first 3 weeks after injection.

except for the cubic parameter. Given the small number of subjects and limited number of data points for each of the subjects, reliability coefficients of .5 or greater reflect adequate model representation in terms of averages. Estimated intercorrelations among the parameters indicatedthat some were moderately to strongly related (au=.688 to .985; Appendix B). Analysis of the fixed effect coefficients of the model (the mean parameters computed over all five subjects, Table 3) indicated that all parameters differed significantly from zero with the possible exception of the slope parameter (t=-1.46; p=0.077). The average intercept (B_0) and quadratic (B_2) terms were positive while the average slope (B_1) and cubic (B_3) terms were negative. The cubic equation that incorporates these average parameter coefficients is shown as the curve labeled T in Figure 5. The function for the early post-injection period is characterized by a decline followed by an increase toward the pre-injection level. Table 3 shows that all variance components differ significantly from zero with the possible exception of the cubic term (p=0.066). This indicates that there is significant variation in the parameter estimates which may be attributable to unknown factors.

The slopes as outcomes model determined that gender was a factor significantly related to all of the parameter estimates except for the intercept. Note in Table 2 the similarity among B_0 parameter values across subjects irrespective of gender. The model for males, however, tended to have a steeper negative slope following the injection when compared to that of the females (t=-2.155; $p < 0.05$). The modelfor males has a more positive quadratic term

Table 4.
Parameter Estimates and Reliabilities for SO9050

Subject	Bo	B1	B2	B3
male 1	-2.69769	-.34126	.37681	-.04829
male 2	-3.34392	-1.31703	.74224	-.07497
male 3	-8.83034	-4.13211	2.69876	-.32033
female 1	-2.47435	.17465	-.05114	-.00000
female 2	-1.94529	.03957	-.56821	.34588
reliability	.984	.978	.949	.776

Table 5.
Fixed and Random Effects for SO9050

effect	fixed coefficient	t	Variance component	chi-square
Bo	-3.85769	-3.15*	7.43198	184.93*
B1	-.99434	-1.18	3.49236	168.26*
B2	.68880	1.136	1.25206	74.17*
B3	-.08053	-1.42	.01581	42.95*

* p<.05

(t=2.747; $p < 0.01$) and a more negative cubic term (t=-3.337; $p < 0.01$). Figure 5 highlights a curvature difference between the male (M) and female (F) models. For EGGW50, the male (M) curve predicts a minimum at approximately 1.3 weeks post-injection, whereas the curve for females is relatively flat during the same time period.

SO9050: The results for S09050 showed the greatest reliabilities of the parameter estimates (Table 4), as well as strong parameter intercorrelations (au =.96 to .99, Appendix B). However, only the fixed intercept coefficient (B_0, Table 5) was significantly different from zero (p=.01). Further, each of the fixed coefficients had significant between-subject variation as reflected by their corresponding variance components (Table 5). This variation is not due to gender difference but due to the relatively large parameter estimate values for subject M3 (Table 4).

Figure 6 presents the total (T), male (M), and female (F) model results for SO9050. Although curves appear to be widely separated, the large degree of variability in the parameter estimates, particularly due to subject M3, preclude the interpretation of a statistically significant difference. It is worthwhile to note, however, that the male and femalemodel trends are similar to those of EGGW50.

SC1090: Parameter and reliability estimates for the model representing the SC1090 over time are given in

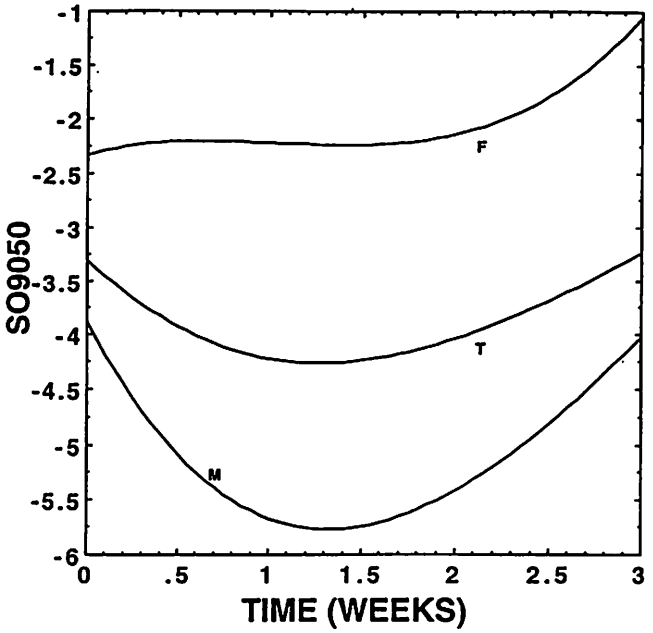


Figure 6. The curves for males (M), females (F), and all subjects combined (T) SO9050 models during the first 3 weeks after injection.

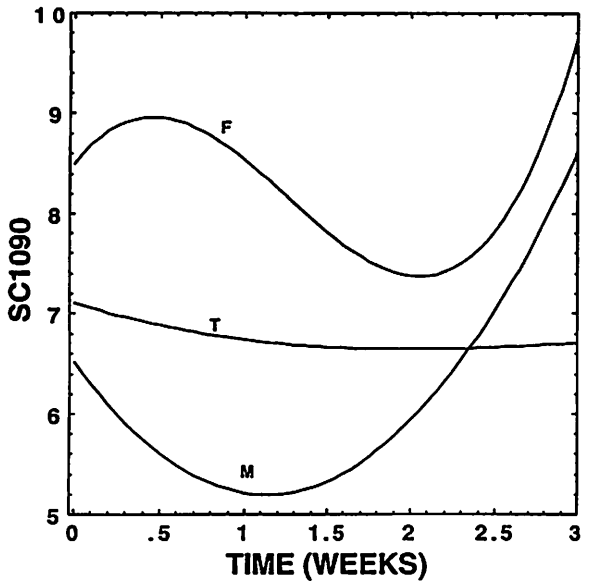


Figure 7. The curves for males (M), females (F), and all subjects combined (T) SC1090 models during the first 3 weeks after injection.

Table 6. All of the reliability coefficients were .49 or above. All parameter estimates also were moderately to strongly correlated (au=.80 to .99; Appendix B). Only the fixed intercept coefficient (B_0) differed significantly from zero (Table 7). The variance components, however, indicated significant subject variation around all of the fixed coefficient estimates, with the possible exception of the cubic term (p=.07).

When the parameter estimates were modeled on gender, there were significant differences between gender

Table 6.
Parameter Estimates and Reliabilities for SC1090

Subject	Bo	B1	B2	B3
male 1	8.02257	-.18238	.52776	-.07595
male 2	7.20959	-1.19082	.66710	-.06018
male 3	1.88052	-2.97655	1.91122	-.22363
female 1	8.34644	.34233	-.54759	.05364
female 2	9.57409	-.13946	-3.34166	1.52204
reliability	.840	.768	.665	.492

Table 7.
Fixed and Random Effects for SC1090

effect	fixed coefficient	t	Variance component	chi-square
Bo	6.908896	5.94*	5.92126	22.27*
B1	-.392238	-.51	2.42074	11.15*
B2	.167757	.34	.98527	14.72*
B3	-.018305	-.37	.00902	8.60**

* p< .05
** p< 10

for the linear term (p<.05), the quadratic term (p<.01), and the cubic term (p<.01). Examination of the predicted curves for males (M) and females (F) in Figure 7 indicates that the males had a relatively steep negative slope immediately following the injection, whereas females had a relatively shallow positive slope. The male curve also predicts a minimum of SC1090 approximately 1 week post-injection. In contrast, the model for females predicts a local maximum SC1090 at approximately 3 days post-injection.

SLQ: The results for SLQ (Table 8) also were similar to those for EGGW50. All reliabilities were above .5 except forthe cubic parameter. Intercorrelations among the fixed parameter coefficients ranged from au =.32 to au =.98 (Appendix B), with relatively weaker relation between the intercept and other parameter coefficients. While only the intercept fixed coefficient differed significantly from zero (p<.05; Table 9), all of the variance components indicated significant variation among subjects (Table 9). Further, the analysis by gender indicated that the linear slope (p < 0.01), the quadratic component (p < 0.010), and the cubic component (p < 0.01) all differed by gender. The male curve, shown in Figure 8, reaches a local maximum at approximately 1.4 weeks. As in the other models for females, the early post-injection period is characterized by a relatively flat curve.

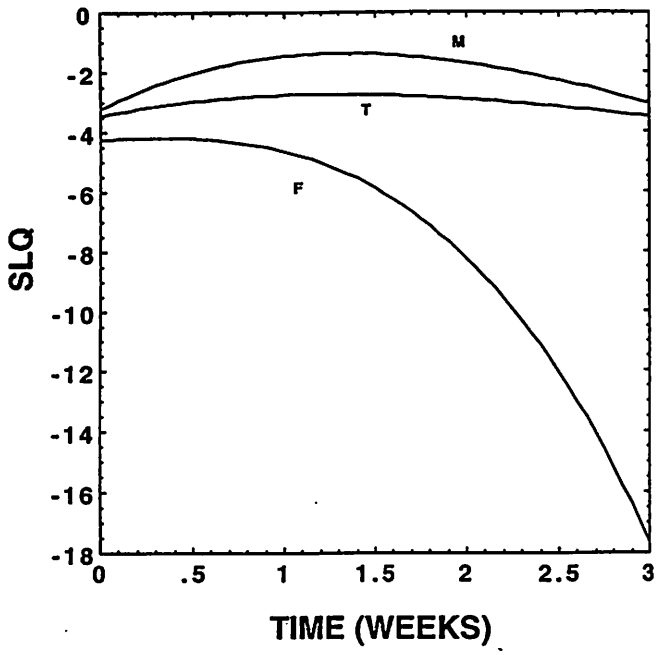


Figure 8. The curves for males (M), females (F), and all subjects combined (T) SLQ models during the first 3 weeks after injection.

Table 8.
Parameter Estimates and Reliabilities for SLQ

Subject	Bo	B1	B2	B3
male 1	-3.27919	.79850	-.96234	12531
male 2	-3.00484	2.89011	-2.44054	.13826
male 3	-.21622	1.99710	-1.36189	.18262
female 1	-3.69757	-.29380	.32205	-.03123
female 2	-4.70840	-.04552	1.33856	-.85685
reliability	.564	.817	.597	.427

Table 9.
Fixed and Random Effects for SLQ

effect	fixed coefficient	t	Variance component	chi-square
Bo	-3.025069	-5.40*	.98565	9.56*
B1	.739719	1.00	2.36627	18.45*
B2	-.515759	-1.44	.49069	16.91*
B3	.059678	1.70**	.00406	12.74*

* p< .05
** p< .10

Relation Between Adductory Change and Symptom Severity

An initial attempt was made to model severity as a function of EGG data over time using the HLM procedure. Both linear and quadratic models were fit within the random coefficients model and then the slopes were tested as functions of gender. In none of the cases did gender predict significant variation in the slopes or curvatures of the relation between EGG measures and symptom severity. There was significant between-subject variation in the parameterestimates for all four of the EGG variables, but the variation was not related to gender. It appeared reasonable, therefore, to evaluate any relation among adduction and symptom severity on an individual patient basis. In order to examine the individual parameter estimates, a General Linear Models (GLM) procedure was used.

Visual inspection of the data suggested minimal, if any, relation between EGG measures and severity ratings for two of the subjects (F1 and M2). The results of the GLM also confirmed that neither cubic, quadratic, nor linear models fit the relation of any EGG measure to severity ratings for these two subjects. For two other subjects, however, results were fairly consistent. Subject M1, for example, showed linear relations between all four predictors and severity, with R^2 values ranging from .49 (SC1090) to .88 (SLQ). Subject M3 had quadratic relations for EGGW50, SO9050, and SLQ to severity, with R^2 values ranging from .98 to 1.0. Subject F2 showed a quadratic relation only for EGGW50 to severity (R^2=.88).

The results for subject M1 indicate positive linear relations for EGGW50, SC1090, and SO9050 with severity and an inverse relation between SLQ and severity. The strongest relation was that of SLQ which predicted 88% of the variance in symptom severity rating. The fact that the relations forM3 and possibly F2 are quadratic indicates that the degree of relation depends on the level. Particularly for subject M2, this might indicate a threshold of adduction above which there is a perceived difference in severity of symptoms and below which there is not.

To address this possibility, an ordered categorical variable was created from severity rating by breaking it at the median value. Severity values greater than 1 were given the class of one; severity values of 0 to 1 were given the class of zero. Using a multivariate GLM, levels of the EGG variables were compared between the severity classes. Because the results might have depended upon subject, a subject factor and subject-by-severity class interaction were included in the models.

The results of this analysis indicated that there were significant differences in EGGW50 between severity classes ($F(4, 25) = 3.63$; $p < .05$) but also a significant severity class by subject interaction ($F(16, 77) = 2.47$; $p < .01$). Examination of the univariate tests for other variables indicated that the interaction was significant for SO9050 ($F(4, 28) = 3.32$; $p = .02$) and approached significance for SC1090 ($F(4, 28) = 2.56$; $p = .06$). Significant differences in severity class

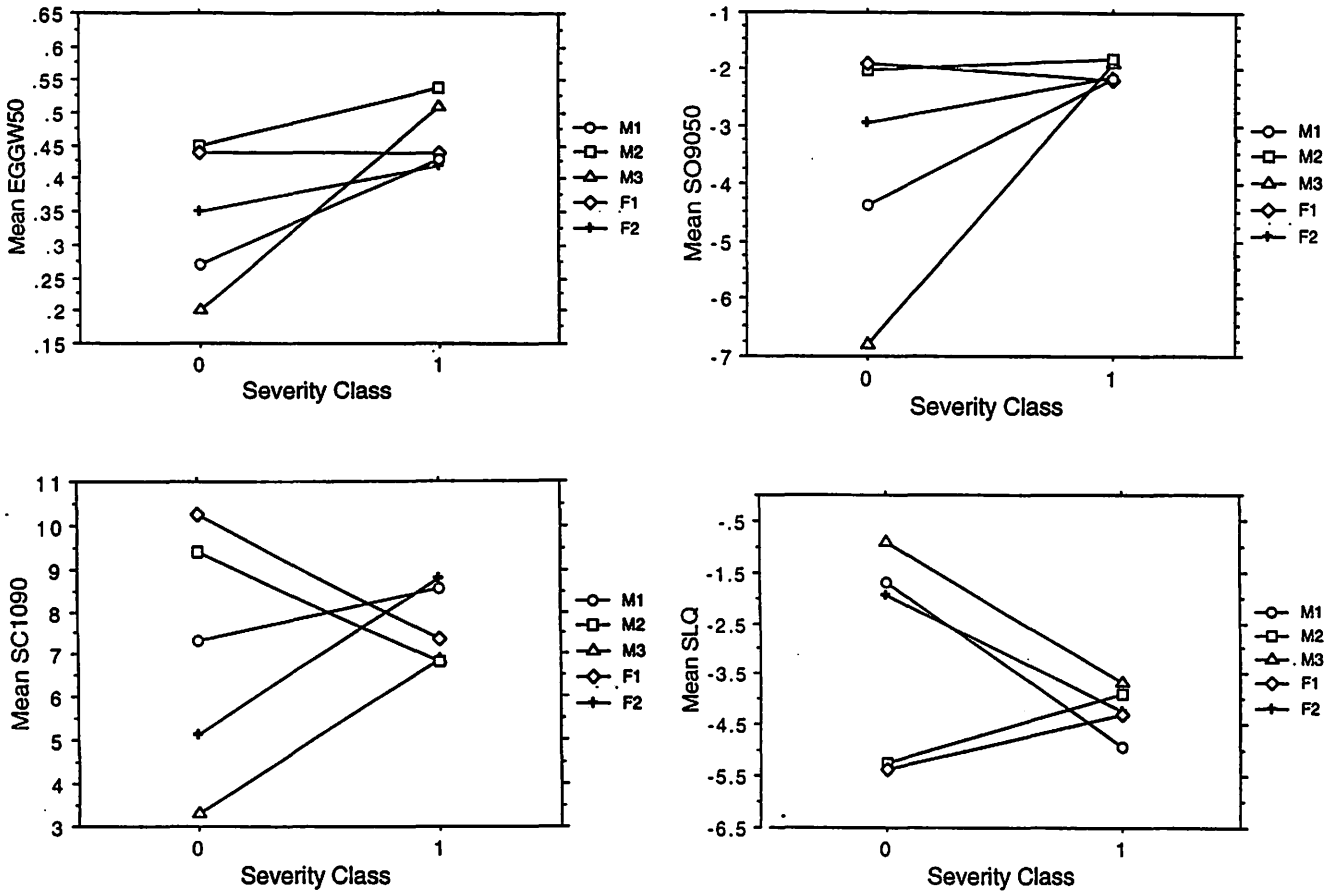


Figure 9 (upper left). Mean EGGW50 for severity class 0 (no impairment or mild impairment) and 1 (moderate to profound impairment) split by individual male (M) and female (F) subjects. Figure 10 (upper right). Mean SO9050 for severity class 0 (no impairment or mild impairment) and 1 (moderate to profound impairment) split by individual male (M) and female (F) subjects. Figure 11 (lower left). Mean SC1090 for severity class 0 (no impairment or mild impairment) and 1 (moderate to profound impairment) split by individual male (M) and female (F) subjects. Figure 12 (lower right). Mean SLQ for severity class 0 (no impairment or mildimpairment) and 1 (moderate to profound impairment) split by individual male (M) and female (F) subjects.

were noted for SC1090 ($F(1,28) = 9.33$; $p < .01$) and SO9050 ($F(1, 28) = 8.04$; $p < .01$).

The mean EGGW50, SC1090, SO9050, and SLQ values for each subject and severity class are shown in Figures 9 through 12, respectively. Examination of the means indicated that for subjects F1 and M2, SC1090 values were higher when their severity values were low, but the result was just the opposite for subjects M1, M3, and F2. The results for SO9050 were somewhat more consistent with larger negative values of SO9050 associated with lower severity ratings for all subjects except F1. In general, only slightly less positive EGGW50 values and more negative SO9050 values were associated with lower severity ratings.

Discussion

Adductory Trends Over Time

The first hypothesis, predicting reduced laryngeal adduction during the first 3 weeks post-injection, followed by a reversal, was supported only by the set of data from male patients. In these data, an early post-injection minimum was consistently observed. Hierarchical Linear Modeling indicated that variance in the pattern of adductory change over time might be attributed to gender. For three male subjects, there was statistically significant negative slope (during the first two weeks post injection), as well assignificant curvature (due to subsequent reversal). This is consistent with clinical observations of temporary post injection hypofunction, with eventual return of spasmodic hyperfunctional symptoms as the paretic effect of Botox® presumably subsides. For the two female subjects, however, the trend during the early post-injection period was relatively flat with a potential maximum. This apparently conflicts with the expected hypofunctional effect of the toxin. On the surface, therefore, a gender difference might seem to account for a binary subgrouping in model variation between subjects during the early post-injection period. Others have postulated gender differences in the duration or magnitude of relief from spasmodic symptoms after Botox® injection, as well as differences in perceived hypofunctional effects (Ludlow et al., 1992; Maloney & Morrison, 1994).

A gender difference in post-injection adduction (if present) might be supported by several explanations. First, it could relate to differences in the ease with which a physician can inject toxin at the intended vocal fold target. The smaller vocal folds of females are possibly a more difficult target than the larger vocal folds of male patients. Thus, the injection procedure may be inherently less accurate for smaller, female larynges. Second, the larger male larynx mayrequire greater force for closure during phonation. It follows that similar toxin doses may result in greater hypofunctional effects for males relative to females. In addition, an equal hypofunctional effect could create a comparatively larger glottal opening for males. Third, females may present lower pre-injection levels of adduction (Bless, Glaze, Lowery, Campos, & Peppard, 1993; Hertegard & Gauffin, 1995) because breathy voice is a behavioral adaptation that might help to minimize spasmodic voice breaks (Blitzer et al., 1988; Izdebski, 1992) during the syllable task, while still maintaining a relatively longer EGG duty cycle that is sometimes present for females (Holmberg et al., 1988; Titze, 1989). Immediately post-injection, the presence of the injected fluid causes the vocal folds to bulge slightly and might contribute disproportionately to over estimation of adduction for female patients. Still, this explanation cannot account for gender differences days after injection. Fourth, it is not known if hormonal change in the female larynx affects calcium equilibrium in thyroarytenoid motor units (Abitbol, de Brux, Millot, Masson, Mimoun, Pau, & Abitbol, 1989), and thereby alters the effects of Botox®. Because the human vocalis muscle shows both estrogen and progesterone receptors (Ferguson, Hudson, & McCarty, 1987), it is possiblethat the vocalis could show variable response under hormonal manipulation.

For several reasons, however, it would be premature to assume a gender-based difference in the patient data. Inspection of the between-subjects HLM suggests some variation in parameter estimates between subjects. Both the initial levels of adduction (intercept parameter), and slope (linear parameter) varied somewhat between subjects of the same gender. Furthermore, the initial level of adduction and slope were moderately to strongly correlated for all of the EGG measures. Thus, it is possible that the greater the level of laryngeal adduction prior to Botox®, injection, the greater the likelihood of strong hypofuctional response (negative slope) during the early post-injection period, irrespective of gender. Also, female patients sometimes report hypofunctional symptoms and can evidence post-injection breathy voice (Blitzer et al., 1988; Ford et al., 1992; Maloney & Morrison, 1994). From these data, therefore, it is not possible to determine whether differences are due to gender or other factors, such as adaptive behavioral compensation, initial levels of adduction, age related changes in vocal fold tissue, or underlying etiology, to name only a few. At the very least, the small number of subjects shouldmandate conservative interpretation of the data. On the basis of these findings, however, it appears likely that between-patient differences in hypofunctional response are not random. Moreover, it seems possible that general subgroups of post-injection trends might exist. In order to determine factors that predict differences in rate change of adduction, frequent measurement of larger subject groups would be necessary.

Individual Differences

Electroglottographic data generally support the second hypothesis that individuals would differ in their adductory variability over time. As previously suggested, parameter estimates, as well as the magnitude of maximum and minimum values varied somewhat among subjects of the same gender. This confirms quantitatively the subjective evidence that individuals differ in adductory response to similar Botox® injection protocols (Ludlow et al., 1992; Maloney & Morrison, 1994; Zwirner et al., 1993). Individual's phonatory trends should be expected to vary for a number of reasons. These might include the potential for individual differences in: (1) the number and distribution of receptor sprouting, (2) theability of poisoned nerves to continue to release minute amounts of transmitter, (3) post-injection sensitivity to cholinergic agonists, (4) behavioral compensations for a histochemically induced vocal fold paresis, and/or (5) the unerdying pathogenesis of ADSD symptoms.

Individual differences also contribute, in part, to the relatively lower reliability of a cubic term within the models. For example, the cubic group model represents a poor fit past three weeks post-injection for one patient (F2). Lower reliability also would be expected for higher-degree parameter estimates because the number of data points (and assessment times) decreased as greater time elapsed after injection. This was due to subject attrition. Additional longitudinal data points over an extended period of time would be helpful to develop models of adductory trends beyond the early post-injection period.

Among the five patients in this study, the individual with minimum post-injection adduction (M3) also suffered from three weeks of intermittent aphonia as well as difficulty swallowing that reportedly lasted more than one week. This was associated with EGGW50 values less than .2 and SO9050 values greater than 6. In 16 of 41 available utterances from this patient (M3), no voicing oscillations could be detected inthe EGG signal, so the measures could not be obtained. By comparison, subjects M1 and M2 reported post-injection breathy voice, but did not evidence the severe hypofunctional symptoms of M3. Voicing oscillations were present in all phonatory attempts for these two patients.

Within-session EGG variability also was not consistent across patients. The most extreme variability within a session was that of SC1090 for patient M2 1.5 weeks after

injection. Relatively large variability of SC1090 also was present for subjects M1 and F1. This may be due to the sensitivity of electroglottography to subtle differences in vertical vocal fold mucosal behavior during closure. Orlikoff (1991) noted previously that for normal subjects, an EGG measure corresponding to rate of contact growth was relatively more variable within a phonation than was a measure of EGG duty cycle.

Relation Between Adduction and ADSD Symptom Severity

Izdebski (1991) speculated that spasmodic voice interruptions are the motor response to faulty or misinterpreted sensory information about lung pressure or vocal fold contact. Specifically, it has been proposed that during oscillation, increased vocal fold contact force orcontact area are mechanisms that activate spasmodic symptoms. Within this framework, it is possible to predict a positive relation between measures of vocal fold contact and ADSD symptom severity. Our third research question (whether trends for change in vocal fold adduction would be related to variation in symptom severity) addressed this hypothesis.

The data generally did not support a strong positive relation between symptom severity and all EGG measures of vocal fold adduction. This pattern occurred uniformly across all EGG measures only for two of the five subjects tested. In contrast, ratings of relatively less severity sometimes occurred when levels of vocal fold contact were relatively high (subjects F1 and M2). This finding would be consistent with the clinical observation that post-injection hypofunction resolves relatively quickly, while ADSD symptoms can be reduced for longer periods of time (Blitzer & Brin, 1992; Brin et al., 1992). Clinicians also report that some patients benefit from injection without noticeable post-injection glottal incompetence (Ford et al., 1992; Jankovic et al., 1990).

It is possible that the ambiguity regarding the relation between glottal competency and symptom severity is due to differences between the tasks during which these variableswere estimated. In this study, glottal competency for voicing was estimated from a non-meaningful syllable string, whereas symptom severity was rated from a variety of speech tasks. Spasmodic symptoms usually are more abundant and severe in connected speech (Bloch, Hirano, & Gould, 1985; Izdebski, 1992). Selection of a syllable task, therefore, was expected to make possible an estimate of glottal competency without the severe, spasmodic changes in adduction that can stop phonation. In contrast, clinical estimates of symptom severity were judged from a wide variety of tasks that included connected speech (Izdebski, 1992; Ludlow et al., 1988, Ludlow et al., 1990).

If symptomatic change is due solely to mechanically reduced glottal competency (and reduced potential for forceful tissue contact), this relation was expected to be evident in spite of task differences that may be cognitively and linguistically based. This is because phonetic skills underlie both syllables and connected speech. Given this assumption, but some findings to the contrary, the results could suggest any of several possibilities. Variation in symptom severity may be due, at least in part, to factors other than changing potential for vocal fold adduction. Since a control group was not followed, it is not possible to ruleout a placebo effect. Furthermore, although Botox® shows relatively strong affinity for presynaptic blockade of acetylcholine release, it has not been demonstrated that ADSD symptom reduction is caused solely by this specific toxin action. Finally, the EGG waveform is a scalar variable that contains relatively greater information about membranous vocal fold closure, but may be less sensitive to compression of the arytenoid vocal processes (Scherer et al., in press). In humans and animals, the arytenoid region is known to show relatively more abundant sensory innervation than the membranous vocal folds(Konig & von Leden, 1961; Larson, Yoshida, & Sessle, 1992; Yoshida, Tanaka, Saito, Shimazaki, & Hirano, 1992). If sensory information is implicated in the pathogenesis of ADSD, then the EGG would be less sensitive to contact and compression in the glottal region from which most abundant sensory input may arise.

Conclusions

Longitudinal electroglottographic measures were used to quantify laryngeal adduction after Botox® injection treatment of five ADSD patients. For three of five patients, the EGG measures revealed a trend for reduced laryngeal adduction during the early post-injection period, followed bya reversal toward greater adduction over time. In addition, individual patients showed idiosyncratic patterns in adductory variability over time. These differences could be reliably assessed and were non-random. Longitudinal application of EGG measures, therefore, could be useful for distinguishing individual patient responses, as well as factors that might predict response for a patient or subgroup of patients.

There are, however, some limitations to the use of EGG measures. First, tremor poses a challenge for the method due to cyclic changes of laryngeal position behind the electrodes. This especially could influence shape sensitive parameters like opening or closing contact slope measures. It also may create an unstable signal to noise ratio throughout the signal. Such influences, when observed in the presence of vocal tremor and laryngeal positioning changes, could not be attributed to adductory changes. Second, because the EGG waveform is a scalar variable, the ability to obtain a relatively noise free EGG signal also relates to the range of difference in vocal fold tissue contact within each glottal cycle. In breathy aphonic patients, minimal tissue contact or lack of tissue contact may result in

minimal signal above the noise floor. In strained, aphonic patients, the constancy oftissue contact also can result in little ac modulation of the signal. It may be possible that, in the case of extreme supraglottic constriction, mechanical coupling of the ventricular folds with the vocal folds contributes to reduced waveform amplitude. Thus, it is important to monitor signal amplitude in order to avoid gross misinterpretation of the waveform.

In the absence of aphonia or significant tremor, EGG measures appear to be one useful method to model oscillatory glottal aspects of injection response. There are, however, other methods which may prove useful or complementary to EGG measures. For example, the glottal volume velocity might serve to distinguish among patients because of the strong dependence of vocal quality and loudness on this variable (Fant, 1979; Gauffin & Sundberg, 1980; Holmberg et al., 1988; Sundberg & Gauffin, 1979). Further, aphonia (resulting in absence of an ac component) does not preclude interpretation of the airflow signal because the remaining dc airflow component can reflect clinically important physiological states of adduction and respiratory effort (Holmberg et al., 1988; Holmberg et al., 1994). Endoscopic measures also may prove useful (Bless et al., 1993; Hertegard & Gauffin, 1995) if a view of the vocal folds is not obscured by spasmodicsupraglottal constriction (Woo, Colton, Casper, & Brewer, 1992; Woodson, Zwirner, Murry, & Swenson, 1992).

It appears possible to model glottal aspects of a laryngeal injection response for individual patients by using electroglottography, an inexpensive, noninvasive methodology. In the future, this quantitative EGG information regarding glottal competency may help clinicians determine dose levels for Botox® that are both efficacious, and devoid (if possible) of significant adverse effects. Given the ability to model patient response curves, it would be profitable to determine factors, or groups of factors, that could account for differences in response among patients.

References

Abitbol, J., de Brux, J., Millot, G., Masson, M. F., Mimoun, O.L., Pau, H., & Abitbol, B. (1989). Does a hormonal vocal cord cycle exist in women? Study of vocal premenstrual syndrome in voice performers by videostroboscopy-glottography and cytology on 38 women. Journal of Voice, 3, 157-162.

Arends, N., Povel, D., Van Os, E., & Speth, L. (1990). Predicting voice quality of deaf speakers on the basis of glottal characteristics. Journal of Speech and Hearing Research, 33, 116-122.

Aronson, A. E. (1985). Clinical voice disorders. An interdisciplinary approach (2nd ed.). New York: Thieme.

Bless, D. M., Glaze, L. E., Lowery, D. B., Campos, G., Peppard, R. C. (1993). Stroboscopic, acoustic, aerodynamic, and perceptual analysis of voice production in normal speaking adults. National Center for Voice and Speech Status and Progress Report, 4, 121-134.

Blitzer, A., & Brin, M. F. (1992). Treatment of spasmodic dysphonia (laryngeal dystonia) with local injections of botulinum toxin. Journal of Voice, 6, 4, 365-369.

Blitzer, A., Brin, M. F., Fahn, S., & Lovelace, R. E. (1988). Localized injections of botulinum toxin for the treatment of focal laryngeal dystonia (spastic dysphonia). Laryngoscope, 98, 193-197.

Bloch, C. S., Hirano, M., & Gould, W. J. (1985). Symptom improvement of spastic dysphonia in response to phonatory tasks. Annals of Otology, Rhinology, and Laryngology, 94, 51-54.

Brennan, R. L. (1983). Elements of generalizability theory. Iowa City, IA: ACT.

Brin, M., Blitzer, A., Fahn, S., Gould, W., & Lovelace, R. (1989). Adductor laryngeal dystonia (spastic dysphonia): Treatment with local injections of botulinum toxin (Botox). Movement Disorders, 4, 287-296.

Brin, M., Blitzer, A., Stewart, C., & Fahn, S. (1992). Treatment of spasmodic dysphonia (laryngeal dystonia) with local injections of botulinum toxin: Review and technical aspects. In Blitzer, A., Brin, M., Sasaki, C., Fahn, S., & Harris, K. (Eds.), Neurologic disorders of the larynx (pp. 214-228). New York: Thieme Medical Publishers.

Bryk, A. S., & Raudenbush, S. W. (1992). Heirarchical linear models: Applications and data analysis methods. Newbury Park, CA: Sage Publications.

Childers, D. G., & Lee, C. K. (1991). Vocal quality factors: Analysis, synthesis and perception. Journal of the Acoustic Society of America, 90, 5, 2394-2410.

Childers, D. G., Smith, A. M., & Moore, G. P. (1984). Relationships between electroglottograph, speech, and vocal cord contact. Folia Phoniatrica, 36, 105-118.

Dejonckere, P. H., & Lebacq, J. (1985). Electroglottography and vocal nodules. Folia Phoniatrica, 37, 195-200.

Dolly, J.O. , Black, J., Williams, R. S., & Melling, J. (1984) . Acceptors for botulinum neurotoxin reside on motor nerve terminals and mediate its internalization. Nature, 307,457-460.

Fant, G. (1979) . Glottal source and excitation analysis. Speech Transmission Laboratory Quarterly Progress and Status Report, (KTH, Stockholm), 1, 85-107.

Ferguson, B. J., Hudson, S. R., & McCarty, K. S. (1987). Sex steroid receptor distribution in the human larynx and laryngeal carcinoma. Archives of Otolaryngology Head Neck Surgery, 113, 1311-1315.

Fisher, K. V., Scherer , R.C., Guo, C.G., & Owen, A. S. (1994). Quantifying phonatory changes after botulinum toxin type a injection. (in review)

Ford, C. N. , Bless, D. M. , & Patel, N. Y. (1992). Botulinum toxin treatment of spasmodic dysphonia : Techniques, indications, efficacy. Journal of Voice, 6, 370-376.

Freeman, F. J. , Cannito , M. P., & Finitzo-Hieber , T. (1985). Classification of spasmodic dysphonia by perceptual-acoustic-visual means. In G. Gates (Ed.) Spasmodic dysphonia: The state of the art, 1984 . New York: The Voice Foundation.

Gauffin, J., & Sundberg, J. (1980) . Spectral correlates of glottal voice source wave form characteristics. Journal of Speech and Hearing Research, 32, 556-65.

Hertegard, S. & Gauffin, J. (1995) . Glottal area and vibratory patterns studied with simultaneous stroboscopy, flow glottography, and electroglottography. Journal of Speech and Hearing Research, 38, 85-100.

Holmberg, E. B., Hillman, R. E., & Perkell, J. S. (1988). Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. Journal of the Acoustical Society of America, 84 (2), 511-529.

Holmberg, E. B., Hillman, R. E., Perkell, J. S., & Gress, C. (1994). Relationships between intra-speaker variation in aerodynamic measures of voice production and variation in SPL across repeated recordings. Journal of Speech and Hearing Research, 3, 484-495.

Izdebski , K . (1992) . Symptomatology of adductor spasmodic dysphonia : A physiologic model . Journal of Voice, 6, 306-319.

Jankovic, J., Schwartz, K., & Donovan , D. T. (1990) . Botulinum toxin treatment of cranial-cervical dystonia, spasmodic dysphonia, other focal dystonias, and hemifacial spasm. Journal of Neurology, Neurosurgery and Psychiatry, 53, 633-639.

Kempster, B., Preston, J., Mack, R., & Larson, C. (1987). A preliminary investigation relating laryngeal muscle activity to changes in EGG waveforms. In T. Baer, C. Sasaki, & K. Harris (Eds.), Laryngeal function in phonation and respiration (pp. 339-364). Boston: Little Brown and Company.

Konig, W. F., & von Leden, H. (1961). The peripheral nervous system of the human larynx. Part 1. The mucous membrane. Archives of Otolaryngology-Head and Neck Surgery, 73, 1-14.

Larson, C. R., Yoshida, Y., & Sessle, B. J. (1993). Higher level motor and sensory organization. In Ingo R. Titze (Ed.), Vocal Fold Physiology: Frontiers in Basic Science. (pp. 227-275). San Diego: Singular Publishing.

Ludlow , C . L . (1990). Treatment of speech and voice disorders with botulinum toxin . Journal of the American MedicalAssociation, 264, 2671-2676.

Ludlow , C . L . , Bagley , J . , Yin , S . G . , and Koda , J . (1992) . A comparison of injection techniques using botulinum toxin injection for treatment of the spasmodic dysphonias. Journal of Voice, 6 , 4 , 380-386 .

Ludlow , C. L., Naunton, R. F., Fujita, M., & Sedory, S . E. (1990) . Spasmodic dysphonia: Botulinum toxin injection after recurrent nerve surgery. Otolaryngology - Head & Neck Surgery , 102, 2, 122-131.

Ludlow, C. L., Naunton, R. F., Sedory, S. E., Schulz, G . M., & Hallett, M . (1988) . Effects of botulinum toxin injections on speech in adductor spasmodic dysphonia. Neurology, 38, 1220-1225.

Maloney , A . P . , & Morrison , M . D. (1994) . A comparison of the efficacy of unilateral versus bilateral botulinum toxin injections in the treatment of adductor spasmodic dysphonia. The Journal of Otolaryngology, 23, 3 , 160-164.

Miller, R. H., Woodson, B. E., & Jankovic, J. (1987).Botulinum toxin injection of the vocal fold for spasmodic dysphonia. A preliminary report. Archives of Otology and Head Neck Surgery. , 113 , 603-605.

Orlikoff, R. F. (1991). Assessment of the dynamics of vocal fold contact from the electroglottogram: Data from normal male subjects. Journal of Speech and Hearing Research, 34, 5, 1066-1072.

Prakash, S., Luschei, E. S., Ramos, C. A., & Bless, D. M. (1994). Botox induced changes in canine laryngeal musculature: Electromyographic and videoendoscopic findings. Journal of the American Speech Language Hearing Association [Abstracts], October, 56.

Rothenberg, M., & Mahshie, J. (1988) Monitoring vocal fold adduction through vocal fold contact area. Journal of Speech and Hearing Research, 31, 38-51.

Scherer, R. C., & Guo, C. G., (1992). "MULTISIG". Monograph, National Center for Voice and Speech, The University of Iowa, Iowa City.

Scherer, R. C., Vail, V. J., & Rockwell, B. (in press). Examination of the laryngeal adduction measure EGGW. In F. Bell-Berti & L. J. Rapheal (Eds.), Producing Speech: A Festschrift for Katherine Safford Harris. American Institute of Physics.

Sellin, L. C., & Thesleff, S. (1981). Pre- and post-synaptic actions of botulinum toxin at the rat neuromuscular junction. Journal of Physiology, 317, 487-495.

Simpson, L. L. (1981). The origin, structure, and pharmacological activity of botulinum toxin. Pharmacological Reviews, 33, 3, 155-188.

Simpson, L. L. (1992). Clinically relevant aspects of the mechanism of action of botulinum neurotoxin. Journal of Voice, 6, 4, 358-364.

Sundberg, J. & Gauffin, J. (1979). Waveform and spectrum of the glottal voice source. In B. Lindblom and S. Ohman (Eds.), Frontiers of speech communication research. Festschrift for Gunnar Fant (pp. 301-320). London: Academic Press.

Thesleff, S., Molgo, J., & Lundh, H. (1983). Botulinum toxin and 4-aminoquinoline induce a similar abnormal type of spontaneous quantal transmitter release at the rat neuromuscular junction. Brain Research, 264, 89-97.

Titze, I. R. (1989). A four-parameter model of the glottis and vocal fold contact area. Speech Communication, 8, 191-201.

Titze, I. R. (1990). Interpretation of the electroglottographic signal. Journal of Voice, 4, 1-9.

Truong, D. D., Rontal, M., Rolnick, M., Aronson, A. E., & Mistura, K. (1991). Double-blind controlled study of botulinum toxin in adductor spasmodic dysphonia. Laryngoscope, 101, 630-634.

Tse, C. K., Dolly, J. O., Hambleton, P., Wray, D., & Melling, J. (1982). Preparation and characterization of homogeneous neurotoxin type a from clostridium botulinum. European Journal of Biochemistry, 122, 493-500.
Valtorta, F., & Arslan, G. (1993). The pharmacology of botulinum toxin. Pharmacological Research, 27, 1, 33-44.

Woo, P., Colton, R., Casper, J., & Brewer, D. (1992). Analysis of spasmodic dysphonia by aerodynamic and laryngostroboscopic measurements. Journal of Voice, 6, 4, 344-351.

Woodson, G. E., Zwirner, P., Murry, T., & Swenson, M. R. (1992). Functional assessment of patients with spasmodic dysphonia. Journal of Voice, 6, 4, 338-343.

Yoshida, Y., Tanaka, Y., Saito, R., Shimazaki, T., & Hirano, M. (1992). Peripheral nervous system in the larynx: An anatomical study of the motor, sensory and autonomic nerve fibers. Folia Phoniatrica, 44, 194-219.

Zwirner, P., Murry, T., Swenson, M., & Woodson, G. E. (1991). Acoustic changes in spasmodic dysphonia after botulinum toxin injection. Journal of Voice, 5, 1, 78-84.

Zwirner, P., Murry, T., & Woodson, G. E. (1993). Perceptual-acoustic relationships in spasmodic dysphonia. Journal of Voice, 7, 2, 165-171.

Acknowledgment

Contract number HN2-031 from the Oklahoma Center for the Advancement of Science and Technology supported equipment, patient recordings, and data analysis. Grant P60 DC00976 from the National Institute on Deafness and Other Communication Disorders partially supported software development.

Appendix A
Speech Tasks

On each recording day, a variety of speech tasks were recorded (with a 6 inch mouth-to-microphone distance) onto the audio portion of SVHS video tape . These recordings were used to obtain the clinical ratings of severity. First, the clinician engaged the patient in brief conversation about the weather and the patient's travel to the center. The patient then was asked to perform each of the following:

1) State their full name and address.
2) Describe how their voice feels. .
3) Describe how their voice sounds.
4) Read the rainbow passage including the title.
5) Sustain the vowel /a/ as long as possible on one breath.
6) Sustain the vowel /a/ as steadily as possible.
7) Read each of the following sentences two times: "Heat lunch every day. Eat lunch every day. We mow our lawn all year. I ambled along the boulevard."
8) Please say /ʔihiʔ ihiʔ ihiʔ ihi/ as quickly as you can.

Appendix B
Correlations Among Parameter Estimates From HLM Analysis

EGGW50
Tau (as correlations)

	β_0	β_1	β_2
β_0	1.000		
β_1	0.856	1.000	
β_2	-0.771	-0.985	1.000
β_3	0.688	0.954	-0.991

SC1090
Tau (as correlations)

	β_0	β_1	β_2
β_0	1.000		
β_1	0.925	1.000	
β_2	-0.824	-0.973	1.000
β_3	0.800	0.960	-0.998

SO9050
Tau (as correlations)

	β_0	β_1	β_2
β_0	1.000		
β_1	0.981	1.000	
β_2	-0.989	-0.998	1.000
β_3	0.994	0.960	-0.999

SLQ
Tau (as correlations)

	β_0	β_1	β_2
β_0	1.000		
β_1	0.730	1.000	
β_2	-0.478	-0.947	1.000
β_3	0.318	0.875	-0.984

NCVS Status and Progress Report - 8
July 1995, 113-130

Formant Trajectory and Segmental Characteristics of Males With Parkinson Disease

Antonia Johnson, B.S.
Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts
Lorraine Olson Ramig, Ph.D.
Wilbur James Gould Voice Research Center, The Denver Center For The Performing Arts
Department of Communication Disorders and Speech Science, The University of Colorado

Abstract

This study describes word segment duration and second formant trajectories produced by males with the neurological disorder Parkinson disease (PD). Formant trajectories of 8 males with PD and 3 nondisordered geriatric males were obtained for 11 words, selected from the speech intelligibility task developed by Kent et al. [J. Speech. Hear. Disord. **54**, 482-499 (1989)]. PD data were compared to those of nondisordered geriatric subjects studied by Weismer et al. [J. Acoust. Soc. Am. **91**, 1085-1098 (1992)]. Speakers with PD produced statistically significantly: (1) longer vowel durations (VD) for 2 words with high monophthongs, (2) shorter transition durations for 3 words with front monophthongs preceded by a liquid or semi-vowel and a front diphthong (3) decreased transition extent for one word and increased transition extent for one word, (4) greater second formant transition rate for 5 words, all front vowels, (5) greater F2 starting frequency for the transition segment for 3 words and (6) greater frequency extents than transition extents for all words collectively and 4 words individually when compared with a nondisordered geriatric control group. PD subjects produced significant increases in vowel and whole word duration for word repetition from day 1 to day 2 for all words collectively. The findings are interpreted in relation to gesture patterns and underlying speech production deficits related to Parkinson disease, including movement duration and initiation, and target undershoot.

Introduction

In recent years there has been increased use of acoustic analysis to explain the perceptual characteristics of dysarthric speech and identify aspects of the disordered speech signal closely related to breakdowns in speech intelligibility (e.g. Weismer et al., 1988, 1992; Ansel & Kent, 1992). Acoustic deviations from given targets have allowed inferences about disturbed motor patterns or articulatory compensations (e.g. Kent & Rosenbek, 1982). This research has generated comparative data on dysarthric motor speech production and contributed to theories of motor speech control (e.g. Ansel & Kent, 1992; Kent & Rosenbek, 1982; Liss et al., 1990).

For dysarthrias in general, and Parkinson disease (PD) dysarthria in particular, much acoustic analysis has focused on a small sample of phonetic segments within a limited variety of sentence productions. For example, articulatory acoustic analysis of PD speech in sentences has indicated shorter second formant transition durations (TD) and transition extents (TE) for /aI/ in "Buy Bobby a poppy" and /Il/ in "Build a big building", in PD subjects with severely disordered speech compared to mildly disordered speech (Forrest et al., 1989) and greater spirantization of stop consonants /p/, /t/, /k/ in four selected sentences (Weismer, 1984). Acoustic measurements of the second formant support the notion of decreased duration of lingual or overall vocal tract movements (TD) and reduced magnitude of vocal tract change (TE) in PD.

From kinematic investigations has come evidence of similar problems in range of movement for lips in PD speech production. These studies also have reported lip movement deficits including decreased speed of movements and reduced initiation (hypokinesia or bradykinesia) associated with rigidity (Leanderson et al., 1972; Yanagisawa et al., 1989) as well as problems of lip acceleration and weakness in the neuromuscular control signals (Netsell et al., 1975). Such disturbed movement patterns could also be

anticipated in lingual or overall vocal tract movement in PD because of research by Munhall, Ostry & Parush (1985) which indicated a uniform basis for the temporal coordination of speech articulators. While the articulatory acoustic data reported by Forrest et al. (1989) support notions of reduced range of vocal tract movement, a more substantial data base involving various types of vocalic trajectories is needed to determine the prevalence of these and other disturbed motor patterns in PD speech production.

A more substantial data base involving a wide variety of F2 trajectory patterns (e.g., gradually rising, sharply falling, and so on) has previously been acquired for the dysarthria associated with amyotrophic lateral sclerosis (ALS) (Kent et al., 1989b, 1992; Weismer et al., 1992). This work has described acoustic characteristics from a standard list of 12 test words derived from a specially designed phonetic intelligibility test (see Kent et al., 1989a). Analysis of the vocalic nuclei of these 12 test words found abnormal segmental and trajectory characteristics across a variety of phonetic environments in ALS compared to nondisordered geriatrics (Weismer et al., 1992). In addition, the slope of the second formant transition for a subset of these words has been found to be fairly well correlated with overall speech intelligibility scores in a study of ALS men (Kent et al., 1989b). It has been proposed that the slope of formant transitions may be a "general index of the severity of speech mechanism involvement and therefore predicts global indices of severity of speech mechanism dysfunction" (Weismer et al., 1992, pp. 1096). Support for this view requires studies beyond ALS to patients having differing types of dysarthria.

While acoustic characteristics for these 12 test words have been obtained for the ALS population (Weismer et al., 1988, 1992), which has significant *lingual* dysfunctioning (and for which vocalic nuclei are perceived as severely distorted, Darley et al., 1975), there has been no comparable study for a population with primary dysfunction of the *larynx. Laryngeal* dysfunction plays a primary role in the dysarthria associated with PD (Hansen et al., 1984; Smith et al., in press; Perez et al., in press). Previous research has documented the role of laryngeal articulation and interarticulator timing with supraglottal articulators for consonants, consonant clusters, pitch control, for morpheme boundary marking, particularly the "glottal stop" (Lindqvist, 1972, Abramson, 1977) and for emphatic stress and in obstruents (Benguerel et al., 1978; Lofqvist & Yoshioka, 1981). Consistent with these studies and the kinematic research (for oral articulator movement) are perceptual studies which have characterized PD dysarthria as having reduced loudness, monotone, disordered rate, and imprecise articulation (Darley et al., 1969, a.b.; Logemann et al., 1978).

The present investigation was initiated to generate comparable articulatory acoustic data for these test words for PD dysarthria. This study extends the acoustic examination to laryngeal-oral interarticulator coordination. The segmental and trajectory data are used to describe differences of acoustic patterns and inferences of motor patterns of vocal tract configurations and movement. Furthermore, because performance variability and a learning effect over repeated productions has been documented in PD voice measures (King et al., 1994), this study also examines day to day acoustic variability in word speech production for subjects with PD dysarthria.

These PD data can be used for three purposes: 1) as raw material for acoustic explanation of intelligibility deficit similarities and differences for dysarthrias of varying etiologies (e.g. PD compared to ALS, brain stem CVA, or MS), 2) as a basis for predictions of acoustic deviations for other speech disordered populations with primary *laryngeal* dysfunction and secondary lingual and labial dysfunction, and 3) as raw material for determining the effects of differing voice and speech treatment programs on acoustic characteristics of PD speech production.

Method

Subjects

Subjects were 8 male patients with Parkinson disease (PD), 7 diagnosed with idiopathic Parkinson disease and one diagnosed with Parkinson disease plus additional neurological involvement. The ages in the PD group ranged from 49 to 76 years. Each PD subject underwent neurological, psychological, and otolaryngological exams prior to his participation in the study. Characteristics of these subjects, including age, length of time since diagnosis, medications, disease stage according to the Hoehn and Yahr Scale (1967), and dysarthria severity rating are summarized in Table 1.

Also participating in this study were three nondisordered control male subjects, aged 62, 67 and 68. They had no past nor current medical problems attributable to neurological disease and exhibited no problems with their speech or voice. Control subject data were obtained to evaluate the similarity of our acoustic measurement to those obtained by Weismer et al., 1992 for a group of 15 nondisordered geriatrics.

Speech Sample

The speech stimuli consisted of 11 monosyllabic words of the form CV, CVC, CCVC, CVCC, CCVCC, and VC, sampled from a larger phonetic intelligibility word task (see Kent et al., 1989a). The test words were *wax, sigh, sip, ship, sew, row, cash, hail, ate, shoot,* and *blend.* Nondisordered geriatric and ALS trajectory data for these words have been reported previously (Weismer et al., 1992). These words contain a variety of F2 trajectories but represent no attempt to sample the full range of vocalic nucleus types in American English.

Table 1.
Descriptive Characteristics of Subjects With Parkinson Disease.

Subject	Sex	Age	Years since diagnosis	Hoehn & Yahr Stage (1967) (a)	Medication	Speech/Voice deficits (severity rating) (b,c)
HW	M	72	5	3	sinemet eldyprel artane	reduced loudness monotone (mild-moderate)
HT	M	76	1	2	sinemet	hoarseness monotone (mild-moderate)
DP	M	49	1	2	sinemet sinemet-cr eldyprel	monotone reduced loudness (mild)
BK	M	71	4	3	sinemet amitriptyline	reduced loudness monotone (moderate-severe)
TB	M	70	4	3	sinemet eldyprel	reduced loudness monotone (moderate)
JL	M	69	1	2	amantadine	monotone reduced loudness (moderate)
LR	M	74	12	2	sinemet sinemet-cr symmetrel eldyprel cogentin	reduced loudness imprecise articulation (moderate)
LHS	M	67	8	4	sinemet florinef artane	reduced loudness imprecise articulation (severe)

(a) The Hoehn and Yahr (1967) Stages range from 1-5; 5 is most severe.
(b) Severity ratings of speech and voice deficits were obtained by consensus of 2 SLP clinicians on scale of 1-5. 1=mild, 2=mild-moderate, 3=moderate, 4=moderate-severe, 5=severe.
(c) Ramig et al., in press.

Procedures

PD subjects were asked to read each word "as clearly as possible" from a series presented individually by videotape at approximately five-second intervals. Nondisordered subjects were asked to read from printed cards displayed one at a time at five-second intervals (Weismer et al., 1992). This controlled timing was designed to minimize the speaking rate effect of words uttered per minute on temporal segment duration (e.g vowel duration, Lindblom, 1963; Browman & Goldstein, 1990). Production of each word was obtained on two separate days within the same week for the PD subjects. Recordings were scheduled to occur at the same time post-medication on each recording date to control for any medication effect on speech production.

Data were collected in the IAC sound-treated booth of the Acoustics Laboratory of The Wilbur James Gould Voice Research Center of The Denver Center For The Performing Arts or a quiet room. Microphone recordings were made of the word samples produced. Each subject was seated in a medical examination chair, and a head mounted microphone (AKG, C451EB at the 75 Hz setting) was placed eight centimeters from the mouth. Data were recorded on a digital audio tape recorder (Sony PC-108M, Denon DTR-80P).

Data analysis

After digitizing the words at 20 KHz on a 486 computer, measures of acoustic variables were made using wide-band spectrographic (300 Hz) and waveform displays generated by CSpeech 4.0. Given an adjustable spectrogram intensity level to a constant 45 dB range, the floor was set at -68 dB, just above the level of background noise to ensure that the low intensity portions of the speech signal were included.

Acoustic analysis of the vocalic nucleus included segmental measurement. Vocalic temporal measurement

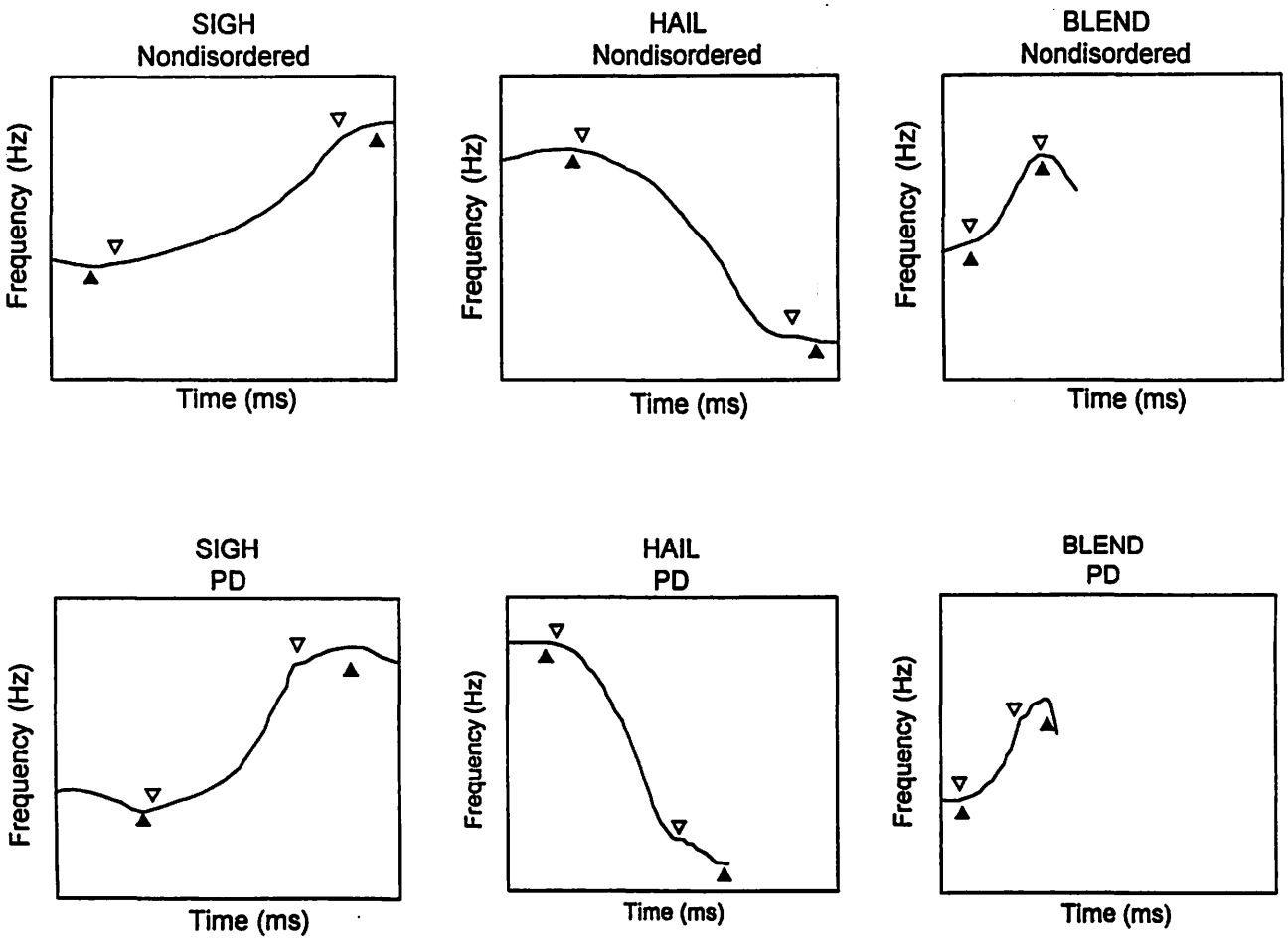


Figure 1. Schematic display of F2 trajectory for sigh, hail, *and* blend. *Trajectories for nondisordered geriatrics are shown at the top and for PD speakers at the bottom. Open arrows (▽) indicate boundaries of the transitional segment for which transition duration, transition extent, transition rate, and starting frequency of F2 were derived. Closed arrows (▲) indicate boundaries of frequency extent.*

followed conventional criteria (Peterson & Lehiste, 1960; Klatt, 1975). Vocalic duration (VD) was measured from first to last large-amplitude complex shape glottal pulse and spectrographic vertical striations in the first and second formants (Weismer et al., 1988, 1992) and was supplemented by audio playback to differentiate vowels from liquids or semivowels.

Acoustic analysis also included measurement of formant transition characteristics. The transitional segment of the vocalic nucleus is associated with the vocalic articulatory gesture demanding the most extensive and rapid change in vocal tract configuration and is reflected most notably in the trajectory of the second formant frequency (F2) (Liss & Weismer, 1992). Trajectory refers to the time-frequency path of the center of the formant band across the entire duration of the vocalic nucleus. Identification of transitional segments of the F2 trajectory used the rule developed by Weismer et al., 1988: onset is the first instant of at least 20 Hz frequency change within 20 ms, and offset is the succeeding instant when 20 Hz/20 ms change is no longer present. When more than one interval met the criterion, conventional criteria of measuring the steepest slope corresponding to that obtained for nondisordered geriatrics was used (Weismer et al., 1992).

Transition measurements for F2 used conventional rules (Weismer et al., 1988, 1992). Transition extent (TE) is the amount of frequency change along the transition segment of a trajectory. Transition duration (TD) is the duration of the transitional segment of the trajectory. Transition rate (TR) or slope is defined as TE/TD and is positive or negative depending on the direction of F2 movement. Starting frequency (SF) is the onset frequency of the transitional segment of the trajectory.

The measurement of frequency extent (FE), defined as amount of frequency change along the trajectory (see Fig. 1), was measured because our previous acoustic analysis revealed multiple instances of a "stairstepping" type of trajectory in which a portion of the trajectory met the 20 Hz/20 ms criterion but was followed by segments with inconsistent or decreased velocity (Johnson, 1993). It was reasoned that it was important to evaluate whether the operationalized definition of TE was reflecting only a portion of the transitional segment, which in turn would foster a faulty interpretation of decreased range of lingual or vocal tract movement. Measurement differences between TE and FE for the PD subjects were compared to the three nondisordered subjects.

Acoustic measurements were obtained for whole word duration since one of our goals is ultimately to link acoustic word data with global indices of speech intelligibility. We chose whole word duration (including pre- and post-word aspiration) as a manageable acoustic measure that would *contribute* to a reasonable overall description of respiratory-laryngeal-oral articulator coordination in single word production. Similarly, ratio of vowel duration to whole word duration (VD/WWD) was used as a *rough* quantitative index for laryngeal/oral interarticulator timing. Previous spectrographic observation had indicated that a distinguishing feature of word production in PD seemed to be multiple cases of prolonged pre- and post-word aspiration and increased fricative duration and rise time (Johnson, 1993; Dromey et al., in press). This was seen as a reflection of inefficient coordination of laryngeal valving gestures and oral closing and opening gestures in contrast to descriptions of nondisordered speech (Lofqvist & Yoshioka, 1981).[1] Whole word duration measures were obtained for the group of 3 nondisordered geriatrics for comparison purposes.

Statistics

Quantitative analysis of selected word and formant trajectory characteristics included computation of group means and standard deviations for segmental and trajectory measures. Statistical analyses by two-tailed independent samples t-tests were carried out to compare articulatory acoustic data from the control group of three nondisordered geriatrics with data from the Weismer et al. (1992) group of 15 nondisordered geriatrics. Because of the small number of subjects in our nondisordered control group, we also compared distributional characteristics of the data of both nondisordered groups. The intent of these comparisons was to assess the legitimacy of using the larger data set as a basis for comparison with experimental subjects. Comparison of the two groups of nondisordered adults was deemed necessary because, although operationalized definitions were the same for the acoustic measures, the comparability of data between our experimental group and data from the Weismer et al. group of 15 nondisordered geriatrics might nevertheless be compromised by unrecognized measurement differences.

Since measures of whole word duration and frequency extent were not obtained by Weismer et al., 1992, comparison of these measures by the PD subjects was made in relation to the group of 3 nondisordered geriatrics. Independent samples t-tests and inspection of distributional data were used to compare differences on whole word duration and frequency extent measures between groups. To compare frequency extent to transition extent for individual words and for all words collectively within each group (PD and NG), paired samples t-tests were used.

To investigate the stability of measure differences over time, paired samples t- tests were used to compare day 1 and day 2 word productions by the PD group.

[1] Multiple acoustic cues of the speech signal are associated with management of the air-flow and interarticulator timing between *laryngeal* and supraglottal articulators. If feasible, it would have been desirable to measure a variety of these (e.g. pre- and post- word aspiration duration, voice onset time (VOT), and fricative duration and rise time).

This study will report PD segmental and trajectory measures which were statistically significantly different from nondisordered geriatric data for both day 1 and day 2 productions. Statistically significant differences for *both* day 1 and day 2 are reported first since together they take account of the greater speech production variability typical of disordered speech (King et al., 1994).

Results

Reliability

All measurements were performed by the first author. In assessing intrajudge reliability, remeasurement of 20% of the waveforms and spectrograms from each word yielded an average measurement difference for vowel segment durations of 3.7 ms (range 0-10ms). The measurement difference tended to be greater for the vowel durations of diphthongs and vowels flanked by the liquid /l/ (ranging from 4-10 ms.), and lower for vowel durations of short duration, closed monophthongs (0 ms)[2]. Remeasurement for whole word durations yielded a measurement difference of 20 ms or less. The different values do not exceed the differences of PD/NG differences reported here. Repeated measurement or calculation of 20% of the trajectory data for TD, TE, TR, FE, VD/WWD and starting frequencies of the transitional segments yielded virtually no differences due to the specificity of the operationalized definitions.

Data summarizing various quantitative aspects of temporal measures and trajectory measures appear in Tables 2 and 3.

[2] Categories of Shriberg and Kent (1982) and from structural linguistics (Bross, 1992) for monophthong and diphthong are used in order to identify commonalities in temporal and trajectory characteristics of vowels produced with a gradually changing articulation (diphthong) compared to relatively static vocal tract configuration (monophthong). One diphthong /aI/ (*sigh*) is phonemic and cannot reduced to a monophthong. The other diphthongs /eI/ (*ate*) and /oU/ (*sew*) are nonphonemic and can be reduced to a monophthong. According to Shriberg and Kent (1982), the diphthongal forms /eI/ and /oU/ occur most commonly in heavily stressed syllables, whereas the monophthongal forms /e/ and /o/ are usually found in weakly stressed syllables. The classification of nonphonemic diphthong is used here because all test words had one syllable.

Table 2.

Summary statistics for second formant frequency of each test word, with the exception of column SF F1, which is starting frequency of transition segment of the first formant. Data reported are for the two subject groups (PD and NG)and for the NG group reported by Weismer et al., 1992. Data reported for the PD group are for Day 1 and Day 2 and for the two NG groups for one day only. VD (ms) = duration of vocalic nucleus, TD (ms) = transition duration, TE (Hz)=transition extent, TR-transition rate, SF F2=starting frequency of second formant. Data reported are group means and standard deviations (in brackets). *= Statistically significant differences ($p<.05$) of PD data from NG-W data. The number of words contributing to the summary statistics is given in parentheses in the column labeled "Group."

Word	Group	VD (ms)	TD (ms)	TE (Hz)	TR (Hz/ms)	SF F2 (Hz)	SF F1 (Hz)
hail	PD (8) Day 1	314 [86]	136 [38]*	1051 [228]	-7.99 [1.69]*	2081 [112]	475 [49]
	PD (8) Day 2	343 [127]	126 [40]*	949 [399]	-7.13 [2.14]*	2071 [82]	467 [89]
	NG (3)	304 [59]	173 [49]	1008 [284]	-5.81 [0.17]	1913 [151]	428 [21]
	NG-W (15)	344 [44]	215 [45]	1073 [198]	-5.12 [1.19]	1984 [168]	428 [87]
ate	PD (8) Day 1	195 [35]	86 [50]	262 [90]	3.63 [1.84]*	1995 [138]	438 [63]
	PD (8) Day 2	217 [81]	86 [31]	244 [99]	3.00 [0.96]*	1986 [69]	495 [55]*
	NG (3)	196 [41]	127 [09]	252 [43]	1.98 [0.26]	1983 [126]	372 [28]
	NG-W (14)	178 [26]	114 [33]	214 [66]	1.86 [0.38]	1948 [206]	408 [82]
sigh	PD (8) Day 1	400 [93]	133 [84]*	475 [382]	3.12 [0.97]	1205 [170]*	661 [152]
	PD (8) Day 2	438 [63]	174 [93]	576 [330]	3.25 [0.55]*	1227 [180]*	682 [118]*
	NG (3)	411 [70]	303 [10]	650 [255]	2.05 [0.50]	1161 [33]	645 [34]
	NG-W (15)	400 [60]	236 [56]	577 [153]	2.44 [0.34]	1391 [142]	578 [78]
sip	PD (8) Day 1	120 [54]*	54 [27]	218 [113]	-4.47 [2.19]	1725 [81]	327 [48]*
	PD (8) Day 2	121 [42]*	54 [13]	319 [151]*	-5.78 [2.27]*	1722 [89]	368 [89]
	NG (3)	101 [39]	77 [05]	268 [32]	-3.87 [0.93]	1477 [61]	340 [66]
	NG-W (15)	72 [15]	51 [17]	177 [68]	-3.46 [1.30]	1630 [148]	374 [45]

Table 2, continued

Word	Group	VD (ms)		TD (ms)		TE (Hz)		TR (Hz/ms)		SF F2 (Hz)		SF F1 (Hz)	
ship	PD (8) Day 1	122	[35]*	93	[34]	578	[202]*	-6.34	[0.97]*	1875	[122]	304	[50]
	PD (8) Day 2	133	[60]*	105	[35]*	571	[182]*	-5.61	[1.38]*	1953	[100]*	307	[48]
	NG (3)	94	[32]	93	[29]	424	[89]	-4.63	[1.46]	1633	[103]	306	[63]
	NG-W (15)	81	[16]	78	[15]	322	[135]	-4.07	[1.33]	1771	[166]	350	[52]
blend	PD (8) Day 1	192	[38]*	86	[33]*	463	[234]	5.10	[1.32]*	1340	[185]*	541	[74]*
	PD (8) Day 2	214	[43]	75	[24]*	430	[248]	5.52	[2.33]*	1334	[149]*	474	[103]
	NG (3)	239	[70]	123	[10]	612	[51]	5.05	[0.87]	945	[31]	445	[56]
	NG-W (15)	240	[37]	150	[27]	601	[152]	4.04	[0.84]	1023	[110]	426	[129]
sew	PD (8) Day 1	324	[62]	186	[61]	477	[146]	-2.64	[0.54]	1306	[124]	507	[67]
	PD (8) Day 2	351	[97]	155	[59]	516	[112]*	-3.72	[1.55]*	1354	[100]*	493	[22]*
	NG (3)	318	[25]	207	[50]	433	[114]	-2.10	[0.25]	1218	[42]	415	[60]
	NG-W (15)	295	[62]	149	[47]	357	[115]	-2.42	[0.48]	1221	[90]	423	[52]
row	PD (8) Day 1	346	[98]	151	[76]	317	[54]*	-3.12	[2.14]*	1131	[69]	428	[62]*
	PD (8) Day 2	379	[103]	124	[51]	274	[125]	-1.95	[0.73]	1163	[106]	440	[63]*
	NG (3)	341	[58]	137	[15]	148	[191]	-1.72	[0.20]	1006	[129]	391	[69]
	NG-W (15)	426	[94]	136	[56]	225	[102]	-1.67	[0.41]	1095	[99]	300	[92]
shoot	PD (8) Day 1	167	[51]*	96	[36]	449	[206]	-4.62	[2.15]	1826	[81]*	289	[52]
	PD (8) Day 2	169	[43]*	111	[18]	551	[149]*	-5.00	[1.31]	1865	[83]*	291	[54]*
	NG (3)	140	[42]	123	[05]	492	[125]	-4.98	[1.22]	1692	[156]	334	[28]
	NG-W (15)	132	[20]	95	[17]	412	[119]	-4.43	[1.40]	1593	[132]	370	[51]
wax	PD (8) Day 1	193	[35]	128	[49]*	651	[334]*	4.71	[1.61]	916	[211]*	484	[107]
	PD (8) Day 2	218	[76]	142	[51]*	767	[330]*	5.86	[2.53]	844	[162]*	491	[118]*
	NG (3)	210	[41]	153	[20]	825	[98]	5.56	[1.23]	665	[26]	394	[18]
	NG-W (15)	274	[49]	213	[47]	1043	[247]	4.93	[0.82]	679	[122]	388	[102]
cash	PD (8) Day 1	233	[56]	43	[22]	137	[115]	-2.75	[1.30]*	1943	[162]	512	[117]
	PD (8) Day 2	259	[34]	56	[23]	133	[63]	-3.30	[1.91]*	1917	[122]	461	[133]
	NG (3)	280	[76]	134	[100]	211	[100]	-1.80	[0.44]	1737	[108]	500	[79]
	NG-W (8)	244	[31]	77	[61]	137	[153]	-1.47	[0.05]	1910	[200]	534	[61]

Table 3.

Summary statistics for selected segmental and trajectory measurements of each test word, for the two subject groups. Data reported for the PD group are for Day 1 and Day 2 and for the NG group for one day only. WWD=whole word duration, VD/WWD=ratio of vowel duration to whole word duration, FE=frequency extent of second formant. Data reported are group means and standard deviations (in brackets). *=statistically significant differences ($p<.05$) of PD data from NG data. The number of words contributing to the summary statistics is given in parentheses in the column labeled "Group."

Word	Group	WWD (ms)	VD/WW (ms/ms)	FE (Hz)
hail	PD (8) Day 1	726 [284]	0.46 [0.10]	1148 [190]
	PD (8) Day 2	677 [119]	0.50 [0.11]	1087 [332]
	NG (3)	626 [195]	0.60 [0.05]	1018 [325]
ate	PD (8) Day 1	552 [110]	0.36 [0.07]	313 [87]
	PD (8) Day 2	579 [127]	0.38 [0.07]	284 [106]
	NG (3)	597 [65]	0.28 [0.02]	269 [38]
sigh	PD (8) Day 1	690 [124]	0.58 [0.08]	574 [372]
	PD (8) Day 2	770 [85]	0.57 [0.06]	692 [279]
	NG (3)	723 [70]	0.56 [0.07]	824 [110]
sip	PD (8) Day 1	589 [105]	0.21 [0.09]	243 [130]
	PD (8) Day 2	668 [144]	0.19 [0.06]	349 [152]
	NG (3)	563 [112]	0.14 [0.01]	255 [59]
ship	PD (8) Day 1	583 [87]	0.21 [0.06]	573 [190]
	PD (8) Day 2	606 [83]	0.21 [0.08]	598 [198]
	NG (3)	599 [102]	0.13 [0.01]	434 [111]
blend	PD (8) Day 1	729 [131]	0.28 [0.09]	590 [303]
	PD (8) Day 2	759 [85]	0.28 [0.04]	500 [301]
	NG (3)	679 [94]	0.29 [0.04]	602 [67]
sew	PD (8) Day 1	745 [109]*	0.44 [0.06]*	615 [180]
	PD (8) Day 2	738 [136]	0.48 [0.10]	586 [138]
	NG (3)	563 [81]	0.57 [0.05]	456 [45]

Table 3, continued

Word	Group	WWD (ms)	VD/WW (ms/ms)	FE (Hz)
row	PD (8) Day 1	636 [125]	0.54 [0.08]	370 [56]
	PD (8) Day 2	734 [202]	0.52 [0.07]	429 [108]
	NG (3)	588 [113]	0.52 [0.06]	336 [185]
shoot	PD (8) Day 1	584 [77]	0.28 [0.07]	485 [198]
	PD (8) Day 2	691 [147]	0.24 [0.04]	590 [140]
	NG (3)	655 [133]	0.19 [0.00]	492 [154]
wax	PD (8) Day 1	692 [124]	0.29 [0.08]	749 [302]
	PD (8) Day 2	742 [137]	0.29 [0.08]	848 [315]
	NG (3)	738 [103]	0.27 [0.02]	746 [358]
cash	PD (8) Day 1	679 [152]	0.36 [0.11]	195 [116]
	PD (8) Day 2	663 [97]	0.39 [0.04]*	175 [70]
	NG (3)	734 [140]	0.34 [0.03]	300 [74]

Nondisordered Speakers

Comparison of the two nondisordered control groups

The formant trajectory characteristics for the 11 words reported here for three nondisordered geriatrics modestly extend the data base published previously by Weismer et al., 1992. For the most part, the nondisordered geriatrics in this study had segmental and formant trajectory measurements similar to those in Weismer et al. data base for the same words. A two-tailed independent samples *t*-test determined that across all measures and words differences from the Weismer data were significant ($p<.05$) for only 1 of 66 cases (6 measures, 11 words). The one case (starting frequency of the second formant for the word *sigh*), which was statistically significantly different from the Weismer data, was excluded from the acoustic analysis. Inspection of distributions revealed that 55 of 66 cases or 86% of the 3 nondisordered geriatric cases were within one standard deviation of the Weismer et al. data. Visual inspection of words indicated that differences between the 3 nondisordered geriatrics and the group of 15 nondisordered geriatrics were primarily due to "idiosyncratic" word productions by one outlier in our nondisordered group (VD for *sip*, TD for *sip, shoot*, TE for *sip*, TR for *blend, cash*, and SFF2 for *sigh, sip)*.

The cumulative evidence from the comparisons of acoustic measurement data of our group of 3 nondisordered controls with the 15 nondisordered geriatrics (Weismer et al., 1992) indicated that it is legitimate to use the Weismer data set as a basis for comparison with the experimental subjects. Therefore, PD acoustic data will be compared to the data reported by Weismer et al. (1992).

Speakers With PD

Vowel duration

The data in Table 2 show that PD speakers had mean VDs that were statistically significantly longer than the nondisordered geriatric group ($p<.05$) for day 1 and day 2 productions of *sip, ship* and *shoot*. VDs were significantly shorter ($p<.05$) for one production of *wax* and *blend* (day 1).

Transition duration

The data in Table 2 show that PD speakers had mean TDs that were statistically significantly shorter than the nondisordered geriatric group ($p<.05$) for both day 1 and day 2 productions of *hail, blend,* and *wax.* TDs were significantly shorter for one production of *sigh* (day 1) and significantly longer ($p<.05$) for one production of *ship* (day 2). The words *hail, blend, wax* all contain front vocalic nuclei, either a diphthong or monophthong preceded by a liquid or semi-vowel.[3]

Transition extent

The data in Table 2 show that PD speakers had a mean TE that was statistically significantly smaller than the nondisordered geriatric group ($p<.05$) for day 1 and day 2 productions of *wax,* and statistically significantly larger ($p<.05$) TE for day 1 and day 2 productions of *ship.* TE was statistically significantly larger ($p<.05$) for one production of the words *sip, sew, shoot* (day 2) and *row* (day 1). All statistically significant differences occurred on words which contain lip rounding features, except for *sip.*

Transition rate

The data in Table 2 show that PD speakers had statistically significantly greater TRs or steeper slopes than the nondisordered group ($p<.05$) for day 1 and day 2 production of *ate, hail, ship, blend, cash.* TR was statistically significantly greater ($p<.05$) for one production of *row* (day 1) and *sigh, sip, sew* (day 2).

Starting frequencies

The data in Table 2 show that PD speakers had statistically significantly higher mean starting frequency for F1 (SFF1) than nondisordered geriatrics ($p<.05$) for both day 1 and day 2 productions of *row* ($p<.05$). Mean SFF1 was statistically significantly higher ($p<.05$) for one production of *sip, blend* (day 1) and *wax, ate, sigh, sew* (day 2). Mean SFF1 was statistically significantly lower ($p<.05$) for one production of *shoot* (day 2).

The data show that PD speakers had statistically significantly higher mean starting frequency of F2 (SFF2) than the nondisordered geriatric group ($p<.05$) for day 1 and day 2 productions of *blend, shoot, wax.* Mean SFF2 was statistically significantly higher for one production of *ship, sew* (day 2).

[3] Physiological characteristics of vowels are based on Shriberg and Kent (1982) and Schulman (1989) categorizations. Open refers to the physiological characteristic of open jaw articulation of monophthongs and the onglide portion of diphthongs. The onglide of the diphthongs in these word productions are either open (/aI/) or mid-open (/eI/, /oU/); the jaw closes slightly during the diphthong (Shriberg & Kent, 1982). For vowel categorization in terms of tongue position, high or low refers to tongue height (relative vertical position of the tongue body) and front and back refers to the tongue advancement or the anterior-posterior tongue position

Frequency extent

Since a FE measure was not made by Weismer et al., 1992, comparison of FE data by the PD subjects was possible only with our group of 3 nondisordered geriatrics (see Table 3). Across all words, data indicate no significant differences between mean FE for the PD group compared to the NG group for both day 1 and day 2. Mean FE was statistically significantly smaller ($p<.05$) for the PD group compared to the NG group for one production of *cash* (day 2).

Comparison of mean FE to mean TE was made within each group (PD, NG). Across all words, there was statistically significantly greater mean FE than mean TE ($p<.001$) for the PD group for both day 1 and 2. In contrast, the NG group of 3 did not show a statistically significant difference between FE and TE for all words collectively.

Comparison of mean FE and TE by word production for individual words show that PD subjects had statistically significantly greater FE than TE ($p<.05$) for day 1 and 2 for *blend, cash, row, shoot.* The PD subject group had statistically significantly greater FE than TE ($p<.05$) for one production of *ate, hail, sigh* (day 2). The NG group of 3 did not show a statistically significant difference between FE and TE for any of the 11 words.

Whole word duration

Since a WWD measure was not made by Weismer et al., 1992, comparison of WWD data by the PD subjects was possible only with the group of 3 nondisordered geriatrics (see Table 3). Across all words, PD speakers had statistically significantly longer mean WWD than the nondisordered geriatric group of 3 subjects ($p<.05$) only for day 2. Individual word data indicate that PD speakers had statistically significantly longer mean WWD ($p<.05$) than the nondisordered geriatric group for one production of the word *sew* (day 1).

Vowel duration/whole word duration

Since a WWD measure was not made by Weismer et al., 1992, comparison of VD/WWD data by the PD subjects was possible only with the group of 3 nondisordered geriatrics (see Table 3). Across all words, data indicate no significant differences between mean VD/WWD for the PD group compared to the NG group for day 1 or day 2. Individual word data indicate that PD mean VD/WWD was significantly smaller ($p<.05$) for one production of the word *sew* (day 1).

Variability

a.First to second word production consistency. Patterns of change from day 1 to day 2 word production of the PD group were examined to evaluate variability in performance and possible testing effect. Data indicate that, for all words collectively, the PD group had statistically

significantly longer mean vowel durations ($p<.001$) and longer mean whole word duration ($p<.018$) for day 2 compared to day 1. When mean measurement changes for individual words were examined, the PD speakers had statistically significantly longer measurements ($p<.05$) of TD for *sigh* and TE for *wax* on day 2 compared to day 1.

The data indicated that VD increases do not wholly account for WWD differences for the PD group for day 2 compared to day 1. This is because 16 of 88 cases (8 subjects, 11 words) had measured VD differences greater than 100 ms compared to 45 of 88 cases for WWD. The words with the greatest measurement *increases* for the PD group were the words *sigh* and *shoot.*

b. Inter- and intra-subject variability. As indicated by the standard deviation data in Tables 2 and 3, PD subjects showed more variability than NG subjects in temporal and trajectory measures. Examination of individual subject data revealed that PD intra- and inter-subject variability appeared highest for transition extent measures in front diphthongs and high front monophthongs, and for vowel and transition durations for consonant-vowel words with diphthongs. Greatest day 1-day 2 variability for VD, TD, TE and WWD were observed in patients rated perceptually as having "severe," "moderately severe," and "moderate" dysarthria (the others were in the "mild" to "moderate" categories).

Discussion

This study examined articulatory acoustic characteristics and consistency of speech production of words containing a variety of formant trajectories spoken by subjects with Parkinson disease (PD) and nondisordered geriatrics (Weismer et al., 1992). It focused on segmental and trajectory measures which allow inferences about lingual or overall vocal tract movement and laryngeal-oral articulator coordination.

In this discussion, we focus only on PD segmental and trajectory measures which were statistically significantly different from the nondisordered geriatric data for both day 1 and day 2 productions. Significant differences in both productions are taken as a valid indicator of essential speech production differences imposed by Parkinson disease; together they take account of the greater speech production variability typical of disordered speech *and* probable testing/learning effects (King et al., 1994). It was reasoned that comparing PD data from only one occasion to nondisordered geriatric data, we would have exaggerated or understated findings because of failure to keep typical PD variability and a testing effect distinct from indicators of essential speech production differences. Because we took a more conservative approach, we can be more confident that the differences documented here are a reflection of essential speech production differences imposed by Parkinson disease.

It should also be noted that, with two exceptions, the segmental and trajectory measures in individual words which exhibited statistically significant differences in both PD word productions compared to NGs were also those in which our nondisordered geriatric group of 3 subjects had measurements within one standard deviation of those in the Weismer et al., 1992 data base. The exceptions were second formant starting frequency for *sigh* and vowel duration in the word *sip.* These were excluded from the following discussion.

There were statistically significant differences in both day 1 and day 2 word productions on segmental and trajectory measures for the following: longer vowel durations for 2 words (*sip* omitted), shorter transition duration for 3 words, smaller transition extent for one word, larger transition extent for one word, greater slope or transition rate for 5 words, higher starting frequency of F1 for one word, and higher starting frequency of F2 for 3 words. (See Table 3.)

In the following discussion, we first will present interpretations of underlying speech production features implied by these differences. Then, through examination of significant trajectory differences *within* their respective phonetic environments, we will postulate several conditions (or acoustic and phonetic patterns) associated with them. We will discuss how the trajectory patterns appear logical when the motor control deficits and constraints imposed by the physical system in Parkinson disease are considered. We will also present possible interpretations for statistically significant differences in acoustic measures from day 1 to day 2 by PD subjects.

This study also examined variability in word production. In our examination of acoustic consistency in the word productions of PD patients from day 1 compared to day 2, we found significant increases by the PD group for all words collectively in vowel duration and whole word duration.

Segmental and Trajectory Characteristics

VD, TD, TE, TR

This section will discuss interpretations of underlying speech production features implied by statistically significant differences in acoustic measures of the PD subjects for both day 1 and day 2 compared to nondisordered geriatrics (see Table 4).

Longer movement durations for lingual or overall vocal tract change for *shoot* and *ship* are implied by the significantly longer vowel durations by the PD group. For the word *shoot*, visual examination of the trajectory data indicated that the vowel duration increase occurred primarily during its transitional segment. For *ship*, the data imply a different speech production pattern -- *abnormal anticipatory (coarticulation) adjustments.* Visual examination of the data for *ship* revealed that its longer vowel duration was due

mostly to longer formant movement at the onset of the vowel nuclei, before the transitional segment.[4] This reduced overlap of successive articulatory gestures or *segmentalization* has been depicted previously as a characteristic of ALS (Weismer et al., 1992), apraxia, (Kent & Rosenbek, 1983), and hearing impaired speakers (Rothman, 1976).

Slowness in initiating lingual or vocal tract movement is an alternate interpretation for the significant vowel duration difference in *ship*. This interpretation is based on occurrences of an essentially flat segment before the transitional segment. This plateau implies a period of relative articulatory immobility and is consistent with previous findings of PD slowness in initiating lip movements (Leanderson et al.,1972) and ballistic hand movement (e.g., Yanagisawa et al., 1989).

When interpreting speech production characteristics for trajectory measurements observed in subjects with

[4] Trajectory characteristics in *ship* were mirrored in *sip*, which was excluded from this discussion

Table 4.
Summary table of PD group segmental and trajectory measures which were statistically significantly different ($p<.05$) for both day 1 and day 2 from the NG-W group of 15.(a) For details of table headings for measures, see caption for Table 2. Implications column refers to PD speech production features for vocal tract movement implied by the significant differences.

	Increased VD	Decreased TD	Decreased TE	Increased TE	Increased TR	Increased SFF2	Implications (b)
hail		x			x		4,5,7
blend		x			x	x	4,5,6,7,8
wax		x	x			x	3,6,8
cash					x		5
ate					x		5
shoot	x						1
sip (c)	x						2
ship	x			x	x		2,7

(a) Weismer et al., 1992.
(b) 1= prolonged transition, 2= slow initiation, 3= target undershoot, 4= acceleration, 5= decreased rapid movement duration, 6= increased anterior constriction, 7= increased anterior tongue movement, 8= overlapping gestural segments (F1+, F2+).
(c) "Sip" is not included in the discussion.

PD, a combination of measures (e.g. TD and TE, or TE and FE) may provide more informed indications of underlying articulatory patterns than one measure alone.

For example, *decreased* lingual or vocal tract *movement change in time and space* was implied by statistically significant shorter transition durations on both occasions for the words *hail, blend,* and *wax* and smaller transition extent for the word *wax*. One interpretation suggested by previous research is that these acoustic differences of shorter TD and TE accompany complex vocal tract movement change (Forrest et al., 1989). The transition duration findings for the diphthong in *hail* is consistent with previous findings for the same measure during sentence production wherein the extent of vocal tract change required for the phonetic segment is relatively large (e.g. that of /aI/ and /Il/ in the Forrest et al., 1989 study). However, the lack of significant differences in PD production compared to NG production for *back* diphthongs and the small movement amplitude associated with the monophthongs in *blend* and *wax* suggest that this "complex movement" interpretation does not adequately explain *shorter* transitions.

There is another interpretation -- *target undershoot*-- suggested by trajectory data for *wax*. Target undershoot occurs when articulators do not reach the requisite position for a particular speech sound before moving on to produce the next sound. This interpretation was implied by the statistically significant shorter transition extent for the word *wax*. The interpretation is also supported by the finding that FE and TE for *wax* were *not* statistically significantly different for the day 1 or day 2 productions. The target undershoot phenomenon for lingual or overall vocal tract movement in this word is consistent with descriptions of undershoot as a characteristic of PD lip movement (Netsell et al., 1975) and with cinefluorgraphic findings for oral articulation in PD subjects described by Netsell et al., 1975. The undershooting phenomenon in vocal tract movement is also consistent with the general finding that Parkinson disease patients have a parallel problem with other motor output. This problem in motor output is found in the reduced amplitude (undershoot) of stride length in walking and letter stroke in handwriting (Stelmach, 1991) and has been suggested as one factor in the reduced vocal fold adduction and loudness reported in PD subjects (Ramig et al., in press).

Unlike *wax*, the words *hail* and *blend* show a trajectory pattern *other* than target undershooting. The data are consistent with three possible interpretations: *an acceleration phenomenon, inconsistent velocity,* and/or *shortened duration of rapid movement* (i.e., the portion of the trajectory meeting the 20 Hz/20 ms criteria for the transition segment). These interpretations are implied because transition durations for *blend* and *hail* are statistically significantly shorter but are *not* accompanied by significantly shorter transition extents or frequency extents. Given the lack of significant differences for the transition and frequency extents *and* the significantly greater transition rates for *blend* and *hail*, the interpretation of an acceleration phenomena seems to be the most logical for these words. An acceleration phenomenon for lingual or vocal tract movement is consistent with kinematic evidence of a lip movement acceleration in PD subjects. PD lip movement acceleration patterns are postulated to come from the same neurogenic basis as tremor, which is uncontrolled due to weakness in neuromuscular control signals (Netsell et al., 1975).

An interpretation of inconsistent velocities and/or shorter durations of rapid vocal tract movement for *ate* and *cash* is implied by their statistically significantly greater transition rates on both day 1 and day 2 which were not accompanied by statistically significant differences for transition duration and extent or frequency extent. This interpretation is also supported for *cash* (and as an alternate interpretation for *blend*) given that the PD group had statistically significantly greater FE than TE on both day 1 and day 2 for these words. This interpretation of disordered velocity for vocal tract change is consistent with findings of inconsistent velocity of lip movement for syllable production previously described for PD subjects (Hirose et al., 1982). Shorter duration of initial rapid vocal tract movement is also reminiscent of descriptions of PD limb movement. Specifically, previous research has found a failure in PD subjects to produce initial quick force or velocity in execution of ballistic movement (Hallett and Khoshbin, 1980; Evarts et al., 1981). This has been postulated to result variously from muscular weakness and rigidity, contracture at limb joints and/or lack of effort (Yanagisawa et al., 1989).

a. Conditions accompanying PD trajectory differences. Significant trajectory differences *and their phonetic contexts* were examined to isolate conditions which may accompany these differences. Our rationale for this detailed examination was that the amount and kind of speech production deficit exhibited by PD subjects could depend upon the kind of physical constraints (e.g. constriction of the anterior oral tract) present during speech production or upon the particular underlying speech production mechanisms (or speech gestures) required. This reasoning was consistent with the generally accepted finding that the motor control deficit in PD movement is not found in all movement, but is best demonstrated in relatively complex movements (Hallet & Khoshbin, 1980; see Forrest et al., 1989, pp. 2621). We wanted to determine *which acoustic and phonetic patterns* (and associated complex movements) in single word production *best demonstrated* deviations in acoustic patterns in PD subjects. We reasoned that addressing this issue would enable us to generate other words which could be used to replicate and verify the acoustic results and interpretations of this study.

One physical condition or constraint-- *greater constriction near the front of the oral tract*-- is implied by the

data for *blend* and *wax*. Acoustic analysis showed that, for both day 1 and day 2, there was statistically significant higher F2 starting frequencies at the onset of the transition segment of these two words. Higher F2 starting frequencies are consistent with an anterior tongue position, elevated jaw position and decreased lip rounding (Pickett, 1980). This interpretation of a constriction condition for these two words is consistent with previous kinematic findings of *smaller* PD lip and jaw displacement during sentence production and repetitive word production (Forrest et al., 1989; Hirose et al., 1982). Though not statistically significant, higher F2 starting frequencies for 10 of 11 words suggest that PD subjects generally have greater constriction near the front of the vocal tract.

Another physical condition or constraint implied by the data is *anterior tongue position* or *movement* (but not posterior or backward movement). This interpretation is indicated by the finding that trajectory differences (TD and TE or TR) reached statistical significance in both occasions only for the *phonetic contexts* of *front* diphthongs or *front* monophthongs preceded by a liquid or semi vowel. The statistically significantly greater ($p<.05$) TE and TR for both productions of *ship* are also consistent with this interpretation. One explanation may be that "tongue tip movements show higher velocities than do either tongue dorsum or lip movements" (Browman & Goldstein, 1990, pp. 361). The PD subject may show more rapid forward tongue movement (e.g., an acceleration phenomenon, rapid initial movement or greater range of movement during tongue advancement) due to a deficit in motor control of a greater "naturally occurring" velocity. This also would be consistent with weakness in neuromuscular control signals, similar to that postulated for accelerated lip movement (Netsell et al., 1975). The more rapid forward tongue movement could also be due to acoustic features associated with the mechanical coupling between the tongue-tip/mandible interaction (Hirose et al., 1982).

An underlying speech production pattern-- *overlapping articulatory gestures*-- is also implied by the acoustic data. Specifically, the data indicate that the words *blend* and *wax* have a unique slope pattern: they are two of only three words of 11 produced with *both* positive F1 and positive F2 slopes. (The third word is *sigh*, which had a statistically significant higher TR for one production.) The F1 transitional segment in these words corresponds to a vocal tract opening gesture; the F2 transitional segment corresponds to probable tongue movement toward the front of the palate.[5] Visual examination of the data indicated that the F1 transitional segment *preceded and overlapped* with the F2 transitional segment. Thus, the data suggest that it is during these kinds of sequential gestures that the motor control deficits in PD may be particularly evident.

This observation is consistent with certain contemporary accounts of successive articulatory overlap -- gesture sliding and blending (Browman & Goldstein, 1990; Saltzman & Kelso, 1987). This is the notion of overlapping activation of several articulators when sequential gestures involve the same articulators. In fact, as the overlapping F1 and F2 trajectories make clear, front diphthongs and front monophthongs preceded by a liquid or semi-vowel use concurrently active underlying gestures involving the same articulator, namely the tongue.[6]

b. TE and FE. The accumulation of findings indicate that transition extent and frequency extent both should be used for interpretation of trajectory findings.[7] The data support this recommendation in several ways.

First, while the amount of FE data we have analyzed is still quite small, the lack of significant differences in frequency extent between the PD group and our nondisordered geriatric group suggests that the magnitude of vocal tract change in space during vocalic nuclei is similar for both. Next, the PD group had statistically significantly greater FE than TE for all words collectively and for four individual words. This suggests that PD subjects have a greater proportion of "stairstepping" type trajectories where only a portion of the transition segment reaches at least the conventional 20 Hz/20 ms standard.

One implication of these findings is that the conventional definition for the transition segment may foster a *faulty interpretation* that PD speech reflects many more instances of lingual or vocal tract target undershoot than is actually the case. In other words, shorter transition extents in PD single word speech production do not necessarily imply smaller range of movement *unless* frequency extent is also smaller.

On the other hand, the conventional definition of the transitional segment *does* quantify the portion of the speech gesture with *sustained rapid change* (at least 20 Hz/

[5] This serial ordering reflects gestures involving the same articulators. Specifically, the gestures involve the body of the tongue in a relatively raised, forward position (probably by the action of the posterior and central parts of the genioglossus muscle) and the tip/blade of the tongue in a raised position (probably by the action of the superior longitudinalis muscle) (Hardcastle & Edwards, 1992). Importantly, lip and tongue movements depend to a certain extent on the movement of the jaw (Hirose et al., 1982).

[6] In more detail: 1) a gesture in which the vocal tract configuration (and jaw) is moving from a closed to an open position (with an accompanying decrease in tongue and vocal tract constriction at the front of the palate) and 2) movement of the constriction of the tongue or vocal tract toward the front of the palate.

[7] One question which arose from the findings was whether the 20 Hz/20 ms operationalized criterion was adequate for defining offset, the point at which the transitional segment ends and the steady state portion begins. The implication which emerges from the findings discussed in this section is that the single point of demarcation resulting from application of this criterion appears inadequate for analysis of the complexities of PD speech, but reveals important speech production characteristics and underlying gesture features.

20 ms change). Indeed, the duration and extent of this acoustic change *within particular phonetic contexts* appear indicative of an essential speech production or gesture feature of Parkinson disease. Therefore, given these findings and their implications, it seems clear that interpretation of transitional segment differences in disordered speech should include frequency extent findings before any assessment of target undershoot and/or acceleration phenomena is made. The transition and frequency extent differences observed in our small group of nondisordered geriatrics, of course, deserve further study with more subjects; the small amount of data limit conclusive statements at this time.

Whole word duration and VD/WWD

The measurements of whole word duration and VD/WWD were used as an initial probe of laryngeal and supraglottal interarticulator coordination. The PD subject data compared to our nondisordered geriatric group of three were not statistically significantly different on these measures for both day 1 and day 2 for words collectively or individually. While these findings may be a preliminary indication that there are no differences in glottal/oral interarticulator coordination in PD compared to NG speakers, it is possible that the lack of significant measurable differences between the two groups was due to the small size of the nondisordered group. However, given the observation of increased fricative duration and rise time in PD word production (Dromey et al., in press), it seems more likely that the whole word duration and VD/WWD measurements were not sufficiently sensitive to differences in coordination of laryngeal valving gestures and oral closing and opening gestures. Therefore, it seems important that examination of PD word production be extended to other acoustic cues associated with management of air-flow and glottal/oral interarticulator timing (e.g. voice onset time, fricative duration and rise time, stop gap duration, spirantized stop closures; see Kent et al., 1989a).

Variability

First to second word production consistency

There are several possible interpretations for PD speakers showing statistically significant increases in VD and WWD for the 11 words collectively for day 2 compared to day 1. The effects of an experimental setting and/or learning, due to repeating words, may encourage certain compensatory speech strategies to produce "clear speech." For example, a reduction in rate could produce the differences observed for the PD repetitions. An interpretation of a learning effect is consistent with the King et al.(1994) findings when studying repeated productions of voice measures produced by patients with PD. It is important to point out, however, that King et al. (1994) reported PD changes in sustained phonatory measures during maximum effort tasks. This differs from the single-word speech production task reported here which was *not* a maximum effort task.

The interpetation of a PD compensatory strategy to produce clear speech on day 2 is consistent with previous findings of longer segmental durations in nondisordered speakers during "clear" speech (Pincheny et al., 1986). However, the nonsignificant differences in transition extent and frequency extent for day 1 to day 2 word productions are *not* consistent with previous findings that slow rates resulted in exaggerated transition extents in nondisordered speakers (see Weismer & Liss, 1991, p. 266). (See "Segmental and trajectory characteristics" for transition duration-frequency duration comments.) Furthermore, the combination of measures (TE, TD and FE which were not significantly different and VD and WWD which were significantly different from day 1 to day 2) appears to highlight an aspect of PD speech production. These data suggest that while PD subjects may have reduced their speech rate at day 2, their prolongation of voicing for vowels was not accompanied by a *systematic increase* in magnitude of *sustained rapid vocal tract change* in *space and time* (TE and/or FE and TD).

The PD day 1 to day 2 word production differences also may reflect an important aspect of PD speech variability. Multiple sampling of words seems to provide a more accurate representation of acoustic differences between disordered and nondisordered speech by limiting the potential for exaggerating differences. For example, using a criteria of statistically significant differences for only one production (rather than both day 1 and day 2 word productions) would have yielded a total of 6 words (compared to 2) for transition extent and a total of 9 words (compared to 5) for transition rate.

Inter- and intra-subject variability

The higher intersubject variability of the PD subjects compared to NG subjects was expected because of the common finding that speech production measures from disordered speakers are more variable than those obtained from a nondisordered group. PD intrasubject variability, particularly for trajectory measures, may indicate decreased articulatory coordination (Ziegler & von Cramon, 1986; Hardcastle & Edwards, 1992).

The higher intrasubject variability of PD subjects compared to NG subjects is also interesting given the argument that high variability is the articulatory characteristic most likely to affect intelligibility. Hardcastle & Edwards (1992) suggested that when speaker errors are consistent, the listener can construct a systematic framework for understanding the speech pattern. But when errors vary from one production to another in several acoustic features, the listener must constantly revise his interpretation and rely more on other cues-- both sound cues and context cues. In fact, individual PD subject data suggested a possible association between highly uneven repetitive productions (where multiple acoustic features differed from day 1 to day 2) and perceptual ratings of severity of dysarthria.

The limited nondisordered geriatric data and the apparent patterns, of course, argue for future study with a larger control sample and multiple repetitions.

Conclusions

These data demonstrate how acoustic analysis can reveal the complexity of PD acoustic deviations from nondisordered speech production. The differences in vocalic acoustic patterns are interesting given the fact that the Parkinson subjects in the current study were quite intelligible when producing the test words (and the vowels). Therefore, acoustic examination revealed acoustic features reflecting lingual or overall vocal tract movement patterns which were not obvious to the listener but which may be as theoretically important.

Our initial probe of laryngeal and supraglottal interarticulator coordination using acoustic measurements of WWD and VD/WWD did not confirm significant measurable differences between the PD group of 8 and the NG group of 3. It is likely that more definitive answers to PD management of air-flow and laryngeal and supraglottal interarticulator coordination patterns in word production must await collection of data from other acoustic cues associated with these kinds of coordinative patterns.

Keeping in mind the very limited data set, the results of the statistically significant PD acoustic differences for both day 1 and day 2 are interpreted as indicating that several speech production patterns characterize PD lingual or vocal tract movement within *selected* word phonetic contexts. The significantly longer vowel durations and associated trajectory characteristics revealed by visual examination imply vocal tract movement prolongation, abnormal anticipatory (coarticulation) adjustment and/or slowness in initiation of vocal tract movement for vocalic production in *shoot* and *ship*. The significant differences in TD, TE, FE to TE, and TR measures, taken together, imply velocity abnormalities (an acceleration phenomenon, inconsistent velocity, and/or shortened sustained rapid movement) for *hail, blend, ate,* and *cash*, and target undershoot for *wax*.

Joint analysis of the acoustic data and phonetic contexts in which significant *trajectory* differences occurred were interpreted as indicating several conditions which seem to accompany these deviations. These conditions included vocalic nuclei or movements associated with: 1) increased vocal tract constriction at the front of the palate at the transitional segment onset (second formant starting frequency), 2) a forward movement of the tongue (front vocalic nuclei), 3) relatively complex and large amplitude change (diphthong), and 4) concurrently active underlying gestures using the same articulator, the tongue (liquid or semi-vowel followed by a front monophthong; front diphthong). (See Table 4.)

The acoustic findings also imply that speech movement features previously found in lip and jaw movement appear to be present in lingual or overall vocal tract movement for a subset of the test words (particularly *hail, blend, wax).*[8] Our findings also suggest that the combination of formant trajectory measurements *and* intra-subject variability in repetitive production may hold promise as a general indicator of speech mechanism involvement in Parkinson disease. While this is somewhat consistent with the similar suggestion (of transition rate or slope) for ALS speech by Weismer et al., (1992), we found that transition duration, transition extent, and frequency extent within certain phonetic environments *jointly* give a more accurate indication of PD trajectory differences than transition rate alone. Further study of the association between trajectory measures and other measures of the speech signal (e.g. fricative duration, rise time, VOT, etc.) would help to determine whether trajectory measurements could be used as a general index of overall severity of speech mechanism involvement in PD. Even more critical would be study of the association between trajectory features (*and* their variability) and overall speech intelligibility scores in PD subjects.

Other clear implications for general PD acoustic study and especially speech treatment efficacy study (Ramig et al., in press) were generated by this study. First, the magnitude of PD intra-subject variability and evidence of a possible testing effect on segment duration (VD and WWD) argue for obtaining multiple productions of words so that researchers do not exaggerate or understate findings because of failure to keep typical PD variability and a testing effect distinct from measured *treatment* effects. This is consistent with PD variability in acoustic parameters of phonation reported by King et al., 1994. Second, the magnitude of PD intra-subject variability underscores the need for analysis of multiple word productions and intrasubject variability in a larger group of nondisordered geriatrics. The formidable task of acoustic analysis of multiple word repetitions could be narrowed by being limited to words which displayed significant differences between PD and nondisordered groups.

If the acoustic results of this study can be replicated in other PD speakers and for other words which reflect the same acoustic and phonetic patterns (e.g. *bland, wed, wicks, lax*, and so forth), a profile of PD speech production will

[8] Given the implications of the acoustic data in the present study, apparently there are impaired movement characteristics which are common to lingual or vocal tract movement and Parkinsonian limb ballistic movements. This commonality has been described previously for motor functioning of labial and limb musculature by Leanderson et al., 1972. Because of important differences in functional anatomy (e.g. see Leanderson et al., 1972, pp. 356), this is somewhat unexpected. However, the accumulating evidence that type of motor control seems similar for Parkinsonian limb and articulatory musculature argues that this should be taken into account when discussing the neuro-muscular basis for motor symptoms in persons with PD.

emerge which will be far richer than currently available. In all cases, the detailed information gained from this study has highlighted speech production features which have not been apparent when a more limited sample or traditional auditory perceptual analysis was used.

Acknowledgements

This research was supported in part by grants NIH-NIDCD #RO1 DC01150 and #P60 DC00976.

References

Abramson, A.S. (1977). "Laryngeal timing in consonant distinctions," Phonetica 34, 295-303.

Ansel, B. and Kent, R. (1992). "Acoustic-phonetic contrasts and intelligibility in dysarthria associated with mixed cerebral palsy," J. Speech Hear. Res. 35, 296-308.

Benguerel, A. P., Hirose, H., Sawashima, M., and Ushijima, T. (1978). "Laryngeal control in French stop production: A fiberscopic, acoustic and electromyographic study," Folia Phoniatrica 30: 175-198.

Bross, R. (1992). "An application of structural linguistics to intelligibility measurement of impaired speakers of English," in Intelligibility in Speech Disorders, edited by R. D. Kent (John Benjamins Publishing Co., Amsterdam) pp. 35-66.

Browman, C. P. and Goldstein, L., (1990). "Tiers in articulatory phonology, with some implications for casual speech," in Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech, edited by J. Kingston and M. E. Beckman (Cambridge University Press: Cambridge) pp. 341-376.

Darley, F. L., Aronson, A. E., and Brown, J. R. (1969a). "Differential diagnostic patterns of dysarthria," J. Speech Hear. Res. 12, 246-269.

Darley, F. L., Aronson, A. E., and Brown, J. R. (1969b). "Clusters of deviant speech dimensions in the dysarthrias," J. Speech Hear. Res. 12, 246-269.

Darley, F., Aronson, A.E., and Brown, J. (1975). Motor Speech Disorders, (Saunders, Philadelphia).

Dromey, C., Ramig, L.O., and Johnson, A. (in press). "Phonatory and articulatory changes associated with increased vocal intensity in Parkinson disease: A case study," J. Speech Hear. Res.

Evarts, E. V., Teravainen, H., and Calne, D. B. (1981). Reaction time in Parkinson disease," Brain 104, 167-186.

Forrest, K., Weismer, G., and Turner, G.S., (1989). "Kinematic, acoustic, and perceptual analyses of connected speech produced by Parkinsonian and normal geriatric adults," J. Acous. Soc. Am. 85, 2608-2622.

Hallet, M. and Khoshbin, S. (1980). "A physiological mechanism of bradykinesia," Brain 103 (2), 301-314.

Hansen, D. G., Gerratt, B. R., and Ward, P. H. (1984). "Cinegraphic observations of laryngeal function in Parkinson's disease," Laryngoscope 94, 348-353.

Hardcastle, W., Edwards, S. (1992). "EPG-based description of apraxic speech errors," in Intelligibility in Speech Disorders, edited by R.D. Kent (John Benjamins, Amsterdam) 287-328.

Hirose, H., Kiritani, S., and Sawashima, M. (1982). "Velocity of articulatory movements in normal and dysarthric subjects," Folia Phoniatrica 34, 210-215.

Hirose, H., Kiritani, S., Ushijima, T, Yoshioka, H., and Sawashima, M. (1981). "Patterns of dysarthric movements in patients with Parkinsonism," Folia Phoniatrica 33, 204-215.

Hoehn, M. and Yahr, M. (1967). "Parkinsonism: Onset, progression and mortality," Neurology 19, 427.

Johnson, A. (1993). The effect of intensive voice therapy on vowel duration and slope of second formant transitions in patients with Parkinson disease. Unpublished master's degree thesis. Boulder: CDSS, University of Colorado, Boulder.

Kent, J. F., Kent, R. D., Rosenbek, J. C., Weismer, G., Martin, R., Sufit, R., and Brooks, B.R., (1992). "Quantitative description of dysarthria in women with amyotrophic lateral sclerosis," J. Speech Hear. Res. 35, 723-733.

Kent, R.D., Weismer, G., Kent, J.F., and Rosenbek, J.C. (1989a). "Toward phonetic intelligibility testing in dysarthria," J. Speech Hear. Disord. 54, 482-499.

Kent, R. D., Kent, J. F., Weismer, G., Martin, R. E., Sufit, R. L., Brooks, B., and Rosenbek, J. C. (1989b). "Relationships between speech intelligibility and the slope of second-formant transitions in dysarthric subjects," Clinical Linguistics & Phonetics 3 (4), 347-358.

Kent, R. D. and Rosenbek, J. C. (1982). "Prosodic disturbance and neurologic lesion," Brain and Language 15, 259-291.

Kent, R. D. and Rosenbek, J. C. (1983). "Acoustic patterns of apraxia of speech," J. Speech Hear. Res. 26, 231-249.

King, J., Ramig, L., Lemke, J., and Horii, Y. (1994). "Variability in acoustic and perceptual parameters of phonation in patients with Parkinson's disease," Journal of Medical Speech-Language Pathology 2, 29-42.

Klatt, D.H. (1975). "Voice onset time, frication, and aspiration in word initial consonant clusters," J. Speech Hear. Res. 18, 686-706.

Leanderson, R., Meyerson, B., A., and Persson, A., (1972). "Lip muscle function in Parkinsonian dysarthria," Acta Oto-laryngol. 74, 350-357.

Lindblom, B. (1963). "A spectrographic study of vowel reduction," J. Acous. Soc. Am. 35, 1773-1781.

Lindqvist, J. (1972). "Laryngeal articulation studied on Swedish subjects," STL- QPRSR 2-3.

Liss, J.M. and Weismer, G. (1992). "Qualitative acoustic analysis in the study of motor speech disorders," J. Acous. Soc. Am. 92, 2984-2987.

Liss, J. M., Weismer, G., and Rosenbek, J. C. (1990). "Selected acoustic characteristics of speech production in very old males," Journal of Gerontology: Psychological Services 45 (2), 35-45.

Lofqvist, A. and Yoshioka, H. (1981). "Interarticulator programming in obstruent production," Phonetica 38, 21-34.

Logemann, J. A., Fisher, H. B., Boshes, B., and Blonsky, E.R. (1978). "Frequency and occurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson's patients," J. Speech Hear. Disord. 42, 47-57.

Munhall, K. G., Ostry, D. J., and Parush, A. (1985). "Characteristics of velocity profiles of speech movements," Journal of Experimental Psychology: Human Perception and Performance 11-4, 457-474.

Netsell, R., Daniel, B., and Celesia, G. G. (1975). "Acceleration and weakness in Parkinsonian dysarthria," J. Speech Hear. Disord. 40, 170-178.

Perez, K., Ramig, L. O., Smith, M., and Dromey, C. (in press). "The Parkinson larynx: Tremor and videostroboscopic findings," Journal of Voice.

Peterson, G. E. and Lehiste, I. (1960). "Duration of syllable nuclei in English," J. Acous. Soc. Am. 32, 693-703.

Pickett, J. M. (1980). The Sounds of Speech Communication: A Primer of Acoustic Phonetics and Speech Perception (University Park Press, Baltimore) pp. 41-57.

Pincheny, M. A., Durlach, N. I., and Braida, L. D. (1986). "Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech," J. Speech Hear. Res. 29, 434-446.

Ramig, L. O., Countryman, S., Thompson, L.L., and Horii, Y., (in press). "A comparison of two forms of intensive speech treatment for Parkinson disease," J. Speech Hear. Res.

Rothman, H. (1976). "A spectrographic investigation of consonant-vowel transitions in the speech of deaf adults," J. Phon. 4, 129-136.

Saltzman, E. and Kelso, J. A. S. (1987). "Skilled actions: A task dynamic approach," Psychological Review 94, 84-106.

Sanes, J. N. and Evarts, E. V. (1985). "Psychomotor performance in Parkinson disease," in Clinical Neuropsychology in Parkinsonism, edited by D. J. Delwarde and A. Agnoli (Elsevier, New York) pp. 117-132.

Schulman, R. (1989). "Articulatory dynamics of loud and normal speech," J. Acous. Soc. Am. 85, 295-312.

Shriberg, L. D. and Kent, R. D. (1982). Clinical Phonetics (Macmillan Publishing Co., New York) pp. 35-73.

Smith, M.E., Ramig, L. O., Dromey, C.D., Perez, K. S., and Samandari, R. (in press). "Videostroboscopic assessment of phonation in Parkinson's disease before and after intensive voice therapy," Journal of Voice.

Stelmach, G. E. (1991). Basal ganglia impairment and force control," in Tutorials in Motor science, edited by J. Requin and G. E. Stelmach (Kluwer Academic Publishers, Netherlands) pp 137-148.

Weismer, G. (1984). "Articulatory characteristics of Parkinsonian dysarthria: Segmental and phrase-level timing, spirantization, and glottal-supraglottal coordination," in The Dysarthrias: Physiology-Acoustics-Perception-Management, edited by M. R. McNeil, J. C. Rosenbek and A. Aronson (College Hill Press, San Diego) pp. 101-130.

Weismer, G., Kent, R. D., Hodge, M., and Martin, R. (1988). "The acoustic signature for intelligibility test words," J. Acous. Soc. Am. 84, 1281-1291.

Weismer, G. and Liss, J. M. (1991). "Acoustic/perceptual taxonomies of disordered speech," in Dysarthria and Apraxia of Motor Speech: Perspectives on Management, edited by C. Moore, K. Yorkston, and D. Beukelman (Brookes, Baltimore) pp. 245.

Weismer, G., Martin, R., Kent, R.D., and Kent, J.F. (1992). "Formant trajectory characteristics of males with amyotrophic lateral sclerosis," J. Acous. Soc. Am. 91, 1085-1098.

Yanagisawa, N., Fujimoto, S., and Tamaru, F. (1989). "Bradykinesia in Parkinson's disease: Disorders of onset and execution of fast movement," European Neurology 29, 19-28.

Ziegler, W. and von Cramon, D. (1986). "Disturbed coarticulation in apraxia of speech: Acoustic evidence," Brain and Language 29, 34-57.

NCVS Status and Progress Report - 8
July 1995, 131-138

Electroglottographic Measures of Vocal Fold Adduction in Parkinson Disease: Pre and Post Intensive Speech Treatment

Gayle Brosovic, B.A.
Lorraine Olson Ramig, Ph.D., CCC-SP
Department of Communication Disorders and Speech Science, The University of Colorado-Boulder
The Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts

Abstract

This study used electroglottography to measure changes in vocal fold adduction in subjects with idiopathic Parkinson disease following speech treatment. Thirty subjects were administered sixteen one-hour sessions of speech therapy over a four-week period. The subjects were randomly assigned to one of two speech treatment groups. One group of subjects received therapy focusing on increasing respiratory (R) support. The other group of subjects received therapy that focused on increasing vocal fold adduction and respiratory support: the Lee Silverman Voice Treatment (LSVT). Measures of electroglottography (EGGW) and Sound Pressure Level (SPL) were made simultaneously pre- and post- treatment during maximum duration sustained vowel phonations. Statistically significant changes were measured in EGGW from the pre to post- treatment condition, supporting increased vocal fold adduction in the LSVT treatment group. EGGW measures did not change in subjects receiving R treatment. There was a statistically significant correlation between magnitude of pre to post-treatment change in EGGW and SPL measures.

Introduction

At least 75% of patients with Parkinson disease have speech and voice disorders (Darley, Aronson, & Brown, 1969a, 1969b; Logemann, Fisher, Boshes, & Blonsky, 1978). Incomplete vocal fold adduction, or bowing, has been observed in Parkinson disease and related to the reduced loudness observed in these patients (Hansen, Gerratt, & Ward, 1984; Smith, Ramig, Dromey Perez, & Samandari, in press). Recent findings have suggested that behavioral treatment (The Lee Silverman Voice Treatment, LSVT) can improve vocal loudness in patients with Parkinson disease through increasing vocal fold adduction and respiratory support (Ramig, 1995; Ramig, Bonitati & Horii, 1991; Ramig, Bonitati, Lemke & Horii, 1994; Ramig, Countryman, Thompson & Horii, in press; Dromey, Ramig, & Johnson, in press).

Previous videoendoscopic reports have documented increased vocal fold adduction in patients with Parkinson disease following the LSVT (Smith et al, in press). However, invasive measures such as endoscopy are often difficult to obtain routinely and to quantify systematically in patients with neurological movement disorders (Smith et al., in press). Therefore, supplemental noninvasive assessments of vocal fold closure should be considered.

One non-invasive measure that may enhance assessment of vocal fold adduction and contribute to objective quantification of pre- to post-treatment change in vocal fold adduction is electroglottography. Electroglottography is a non-invasive procedure that provides a signal related to glottal kinematics by measuring vocal fold contact through electrical impedance changes (Haji, Horiguchi, Baer, & Gould, 1986; Scherer et al., 1993). The underlying principle is that the opening and closing of the vocal folds provides a variable resistance to the flow of the current. The amplitude variations are characteristic of the amount of vocal fold contact. The waveform is the outcome of a change in the conductance of the electrical signal that is brought about by a change in the degree of vocal fold contact.

While the non-invasive aspect of electroglottography is a positive feature in measurement feasibility, questions have been raised about the validity and reliability of the EGG signal. The validity of the EGG signal has been

investigated in several studies. Rothenberg and Mahshie (1988) concluded that with a strong signal, EGG output can be used to monitor the degree of vocal fold adduction during voiced speech where there is no supraglottal constriction. Childers, Hicks, Moore, & Alsaka (1986) determined that the electroglottogram is related to vocal fold contact area. However, they noted that mucous strands between the vocal folds can have slight effects on the EGG waveform. During abduction, the EGG waveform rises slowly as the mucous strand is stretched, but still continues to bridge the glottis. During this opening stage, the mucous strand will eventually break, resulting in a steep rise in the EGG waveform (Childers et al., 1986).

Others have questioned the usefulness of electroglottographic measures in the assessment of vocal fold adduction. Karnell (1989) concluded that other glottic activity measures should besynchronized with the EGG data to minimize or avoid misinterpretation due to variations in cycle-to-cycle glottal activity. Kitzing (1990) questioned the interpretation of the individual EGG waveform because the different phases of the period are not securely defined and there is insufficient knowledge about what they correspond toacoustically and physiologically. However, Kitzing also concluded that valuable information can be obtained when EGG is combined with other methods like videostroboscopy.

Recently, Scherer et al. (1993) have supported the usefulness of electroglottographic measures and have documented the ability of the EGG signal to discriminate among the voicequalities along the adductory continuum of breathy-normal- pressed. These findings are consistent with our preliminary work (Ramig, Fazoli, & Scherer, 1990) supporting the usefulness of EGG in documenting treatment-related adductory changes in patients with Parkinson disease. Scherer et al. have proposed EGGW-25 as one way to quantify the EGG waveform. The measure EGGW-25 is derived from the EGG waveform by taking the ratio of distances (or times) obtained by an intersection line through the signal waveform at the 25% height location.

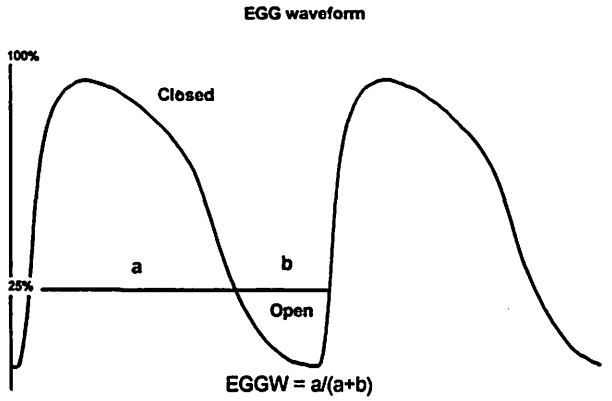


Figure 1. Explanation of the EGGW measurement taken at the 25% height location from the electroglottograph waveform. The top of the waveform relates to maximum glottal adduction or closure. The bottom of the waveform relates to maximum glottal abduction or glottal opening.

The goal of this study is to evaluate the usefulness of the electroglottographic signal as a measure reflecting treatment-related change in vocal fold adduction in idiopathic Parkinson disease following two forms of intensive treatment. It is hypothesized that the electroglottographic signal will reflect increased vocal fold adduction in patients with Parkinson disease who received treatment designed to increase adduction (LSVT) and that the increased adduction will not be observed in patients who received treatment focusing on respiratory support only. It is further hypothesized that these measures will relate significantly to treatment-related increases in SPL.

Methodology

Subjects

Thirty patients with idiopathic Parkinson disease were the subjects in this experiment. All participants were volunteers and residents of Colorado. Subjects completed otolaryngological, neurological, and neuropsychological assessments before participating in the study; each subject was given arating on the Hoehn and Yahr scale (1967) which evaluates the severity of Parkinson disease symptoms. The subjects were neuropharmacologically stable and did not change medication during the course of the treatment. Subjects were randomly assigned to one of two speech treatment groups. The first group of eleven subjects received therapy focusing on respiration only. The second group of thirteen subjects received therapy that included both respiratory and laryngeal adduction treatments (LSVT). Analysis of intensity revealed that the two treatment groups were comparable pre-treatment ($t(22) = .52$, $p=.608$). Analysis further revealed that the two groups were also comparable for EGGW pre-treatment ($t(22)= .76$, $p=.454$).

Six subjects were eliminated from the study due to difficulty in analyzing the EGG signal. Of the twenty-four remaining subjects, their ages ranged from 49 to 81 years. The R treatment group ages ranged from 51 to 81 years, with a mean of 63.1 years. The LSVT group ages ranged from 49 to 79 years, with a mean of 62.5 years. Subject profiles are summarized in Table 1. Detailed subject descriptions have been reported previously (Ramig, Countryman, Thompson & Horii, in press).

Table 1.
Profiles of Subjects With Parkinson Disease by Treatment Group

Respiratory Only						Voice and Respiratory (LSVT)					
Subject	Age in years	Gender	Years since diagnosis	Stage	Medication	Subject	Age in years	Gender	Years since diagnosis	Stage	Medication
TB	70	M	4	3	a,c	DB	62	M	2	2	a,c
PC	64	F	2	3	a,c,p	ED	68	M	5	3	a,k
JF	52	M	12	2	a,d,e,p	DH	61	M	1	2	a,c
CH	67	M	18	3.5	b,c,j	EI	67	M	8	3	a
BK	71	M	4	3	a,n	NJ	49	F	9	2	a,d,g
JL	69	M	1	2	i	RL	64	M	7	2	a
FN	51	M	6	1.5	a,c	MM	79	M	1	3	a,c
KR	81	M	2	2	a,c,n	DP	49	M	2	2	a,b,c
SS	51	F	3	1	none	GS	52	M	18	3	a,c,d,e
TS	64	M	10	3	a	GSE	54	M	14	3.5	b,i,k
BT	54	M	3	1	b,c,h	PS	56	M	8	2	a,c,h
						HT	76	M	1	2	a
						HW	76	M	9	3	b,c,l
Mean:	63.1		5.9	2.3			62.5		6.5	2.5	
Median:	64.0		4.0	2.0			62.0		7.0	2.0	
SD:	9.9		5.3	0.9			10.3		5.3	0.6	

a=sinemet; b=sinemet-cr; c=eldyprel; d=parlodel; e=symmetrel; f=pergolide; g=cogentin; h=artane; i=amantadine; j=permax, k=deprenyl; l=parsidol
Anti-depressants: n=amitriptyline; o=tofranil; p=elavil

Treatment

Both groups received intensive treatment (16 sessions in one month) with the subject being pushed to maximum effort levels in each session. The R treatment group received training that stimulated increased respiratory effort and sustained inspiration and expiration to provide increased volume and subglottal air pressure for speech. The LSVT group focused on increasing intensity through increased vocal fold adduction. Isometric pushing or lifting while phonating stimulated greater adduction. These maximum effort tasks also stimulated respiratory functioning, so the treatment addressed both increased vocal fold adduction and respiratory support (Ramig,Countryman, Thompson & Horii, in press). Detailed descriptions of treatment have been provided elsewhere (Ramig, 1995; Ramig, Countryman & Pawlas, in press).

Data Collection

For all subjects, the data were collected in an IAC sound-treated booth pre and post treatment.The recordings were scheduled at the same time post-medication for each date.The task protocol included several respiratory, phonatory, and articulatory tasks. Multi-channel data were stored on a Sony PC-108M 8 channel digital audio tape (DAT).

The EGG data were collected by placing two electrodes bilaterally on the neck anteriorly at the level of the thyroid cartilage perpendicular to the glottis. The current passes from one electrode through the laryngeal tissue to the receiving electrode. The electrodes of the Synchrovoice Research Electroglottograph were held in place by a velcro strap; this was used for comfort and to maintain the position of the electrodes throughout the session. After amplification (Tektronix Amplifier 502 TM 506), the signal was recorded into one channel of a Sony, PC-108M 8-channel DAT recorder.

A sound level meter (Bruel and Kjaer Type 2230) was positioned 50 cm from the subject to record sound pressure level and a head-mounted microphone (AKG C410) was located 8 cm from the lips.

Each subject was asked to Stake a deep breath and sustain /a/ for as long as you can. T During this task, the subjects were provided a clock with a second hand for monitoring duration of phonation. No instructions were given concerning intensity. This task was repeated three to four times at the beginning of the recording session (depending on subject fatigue) and two times near the end of the session.

Data Analysis-EGGW-25

The EGG signal was digitized using a 16 bit Digital Sound Corporation A/D converter. The EGG signal was analyzed with an in-house software program (EGGW) on a VAX 4000/200 computer to derive a measure called EGGW-25. This measure is based on the width of the EGG duty cycle at 25% of it height (Scherer & Vail, 1988).

EGGW-25 measures were obtained for nine to ten cycles within the middle two seconds of each of the five to six sustained vowels for each subject. Values for each of the

vowels were averaged together for each session. Intrasubject and measurement reliability were evaluated through Pearson correlation coefficient.

Data Analysis-SPL

The signals from the sound level meter were digitized and analyzed by custom software on a VAX 4000/200 computer to obtain mean and standard deviations of intensity during sustained phonation. In order to eliminate the effects of any silent segments to the SPL analysis, a cursor system was employed to set a floor criterion corresponding to the lowest SPL level during any phase in the recording. The data points from the sound level meter which fell above this level were analyzed for central tendency and variation.

Results

Reliability

Correlation coefficients reflecting intrasubject test-retest reliability for EGGW-25 revealed pre1 to pre 2 correlations of .68 and post 1 to post 2 correlations of .94. For SPL, pre1 to pre2 correlation values were .74 and post1 to post 2 correlations were .98. These correlations are based upon seventy-five percent of the subjects completing a second pre-treatment voice recording and forty-one percent of the subjects completed a second post-treatment voice recording. Twenty percent of the EGGW-25 data were reanalyzed to assess measurement reliability. Interexaminer reliability was .99 for pre and post values. Intraexaminer reliability was 1.0 between original and reanalyzed data. Twenty percent of the SPL data were reanalyzed to assess measurement reliability. Intraexaminer reliability was .99 for pre and post values.

Variables

EGGW-25: The goal of this study was to evaluate treatment-related changes in EGGW, a measure of glottal adduction. Means and standard deviations for EGGW-25 are presented in Table 2. A repeated measures analysis of variance of EGGW-25 (with treatment group as a between-

Table 2.
Means and Standard Deviations of EGGW Values for Subjects Pre- and Post-Treatment

		PRE 1		PRE 2		POST 1		POST 2	
subject	treatment	mean	SD	mean	SD	mean	SD	mean	SD
TB	Respiratory	0.553	0.03	0.523	0.02	0.543	0.04	0.519	0.06
PC	Respiratory	0.508	0.03	0.485	0.03	0.493	0.02		
JF	Respiratory	0.629	0.04			0.650	0.04		
CH	Respiratory	0.599	0.01	0.521	0.03	0.634	0.07		
BK	Respiratory	0.560	0.00	0.537	0.00	0.534	0.00		
JL	Respiratory	0.627	0.03	0.513	0.02	0.489	0.03	0.495	0.04
FN	Respiratory	0.492	0.06	0.581	0.01	0.530	0.04		
KR	Respiratory	0.514	0.01	0.466	0.01	0.511	0.04		
SS	Respiratory	0.776	0.05	0.749	0.02	0.640	0.04	0.688	0.02
TS	Respiratory	0.588	0.01	0.552	0.02	0.526	0.03	0.538	0.01
BT	Respiratory	0.528	0.02	0.519	0.03	0.499	0.01		
DB	Voice & Resp.	0.593	0.02			0.538	0.02		
ED	Voice & Resp.	0.451	0.02			0.617	0.02	0.576	
DH	Voice & Resp.	0.502	0.03	0.623	0.02	0.574	0.02		
EI	Voice & Resp.	0.582	0.06			0.708	0.02		
NJ	Voice & Resp.	0.570	0.03	0.546	0.01	0.653	0.02	0.622	0.01
RL	Voice & Resp.	0.488	0.06			0.538	0.01		
MM	Voice & Resp.	0.517	0.04	0.492	0.04	0.587	0.02		
DP	Voice & Resp.	0.590	0.09	0.624	0.04	0.628	0.03	0.660	0.01
GS	Voice & Resp.	0.450	0.05	0.550	0.10	0.452	0.04	0.429	0.04
GSE	Voice & Resp.	0.570	0.03			0.669	0.01		
PS	Voice & Resp.	0.676	0.05	0.709	0.02	0.733	0.03		
HT	Voice & Resp.	0.603	0.06	0.602	0.06	0.641	0.02	0.673	0.03
HW	Voice & Resp.	0.575	0.02	0.558	0.03	0.632	0.02	0.631	0.00

subjects factor) revealed significant pre- to post-treatment group interaction (F(1,22) =12.00, p=.002). The average change for the LSVT group was .062 (S.D. = .054). The average EGGW-measured change for the R group pre to post-treatment was -.030 (S.D. = .060). There was no main effect for pre-to post-treatment. Increased values for pre to post treatment changes in vocal fold adduction for the LSVT group (reflecting more adduction) are shown in Figure 2. One subject in the LSVT, GS, did not show any notable change in adduction pre- to post-therapy. In the R group, the values for EGGW-25 generally decreased (reflecting less adduction) from pre-to post-treatment. Three subjects (JF, CH and FN) showed increases in adduction and did not follow the general trends of the R treatment group.

SPL: Means and standard deviations for SPL are presented in Table 3 (see following page). A repeated measures of analysis of variance showed a significant effect for treatment type (F (1,22) =89.16, p=.000). The average change in SPL for the R group pre to post-treatment was -1.71 dB (S.D. = 2.92); the average change for the LSVT group was 13.2 dB (S.D. = 4.59). Increased values for pre to post treatment changes in mean sound pressure level for the LSVT group (reflecting more adduction) are shown in Figure 3. Values for the R only group generally decreased or remained the same following therapy. Three subjects (PC, JF and FN) showed increases in intensity and did not follow the general trends of the R treatment group.

A statistically significant correlation was measured between change in SPL, the primary treatment variable, and change in EGGW (r = .7185, p=.000). Subjects who had greater increases in vocal fold adduction had increases in SPL.

Discussion

This study evaluated changes in an electroglottographic measure of vocal fold adduction (EGGW) following two forms of speech treatment in patients with Parkinson disease. EGGW data demonstrated that adduction increased in the LSVT group and suggest that adduction remained constant or decreased in the R treatment group. SPL demonstrated similar findings in that intensity increased statistically significantly in the LSVT group and generally decreased in the R only group. Also, both EGGW-25 and SPL measures correlated highly for both the LSVT and R treatment groups. This supports a significant relationship between pre- to post-treatment changes in intensity and vocal fold adduction, measured by electroglottography.

Reduced loudness in Parkinson disease subjects is generally attributed to bowed vocal folds and/or hypoadduction. The results indicate that the increases in intensity in the LSVT group could be associated with increases in glottal closure measured by EGGW-25. Similarly, increases in adduction were generally not found in the R group; increases in intensity were also not found.

The results of this study support the findings of Ramig et al. (1991) and Ramig et al. (in press) that intensive voice therapy aimed at increasing vocal fold adduction and respiratory support will increase vocal intensity.

The results of this study contribute to the understanding of vocal fold function in patients with Parkinson disease. Invasive measures, like videostroboscopy, are difficult not only to obtain, but also to analyze, in patients with movement disorders. These measures can also be uncomfortable to the subject. SPL and EGGW, two non-invasive measures, correlated highly with regard to changes in vocal fold adduction pre- to post-treatment. This study supports the use of non-invasive means to examine and assist in more detailed verification of laryngeal function in patients with neurologic disorders.

The possible applications of electroglottography should not be overlooked. Quantifying laryngeal function is an important process in the clinic as well as in research; the less intrusive, simple measurement of the electroglottogram should be included in this process. The role of electroglottography, possibly combined with other measures like SPL, can strengthen the documentation of changes in vocal fold adduction .

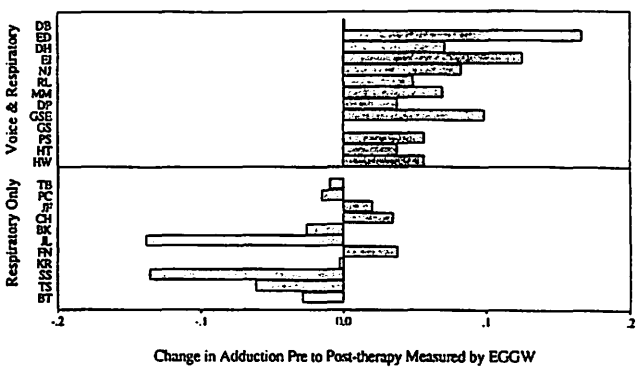


Figure 2. Difference in post 1 and pre 1 EGGW values obtained for each subject.

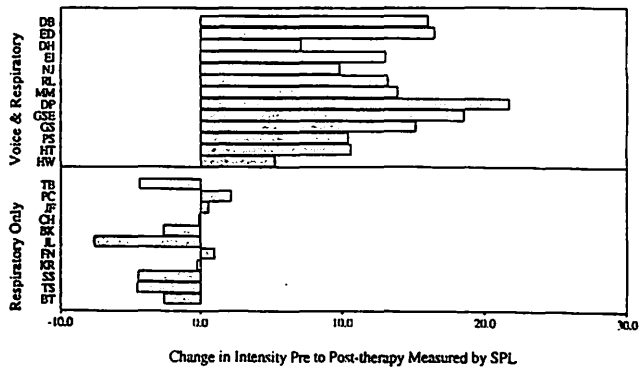


Figure 3. Difference in post 1 and pre 1 SPL values obtained for each subject.

Table 3.
Pre- and Post-Treatment SPL Values for Subjects by Treatment Group

		PRE 1		PRE 2		POST 1		POST 2	
subject	treatment group	mean	SD	mean	SD	mean	SD	mean	SD
TB	Respiratory	74.24	1.16	73.34	1.13	69.89	1.05	69.85	0.72
PC	Respiratory	62.09	0.81			64.23	1.67		
JF	Respiratory	70.91	3.23			71.48	0.4		
CH	Respiratory	65.93	1.47	67.69	0.65	65.84	1.35		
BK	Respiratory	64.84	1.8	66.52	1.27	62.22	1.28	62.54	0.95
JL	Respiratory	72.13	2.12	67.73	1.44	64.51	1.93	66.29	0.82
FN	Respiratory	71.87	1.19	70.11	2.07	72.84	2.47		
KR	Respiratory	66.61	1.18			66.37	1.77		
SS	Respiratory	73.08	0.83	82.73	5.06	68.65	2.28	69.97	1.03
TS	Respiratory	70.12	2.35	67.71	1.53	65.62	0.72		
BT	Respiratory	71.75	0.95	69.47	1.39	69.15	1.19		
DB	Voice & Resp.	65.29	1.37			81.36	0.46		
ED	Voice & Resp.	67.33	0.67			83.84	2.59		
DH	Voice & Resp.	70.89	1.92			77.95	1.49		
EI	Voice & Resp.	63.55	1.19			76.62	0.85		
NJ	Voice & Resp.	75.95	2.47	77.07	2.05	85.79	1.21	83.24	0.84
RL	Voice & Resp.	70.36	2.77			83.56	1.13		
MM	Voice & Resp.	66.44	0.69	69.50	0.81	80.33	1.16		
DP	Voice & Resp.	61.09	0.65	62.96	0.35	82.83	1.73	83.72	1.33
GS	Voice & Resp.	65.59	1.76	63.85	0.68	80.76	0.62	80.57	1.39
GSE	Voice & Resp.	66.90	1.67			85.48	1.44		
PS	Voice & Resp.	72.96	0.84	80.73	0.94	83.4	0.94		
HT	Voice & Resp.	69.07	1.09	70.07	1.51	79.71	1.2	78.33	2.44
HW	Voice & Resp.	75.31	0.42	73.27	1.02	80.5	0.9	84.05	1.73

References

Anastaplo, S., Karnell, M. P. (1988). Synchronized videostroboscopic and electroglottographic examination of glottal opening. Journal of Acoustical Society of America, 83 (5), 1883-1890.

Aronson, A. E. (1985). Clinical Voice Disorders. New York: Thieme-Stratton, 104-106.

Baer, T., Lofquist, A., McGarr, N. S. (1983). Laryngeal vibrations: a comparison between high-speed filming and glottographic techniques. Journal of Acoustical Society of America, 73 (4), 1304-1308.

Biever, D. M., Bless, D. M. (1989). Vibratory characteristics of the vocal folds in young adult and geriatric women. Journal of Voice, 3 (2), 120-131.

Bless, D. M., Hirano, M., Feder, R. (1987). Videostroboscopic examination of the larynx. Ear, Nose, and Throat J., 66 (7), 289-296.

Bless, D. M., Hirano, M., Ford, C. N. Introduction to reading and interpreting videostroboscopic images of the larynx. An instructional videotape.

Childers, D. G., Hicks, D. M., Moore, G. P., Alsaka, Y. A. (1986). A model for vocal fold vibratory motion, contact area, and the electroglottogram. Journal of Acoustical Society of America, 80 (5), 1309-1320.

Childers, D. G., Larar, J. N. (1984). Electroglottography for laryngeal function assessment and speech analysis. IEEE Transactions for Bio. Med.; BME-31 (12), 807-816.

Colton, R. H., Conture, E. G. (1990). Problems and Pitfalls of Electroglottography. Journal of Voice, 4 (1), 10-24.

Critchley, E. M. R. (1981). Speech disorders of Parkinsonism: a review. Journal of Neurology, Neurosurgery and Psychiatry, 44, 751-758.

Darley, F. L., Aronson, A., Brown, J. (1969a). Differential patterns of dysarthria. Journal of Speech and Hearing Research, 12, 246-249.

Darley, F. L., Aronson, A., Brown, J. (1969b). Clusters of deviant speech dimensions in the dysarthrias. Journal of Speech and Hearing Research, 12, 462-469.

Dromey, C., Ramig, L. O., Johnson, A. (in press). Phonatory and articulatory changes associated with increased vocal intensity in Parkinson disease: A case study. Journal of Speech and Hearing Research.

Fex, S., Fex, B., Hirano, M. (1991). A clinical procedure for linear measurement at the vocal fold level. Journal of Voice, 5 (4), 328-331.

Gilbert, H. R., Potter, C. R., Hoodin, R. (1984). Laryngograph as a measure of vocal fold contact area. Journal of Speech and Hearing Research, 27, 173-178.

Haji, T., Horiguchi, S., Baer, T., Gould, W. J. (1986). Frequency and amplitude perturbation analysis of electroglottograph during sustained phonation. Journal of Acoustical Society of America, 80 (1), 58-62.

Hanson, D. G., Gerratt, B. R., Ward, P. H. (1984). Cinegraphic observation of laryngeal function in Parkinson Rs disease. Laryngoscope, 94, 348-353.

Hirano, M., Bless, D. M. (1993). Videostroboscopic Examination of the Larynx. San Diego: Singular Publishing Group, Inc.

Hoehn, M., Yahr, M. (1967). Parkinsonism: onset, progression and mortality. Neurology, 17, 427.

Karnell, M. P. (1989). Synchronized videostroboscopy and electroglottography. Journal of Voice, 3 (1), 68-75.

Kitzing, P. (1990). Clinical applications of electroglottography. Journal of Voice, 4 (3), 238-249.

Lecluse, F. L. E., Brocaar, M. P., Verschuure, J. (1974). The electroglottography and its relation to glottal activity. Folia Phoniat., 27, 215-224.

Logemann, J. A., Fisher, H. B., Boshes, B., Blonskly, E. R. (1978). Frequency and concurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson Rs disease patients. Journal of Speech and Hearing Disorders, 42, 47-57.

Potter, C. R., Gilbert, H. R., Hoodin, R. (1982). Validation of the Laryngograph: a measure of vocal fold contact area. A paper presented at ASHA (Toronto).

Ramig, L. (1995). Speech therapy for Parkinson Rs disease. In W. C. Koller and G. Paulson (Eds.), Therapy of Parkinson Rs Disease (pp. 539-548) New York: Marcel-Dekker.

Ramig, L. O. (1992). The role of phonation in speech intelligibility: a review and preliminary data from patients with Parkinson Rs disease. In R. D. Kent (Ed.), Intelligibility in Speech Disorders: Theory, Measurement and Management (pp. 119-156) Amsterdam: John Benjamin.

Ramig, L., Bonitati, C., Horii, Y. (1991). The effectiveness of speech therapy for patients with Parkinson Rs disease. National Center for Voice and Speech Status and Progress Report, 1, 61-86.

Ramig, L. O., Bonitati, C., Lemke, J., Horii, Y. (1994). Voice treatment for patients with Parkinson disease: Development of an approach and preliminary efficacy data. Journal of Medical Speech-Language Pathology, 2 (3), 191-209.

Ramig, L. O., Countryman, S., Pawlas, A. (in press). The Lee Silverman Voice Treatment: A Practical Guidebook for the Treatment of Speech and Voice Disorders in Parkinson Disease. Iowa City, IA: National Center for Voice and Speech.

Ramig, L., Countryman, S., Thompson, L., Horii, Y. (in press). A comparison of two forms of intensive speech treatment for Parkinson Disease. Journal of Speech and Hearing Research.

Ramig, L., Fazoli, K., Scherer, R. (1990, January). Vocal fold adduction in idiopathic Parkinson disease. A paper presented at the Conference on Speech Disorders, San Antonio, TX.

Ramig, L., Gould, W. J. (1986). Speech characteristics in Parkinson Rs disease. Neurologic Consultant, 4, 1-8. Ramig, L., Mead, C., DeSanto, Horii, Y. (1988a). Voice therapy and Parkinson Rs disease. Journal of the American Speech and Hearing Association, 30 (10), 128.

Ramig, L. O., Mead, C., Scherer, R., Horii, Y., Larson, K., Koehler, D. (1988, February). Voice therapy and Parkinson Rs disease: a longitudinal study of efficacy. A paper presented at the Clinical Dysarthria Conference, San Diego.

Ramig, L. O., Scherer, R. C. (1992). Speech therapy for neurologic disorders of the larynx. In A. Blitzer, M. F. Brin, C. T. Sasaki, S. Fahn, K. S. Harris (Eds.), Neurologic Disorders of the Larynx (pp. 163-180). New York: Thieme Medical Publishers, Inc.

Rothenberg, M., Mahshie, J. J. (1988). Monitoring vocal fold abduction through vocal fold contact area. Journal of Speech and Hearing Research, 31, 338-351.

Scherer, R. C. (1991). Physiology of Phonation: A Review of Basic Mechanics. In C. N. Ford, D. M. Bless, (Eds.), Phonosurgery: Assessment and Surgical Management of Voice Disorders (pp. 77-93). New York: Raven Press, Ltd.

Scherer, R. C., Druker, D. G., Titze, I. R. (1988). Electroglottography and direct measurement of vocal fold contact area. In O. Fujimura (Ed.), Vocal Physiology: Voice Production Mechanisms and Function. New York: Raven Press, 279-291.

Scherer, R. C., Gould W. J., Titze, I., Meyers, A. Sataloff, R. (1988). Preliminary evaluation of selected acoustic and glottographic measures of clinical phonatory function analysis. Journal of Voice, 1, 230-244.

Scherer, R. C., Vail, V. (1988). Measures of laryngeal adduction. Journal of Acoustical Society of America, sl, S81 (A).

Scherer, R. C., Vail, V., Rockwell, B. (1993). Examination of the laryngeal adduction measure EGGW. National Center for Voice and Speech Status and Progress Report, 5, 73-82.

Smith, M. E., Ramig, L. O., Dromey, C., Perez, K. S., Samandari, R. (in press). Intensive voice treatment in Parkinson disease : Laryngostroboscopic findings. Journal of Voice.

Titze, I. R. (1984). Parameterization of glottal area, glottal flow, and vocal fold contact area. Journal of Acoustical Society of America, 75 (2), 570-580.

Titze, I. R. (1990). Interpretation of the electroglottographic signal. Journal of Voice, 4 (1), 1-9.

Yahr, M. D. (1992). Parkinsonism. In L. P. Rowland (Ed.), Merritt Rs Textbook of Neurology. Philadelphia: Lea and Febiger, 526.

Acknowledgments

The research presented here was supported in part by NIH grant R01CD01150. Our appreciation is extended to the subjects who participated in this study. The authors also gratefully acknowledge the contributions Ron Scherer and Christopher Dromey for their assistance in this research.

NCVS Status and Progress Report - 8
July 1995, 139-151

Sense of Effort and the Effects of Fatigue in the Tongue and Hand

Nancy Pearl Solomon, Ph.D.
Donald Robin, Ph.D.
Sara Mitchinson, M.A.
Douglas VanDaele, B.S.E.
Erich Luschei, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa

Abstract

Fatigue and increased effort are common symptoms for people with movement disorders and dysarthria, but they are rarely quantified. In an attempt to develop a clinically useful and physiologically meaningful measure of fatigue, we used a task that involves sustaining a target effort level without visual feedback while squeezing a bulb connected to a pressure transducer. In the first experiment, 12 healthy young adults performed the constant-effort task with the tongue and the preferred hand at 3 submaximal levels of effort. The resulting pressure declined over time as a negative exponential function with a nonzero asymptote. In the second experiment, 6 subjects performed the constant-effort task before and after acutely fatiguing the tongue and hand. The rate of pressure decline was significantly greater after fatigue. One possible mechanism for the characteristic negative exponential function is that it reflects a constant descending drive from higher centers in the CNS to the appropriate motoneuron pools. Thus, this technique may elucidate the contribution of central fatigue to normal and disordered speech.

People with neurologically based movement disorders often report fatigue and increased effort associated with speech. Kent, Kent, & Rosenbek (1987) have argued that understanding fatigue in relation to speech may be important for improving our ability to manage patients with neuromotor speech disorders. However, few studies have systematically described or quantified fatigue related to speech. A few studies on fatigue of the muscles of speech articulation, using an endurance (or fatigue) paradigm are available (lips: Amerman, 1992; Van Boxtel, Goudswaard, Van Der Model, & Van Den Bosch, 1983; jaw: Lyons, Rouse, & Baxendale, 1993; Mao, Stein, & Osborn, 1993; Van Boxtel et al., 1983; tongue: Scardella, Krawciw, Petrozzino, Co, Santiago, & Edelman, 1993). Studies conducted in our laboratory have added to the literature on endurance of the tongue in a variety of subject groups, including normally speaking children (Robin, Somodi, & Luschei, 1991) and elderly adults (Lorell, Robin, Somodi, Solomon, & Luschei, 1992), subjects who had extraordinary speaking and "tonguing" skills (Robin, Goel, Somodi, & Luschei, 1992), and subjects with motor speech disorders (Lorell, Robin, Solomon, & Luschei, 1993; Robin et al., 1991; Solomon, Lorell, Robin, Rodnitzky, & Luschei, 1995; Solomon & Robin, 1994; Solomon, Robin, Lorell, Rodnitzky, & Luschei, 1994; Solomon, Robin, & Luschei, 1994; Stierwalt, Robin, Solomon, Weiss, & Max, in press).

The scarcity of studies on speech-related fatigue is understandable; this is a difficult topic to study for a variety of reasons. The muscles of interest often are inaccessible or difficult to isolate, physiologic needs (such as ventilation or swallowing) may interfere with the task, and noninvasive, comfortable techniques have been lacking. We have used instrumentation that eliminates most of these limitations: the Iowa Oral Performance Instrument (IOPI, Breakthrough, Inc.). The IOPI transduces the pressure exerted on a pliable air-filled bulb by a given body part. The IOPI is more comfortable for sustained tasks with the tongue than, for example, a strain-gauge force transducer which involves pressing the tongue against a thin rigid "arm" (Abbs, 1990; Barlow & Abbs, 1983). In addition to the tongue, we routinely assess endurance in the hand. Although there are important biomechanical and physiological differences between the tongue and the hand, we have found assessment of the hand to be useful for comparative and training purposes. Interpretations of differences between the structures, however, must account for these inherent differences.

Because fatigue is typically defined as a failure to maintain the required or expected force (Edwards, 1981),

the endurance task we (and others; e.g., Fuglevand, Zackowski, Huey, & Enoka, 1993; Lyons et al., 1993) have used should be an accurate index of fatigue. It involves providing subjects with visual feedback (via the LED display on the IOPI) and instructing them to maintain a target pressure as long as they can. We have used a submaximal pressure level for this task, usually 50% of the pressure attained during a brief maximal effort. Most neurologically normal adult subjects terminate tongue endurance trials at this level after approximately 30 to 40 seconds, and hand endurance trials after 50 to 60 seconds. Trials are terminated because subjects are unwilling or unable to continue due to a perception of extreme effort. Other feelings of discomfort (soreness, tingling, dull pain) may be experienced as well. Mild muscle soreness for 1 to 2 days following testing occurs on occasion.

For a quantitative assessment of fatigue to be useful clinically, it needs to be valid, reliable, noninvasive, and, preferably, nonaversive to the patient. The endurance task we have used appears to meet all of the criteria listed (Robin et al., 1991) except possibly the last. Not all patients like the endurance task (although some enjoy the competition), but it is tolerable and symptoms of discomfort are quite transient. Another drawback to this task is that it does fatigue the tongue and hand, and no repeated trials or other measures of these structures can be made until after a substantial rest period.

The task reported in this article resulted from our attempt to develop a measure of fatigue that would be clinically useful, non-aversive, and physiologically meaningful. We hypothesized that because effort increases when force or pressure is held constant (Enoka & Stuart, 1985, 1992; Jones, 1983), force or pressure should decrease when effort is held constant. In fact, it has been demonstrated previously that this is the case (Cain & Stevens, 1971, 1973; Eason, 1959; Jones & Hunter, 1983b). The characteristics of the decrease in force or pressure (e.g., rate, shape of curve) should reflect the process or processes responsible for fatigue.

Fatigue processes usually are dichotomized as peripheral or central, referring to the site at which failure to maintain a certain output (e.g., force or pressure) occurs. The term peripheral fatigue is used to characterize failure to maintain an output resulting from sites below the spinal cord, and central fatigue is used to describe the mechanisms for fatigue resulting from failure in the spinal cord and supraspinal centers (McCloskey, 1981; McCloskey, Gandevia, Potter, & Colebatch, 1983; see reviews in Gandevia, 1987, 1992). The sense of effort is thought to be a supraspinal phenomenon, apparently related to an awareness of the descending motor drive by central feedback or efference copy (Enoka & Stuart, 1985; Gandevia, 1982, 1987; Matthews, 1982; McCloskey, 1981). Thus, the increased effort perceived during endurance tasks may reflect an increased descending drive to motoneurons involved in performing the task. It is important to note that subjects also receive feedback from peripheral somatic afferents (e.g., deep muscle afferents, cutaneous nerve endings) which may have affect performance on fatigue tasks as well.

If sense of effort is a reflection of descending motor drive, then it is possible that a constant-effort task is akin to maintaining a constant descending motor drive to the motoneuron pool. Assuming that a constant descending drive results in constant synaptic depolarizing current to motoneurons, an exponential function is expected. Evidence for this was provided in a study by Kernell and Monster (1982). They delivered a constant intracellular injection of depolarizing current to motoneurons of the cat spinal cord. The firing rate of the motoneurons decreased exponentially with time but did not cease firing completely. If human subjects mimic this experimental situation by keeping a constant sense of effort, one might expect an exponentially declining pressure as a consequence of decreases in motoneuron firing rate. Indeed, previous studies of constant-effort contractions of the hand and arm starting at a variety of submaximal force levels have described the resulting force curves as declining exponentially (Cain & Stevens, 1971, 1973; Eason, 1959; Jones & Hunter, 1983b). From our pilot studies, we also found that pressure curves decreased exponentially to a non-zero asymptote when submaximal effort was held constant by the tongue as well as the hand.

We conducted two experiments to examine the effects of constant effort on exerted pressure. The purpose of the first experiment was to model the pressure function generated by two body structures (tongue and hand) during constant-effort tasks across three starting levels (25%, 50%, and 75% of maximal pressure) by an exponentially declining curve to a nonzero asymptote. To explore whether characteristics of the pressure curves would reflect fatigue, the second study was conducted to examine the effects of acutely fatiguing the muscles.

Experiment 1: Model of Constant Effort

Method

Subjects

Twelve healthy adults, 6 men and 6 women, served as subjects. They were between the ages of 18 and 23, had negative histories for speech or movement disorders, and did not participate in activities that required excessive use of the hand or tongue. They had no previous experience with the instrumentation or procedures used in this study. They were recruited from an undergraduate dormitory at the University of Iowa, provided informed consent, and were paid for their participation.

Instrumentation

To measure the pressure exerted by the hand and the tongue, we used the IOPI, described previously in this journal (Robin et al., 1991). Briefly, the IOPI consists of a

pressure transducer with a digital display and an LED display. Pressure is exerted on an air-filled bulb.[1] The tongue bulb was placed on the hard palate just behind the alveolar ridge. The subject was instructed to push against the bulb in a rostral direction with the anterior dorsum of the tongue. The teeth rested lightly on the shaft of the bulb. The hand bulb was placed in the palm of the hand and the subject was instructed to squeeze it with a fist configuration.

Maximal pressure data were determined from the digital display on the IOPI. The pressure-sensing and peak-holding circuitry of the IOPI can follow pressure pulses that rise from 0 to 100 kPa in 30 ms, dynamics which are more than adequate for registering peak pressures generated by normal subjects. Pressure data for the constant-effort trials were digitized to a personal computer using data acquisition software (DATAQ Instruments, Inc., Akron, OH) at a sampling rate of 4 Hz, a rate deemed adequate for tracking these relatively slowly changing functions.

Protocol

Each subject participated in 3 data collection sessions within approximately one week. Only one examiner (NPS) conducted all sessions for all subjects so that consistency in data collection was maintained. Typically, the examiner and two subjects were in the room.

During each session, measures of maximal strength for the tongue and the hand (order alternated between subjects and sessions) were collected first. Subjects were instructed to "squeeze the bulb as hard as you can." The greater pressure generated in 2 trials was taken as the maximum. Next, constant-effort trials for the tongue and hand, again with the order alternated for subjects and sessions, were performed. The IOPI's LED display was adjusted so that a green light in the middle of the display represented either 25%, 50%, or 75% of the maximum pressure. Instructions were: "Get the green light to go on, then close your eyes and keep the effort the same. Think about how difficult it was to make the green light go on, and then don't let it get any easier or any harder." Note that visual feedback was not provided. Subjects were reminded several times to "keep your effort the same; don't let it get harder or easier" during the trial. Trials were terminated after a constant pressure appeared to be maintained for 5 to 10 s. This was considered to represent the asymptote for the curve. Trials typically lasted 20-30 s for the tongue, and 40-60 s for the hand. Rest periods of 3 to 5 minutes were required after each trial. The order of pressure levels was counterbalanced for each subject and for each session. Subjects performed the task once at each of the 3 effort levels for each structure (tongue and hand) during each of the 3 sessions.

[1] First generation clear silastic tongue bulbs were used in this and all previously published studies from our laboratory.

Data Analysis

The digitized pressure curves resulting from the constant-effort task were transferred to MATLAB, a computer program for mathematical analysis. Each pressure curve was optimally fit to the exponential model:

$$F(t) = \exp(b - a*t) + c$$

where F(t) represents the pressure resulting from IOPI bulb compression as a function of time, t is time, a is the rate of pressure decrease, b is a parameter that represents the natural log of the y-intercept (value of F(t) at t=0) minus the asymptote value, and c is the asymptote or residual pressure. Thus, the y-intercept of the fitted model can be determined by taking the exp(b) and adding the asymptote (c). An idealized curve of this function is illustrated in Figure 1.

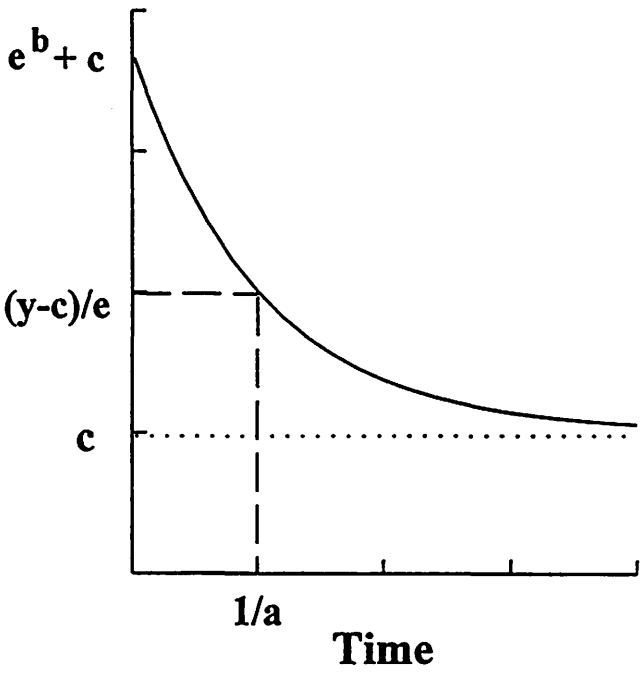


Figure 1. Equation used for the exponential model and idealized function. See text for explanation of terms and measures.

Initial estimates of the parameters in this model were provided for the curve fitting procedure (fmins in MATLAB). The c parameter was estimated as the mean of the last 5 s of the pressure curve. To estimate the a and b parameters, the estimated c parameter was subtracted from the pressure curve and the natural logarithm of that data was taken to generate the model:

$$\ln(F(t)-c) = b - a*t.$$

Four points were selected to estimate the general shape of the pressure curves: the second and third data point values for the initial phase of the trial, and the median and

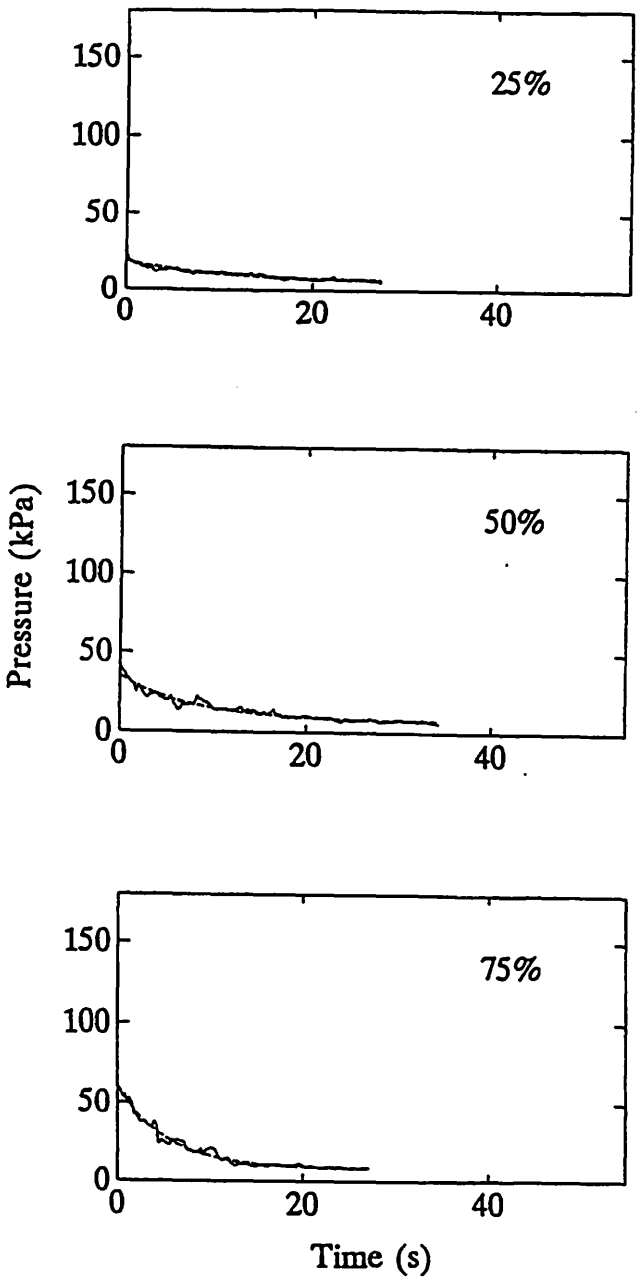


Figure 2. Constant-effort trials by Subject M2 with the tongue for initial levels of 25%, 50%, and 75% of maximum pressure. The solid lines are the actual data; the dashed lines are the fitted exponential curves.

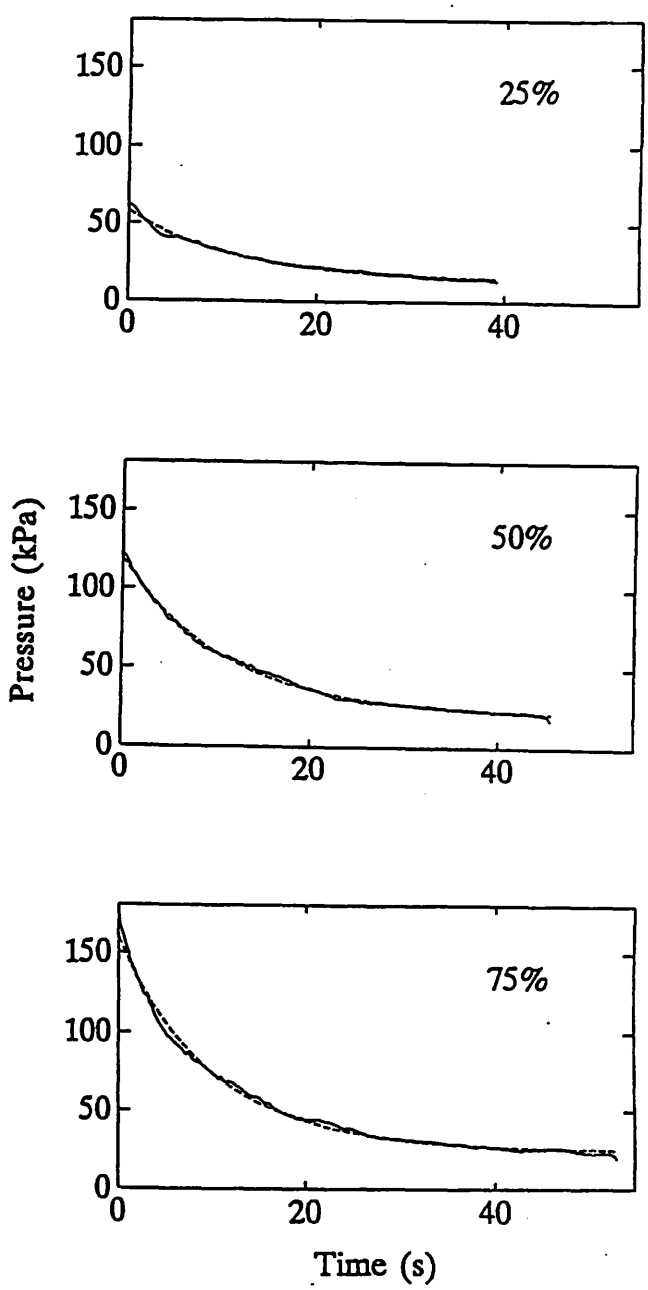


Figure 3. Constant-effort trials by Subject M2 with the hand for initial levels of 25%, 50%, and 75% of maximum pressure. The solid lines are the actual data; the dashed lines are the fitted exponential curves.

median-plus-one data points for the middle phase of the trial. In each case, two points rather than one were used to account for variability. The late phase of the trial was estimated previously by the c parameter.

The resulting estimates of the a and b parameters, along with their associated time values, as well as the c parameter estimate were substituted into the exponential model to generate a fitted curve with a time base identical to that of the pressure curve. The sum of squares function was then used as a measure of the difference between the pressure curve and the fitted curve generated by the parameter estimates. The sum of squares function was then iteratively optimized by varying the values of a, b, and c until the difference between the pressure and fitted curves was at a mathematical minimum. The correlation coefficient was then calculated between the pressure curves and the fitted curves. This optimization method provided the mathematically optimal fit of the exponential model to the data, and as such, the correlation coefficient between the pressure curve and the fitted curve could be utilized as a measure of how

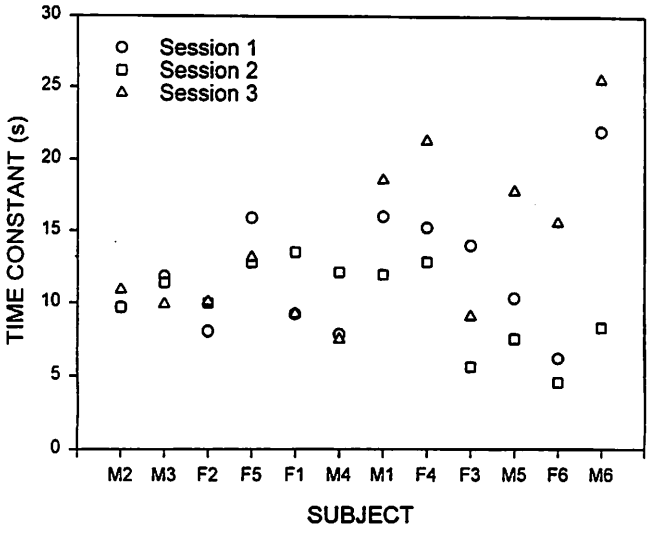


Figure 4. Variability in the time constants. Each symbol represents the time constant for the fitted exponential function for each of 3 constant-effort trials per subject. These trials were performed by the hand at the initial level of 50% of maximum pressure ("50% effort").

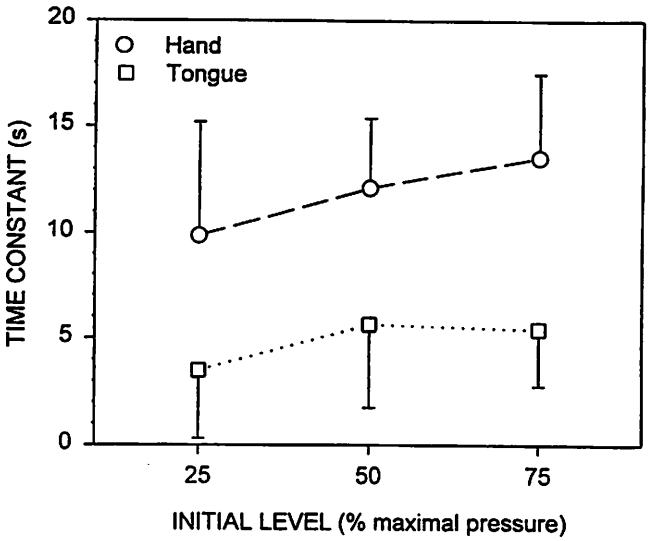


Figure 5. Average values (and -1 or +1 standard deviation across subjects represented by the error bars) for the time constants of the functions fitted to constant-effort curves for the tongue (squares) and hand (circles) at initial levels of 25%, 50%, and 75% of maximum pressure.

well the pressure curve fit the exponential model. An $r^2 > 0.70$ was considered to indicate a good fit.

For this study, the dependent variables of interest were the rate of pressure decay which is represented by time constant (1/a), and the residual pressure (c) as a percentage of the y-intercept (i.e., c normalized for starting level of the fitted curve, c_n). Repeated-measures analyses of variance with session (1, 2, 3) and level (25%, 50%, 75%) as within-subjects factors were conducted for each structure to examine for order effects. Because no significant order effects were detected (see Results for details), data were averaged over the three sessions. Then, repeated-measures analyses of variance with 1 between-subjects factor (gender) and 2 within-subjects factors (structure and level) were conducted. Tukey's studentized range test was used (95% confidence intervals) for post-hoc comparisons. A probability level less than 0.05 was considered statistically significant.

Results

A total of 216 curves (12 subjects, 2 structures, 3 starting levels, 3 trials) were generated from the constant-effort task, and 80% (n=173) of these were described well by the exponential model. Examples of data that were fit nearly perfectly by the model are provided in Figures 2 (tongue) and 3 (hand). These curves were generated by one subject at each of the 3 starting levels (25%, 50%, 75%).

Of the 20% (n=43) of the curves that were eliminated from further analysis, 33 did not meet the criterion of $r^2>0.70$. The additional curves were rejected because the fitted curve resulted in a negative asymptote or an asymptote that was substantially lower than the actual data, and/or the time constant was greater than the duration of the trial. Failed optimization attempts resulted from trials that were terminated too soon (i.e., before an asymptote was established), poor initial estimates of the parameters a, b, c, or curves that were not exponential in shape. Most of the eliminated data (n=24) was for the tongue at the 25% level; 11 of the eliminated curves were for the tongue at 50%; 3 were for the tongue at 75%. For the hand, 4 curves were eliminated at the 25% level, 0 at 50%, and 1 at 75%.

Individual time constants for each subject and trial for the hand at the 50% level are illustrated in Figure 4. This graph was chosen because there were no eliminated trials. About half of the subjects were remarkably consistent across the 3 days of data collection. The others showed more variability. The apparent trend for time constants to be greater for session 3 did not occur for the other levels of initiation for the hand or for the tongue. Analysis of the time-constant data revealed no systematic order effect for the tongue [$F(2,50)=0.14$, $p=.8710$] or hand [$F(2,82)=1.74$, $p=.1817$). Therefore, data for each subject were averaged across the 3 sessions for further analyses.

Average time constants for the 12 subjects for each structure and each effort level are illustrated in Figure 5; error bars represent 1 standard deviation either above or below the mean. Statistical analysis examined the effect of gender, structure and level. No significant interactions were found between any variables. Likewise, no statistically significant differences were detected for gender [$F(1,10)=1.063$, $p=.3269$] or level [$F(2,44]=2.55$, $p=.0896$]. A significant difference between structures was found such that the time constant was greater for the hand than for the tongue [$F(1,44)=56.57$, $p<.0001$; see Table 1].

Table 1.
Results from Experiment 1.
Means (and standard deviations) for the two variables time constant (TC; 1/a in the exponential model) and normalized asymptote (c_n; c adjusted as a percentage of the y-intercept in the exponential model) for the tongue and hand at the 3 starting levels (expressed as a percentage of maximal pressure). The exponential models were obtained by fitting pressure curves generated during constant-effort tasks to a negative exponential equation with a nonzero asymptote (see text for details).

TONGUE

Level:	25%				50%				75%			
Variable:	**TC**		c_n		**TC**		c_n		**TC**		c_n	
Women	2.89	(1.03)	37.52	(11.67)	5.67	(4.05)	36.18	(11.53)	5.04	(3.07)	37.04	(17.90)
Men	4.06	(4.80)	56.69	(14.19)	5.61	(4.13)	41.44	(11.86)	5.75	(2.39)	41.60	(12.37)
Combined	3.48	(3.17)	47.10	(15.66)	5.64	(3.90)	38.81	(11.48)	5.40	(2.65)	39.32	(14.86)

HAND

Level:	25%				50%				75%			
Variable:	**TC**		c_n		**TC**		c_n		**TC**		c_n	
Women	7.75	(2.06)	54.12	(12.99)	11.47	(3.07)	37.48	(8.66)	12.54	(4.93)	29.16	(3.13)
Men	11.98	(6.91)	54.69	(11.15)	12.73	(3.62)	42.14	(8.96)	14.47	(2.81)	37.21	(9.57)
Combined	9.87	(5.34)	54.41	(11.54)	12.10	(3.27)	39.81	(8.75)	13.50	(3.96)	33.18	(7.98)

Note. Summary statistics are based on 6 men and 6 women, except n=3 women and n=3 men for the tongue at the 25% level.

Summary statistics for the asymptotes normalized to the y-intercepts of the fitted curves (c_n) are provided in Table 1. Data from the three sessions were averaged for each subject because statistical analysis revealed no significant order effect across sessions for the tongue [$F(2,50)=0.23$, $p=.7975$] or hand [$F(2,82)=0.12$, $p=.8879$). No significant interactions were found between structure, gender, and level, nor were significant main effects found for for structure [$F(1,44)=0.00$, $p=.9648$] or gender [$F(1,10)=1.3929$, $p=.2652$]. Statistically significant differences for c_n were found as a function of level [$F(2,44)=19.95$, $p < .0001$]. Post-hoc testing revealed that c_n was significantly greater at the 25% level than at the 50% and 75% levels.

Summary and Discussion of Experiment 1

Twelve young adult subjects, unremarkable for tongue use or speech, generally reduced pressure exponentially from some predetermined level to a lesser pressure when asked to keep effort constant. This characteristic shape of the response occurred for 80% of the curves generated.

The rate of decay of the exponential function can be described by a time constant (the reciprocal of the variable a in the exponential model). The time constant indicates the duration from the beginning of the trial (specifically, the y-intercept of the modelled curve) to approximately one-third of the total pressure decay to the asymptote. In general, the time constant approximated 12 s for the hand and 5 s for the tongue. The result for the hand is in remarkable agreement with those published for handgrip by Cain and Stevens (1973). The time constants (the reciprocal of their parameter _) reported for their constant-effort levels 2-4 (which appear to represent middle-to-high effort levels) were 12.3, 10.2, and 11.2 s, respectively. It is encouraging that we were able to replicate these time constants even with different instrumentation (IOPI vs. hand dynamometer), measures (i.e., pressure vs. force), and somewhat different initiation levels (Cain & Stevens, 1973, used 4 force levels ranging from 55 to 220 N). Cain and Stevens (1971, 1973) found, however, that a single exponential function was not adequate to describe constant-effort pressure curves, and provided a second exponential term _ to describe the more slowly decaying portion of the curves (1/_ = 67 to 77 s). The present data were fit well with a model including only one exponential term. The difference between models may relate to Cain and Stevens' observation that force continued to decrease over their 2-min trial whereas we were able to determine an asymptote. The kinesthetic feedback afforded by the hand dynamometer versus the IOPI may account for the difference in residual pressure between the studies. Nevertheless, it is noteworthy that the first exponential term in the model by Cain and Stevens agrees with the term used in our model.

Jones and Hunter (1983b) were able to describe their "constant-effort" curves with a single exponential term, and attributed this to the reduction in cutaneous contact during their task (elbow flexion with the transducer attached by a strap to the wrist) in comparison to the

handgrip task used by Cain and Stevens. However, the time constant listed for their model (reciprocal of their term b) was 28 s, reflecting the markedly slower initial decay in their curves than in those published by Cain and Stevens (1971, 1973) and those in the present study. It is unlikely that the differences in time constants between studies was due to initiation levels, because trials by Jones and Hunter (1983b) started at 35%, 50%, and 65% of maximum. These levels are similar to those used in the present study, and we found no statistically significant difference in time constant between the levels. Instead, the difference between studies probably relates to the instructions given. Our instructions were essentially identical to those used by Cain and Stevens ("the contraction never felt any less nor any more difficult than it had at the outset," 1973, p. 122), but not those used by Jones and Hunter ("maintain the contraction so that it felt the same," 1983b, p. 78). In addition, EMG data in the latter study revealed that effort increased, especially during the lowest initial level of force (35% of maximum). Therefore, the task reported by Jones and Hunter (1983b) cannot be considered to reflect constant effort, especially at the beginnings of the trials.

Less consistency of the present data with previous research is apparent for the values of the asymptotes. Eason (1959) was the first to report constant effort trials, and did not model the resulting pressure curves mathematically. However, inspection of the figure provided reveals that constant-effort trials using elbow flexion terminated after 2 min at approximately 65% of the initial force value. In contrast, Cain and Stevens reported a residual factor c that was approximately 20% of initial values for low initial forces and 10% for high initial forces. In the present study, asymptotes typically were one-third to one-half of initial pressure values for the hand or tongue. Differences between studies may relate to instrumentation (specifically, differences in hand configuration and comfort) and the specific muscles involved (between elbow flexion and handgrip).

There are multiple mechanisms, both peripheral and central, that could contribute to the characteristic exponential function that we have observed. Peripherally, fatigue could involve propagation failure along axons to the muscles, blockage at the neuromuscular junction, or problems with contractile properties of the muscles themselves. However, some literature has suggested that these processes take a longer time to develop than the short durations (generally less than 60 s) of the endurance and constant-effort tasks we have used to assess fatigue (Bigland-Ritchie, Johanson, Lippold, & Woods, 1982; Bigland-Ritchie, Johanson, Lippold, Smith, & Woods, 1983; Borg, Grimby, & Hannerz, 1983; Burke, Levine, Tsairis, and Zajac, 1973; Fuglevand et al., 1993), and therefore may not contribute substantially to the present results. Because these studies involved limb muscles, similar experiments involving the tongue muscles are needed to validate the generalizability of this conclusion.

In contrast, some biochemical changes in or related to muscle can occur rapidly during static contractions. These include decreases in blood flow (Sjøgaard, Savard, & Juel, 1988), and breakdown of phosphocreatine (Hultman & Sjöholm, 1986). Whether such metabolic changes in muscle affected performance in the present constant-effort task cannot be determined.

Modifications in central mechanisms that would affect the pressure curve have been discussed in terms of an exponential decrease in motoneuron firing rate due to motoneuronal adaptation to a constant excitatory input (Kernell & Monster, 1982). An additional mechanism that might contribute to pressure decline over time is the derecruitment of motor units, although we know of no data demonstrating motor unit derecruitment occurring in an exponential fashion. Direct evidence for central mechanisms explaining these results is slight, especially for the tongue, but they cannot be ruled out. Cain and Stevens (1973), in a study that examined changes in EMG activity during a constant-effort task, speculated that the mechanism for fast exponential decay at the beginning of trials "does not appear to be a response to the development of fatigue in forearm muscles. Instead, it is probably a response to the buildup of fatigue in some other locus" (p. 126). Jones and Hunter (1983b) suggested that the exponential decay in force related to "some variable proportional to the centrally-generated motor command in estimating effort" (p. 82). Studies of muscle activity, particularly single motor unit activity, during fatigue tasks involving the tongue are needed to address the issue of central fatigue.

Effects of Structure, Gender, and Starting Level

The time constants of the exponential functions were greater for the hand than the tongue. If the time constant reflects fatigue, then this finding is consistent with our previous data on tongue and hand endurance (Lorell et al., 1992, 1993; Robin et al., 1991, 1992; Solomon et al., 1994, 1995; Solomon & Robin, 1994; Stierwalt et al., in press). In those studies, normal and disordered subjects maintained 50% of their maximal strength for nearly twice as long with the hand as with the tongue. Apparently the tongue is more fatigable than the hand for the endurance and constant-effort tasks we have utilized.

A comparison of strength and fatigue of the upper and lower lips by Amerman (1992) revealed that the lower lip was stronger and less fatigable than the upper lip in one normal subject. He speculated that the difference in fatigability may relate to central processes (e.g., "learning") involving activation of motor units within (Krantz, 1981) and between (Lindström & Hellsing, 1983) muscles to compensate for fatigue. Perhaps the handgrip, with its relatively great strength and the involvement of various muscles of the arm, hand, and fingers, is more likely to demonstrate longer time constants in our model than tongue

elevation because of these mechanisms. That is, although histochemical studies of the tongue in rats and cats indicate that it is composed of fatigue resistant motor units (Hellstrand, 1980; Sato, Suzuki, Sato, Sato, & Inokuchi, 1990), the tongue may have fewer motor units and synergistic muscles to share the task. It should be recognized that the biomechanics of the hand flexor and tongue elevator muscles are so different that it may not be possible to compare muscle contractions on the basis of muscle physiology alone. The serial arrangement (i.e., tendon-muscle-tendon) of the forearm and hand flexor muscles contrasts sharply with the tongue as a muscular hydrostat. However, the finding of negative exponential functions for both structures suggests that the neural drive to the muscles for the constant-effort task is similar.

In every case, the men in this study had greater hand strength than the women. A similar difference has been reported in other studies of handgrip (e.g., Clarke, 1986). However, no gender difference was found for time constant or the normalized asymptote. Therefore, the effect of absolute strength on the variables derived from the constant-effort task may not be a concern, at least for neuromuscularly normal adults. This observation is encouraging, especially when viewing this task as a clinical tool. That is to say, the development of normative data will not have to account for gender or strength.

We expected to find a difference in the time constant across the three starting levels such that it would be smaller for the higher pressure levels. This hypothesis followed from the work of Kernell and Monster (1982), who demonstrated by injecting current into motoneurons with different muscle fiber types for the cat gastrocnemius, that fast-twitch fatigable (FF) motor units had smaller time constants than slow-twitch (S) units. At the highest starting level used in the present study, we assume that FF units are being recruited. Thus, the decay of the constant-effort curve would be expected to be faster. However, no statistically significant differences across the three levels were found. From inspecting the summary statistics, the only apparent difference was that the time constant tended to be smaller for the 25% level in comparison to the two higher levels. Unfortunately, most of the data that were not fitted by the exponential model were for the 25% level, especially for the tongue. Because the physiologic range of tongue function during speech probably encompasses 25% or less of the maximum physiologic capacity, it behooves us to test relatively low pressures produced by the tongue to examine the relevance of this task to speech dynamics.

The finding that c_0 was greater at the 25% level than at the 2 higher levels probably relates to the type of motor unit active during the task. The starting pressure for the 25% level was so low, perhaps only slow nonfatigable motor units were recruited. At this level, the maintenance of the amount of effort needed may have resulted in a pressure level that was relatively close to the initial pressure. Because the asymptote was normalized to the initial pressure, the variable c_0 was relatively high at the 25% starting level.

Although variability across sessions was evident, this variability was not found to be systematic. This implies that there were no apparent training or fatigue effects. The study was carefully designed to randomize trial order, and to allow for adequate rest periods within each session and adequate time between sessions to avoid fatigue. Thus, we were encouraged by not finding a trial effect. However, in situations where time constraints are more rigid, it may be necessary to perform several trials of this task within a relatively short period of time. In such situations the effects of fatigue need to be considered carefully. Future studies may determine that a certain number of trials conducted within a certain period of time would not fatigue a subject, and that averaging the results from these trials might result in a meaningful and valid measure.

One way to limit the number of trials needed to obtain useful results would be to limit the number of starting levels used. In Experiment 1, there was no significant differences in the time constant across the 3 levels tested. In addition, the 50% level worked well in that subjects appeared most comfortable with that level (it was easier to "sense" than the 25% level, and less tiring than the 75% level). Finally, there were few "failures" of the curves produced at the 50% starting level to be fit by the exponential model. Therefore, this middle level, although not necessarily representative of the physiologic operating range for typical motor activities, may be considered the most useful of the three levels used in Experiment 1 for experimental and, possibly, clinical purposes.

Data that Did Not Fit the Exponential Model

As stated previously, the demonstrated relationship between effort and pressure, that when the former is maintained the latter will decrease exponentially to a nonzero asymptote, was true for 80% of our data. When considering the 20% of the data that could not be fit by the model, we could ascertain no specific explanation. It is possible that subjects inadvertently released pressure, swallowed, or performed some other interfering movement of the tongue. However, when such interfering behaviors were detected, trials were repeated. There was no particular mathematical function that was a good fit for the eliminated curves. Some of these curves appeared linear, others appeared essentially random (i.e., "noisy"). Some of the data may not have fit the model due to early termination of the trial by the experimenter. In later experiments, we have been careful to allow the trials to continue until an adequate asymptote is obtained.

Although a portion of the data could not be fit by the exponential model, it should be emphasized that the majority of curves produced by every subject was described by the model. Thus, it is not the case that constant effort held different meanings for certain subjects. It is true, however, that some subjects appeared to have more difficulty with the task than others (e.g., 6 out of 18 curves were eliminated for subject F6 whereas none were eliminated for Subjects F3 and M2).

Experiment 2

The purpose of Experiment 2 was to characterize differences in the pressure curve resulting from constant efforts by the tongue and the hand before and after acute fatigue. First, we hypothesized that the time constant of the resulting exponential functions would be smaller for the pressure curves generated after fatigue because central drive to the motoneuron pool would be greater (McCloskey et al., 1983). Second, if force generation is affected by acute fatigue, then given the same starting level, we might expect the asymptote to be smaller after fatigue.

Method

Six of the subjects from Experiment 1, 3 women (subjects F1, F2, and F3) and 3 men (subjects M1, M2, and M3), participated in Experiment 2 one month later. Subjects were, therefore, familiar with the constant-effort task. One investigator (SIM) provided instructions and collected the data for Experiment 2. The instrumentation was the same as before, as were the procedures for obtaining strength measures for the tongue and hand.

Maximum pressure for each structure was determined at the beginning of the session. The constant-effort task at an initial level of 50% of maximum pressure, as described previously, was performed first by the hand and then the tongue. Next, the hand was fatigued by rapid (approximately 1/s) repetitive maximal efforts of squeezing the hand bulb until the subject failed to achieve at least 70% of the maximal referent pressure on 3 successive trials. This criterion level was selected to ensure fatigue without compromising the subjects' ability to perform the subsequent constant-effort task. Loud verbal encouragement and visual feedback were provided. The subject was then told to stop repetitive efforts, to immediately generate 50% of the original maximal pressure, close his or her eyes, and begin the constant effort trial. Finally, after a rest period, the fatiguing procedure and post-fatigue constant-effort trial were performed with the tongue.

Pre- and post-fatigue pressure curves were fit with the negative exponential function described in Experiment 1. The resulting time constants and asymptotes normalized to the y-intercept of the fitted curves were analyzed with a repeated-measures analysis of variance with two within-subjects factors (fatigue and structure).

Results

All subjects were able to perform the constant effort tasks and achieve fatigue. All curves but one were described well by the negative exponential function. The one rejected curve was from a pre-fatigue trial with the tongue. Instead, we used a curve generated by this subject (F3) during Experiment 1 as the pre-fatigue trial.

Individual subject data and summary statistics for the time constant and normalized asymptote are listed in Table 2. In addition, the mean time constants (and standard deviations as error bars) for each structure before and after fatigue are plotted in Figure 6. No interaction between fatigue and structure [$F(1,15)=0.23, p=.6351$] was found for time constant. The time constant was significantly smaller after fatigue [$F(1,15)=9.13; p=.0086$]. This difference held for all 6 subjects for the tongue and 5 of the 6 subjects for the hand. The time constant was significantly smaller for the tongue than the hand [$F(1,15)=16.06, p=.0011$].

Analysis of the normalized asymptote revealed no significant interaction [$F(1,15)=3.44; p=.0833$] or main effects [fatigue: $F(1,15)=3.03, p=.1021$; structure: $F(1,15)=1.85, p=.1937$]. It is clear from the individual data (Table 2) that there was no systematic relationship for the normalized asymptote to change with fatigue for the tongue. However, 5 out of 6 subjects had lower asymptotes after fatigue with the hand. Perhaps if a larger number of subjects were included in the study, a statistically significant difference would have been revealed.

Summary and Discussion of Experiment 2

After acutely fatiguing the tongue and hand, pressure produced during the constant-effort task (starting at 50% of the maximum pressure) decreased more quickly than before fatigue (i.e., the time constant of the exponential function was smaller). This observation was extremely consistent; it held for 11/12 pre- versus post-fatigue comparisons.

In the few studies that have examined the effect of fatigue on sense of effort, a constant relationship between perceived effort before and after fatigue has been shown (Cain & Stevens, 1971; Jones & Hunter, 1983a). That is, the perceived intensity of contractions is heightened (by a simple factor of magnitude) after fatigue, but the nature of the relationship is the same as it was before fatigue (e.g., exponential). Although Cain and Stevens (1971) used constant-effort trials to fatigue the hand, they did not examine the effect of fatigue on constant-effort trials as was done here.

It should be noted that the starting level used for the constant-effort task corresponded to 50% of the maximum pressure generated with rested muscles (i.e., before fatigue). We have assumed implicitly, at least in the terminology that has been used thus far, that this level is comparable to 50% of maximal effort. However, it has been demonstrated

Table 2.
Results from Experiment 2.
Individual subject data for the variables time constant (TC; 1/a in the exponential model) and normalized asymptote (c_n; c adjusted as a percentage of the -intercept in the exponential model) for constant-effort curves performed by the tongue and hand before and after acute fatigue.

TONGUE

Subject	TC-before	TC-after	c_n-before	c_n-after
F1	3.96	2.38	40.05	36.22
F2	2.87	1.47	25.37	36.91
F3	2.35	1.02	34.03	40.13
M1	11.88	3.14	23.89	15.29
M2	7.67	1.20	52.53	46.39
M3	11.51	6.97	32.55	35.55

HAND

Subject	TC-before	TC-after	c_n-before	c_n-after
F1	6.04	7.62	31.70	21.77
F2	7.30	5.63	39.98	30.48
F3	9.66	7.55	41.37	28.69
M1	10.85	8.35	7.36	3.51
M2	17.62	12.36	40.38	41.93
M3	12.98	5.52	56.61	25.55

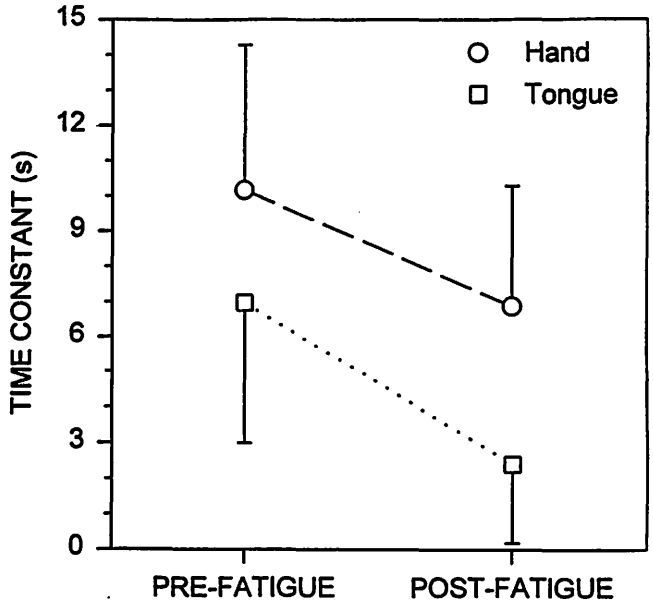


Figure 6. Average values (and -1 or +1 standard deviation across subjects represented by the error bars) for the time constants of the functions fitted to constant-effort curves performed before and after a fatiguing task. Trials were performed at an initial level of 50% of maximum pressure (determined before fatigue), and are shown for the tongue (squares) and the hand (circles).

previously that there is not a linear, absolute relationship between effort and pressure (Somodi, Robin, & Luschei, in press). In fact, based on average data from healthy young adults, 25%, 50%, and 75% of maximum pressure corresponds to approximately 15-20%, 60-65%, and 85-90% effort, respectively, for the hand, and less than 5%, 25-30%, and 80-85% effort, respectively, for the tongue (Somodi et al., in press). Because the perception of effort required to perform the same amount of work increases after fatigue, the effort level at 50% of maximum pressure was probably even higher after our subjects were fatigued. To understand how this potential difference in sense of effort before and after fatigue affected the results, it would have been instructive to ask subjects to rate effort after performing the constant-effort tasks.

We predicted that the residual pressure, or asymptote, of the pressure curves would be smaller after fatigue because of a reduction in force generation capacity. However, the asymptote of the exponential function did not change systematically after fatigue. It is possible that the residual pressure was low enough that it was within the capacity of the muscles even when fatigued (recall that the muscles were fatigued only to the inability to generate 70% of the maximum pressure so that the task starting at 50% of maximum pressure was still possible). It appears from these preliminary findings that the asymptote obtained from this task may not be a sensitive indicator of fatigue.

The type of fatigue test used can clarify where sites of fatigue may occur. For example, fatigue induced by electrical stimulation to nerves (e.g., Burke et al., 1973) involves only peripheral processes. Voluntary contractions may invoke breakdown at central and peripheral cites. Combinations of electrical stimulation and volitional contractions have been used to induce fatigue (Bigland-Ritchie, Furbush, & Woods, 1986) in order to parse out the effects of central and peripheral processes. Although it is difficult to isolate the site and mechanism of fatigue, Binder-Macleod and Snyder-Mackler (1993) point out that central fatigue will occur along with peripheral fatigue in subjects who are not motivated or those who have a disorder that affects central drive. They assert, appropriately, that "clinically, however, volitional effort may be most representative of the phenomenon we are trying to evaluate" (p. 904).

The technique described in this report may prove to be useful for assessing fatigue in people with a variety of neuromuscular disorders. The results from this experiment have indicated that the time constant from the constant-effort task is smaller after than before fatigue in neurologically normal subjects. It may follow that the time constant will be smaller than normal in subjects who experience fatigue chronically. Certain neurological disorders may be used to differentially assess sites and mechanisms of fatigue. For example, myasthenia gravis is characterized clinically by muscle weakness and fatigue that worsens as the muscle is used and recovers with rest or with a dose of an acetylcholinesterase inhibitor. Performance of the constant-effort task by a person with myasthenia gravis would reveal the

contribution of rapid peripheral fatigue. Additionally, assessing performance before and after administration of a short-acting AChesterace inhibitor would reveal the extent of contribution of the neuromuscular junction to the task. Other disorders, such as amyotrophic lateral sclerosis, that are characterized by marked muscle weakness, or chronic fatigue syndrome, apparently a disorder of central origin (Kent-Braun, Sharma, Weiner, Massie, and Miller, 1993; Stokes, Cooper, & Edwards, 1988), should be studied with this and other techniques to reveal the effects of fatigue arising from various mechanisms. It is our ultimate goal to develop a volitional task that will reflect central fatigue without inducing appreciable peripheral fatigue.

Conclusions

The time constant of the exponential curve produced during a sustained constant-effort task may be a useful indicator of fatigue. From the first experiment described in this paper, we provide data from young neurologically normal adults regarding this rate-of-decay variable. The time constant was smaller for the tongue than the hand, but was similar across gender and 3 starting levels (25%, 50%, and 75% of maximum pressure). The residual pressure, or asymptote of the exponential curve normalized to the initial pressure, varied with starting level but not gender or structure. Therefore, if these measures are to be used for the comparisons of various states of fatigue or populations of subjects, it would be prudent to control for structure and level.

The second experiment examined the effect of acutely fatiguing the structures of interest, and comparing performance of the constant-effort task before and after fatigue. The time constant of the modelled exponential curve was smaller (faster rate of decay) after fatigue for both the tongue and hand, but the residual pressure did not change systematically. Whether similar differences in the exponential decay of pressure produced by the tongue and hand would be found in people with fatiguing disorders of a neurologic basis remains to be seen. We are hopeful that the constant-effort task at submaximal levels may prove useful as a clinical assessment technique, because it is completely noninvasive, rarely perceived as fatiguing by the subject, and may reveal central-fatigue processes. Especially in combination with valid and reliable self-rating scales of fatigue (e.g., Chalder et al., 1993; Schwartz, Jandorf, & Krupp, 1993), the constant-effort task may supplement the assessment of fatigue. Thus, fatigue can be monitored in people with neuromotor diseases in terms of disease progression and treatment efficacy.

Acknowledgments

This research was supported by research grants 1-R03-DC01182 and P60-DC00976 from the National Institutes on Deafness and Other Communication Disorders. A portion of this work was presented at the 1992 Conference on Motor Speech in Boulder, CO. Author Sara I. Mitchinson currently is affiliated with Cognitive and Rehabilitation Specialists, Westchester, IL 60154. We would like to thank the three anonymous reviewers who provided helpful suggestions for the final version of this manuscript.

References

Abbs, J. H. (1990). Orofacial impairment in Parkinson's disease: Reply from the author. Neurology, 40, 192-193.

Amerman, J. D. (1991, April). Fatigue indexing of the perioral musculature. Paper presented at the Speech Motor Control Conference, Boulder, CO.

Barlow, S. M., & Abbs, J. H. (1983). Force transducers for the evaluation of labial, lingual, and mandibular motor impairments. Journal of Speech and Hearing Research, 26, 616-621.

Bigland-Ritchie, B., Furbush, F., & Woods, J. J. (1986). Fatigue of intermittent submaximal voluntary contractions: Central and peripheral factors. Journal of Applied Physiology, 61(2), 421-429.

Bigland-Ritchie, B., Johanson, R., Lippold, O. C. J., & Woods, J. J. (1982). Changes of single motor unit firing rates during sustained maximal voluntary contractions. Journal of Physiology, 328, 27P-28P.

Bigland-Ritchie, B., Johanson, R., Lippold, O. C. J., Smith, S., & Woods, J. J. (1983). Changes in motoneurone firing rates during sustained maximal voluntary contractions. Journal of Physiology, 340, 335-346.

Binder-Macleod, S. A., & Snyder-Mackler, L. (1993). Muscle fatigue: Clinical implications for fatigue assessment and neuromuscular electrical stimulation. Physical Therapy, 73(12), 902-910.

Borg, J., Grimby, L., & Hannerz, J. (1983). The fatigue of voluntary contraction and the peripheral electrical propogation of single motor units in man. Journal of Physiology, 340, 435-444.

Burke, R. E., Levine, D. N., Tsairis, P., & Zajac, F. E. (1973). Physiological types and histochemical profiles in motor units of the cat gastrocnemius. Journal of Physiology, 234, 723-748.

Cain, W. S., & Stevens, J. C. (1971). Effort in sustained and phasic handgrip contractions. American Journal of Psychology, 84, 52-65.

Cain, W. S., & Stevens, J. C. (1973). Constant-effort contractions related to the electromyogram. Medicine and Science in Sports, 5(2), 121-127.

Chalder, T., Berelowitz, G., Pawlikowska, T., Watts, L., Wessely, S., Wright, D., & Wallace, E. P. (1993). Development of a fatigue scale. Journal of Psychosomatic Research, 37(2), 147-153.

Clarke, D. H. (1986). Sex differences in strength and fatigability. Research Quarterly for Exercise and Sport, 57(2), 144-149.

Eason, R. G. (1959). The surface electromyogram (EMG) gauges subjective effort. Perceptual and Motor Skills, 9, 359-361.

Edwards, R. H. T. (1981). Human muscle function and fatigue. In R. Porter & J. Whelan (Eds.), Human muscle fatigue: Physiological mechanisms (pp. 1-18). London: Pitman.

Enoka, R. M., & Stuart, D. G. (1985). The contribution of neuroscience to exercise studies. Federation Proceedings, 44, 2279-2285.

Enoka, R. M., & Stuart, D. G. (1992). Neurobiology of muscle fatigue. Journal of Applied Physiology, 72, 1631-1648.

Fuglevand, A. J., Zackowski, K. M., Huey, K. A., & Enoka, R. M. (1993). Impairment of neuromuscular propogation during human fatiguing contractions at submaximal forces. Journal of Physiology, 460, 549-572.

Gandevia, S. C. (1982). The perception of motor commands or effort during muscular paralysis. Brain, 105, 151-159.

Gandevia, S. C. (1987). Roles for perceived voluntary motor commands in motor control. Trends in NeuroSciences, 10(2), 81-85.

Gandevia, S. C. (1992). Some central and peripheral factors affecting human motorneuronal output in neuromuscular fatigue. Sports Medicine, 13, 93-98.

Hellstrand, E. (1980). Morphological and histochemical properties of tongue muscles in cat. Acta Physiologica Scandanavia, 110, 187-198.

Hultman, E., & Sjöholm, H. (1986). Biochemical causes of fatigue. In N. L. Jones, N. McCarney, A. J. McComas (Eds.), Human muscle power (pp. 215-238). Champaign, IL: Human Kinetics Publishers.

Jones, L. A. (1983). Role of central and peripheral signals in force sensation during fatigue. Experimental neurology, 81, 497-503.

Jones, L. A., & Hunter, I. W. (1983a). Effect of fatigue on force sensation. Experimental Neurology, 81, 640-650.

Jones, L. A., & Hunter, I. W. (1983b). Force and EMG correlates of constant effort contractions. European Journal of Applied Physiology and Occupational Physiology, 51, 75-83.

Kent, R. D., Kent, J. F., & Rosenbek, J. C. (1987). Maximum performance tests of speech production. Journal of Speech and Hearing Disorders, 52(4), 367-387.

Kent-Braun, J. A., Sharma, K. R., Weiner, M. W., Massie, B., & Miller, R. G. (1993). Neurology, 43, 125-131.

Kernell, D., & Monster, A. W. (1982). Time course and properties of late adaptation in spinal motoneurones of the cat. Experimental Brain Research, 46, 191-196.

Krantz, H. (1981). Control of motoneuron firing during maintained voluntary contraction in normals and in patients with central lesions. In J. E. Desmedt (Ed.), Motor unit types, recruitment and plasticity in health and disease. Progress in clinical neurophysiology (pp. 358-367), vol. 9. Basel: Karger.

Lindström, L., & Hellsing, G. (1983). Masseter muscle fatigue in man objectively quantified by analysis of myoelectric signals. Archives of Oral Biology, 28(4), 297-301.

Lorell, D. M., Robin, D. A., Solomon, N. P., & Luschei, E. S. (1993). Tongue strength and endurance training in a woman with ALS [Abstract]. Asha, 35, 180.

Lorell, D. M., Robin, D. A., Somodi, L. B., Solomon, N. P., & Luschei, E. S. (1992). Tongue strength and endurance in the elderly [Abstract]. Asha, 34, 175.

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.

Matthews, P. B. C. (1982). Where does Sherrington's "muscular sense" originate? Muscles, joints, corollary discharges? Annual Review of Neuroscience, 5, 189-218.

McCloskey, D. I. (1981). Corollary discharges: Motor commands and perception. In V. B. Brookes (Ed.), Handbook of physiology: A critical comprehensive presentation of physiological knowledge and concepts. Section I: The nervous system (pp. 1415-1448). Bethesda, MD: American Physiological Society.

McCloskey, D. I., Gandevia, S., Potter, E. K., & Colebatch, J. G. (1983). Muscle sense and effort: Motor commands and judgments about muscular contractions. In J. E. Desmedt (Ed.), Motor control mechanisms in health and disease (pp. 151-167). New York: Raven Press.

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.

Sato, I., Suzuki, M., Sato, M., Sato, T., & Inokuchi, S. (1990). A histochemical study of lingual muscle fibers in rat. Okajimas Folia Anatomica Japonica, 66(6), 405-416.

Scardella, A. T., Krawciw, N., Petrozzino, J. J., Co, M. A., Santiago, T. V., & Edelman, N. H. (1993). Strength and endurance characteristics of the normal human genioglossus, American Review of Respiratory Disease, 148, 179-184.

Schwartz, J. E., Jandorf, L., & Krupp, L. B. (1993). The measurement of fatigue: A new instrument. Journal of Psychosomatic Research. 37(7), 753-762.

Sjøgaard, G., Savard, G., & Juel, C. (1988). Muscle blood flow during isometric activity and its relation to muscle fatigue. European Journal of Applied Physiology and Occupational Physiology, 57, 327-335.

Solomon, N. P., Lorell, D. M., 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. (1994). Tongue function in Parkinson's disease: Speech and nonspeech measures [Abstract]. Asha, 36, 129.

Solomon, N. P., Robin, D. A., Lorell, D. M., Rodnitzky, R. L., & Luschei, E. S. (1994). Tongue function testing in Parkinson's disease: Indications of fatigue. In J. Till, K. Yorkston, & D. Beukelman (Eds.), Motor speech disorders: Advances in assessment and treatment (pp. 147-160). Baltimore: Paul H. Brookes.

Solomon, N. P., Robin, D. A., & Luschei, E. S. (1994, March). Strength, endurance and sense of effort: Studies of the tongue and hand in people with Parkinson's disease and accompanying dysarthria. Paper presented at the Conference on Motor Speech, Sedona, AZ.

Somodi, L. B., Robin, D. A., & Luschei, E. S. (in press). A model of "sense of effort" during maximal and submaximal contractions of the tongue. Brain and Language.

Stierwalt, J. A. G., Robin, D. A., Solomon, N. P., Weiss, A. L., & Max, J. L. (in press). Dysarthria following traumatic brain injury: Strength, endurance, and speech ability. In D. A. Robin, K. M. Yorkston, & D. R. Beukelman (Eds.), Disorders of motor speech, Baltimore: Paul H. Brookes.

Stokes, M. J., Cooper, R. G., & Edwards, R. H. (1988). Normal muscle strength and fatigability in patients with effort syndromes. British Medical Journal, 297, 1014-1017.

Van Boxtel, A., Goudswaard, P., Van Der Molen, G. M., & Van Den Bosch, W. E. J. (1983). Changes in electromyogram power spectra of facial and jaw-elevator muscles during fatigue. Journal of Applied Physiology, 54, 51-58.

NCVS Status and Progress Report - 8
July 1995, 153-156

Suitability of Minidisc (MD) Recordings for Voice Perturbation Analysis

William Winholtz, A.A.S.
The Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts
Ingo Titze, Ph.D.
The Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts
Department of Speech Pathology and Audiology, The University of Iowa
Robert Lange, Ph.D.
The Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts

Abstract

A new digital recording format, Minidisc (MD), shows promise for high quality voice recordings. It is available in a portable size and uses magneto-optical recording techniques on a miniature compact disc. The disc can be recorded an unlimited number of times with essentially the same playback life span. However, the digital recording technique uses a data compression algorithm that may interfere with acoustic voice perturbation analysis. This study investigated what effects this compression may have and whether the MD format is viable for use in this application.

The MD format was evaluated by traditional synthetic test signals used on recording devices. In addition, human phonation recorded on DAT was used as input to the MD. The output of the MD was then compared to the original DAT recording. The two signals were analyzed for long and short term perturbation measures, and a visual inspection of their waveforms was done. The results indicated that the MD format performed as well as the DAT in all areas of standard tests, with the exception of signal-to-noise ratio. S/N for the MD was approximately 10 dB less than for the DAT under normal operating conditions. However, in comparing perturbation measures on normal human vowels, there were no significant differences between the two formats; i.e., no distortions in voice perturbation were introduced by the MD record/playback process.

Introduction

Recent studies have shown that the entire electroacoustic link (microphone, amplifier, recorder and filters, if used) is critical for high precision F_0 and amplitude extraction in voice signals (Titze & Winholtz, 1993). A significant contributor to extraction error is the signal recorder (Doherty & Shipp, 1988). The present study evaluates the Minidisc (MD) format, a new digital recording technology, for high quality voice recordings.

The Minidisc is similar to the compact disc and shares some of its capabilities. Random access to any location with this disc format is usually less than two seconds. The transport is relatively simple and its reliability may be better than the more complex tape transports of Digital Audio Tape (DAT) machines. The disc itself is a miniature 2.5 inch magneto-optical compact disc that allows up to 74 minutes of recording time. Its capacity for repeated record/playback is indefinite. Organization of the disc is in sequential tracks, or recordings. The beginning of a track is identified by a track marker that can be edited during or after recording; up to 256 track markers may be assigned per disc. Unlike DAT recordings, which can be inserted anywhere in a previous recording, MD recordings are made in a sequential fashion. Recordings can be added to the end of the last recording, but cannot be inserted between existing recordings. There is an 8 second playback audio buffer to minimize effects from physical shock and facilitate a seamless cuing system. Tracks that are 8 seconds or shorter are played from memory, making repeat functions possible with little wear on the transport or recording media. Repeating a track longer than 8 seconds is seamless, without interruption from the end of the track to its start. The unit accepts line or microphone level inputs and has a record-level function which is manual or automatic. Currently, the MD recorder is available as a battery powered portable unit the size of an audio-cassette "walkman" or as an AC powered, conven-

tional sized unit similar to a compact disc player. Price is comparable to a consumer grade DAT recorder.

The MD format uses a data compression algorithm developed by Sony, known as Adaptive TRansform Acoustic Coding (ATRAC). The algorithm codes and decodes the signal before recording and playing from disc. ATRAC is an auditory perceptual coder that achieves a 5 to 1 reduction in data, based on the psychoacoustic principles of the human ear. Quantization is 16 bits sampled at 44.1 KHz. Upon playback, the data are input to a buffer and clocked out at a highly stable rate. Any wow or flutter associated with transport instability is essentially eliminated.

The primary goal of this investigation is to evaluate the overall performance of the MD with regards to its ease of use for voice recording/playback and its reliability for extracting voice perturbation measures. Of particular interest, however, is the potential for undesirable effects due to the ATRAC data compression step. The ATRAC algorithm constitutes an innovation not used in previous digital formants, such as direct A/D/A sampling, pulse code modulation, compact disc, and DAT.

It is true that differences in performance may exist among these (older) digital formants. Variables such as sampling rate, resolution of quantization, and method of quantization may vary between specific implementations. Also, the process of re-sampling a signal (sampling initially from analog to digital, converting back to analog, and re-sampling a second time to digital) is something which happens whenever a voice is initially recorded onto MD or DAT, played back through its standard analog output, and re-digitized for computer analysis. Re-sampling can introduce very slight differences in the digital representation of the signal, on account of the phase difference between clocks at each sampling pass.

For the task of voice perturbation analysis, it is our view that a high level of performance can be expected from a commercially available recorder which uses any of these older digital formants. (It should be noted that Doherty & Shipp (1988) demonstrated no significant differences in the jitter and shimmer measures obtained using two of these digital formats: direct A/D/A and pulse code modulation on video cassette.)

For the purposes of the present study, the DAT recorder has been employed as an example of a common digital archive against which the new MD format and its ATRAC could be compared. In some cases, the MD has been tested alongside the DAT; whereby, the two share a common (input) source. The performance of the MD is then evaluated relative to the DAT's own performance. In other cases (which include the data obtained from the human subjects), the DAT has served as the source to the MD. In these tests, the performance of the MD is evaluated absolutely; i.e., on the basis of whether the MD can pass a signal from input to output without introducing distortions manifested in the measured levels of perturbation.

Method

As stated previously, the MD format was evaluated using two methods. In the first method, standard industry tests were used to compare the MD to a 16 bit linear DAT. The second method was similar to the first, but human and synthetic voice signals were used as input.

For the industry standard signals, a Sound Technology audio analyzer (ST1500A) was used to generate sinewave inputs. The same ST1500A was then used to analyze the output signals from a MD recorder (Sony MZ-1) and DAT recorder (Panasonic SV-3700). The tests included frequency response, distortion, and signal-to-noise ratio. The signals were recorded at 10 dB below the maximum operating level of the recorders to allow for necessary headroom and simulate typical record levels. The signals were then played back and analyzed directly.

In preparation for the phonation test, a synthesized vowel with no imposed modulation was generated as a reference for baseline perturbation measures. A formant synthesizer was used to produce an /a/ vowel at 150 Hz, sampled at 20 KHz. It was analyzed by a high precision software program GLIMPES (Titze & Liang, 1993) to obtain long term and short term amplitude and frequency perturbation measures (coefficient of variation and perturbation factor). A comparison was made between the analysis of the synthetic vowel retrieved from MD versus its counterpart retrieved from DAT.

In addition to the use of synthetic vowels, 20 normal subjects (10 male and 10 female) were asked to sustain the vowel /a/ at a comfortable intensity level and target pitch (100 Hz for the males and 200 Hz for the females). A headmount microphone (AKG-C410) was used to transduce the signals, which were amplified (ATI-1000) and recorded on the DAT recorder. The vowels were then played back from the DAT and re-recorded with the MD recorder. The entire utterance was played from each recorder, amplified (Tektronix 501), and then digitized (DSC-200 16 bit 20K samples/sec) into a VAX computer. The computer files were edited to eliminate onset and offset transitions; a two second duration of vowel was retained. The files were time aligned to within 50 ms so that the same section of phonation would be analyzed.

For the human phonations, a visual waveform comparison was made to expose any apparent distortions that may have been caused by signal processing (ATRAC) of the MD format. The source and output signals of the MD over the same section of phonation were used in the comparison.

Results

Table 1 presents the results of conventional sinewave tests. The performance of the MD recorder was similar to that of the DAT, with the exception of the dynamic

Table 1.
Results of Standard Test Signal Analysis
for the MD and DAT Recorders

TEST	MD	DAT
DYNAMIC RANGE	82 dB	92 dB
S/N	72 dB	82 dB
DISTORTION	.015%	.015%
FREQUENCY RESPONSE (±3 dB to 3 dB down points)	7 Hz-20 KHz	5 Hz-20 KHz

range and signal-to-noise tests. The S/N ratio of the MD was approximately 10 dB less than that of the DAT.

It should be noted that the ATRAC compression algorithm constitutes a nonlinear process whereby the amplification of a given frequency band may vary, depending on the signal activity in a different frequency band. This implies that the frequency response test, a meaningful spectral characterization for linear systems, may not be interpretable as a complete description of how a nonlinear system behaves. However, the industry-standard frequency response test, performed here as a sweep-tone measurement, is the only means available at this time for the standard assessment of recorders like the MD.

Table 2 gives perturbation measures for a synthesized vowel (for details, see Titze & Winholtz, 1993). CVA is the coefficient of variation of amplitude (standard deviation of cycle amplitude divided by the mean amplitude, expressed in percent). PFA is the amplitude perturbation factor (mean of the absolute cycle-to-cycle difference, divided by the mean amplitude, expressed in percent). CVF and PFF are analogous functions which measure pitch period rather than amplitude. The MD yielded slightly higher baseline amplitude perturbation measures than the DAT, a predictable result on account of its inferior S/N ratio (see discussion below). The frequency perturbation measures, however, were similar to those of the DAT.

Table 2.
Results of baseline amplitude and frequency perturbation measures obtained from a synthesized vowel. CVA is the coefficient of variation of amplitude, PFA is the amplitude perturbation factor, and CVF and PFF are similar fundamental frequency measures. All measures are expressed in percent.

INPUT	CVA	PFA	CVF	PFF
DIRECT	.014	.018	.001	.000
DAT	.045	.030	.099	.007
MD	.150	.045	.092	.008

Table 3 shows correlations, absolute differences, and signed differences between the MD output and its source for perturbation measures from human subjects. The correlations (Pearson Product Moment) between the two sets of scores as a function of the subject were all above .95 ($P<.05$), indicating no statistically significant differences. The average absolute differences for both male and female subjects were less than .2 % for the CVA measure and less than .06% for the PFA, CVF, and PFF measures. For the signed difference, the MD output gave slightly lower average values in amplitude measures (CVA and PFA) for the male group and slightly higher average values for the female group. For frequency measures (CVF and PFF), the differences for both groups were almost unmeasurable.

Table 3.
Results of correlation (as a function of the subject), average absolute difference, and average signed difference between the MD output and its source for amplitude and frequency measures from twenty normal subjects. Differences are expressed in percent.

	CVA	PFA	CVF	PFF
		MALE		
CORRELATION	.999	.999	1.000	1.000
ABSOLUTE DIFFERENCE	.19	.05	.01	.01
SIGNED DIFFERENCE	.16	.01	.00	.00
		FEMALE		
CORRELATION	.998	.997	.999	.992
ABSOLUTE DIFFERENCE	.16	.05	.02	.00
SIGNED DIFFERENCE	-.13	-.03	-.01	.00

One possible explanation for why the amplitude measures in Table 3 (CVA and PFA; absolute difference) indicate consistently greater distortions than the frequency measures (CVF and PFF) is that the signal-to-noise ratio of the MD system, though very high, is limited. A recent study in our lab (Titze & Winholtz, unpublished data) indicated that, in the case of an ambient-noise contaminated voice signal, both amplitude and frequency perturbation measures are adversely affected in proportion to increasing levels of noise. For a given level of noise, however, the amplitude measure appears to suffer more than the frequency measure. If the noise imposed by the circuitry and signal processing of the MD is assumed to have a similar effect on perturbation measures, then the residual noise specified by the MD's 72 dB S/N ratio may be sufficient to distort the amplitude measures more readily than the frequency measures. This is an indication that the superior S/N ratio of the DAT (82 dB) may constitute an appreciable advantage (over MD) when

performing sensitive amplitude perturbation measures on already normal phonations.

On the other hand, the 72 dB S/N ratio of the MD is still superior to what can be obtained from many analog recording formats. Likewise, the MD's distortions of CVA and PFA (reported in both Tables 2 and 3) are very low (on the order of 0.2%), considering that they are measures for baseline perturbations. In other words, these small but consistent distortions are reported for stable phonations which exhibit only residual levels of shimmer. Using these methods of calculating perturbation levels, it appears, therefore, that no significant distortions in voice stability have been introduced by the MD through-put and, in particular, the ATRAC compression algorithm.

Figure 1 illustrates the time waveform for male subject 1. The top trace (a) is the input to the MD recorder, and the bottom trace (b) is the output of the MD recorder over the same 4-cycle interval. The waveforms are visually identical, showing no sign of distortion or degradation attributable to the MD storage/retrieval step.

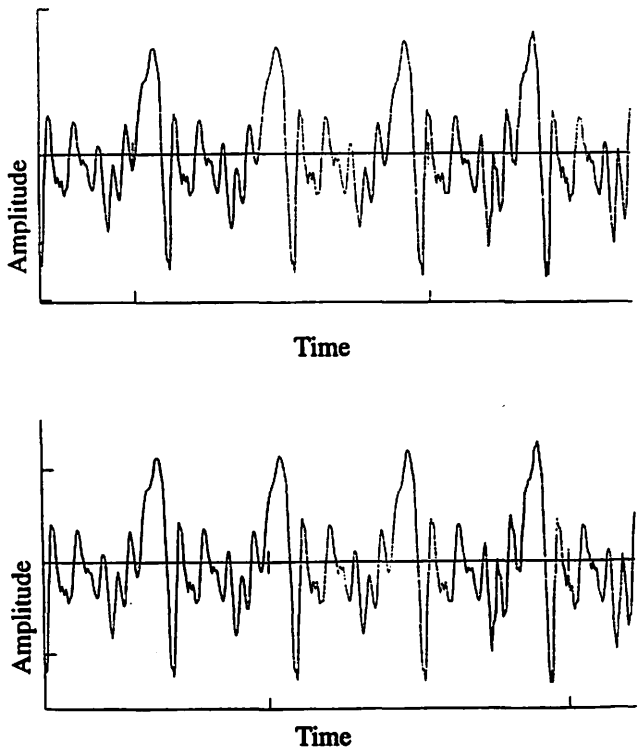


Figure 1. Waveform comparison of MD output (a-top) versus MD output (b-bottom) for the same section of human phonation (/a/ vowel).

Conclusion

A new digital recording format has been evaluated for high quality voice recording. The results indicate a level of performance for the MD similar to that of a DAT recorder, but with a slightly lower signal-to-noise ratio. Long and short term perturbation analysis of human vowels stored and retrieved from the MD yielded results which were indistinguishable from those obtained by analysis of its source. The MD format offers some advantages over the DAT format. The disc is less susceptible to damage and wear than DAT tape and provides quick random access. The MD transport is potentially more reliable because there is no intricate threading mechanism to load the tape, no rotating head to wear against the tape, and no tape ballistics. The new MD format may provide a more convenient means for the high quality recordings needed in voice perturbation analysis.

Acknowledgement

This study was supported by a grant from the National Institutes of Health, grant No. R01 DC00387-04.

References

Doherty, E.T., & Shipp, T. (1988). Tape recorder effects of jitter and shimmer extraction. *Journal of Speech and Hearing Research, 31*, 485-490.

Titze, I.R., & Liang, H. (1993). Comparison of F_0 extraction methods for high precision voice perturbation measurements. *Journal of Speech and Hearing Research, 36*, 1120-1133.

Titze, I.R., & Winholtz, W.S. (1993). The effect of microphone type and placement on voice perturbation measurements. *Journal of Speech and Hearing Research, 36*, 1177-1190.

Titze, I.R., & Winholtz, W.S. (in preparation). The effects of reverberation and ambient noise on voice perturbation measurements.

Ward, A. (1990). *A Manual of Sound Archive Administration*. Brookfield, Vermont: Grover Publishing Company.

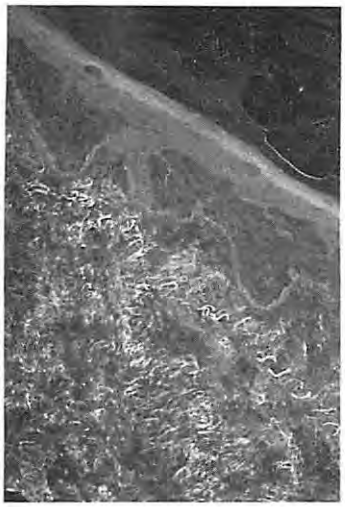
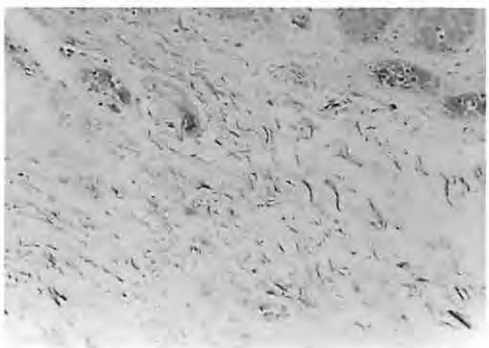
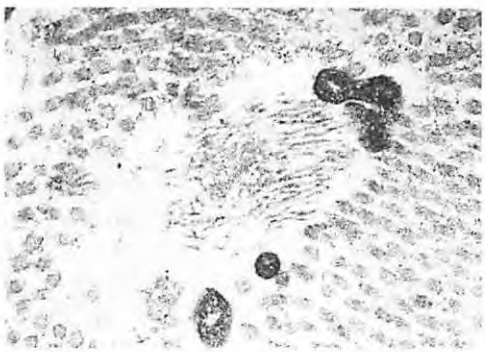
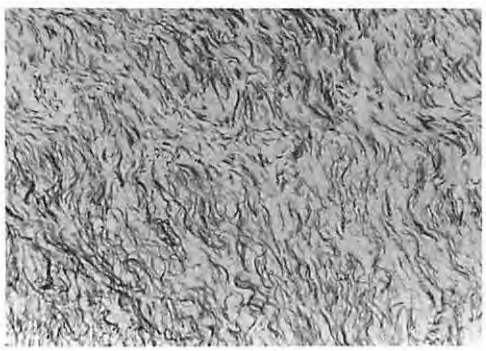


Photo 1 - left). Frozen section of normal vocal fold stained with decorin detected using an indirect immunofluorescence method. Fibrillar material in the superficial lamina propria is stained green. Section was photographed using a 15 second manual exposure time. Magnification 200X. Photo 2 - center). Frozen section of human vocal fold stained with antibody directed against keratan sulfate by the immunoperoxidase method. Fibrillar material in the lamina propria shows strong staining as does the region of the vocal ligament. Section is counterstained with hematoxylin. Original slide magnification 40X. Photo 3 - right). Frozen section of normal vocal fold stained with chondroi tin sulfate using an immunoperoxidase method. Brown staining of the cytoplasm of scattered cells in the lamina propria is seen (arrows). Some of the cells which stain positively also stain with an antibody directed against CD68. Section is counterstained with hematoxylin. Original slide magnification 40X.







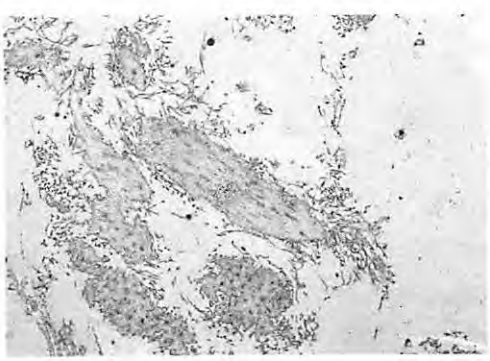
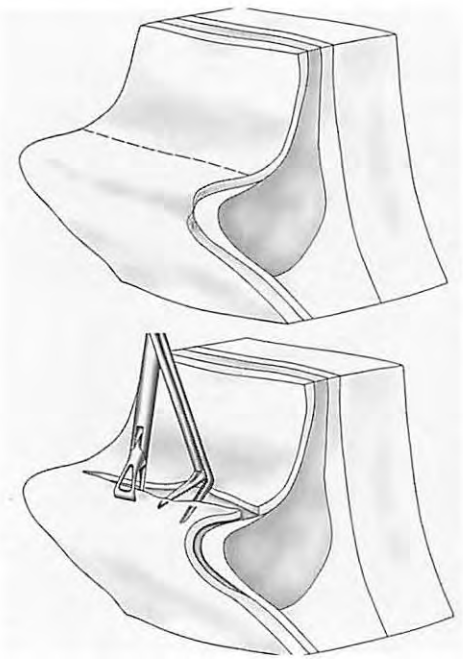


Photo 4 - top left). In this photomicrograph, a portion of the superficial lamina propria is seen. Tiny black fibrils are haphazardly arranged throughout the photograph. Verhoeff's elastin stain, 400 X magnification. Photo 5 - top right). This electron micrograph of the same region shown in Photo 4 illustrates the ultrastructural appearance of oxytalan, the fibrillar form of elastin, in the center of the picture. The small bundle is surrounded by collagen fibers cut tangentially. Dark circular profiles are cellular fragments. Section is stained with Uranyl acetate and lead citrate. Original print magnification: 125,000X. Photo 6 - middle left). An electron micrograph of the superficial region of the lamina propria shows a central aggregate of elaunin. Tangentially cut collagen is seen at the left edge of the photograph. Dark profiles are cell fragments. Section is stained with Uranyl acetate and lead citrate. Original print magnification: 125,000X. Photo 7 - middle right). This electron micrograph shows the ultrastructural appearance of the mature elastin fibers from the mid portion of the vocal fold. The amorphous component is much more abundant than in elaunin. Section is stained with Uranyl acetate and lead citrate. Original print magnification :50,000X. Photo 8 - left). Abundant elastin fibers are seen in this Verhoeff's elastin stained section of the mid portion of the vocal fold. Magnification 400X.



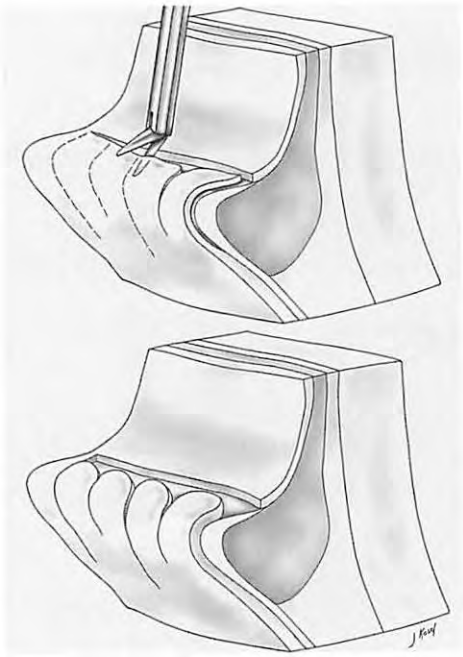
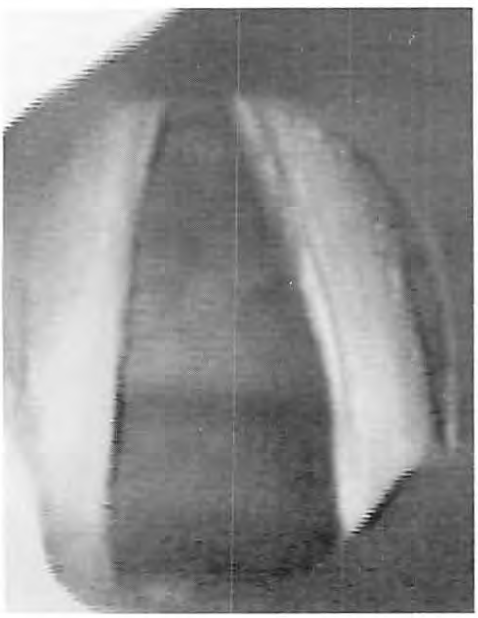
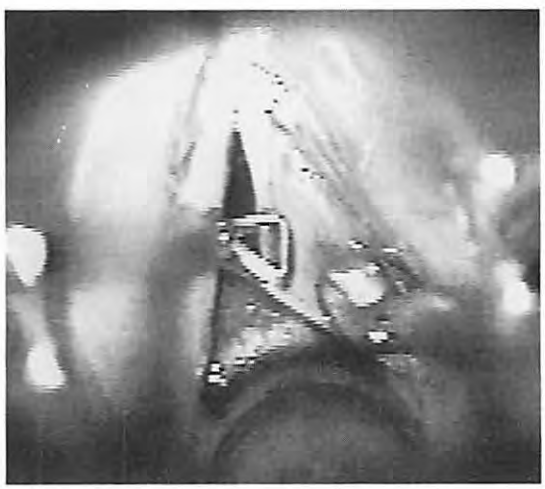
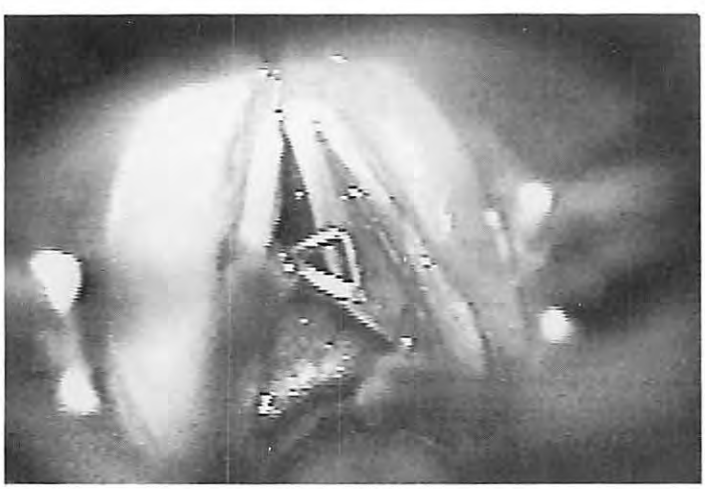


Photo 9A-D. Drawings showing steps in surgical management of Type II sulcus (vergeture) based on technique of Pontes and Behlau: (A - far left[top]) superior vocal fold incision, (B - far left [bottom]) separating epithelial layer from intermediate and deep lamina propria attachments, (C - immediate left [top]) "slicing" the inferior flap that has been undermined 2 mm inferior to sulcus into segments of variable length, and (D - immediate left [bottom]) rough appearance of mucosa after relaxation of contracture. Each diagram is a simplified 3-D medial-sagittal construct with the vocal ligament in white. (E-H - left to right). Operative videoprints demonstrating each of the same steps.







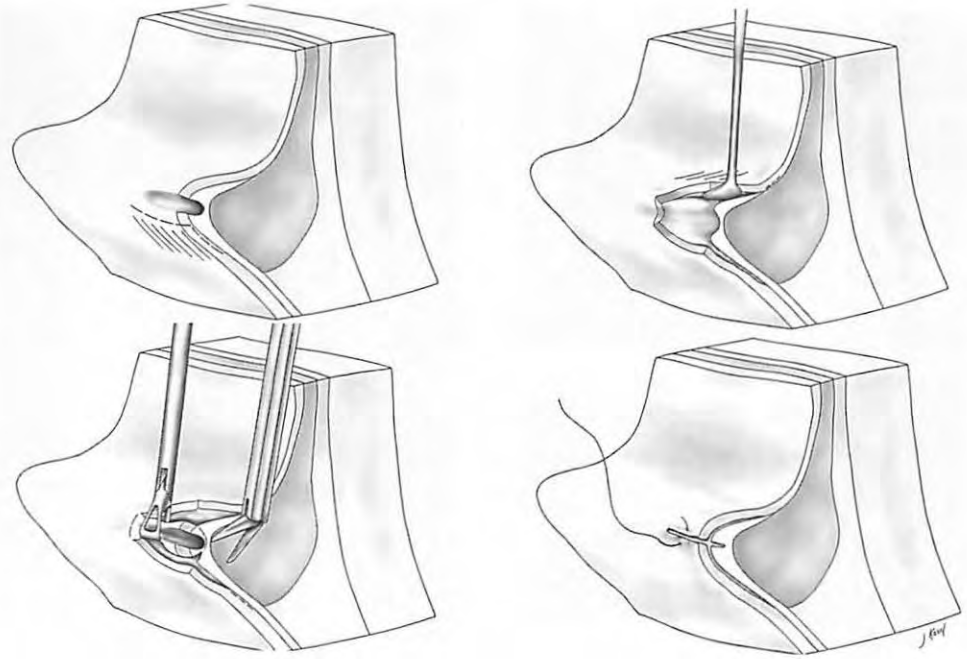


Photo 10A-D. Drawings showing steps in surgical management of Type III sulcus (vocalis): (A - far left [top]) mucosal incision is placed inferior to lip of sulcus and flap developed, (B - far left [bottom]) after a similar step superiorly, the sulcus (and contiguous cyst when present) is separated from the vocal ligament (and vocalis if necessary) and excised, (C - immediate left [top]) flaps are further elevated to relieve tension and allow passive approximation, (D - immediate left [bottom]) and further approximation is accomplished, if necessary, with 1 - 2 absorbable sutures. Again, each diagram is a simplified 3-D medial-sagittal construct with the vocal ligament in white.

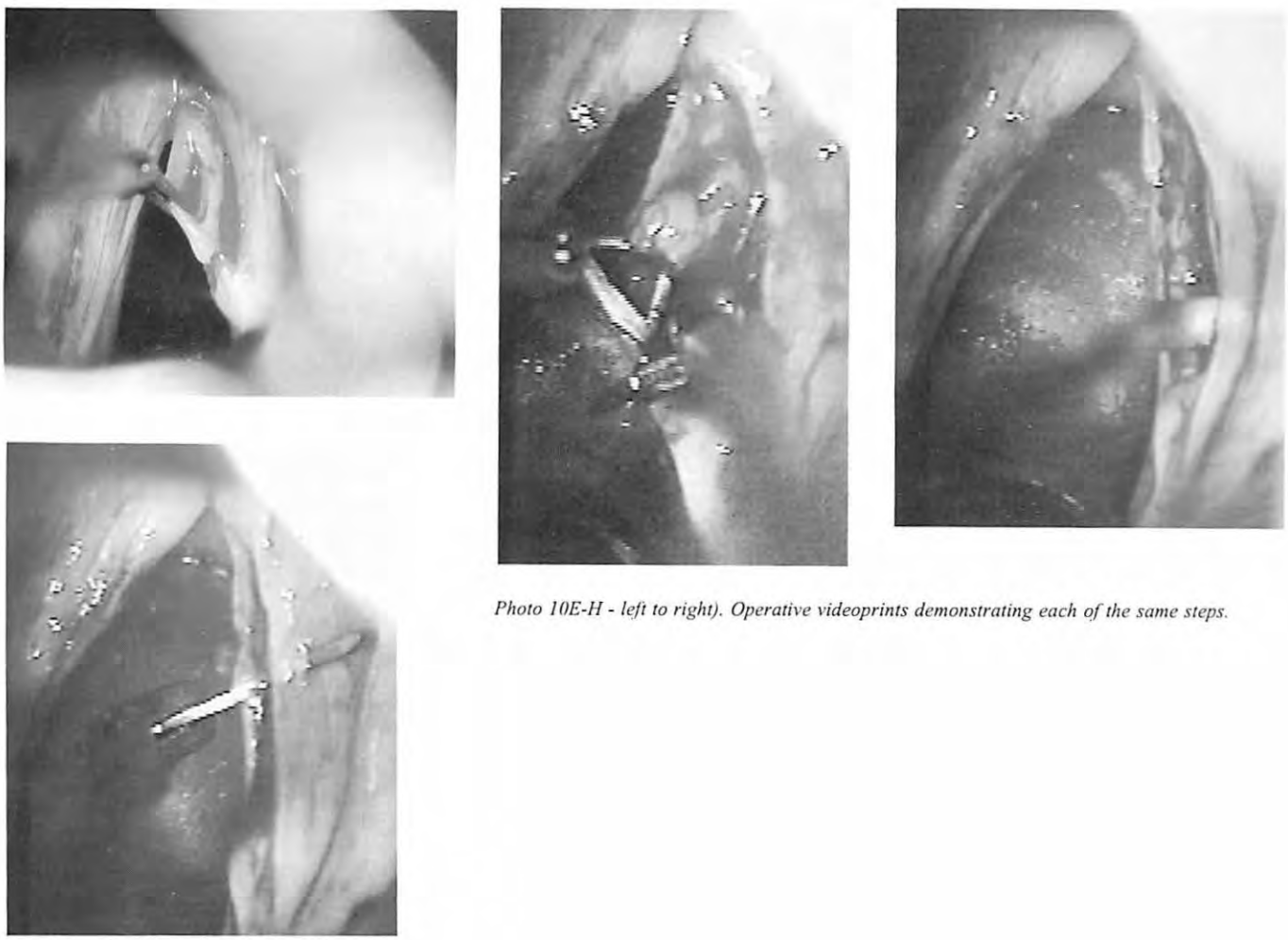


Photo 10E-H - left to right). Operative videoprints demonstrating each of the same steps.

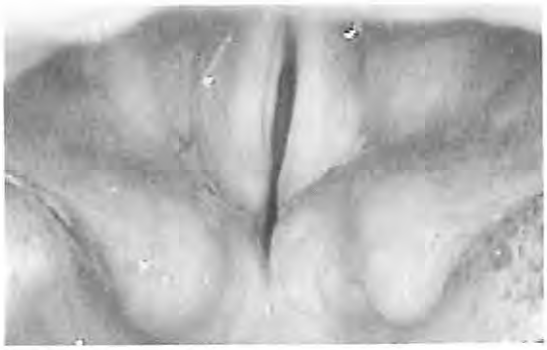
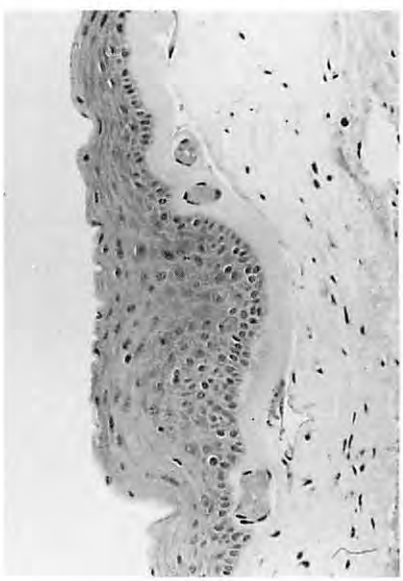


Photo 11. Physiological sulcus in (A - upper left) patient with vocal fold paralysis, and (B - upper center) a volunteer subject with no vocal complaints. Photo 12A - upper right) Operative microscopic view of Type II sulcus vergeture; patient of M. Bouchayer, photo reprinted with permission of author and Raven Press. Photo 12B - immediate left) Operative microscopic view of Type III sulcus vocalis in patient #9 from pathological sulcus treatment group.



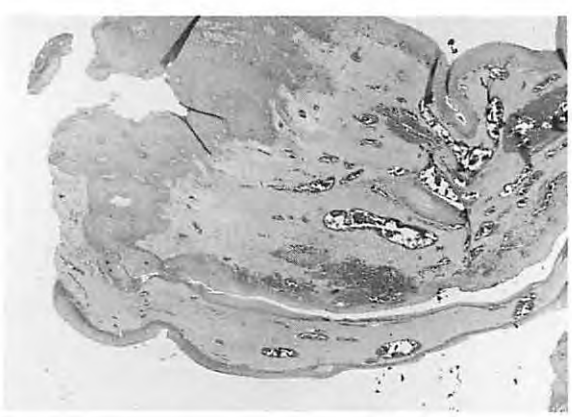


Photo 13A - upper left) Photomicrograph of sulcus vergeture excised from patient #20. Hematoxylin and eosin (H & E) stain at x10 magnification. B - upper center) Whole mount laryngeal specimen from M. Hirano, reproduced with permission of author, Raven Press, and Otologia Fukuoka. Photo 14A - upper right) Photomicrograph of excised Type III sulcus vocalis (patient #12) showing some inflammation and neovascularization. H & E stain at x20 magnification. Photo 14B - immediate left) Similar preparation of Type III lesion (patient #15) with massive inflammation, neovascularization, and fibrosis. H & E stain at x10 magnification.

Photo 15 - left). Display shows the shipping containers for excised dermis to be sent to laboratory for processing. Specimen jar and container are labelled with patient identification information. Photo 16 - right). The tuberculin syringe with autologous collagen is shown prior to insertion into the injector device.

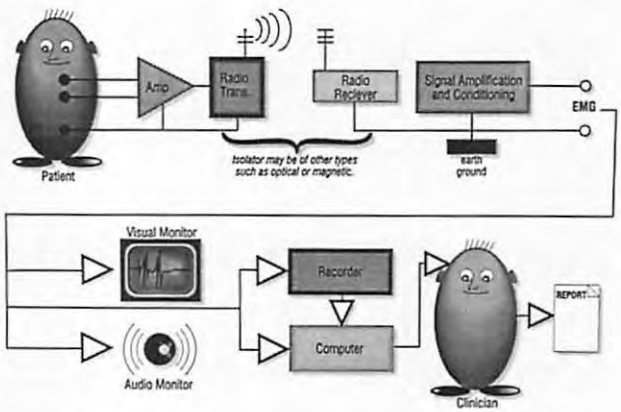


Photo 17. Major components of an EMG instrumentation system.

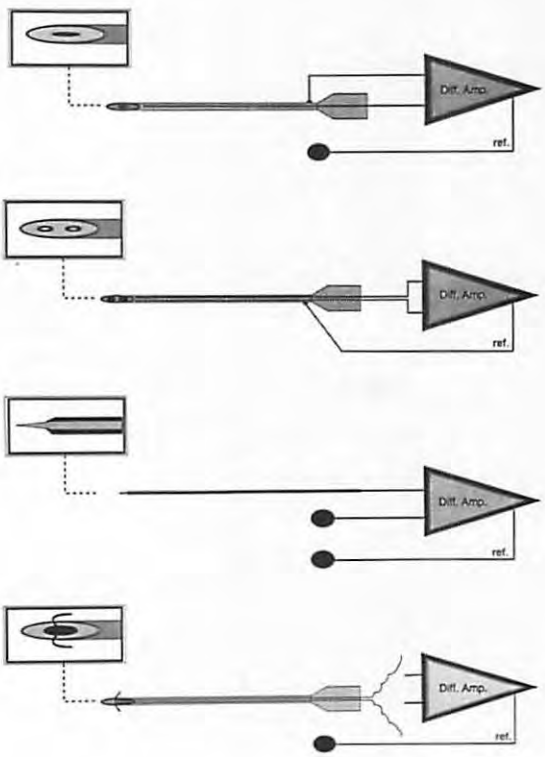


Photo 18. Four types of needle electrodes, numbered from top down, respectively: 1) concentric needle electrode; 2) bipolar concentric needle electrode; 3) monopolar needle electrode; and 4) bipolar hooked-wire electrodes.

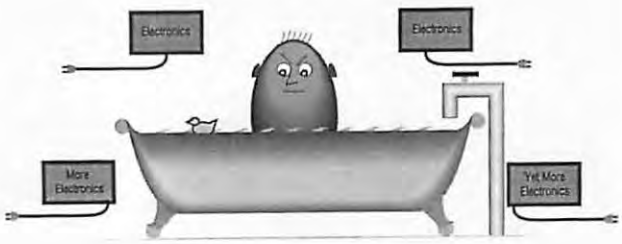


Photo 19. Earth-grounding a person is like putting them in a bath tub, from a standpoint of electrical safety.

Part II

Tutorial reports and updates of Training, Continuing Education and Dissemination of Information

NCVS Status and Progress Report - 8
July 1995, 157-174

Electromyographic Techniques for the Assessment of Motor Speech Disorders

Erich Luschei, Ph.D.
Eileen Finnegan, M.A.
Department of Speech Pathology and Audiology, The University of Iowa

Introduction

Scientists discovered, during the mid-19th century, that very small electrical currents were generated by contracting muscles. By 1912[1] instruments used for detecting these "action currents" had become sensitive enough to record voluntary muscle activity in humans. Limitations in the photographic recording devices available at that time provided records that were only a second or two in duration, but these records were sufficient to provide our first insights into the control of muscle by the nervous system. The signal recorded by this method was called the "electromyogram," and is still known by that name today. It is often abbreviated to "EMG." Although the study of the EMG signal has been widely used as a research tool for studying muscle and the general principles of motor control in the body, it has also evolved into a medical procedure that is used in hospitals and clinics throughout the world.

The primary purpose of this chapter is to help the reader better understand EMG procedures: where the signal comes from, the instrumentation used to record it, and how the signals may be processed and interpreted. Before considering those topics, however, it may be of interest to consider a few examples of how EMG recording is used in modern medical settings.

Diagnosis of Systemic Neurological Diseases

Suppose a patient comes to clinic complaining of very significant weakness. This could result from diseases at several sites in the peripheral or central nervous system. One site is the neuromuscular junction, where action potentials in motor nerve fibers cause action potentials in muscle cells. Failure of the neuromuscular junction is the cause of diseases such as myasthenia gravis, or the result of poisoning by toxins (botulinum toxin for example). Another potential cause is failure of conduction in peripheral nerves (peripheral neuropathy). Another chilling possibility is amyotropic lateral sclerosis (ALS), in which motoneurons slowly die. In considering this problem, it should be emphasized that an experienced clinician would probably have a good idea about which of these possibilities was the most likely cause of the weakness without using an EMG exam, based upon the history and a physical exam. The prognosis and treatment of the weakness would be very different, however, for the different possible causes, therefore any responsible clinician would want to have as much reliable information as possible before informing the patient and/or starting treatment. In this case, as in all others, EMG recordings yield additional information that has to be interpreted within the context of all other medical information. It is likely, in this hypothetical case of weakness, that a neurologist would order an EMG exam from a special diagnostic lab. Peripheral nerves in the forearm containing motor nerves to muscles in the hand may be percutaneously stimulated while recording the EMG of the hand muscles. The synchrony of the electrical nerve stimulation produces a large evoked EMG potential that can document the viability of the neuromuscular junction and also measure the conduction velocity of the action potentials in the peripheral nerves. The types of activity and waveforms of EMG potentials, recorded with needle electrodes inserted into the muscles, may definitively diagnose the disease of ALS.

Generally speaking, the diagnosis of motor disorders in neurological clinics is currently the main well-established clinical use of EMG recording and analysis. Texts such as that by Kimura[2] cover the substantive issues in detail, so will not be repeated in the current chapter. The types of electrodes used in neurological EMG analysis will be covered later, however, since the terminology can be a matter of significant confusion.

Evaluation of Paralyzed or Spastic Muscles

When a person cannot move a limb or articulator, it can be surprisingly difficult to know the exact cause. Damage to a particular nerve, either as the result of trauma, a surgical procedure, or simply as a "spontaneous" event can produce what appears to be "paralysis." One such problem that is familiar to speech-language pathologists who work in the field of voice, and their colleagues in the medical field of otolaryngology, is unilateral vocal fold paralysis. The condition is given this name because one of the vocal folds does not appear to move. The presumptive cause is damage to or cutting of the recurrent laryngeal nerve. One of the pioneers in the use of electromyography of laryngeal muscles, Faaborg-Andersen[3], studied EMG signals from "paralyzed" laryngeal muscles of patients with a presumptive diagnosis of unilateral vocal fold paralysis, and observed that many of them in fact exhibit EMG activity. Other otolaryngology studies have confirmed this observation[4-10]. Failure of a limb or articulator to move, or exhibit a normal range of motion, can have several causes: damage to motoneurons or motor nerves, mechanical changes in joints, co-contraction of antagonist muscle groups, or inappropriate reinnervation of muscles following nerve damage. Laryngeal muscle EMG recording is beginning to be used to help resolve some of these possibilities as they relate to unilateral vocal fold paralysis[11-16], but the procedure is far less widespread and "mature" than the procedures used in neurological diagnostic EMG labs.

Many readers may wonder, at this point, "Why don't the voice specialists with a patient with an immobile vocal fold just send the patient to the neurology EMG lab?" This could very well be done. The premise of this chapter is, however, that the interests of patients with dysarthrias will be best served by speech-language pathologists who are directly involved in EMG procedures. An appreciation of the anatomy and physiology of the articulators, the characteristics of speech and swallowing, and certain technical aspects of doing EMG recordings from inaccessible muscles that move vigorously during function will greatly improve any EMG procedures that are done. A speech-language pathologist may have more knowledge in these areas than a neurologist who has not specialized in the study of the articulators. Thus the speech-language pathologist has much to offer even if an EMG recording were conducted in a neurology EMG lab.

Many dysarthrias result from central nervous system disorders that produce abnormally high levels of activity (spasticity) in groups of muscles of the limbs and articulators. In some cases the ability of patients to walk, for instance, can be improved by surgically cutting the tendon of certain spastic muscles that are causing the most interference with the function. It may be difficult to know, ahead of time, which muscle is the "culprit" just using physical examination. In these cases, surgeons can use EMG recording to help make their decision about which muscle to tenotomize[17-19]. Although surgery of this type is not used to treat the dysarthrias, a related procedure, reversible paralysis of muscles by injection of botulinum toxin has been widely used to treat certain dystonias[20-27]. In this procedure, EMG recording can be used as an aid to verify the injection site[28-33].

The goal of this chapter is provide a resource for anyone who wishes to use or better understand how to record and interpret the EMG. It seems likely that it has many uses in addition to the examples given above, including the study and treatment of motor speech disorders. It is appreciated by the authors that technical procedures of this type are somewhat foreign to many clinical speech pathologists, not only because they appear to interfere with the person-to-person relationship, which is the heart and soul of the profession, but also because the instrumentation has too many knobs and wires. On the first point, it may be helpful to consider the fact that EMG recording is merely a way of extending the clinician's senses to obtain information that we can't ordinarily sense. As for the second point, it seems likely that the problems of "too many knobs and wires" will be eased a great deal in the future by EMG instrumentation designed for clinical use rather than for research. In addition, however, most clinicians will find the knobs and wires relatively simple once they are understood.

Principles of EMG Recording

Source of the EMG Signal

An action potential must sweep along muscles cells in order for a muscle to contract. These action potentials create, as they move along, minute electrical currents flowing outside the muscle cells (Fig. 1). These tiny currents don't really do anything. They are just there as a part of the process, like the bow wave on a ship going through the water.

The nervous system activates muscles in terms of motor units, defined as "a motoneuron and all the muscle cells it innervates." Because all the muscle cells in one motor unit "fire" (have action potentials) at the same time, the tiny extracellular currents from individual muscle cells add up to produce a current which is rather easily detected by recording electrodes. Since these currents are going in the same direction at the same time, the total waveform is stereotyped. Such a potential is called a "single motor unit potential" (Fig. 2). This potential is equivalent to a group of ships going through the water in precise synchrony; their bow waves would add up and produce quite a wave!

Because the firing of the motoneurons of a naturally-activated muscle is not synchronized, the single motor unit potentials interact with one another in a complex manner, and produce an interference pattern. Referring to our earlier analogy, one may imagine hundreds of groups of ships racing up and down the waterway. Their bow waves

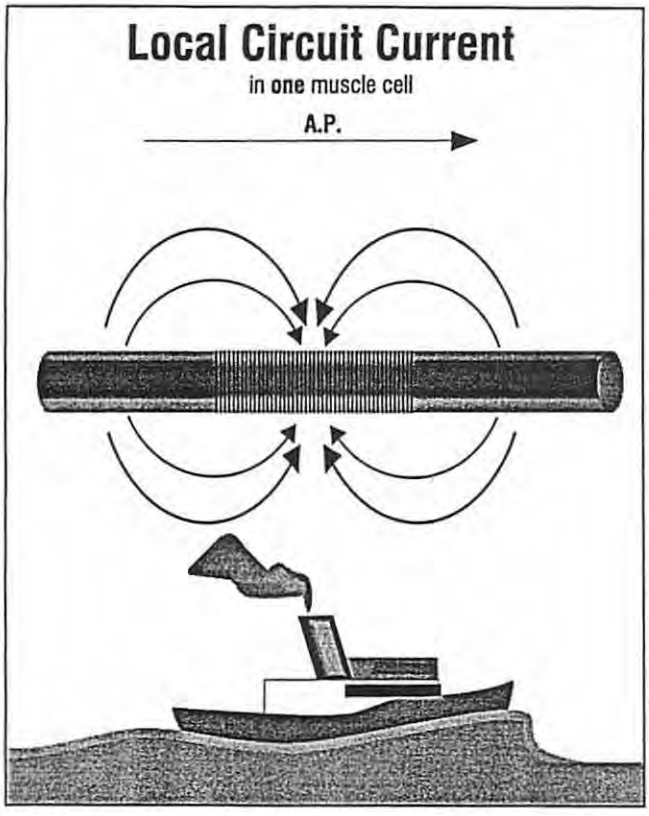


Figure 1. The extracellular current caused by an action potential sweeping along the muscle cell is analogous to the bow and stern wave caused by a ship going through the water.

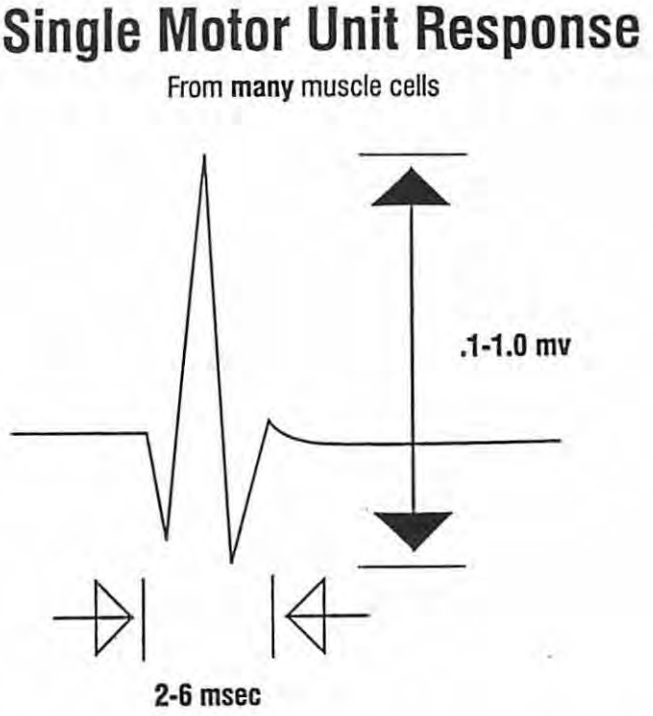


Figure 2. Drawing of a typical "single motor unit potential" caused by synchronous firing of muscle action potentials in all the muscle fibers innervated by a single motor unit.

would interact and produce a "chop," which is just an hydraulic interference pattern. One may intuitively sense that the size of the "chop" will be related to the number of groups of ships that are traveling about. Obviously the more ships, the higher the "chop." The EMG interference pattern is no different; the more motor units that are active, the "higher" the EMG signal. Typical records showing such an interference pattern are illustrated in Fig. 4a. The metric to be used for describing the "height" of the EMG is a complicated matter, however, which will be considered later. For the moment, it can be taken as a fact that the size of the EMG signal bears a monotonic relationship to the degree to which the muscle has been activated.

Electrode Configuration in EMG Recording

The EMG signal is almost always recorded with a differential amplifier, which creates an output that is proportional (and much larger) than the difference in the voltage of its two inputs. How these two inputs are spaced with respect to the active muscle has a large effect on the nature of the EMG signal and the interpretation one can apply to it. Let's go back to the hydraulic analogy. Suppose we wanted to measure the "chop" on the water created by the ships dashing about. The use of a differential amplifier is equivalent to measuring the height of the water between two points. Suppose these two points are close together, say a foot apart. If a bunch of ships were steaming around miles away from our two recording points, the size of the chop would be much attenuated by the time it got to us. Also, all the little peaky white caps would be gone; we would just see small slow "rollers" going by. So our signal would be small, and it would be composed of low frequencies. However if we had our two recording points right out in the midst of all those ships, we would detect large water height differences, And many of the waves would have sharp peaks. In this case, our signal would be large and be composed of much higher frequencies than when the recording points were distant from "the action."

When the electrodes are far apart, they are able to "survey" activity over large areas. This has some advantages in some circumstances, but in this case the source of the activity is not known with precision. This can be a major diagnostic problem. Suppose your electrodes are far apart in a paralyzed muscle. Adjacent muscles that are active may produce a sizable EMG signal. Such "distant" activity would be relatively small compared to what one would expect of a normally active muscle, and would be composed of low frequency waveforms. Thus one would suspect (hopefully) that the activity was not from the muscle under study. Close-

spaced electrodes tend to "reject" muscle activity from distant active muscles, however, and would be preferable, therefore, in assessing a muscle suspected of being paralyzed.

Four generalizations about amplification and electrodes spacing may be offered at this points:

1. Differential amplifiers are essential.
2. Close electrode spacing is the best in the majority of cases.
3. Muscle activity close to the electrodes is represented by a signal having high frequencies, often showing characteristic single motor unit waveforms. Such signals sound "crisp" in the audio monitor (see below).
4. Signals from distant active muscles may be picked up even by electrodes that are very close together. These signals are composed predominantly of low frequencies, however, and sound "dull" and "muffled" compared to the signals of nearby muscle fibers.

For those desiring a more rigorous treatment of this topic, an excellent review article is that by De Luca[34]. *Another excellent source, covering myoelectric theory and many very pragmatic issues in EMG recording, is that by Loeb and Gans*[35].

Instrumentation

Overview

The scheme represented in Photo 17 (see center plate) identifies some of the major aspects of an overall instrumentation system. Three electrodes are attached to the patient. One of these, the "reference" lead, is attached to the skin with a conductive gel electrode, and serves to hold the subject's body, as a whole, to the electrical reference voltage of the EMG amplifier and input stage of the isolator. The other two electrodes are the inputs to the differential amplifier, and their difference in voltage is amplified by the gain of the amplifier. This amplified voltage difference is applied to the input stage of the isolator. The input stage of the isolator modulates a very low-power radio transmitter as a simple linear function of the signal. It should be realized, however, that radio transmission is only one of several ways of sending the information in the signal across an insulating gap. The output stage of the isolator is a radio receiver that demodulates the radio signal to recover an exact replica of the EMG signal which now can be safely referenced to earth ground. The rationale for use of the isolator stage will be discussed in detail below.

The EMG signal may be further amplified and processed, such as being bandpass filtered, before it is presented to the eye and/or ear of the clinician. On-line monitors are very useful to make sure the recordings are proceeding appropriately, but documentation and objective analysis requires that the signal be permanently recorded. This can be done by tape recorders, which then may be played back off-line to a computer for analysis. Alternatively, state-of-the-art computers may now be used as the data acquisition recording device, as well as serve for data analysis. The last stage of this "system," perhaps the most essential, is the experience and intelligence of the clinician in interpreting the EMG signal.

The instrumentation system just described may be usefully regarded as being like a "chain" insofar that it is only as strong as its weakest link. Inappropriate electrodes, amplifiers or recorders that are noisy or lacking in fidelity, or lack of experience by the clinician in recognizing artifactual signals, for example, can all defeat an otherwise adequate system. It will, therefore, be useful to consider these various parts of the system in detail.

Electrodes

There are basically three types of electrodes that have been widely used in EMG recording: Skin electrodes, intramuscular fine-wire (hooked-wire) electrodes, and intramuscular rigid needle electrodes. The latter two types are schematically illustrated, along with their usual connections to a differential amplifier, in Photo 18 (see center plate).

Skin electrodes: Muscle activity from large muscles directly under the skin may be easily detected and quantified by simply attaching metal disk electrodes to the skin surface. This approach often seems most desirably because it is non-invasive. No one seems to like needles! The fact of the matter is, however, than there are major limitation to the use of skin electrodes by speech-language pathologists; most of the muscles involved in speech and oral functions such as swallowing are not directly under the skin. These muscles are also relatively small and complex in their anatomy, so records obtained with skin electrodes provide only a limited measure of how these muscle systems are operating. The lips are an exception, however. Lip EMG during speech may be easily recorded with skin electrodes[36-39]. The surface lip EMG signal can sometimes be misleading, however, because it may reflect the activity of different muscles or functionally different motor units within what is anatomically one muscle[40]. Surface electrodes placed over the masseter muscles can record an EMG signal related to vigorous jaw motion or relatively forcible biting. These electrodes may detect some low-level activity during speech tasks[41], but the exact source of this activity is uncertain. It could be from remote jaw muscles or from small facial muscles located under the skin close to or overlying the masseter muscle, i.e.. buccinator or platysma.

One skin electrode that is common to all of the recording schemes to be discussed is the electrode attached to the reference (common) input to the differential amplifier. This is a very important electrode because a stable low-

resistance contact between the skin and metal wire carrying the signal to the amplifier is critical to good noise-free recording conditions. This is facilitated by carefully cleaning the skin before the electrode is secured in place by an adhesive collar. After cleaning the skin, a conductive electrode gel is then applied to the "cup" of the electrode containing the metal surface, and the electrode attached to the skin. The metal inside the electrode "cup" is a complicated material generally made of compressed silver powder and "special ingredients" known only to the manufacturer. In the experience of the author, some commercial skin electrodes are better than others, and all of the good ones are surprisingly expensive (but well worth the cost). Large commercial disposable electrodes made for hospital recording of the electrocardiogram (ECG) are too large for placing over most muscles, but make excellent skin electrodes for the reference electrode. They come with the gel and adhesive collar in place, and because of their large size, provide an excellent electrical contact without excessive skin preparation.

Rigid Needle Electrodes: Needle electrodes for recording the EMG are basically a long thin hypodermic needle with one or two insulated wires in the lumen. If there is one wire, the needle is usually called a "concentric" EMG electrode. If there are two wires, then it is called a "concentric bipolar" electrode. When using a concentric electrode, one input of the differential amplifier is connected to the center wire of the electrode, and the shaft of the hypodermic needle (which is uninsulated) is connected to the other amplifier input. The signal that is picked up by this electrode is any voltage difference that exists between the exposed center wire at the tip and the shaft, which is essentially held at a neutral "reference" voltage by its large uninsulated area. This electrode can be quite "selective" in recording single motor units. The third wire of the electrode is connected to the reference point on the patient (the skin of the forehead is often convenient when studying cranial muscles).

When using concentric bipolar electrodes, the inputs to the differential amplifier are connected to the two wires in the lumen of the needle. The shaft of the needle may be connected to the reference input of the amplifier, but this is not necessary. The reference input may be connected to a distant point on the skin. The concentric bipolar electrode is more "selective" than the concentric electrode, particularly if the needle can be rotated when in place. This can change the alignment of the two exposed wire tips at the lumen of the needle with respect to the direction of the muscle fibers; such alignment can have a dramatic effect on the degree to which a particular muscle fiber action potential is picked up. From a practical standpoint, concentric bipolar electrodes are more expensive than concentric electrodes, and thus used mainly for research or specialized studies.

Monopolar EMG electrodes are insulated solid needles whose tip has been bared of the insulating material. The wire from the needle is connected to one input of the differential amplifier, and the other input of the amplifier is connected to a second skin electrode (in addition to the electrode connected to the reference input to the amplifier). One might suppose it would be simpler to connect the second input of the amplifier directly to the reference input. If this is done, however, the resistance to ground of the two inputs will be greatly unbalanced. This condition increases the response of the system to external noise. Monopolar electrodes record muscle activity just as well as the other types, but they are also more responsive to activity generated in muscles distant to the one being studied, i.e. they are less selective. The monopolar arrangement is, in effect, the type of recording configuration used when one records from the tip of an insulated hypodermic needle used to inject botulinum toxin in a laryngeal muscle .

Rigid needle electrodes are used almost exclusively in most EMG diagnostic clinics, such as those found in departments of neurology. This is because the waveform and firing patterns of single motor units are of primary interest in diagnosing most neurological diseases that affect muscle activity. These electrodes can be advanced and withdrawn by the clinician to isolate particular units. When used in muscles directly under the skin during isometric contraction, in which there is relatively little actual movement of the muscle tissue, these electrodes are relatively painless to the patient, and do little damage to the muscle. When used to record from small and/or thin muscles, particularly during functions that involve significant movements, this type of electrode has serious problems. First of all, an EMG interference pattern, as distinct from single motor unit activity, is useful for determining if and how much a muscle can be activated. Concentric bipolar electrodes are very selective for single motor units, so one might see little, if any, evidence of activity in a normal muscle, even at fairly high levels of activation. Another serious problem is that movement of a muscle with a rigid needle in it can be quite painful. Pain from this source may have large effects on the motoneuron pool that one is attempting to study, as well as hurt the patient. If the larynx, pharynx, or other oral structure remained very still, this probably wouldn't be a significant factor, but large, vigorous, movements occur during speaking and swallowing. Movements of the very sharp tip of a rigid needle electrode within a small muscle could cause damage. Furthermore, these movements could cause the population of motor units under observation to change each time they occur. This would make it difficult to quantitatively compare a current EMG record with one taken before the movement occurred. There are currently no empirical studies to show that these possible problems actually occur, although this is the subject of an on-going study in one of the authors' laboratory. To start with, however, it is worth keeping in mind that rigid EMG electrodes are not the only or necessarily the best electrodes

when one is attempting to study natural movements, particularly of the articulators.

The main advantage of rigid needle electrodes is that they may be advanced and withdrawn in the muscle to explore various parts of a muscle, or to be redirected to other muscles. The potentials they record during short period of relative immobility of muscles can provide valuable information. Such information has to do with the shape of action potentials. These potentials can confirm diagnoses of such diseases as ALS, or indicate whether a muscle is being reinnervated. When these are the main diagnostic questions, the EMG recording and interpretation is obviously best done by an experienced neurologist.

Bipolar Hooked-Wire Electrodes: Hooked wire electrodes were first used by Basmajian[42], who, in collaboration with several colleagues, described the activity of many muscles of the body during natural movements. These observations are summarized in a classic book Muscles Alive[43]. This type of intramuscular electrode has also been used for study and assessment of laryngeal muscles[44-50], tongue[51-55], velum[56-68], pharynx[69-72] as well as some combinations of these muscles[73-77]. They are made with two very fine, flexible, insulated wires. For recording gross EMG, about one millimeter of insulation is removed from each wire. These two wires are placed in the lumen of a suitable hypodermic needle so that the bared portions extend just beyond the tip, and then the ends of the wire are bent over to form a "hook." When the needle is placed into the muscle and then withdrawn, the hooks catch in the muscle and the wires are left behind in the muscle. Because of the flexibility of the wires, the muscle may make large movements without causing pain. Most people find these electrodes quite comfortable. They may be left in place during repeated measures of speech and swallowing.

While there are several good reasons for using hook-wire electrodes, they have the major disadvantage of not permitting adjustments of the electrode position once the hypodermic needle has been withdrawn. Obviously the electrode cannot be pushed further into the muscle. One could imagine (naively) that one could gently tug on the wires and pull them along the route of the wires through the muscle. To do this, however, the hooks have to be straightened out, and once this is done, the wires do not generally remain in the muscle. For all practical purposes, one gets one attempt per needle. When recording from laryngeal muscles, it is advisable to have extra electrode assemblies in order to try a second or third time to get the wires into the desired muscle.

Choosing the wire for these electrodes is a compromise between using wires that are too ductile and wires that are too stiff. Very ductile wires can be inserted with very fine needles, and cause no sensation in the muscle. But their hooks are very weak, and they often do not "hook" into the muscle, and get pulled out with the needle. On the other hand, their weak hooks probably do minimal damage when these wires have to pulled out after the recording is over. Stainless steel wire with an uninsulated diameter of 50 μm or 75 μm (0.002 inch or 0.003 inch) are the most widely used types of wire. They generally hold the "hook," and produce only a slight "funny feeling" in the muscle. They probably damage some muscle cells when the hook is pulled out, but there is no pain associated with this event, so the damage is quite limited. Single-strand 75 μm stainless steel wire with various types of insulation is available in small quantities from several vendors in the United States. Some companies will fabricate special configurations of fine wire (usually for a minimum order of 1000 ft). One such configuration, "bifiler," where two strands of wire are bonded together along their entire length, has two advantages to using two wires that are not bonded together. First of all, the bared ends of bonded wire remain in a fixed relationship, and in particular cannot "short together" in the muscle as it moves. Second, the fact that the two wires have to move together greatly reduces their tendency to produce movement artifacts and microphonics when they are mechanically disturbed. An explanation of this property is offered later in the chapter.

Once the wire is chosen, one has to select a needle size. A 25 gauge hypodermic needle is suitable for a pair of 75 μm diameter wires. In some cases, one can choose a smaller needle that will accept the wires, but the fit is so close that the wires often "hang up" when the needle is withdrawn, so the wires get pulled out; the wires have to have enough room in the lumen of the needle to accommodate slight bends in the wire. A single strand of insulated 75 μm wire can be inserted with a 30 gauge needle, but this requires two needle insertions for differential recording . The larger electrode spacing obtained in this case may be advantageous, however, if the goal is to sample a large area of a muscle.

When placing the electrode in a package for sterilization, it is important to use a relatively large package so the wires may be loosely looped in a way that allows them to be straightened out without forming kinks when the package is opened. Once formed, these kinks are impossible to remove. Because the hypodermic needle cannot be pulled passed a kink in the wire, such an electrode might as well be discarded before the electrode is inserted into the muscle.

Connecting the amplifier to fine wires can be more troublesome than one might suppose. If one uses very long pieces of fine wire, then these leads may be taken directly to the inputs to the amplifier and put under a screw- type terminal (gold-plated surfaces are a VERY good idea). However, these long wires have a tendency to become tangled and form kinks when removed from the sterilization packages. The wires may be kept much shorter if the amplifier has a preamplifier stage in a small box that can be placed close to the patient. Such a preamplifier stage may

have screw-type terminals as well, but spacing usually makes it desirable to have a wire "grabber" that requires less manipulation than a screw terminal. One may use spring-loaded commercial "grabbers", but keep in mind that the very small size of the wire requires a close-tolerance fit of the "jaws," and that an oxide coating on either the wires or the grabber surfaces can cause electrical "noises" that can interfere with recording. Grabbers with gold-plated surfaces are, for this reason, much to be preferred. We currently use gold-plated extension springs to make contact with the fine wires. The springs are bent to open a few loops, the wire inserted between them, and the spring allowed to straighten out. They seem to be a vast improvement over other methods we have used previously.

Amplifiers

To detect and study the activity of muscles requires one or more differential amplifiers that have a selection of gains between about 1000 and 50,000 (60 dB to 94 dB) with a bandpass between 30 Hz and 5 kHz. These characteristics are easily obtained with modern electronics. In a clinical situation, it is often difficult to control external sources of 60-cycle interference, so a 60-Hz notch filter in the amplifier can be quite helpful. A very important feature of the amplifier is its "common mode rejection ratio" (CMRR). When used with the proper electrode configuration, an amplifier with a high CMRR is the most effective way to prevent amplification of the 60 cycle electrical fields. There are companies that manufacture amplifiers that are marketed specifically for recording EMG activity. One pays a premium for such equipment, however, and the fact of the matter is that the electronic requirements are easily met by devices made from a few integrated circuits, resistors, and capacitors. Whatever amplifier is purchased or fabricated, it is crucial that it include an isolation stage. The function of an isolation stage is to avoid having the patient connected to earth ground. Connecting a subject to earth ground with a low resistance electrode (the large skin electrode with conductive gel) is essentially like standing the patient in a bathtub. In this condition, the patient could be electrocuted by coming into contact with any equipment that is electrically "hot."

Isolation Amplifiers Must be Used for Electrical Safety!!!

Please refer to Photo 19 (see center plate).

Earth-Grounding a patient is equivalent to standing them in a bathtub full of water.

In the scheme illustrated earlier, such isolation is represented by using radio signals to transmit the output of the first stage of amplification to the earth-grounded part of the system. In this case, the patient could possibly be on the moon! It is impossible to accidentally electrocute anyone who has no electrical connection to earth ground! Use of the proper equipment eliminates this danger, **BUT** it is possible for a person to unwittingly defeat this safety feature by connecting the reference (indifferent) electrode to earth ground, or, for example, by having the patient sitting on a grounded chair. It should be noted that radio transmission is certainly not the most common way of achieving isolation. It was chosen for illustration because it is intuitively easy to understand. The same effect can be achieved by using modulated light or magnetic fields to couple the signal across a "gap" that keeps the patient totally isolated from earth ground.

Setting the appropriate gain of an amplifier is a matter of some judgment. Usually, the gain is set so the largest potentials observed stay within the input range of the recording instrument. If one suspects muscle paralysis, or is seeking evidence of low-level activity, one should not hesitate to try the highest gains on the amplifier, even if some other events and/or noise go outside the range. Some motor unit activity in a muscle is very small, and can "hide" within the noise of the baseline if gains of 1K-5K are the only ones used. Unless higher gains are used, the "failure to observe activity" in a muscle is questionable.

Monitors

A type of monitor that has traditionally been used with the EMG is an audio amplifier connected to a loudspeaker. This type of monitor has several advantages. It is inexpensive, you don't have to watch it, and the human can, with some experience, use his or her "ears" to perform an acoustic analysis that can provide critical on-line information. For example, the sound of muscle activity that is close to an intramuscular electrode sounds "crisp" because it has mainly high- frequency components. EMG from muscles located at greater distance has a "dull" sound. This is because the tissue that intervenes between the EMG source and the electrodes acts like a low-pass filter (DeLuca, 1979)[34]. Single motor unit "spikes", which are produced by the near- simultaneous action potentials in the muscle cells innervated by the same motoneuron, produce a "crack" or "pop" which has a unitary sound. The ear can quickly detect the pattern and rate of firing of a single motor unit. In fact, a subject who can control his or her muscles can quickly learn to turn a unit on and off by listening to its discharges. A inexpensive receiver amplifier, which can be purchased from a commercial electronics store for a little over $100, is quite adequate for this purpose.

The limitation of an audio monitor is that it provides no permanent record, and the amplitude of the EMG cannot be determined with any precision. It also cannot reveal the presence of some problems such as high frequency interference (as from a near-by video monitor) and large low-frequency movement artifacts. Another problem with audio monitors is that they produce a terrible 60 cycle din when the electrode is removed or a wire becomes

detached. A very good idea is to have an audio monitor with a handy volume control, or an assistant who can anticipate the "blast" and turn the volume down before it happens.

A visual monitor such as an oscilloscope or a computer capable of providing on-line display of waveforms is very useful. They can be used to reveal the many types of artifacts that interfere with EMG recording, and by knowing the gain of the amplifier which is being used, it can be used to estimate the actual amplitude of the signals. Waveforms that can't be discriminated by the audio monitor carry potentially important information. If an oscilloscope were to be purchased for this use, then a very basic digital oscilloscope, costing about $1000, would be quite adequate. A digital oscilloscope can "freeze" a picture "on-line", and thus the EMG signal can be studied in order to make a decision during the procedure. Older analog oscilloscopes are still quite useful, however, for verification of the signal's basic characteristics. There are also computer hardware/ software systems (see below) that are very useful for on-line monitoring.

Recorders/ Data Acquisition and Storage Devices

It will be very important for any person or group of people making decisions based upon EMG recordings to be able to directly compare records from different patients. Historically, polygraph records or tape recorders have been used for permanent storage of EMG records. Both of these instruments have serious problems, both for storing data and for its subsequent display and analysis. Another approach, made possible by the recent development of powerful but inexpensive computer systems, is direct digitization of EMG records and "streaming" of these data to computer disk. Commercial software systems operating on personal computers are currently capable of continuously acquiring and storing 80,000 samples per second. This would make it possible, for example, to acquire data on 4 channels of EMG at a sampling rate of 20,000 sample per second, quite fast enough to resolve the finest details of the waveform of a motor unit response. Some software/hardware data acquisition systems provide a continuous display of the EMG traces, whether or not data is actually being stored to disk, so they provide a very good visual monitor of the records as well as a means of data storage. These same computer programs can also provide a method of quantitatively analyzing the EMG records. The use of one such data acquisition and analysis system, that developed by DATAQ Instruments, Inc., Akron, Ohio, will be illustrated later in conjunction with the section on data analysis.

Direct digitization and storage of EMG records will unquestionably produce very large data files. Consider, for example, the need to store 10 minutes of data at 80,000 sample per second. Remember each sample needs two bytes of storage (assuming 12 bit A/D conversion). This would produce a record of almost 100 megabytes. If this record were left on a computer's "hard disk," it would be a major problem. Such a record could be moved, off-line, however, to another storage medium such as a magneto-optical disk. Currently, each removable magneto-optical disk stores 230 MB of data, so the recording sessions would require less than half of such a disk for storage. The storage cost would thus be less than $25, and it is very likely that this cost will go down over time.

Although it is natural to look at data storage systems in terms of technical problems and costs, the most important factors, in the long run, are those related to ease of access. Simply put, stored data that doesn't get looked at and analyzed is absolutely worthless. If data is stored in a way that requires a great deal of human time for it to be accessed and/or analyzed, it may never be used, and if it is, wages provided to assistants for this process may easily wipe out any savings associated with a less expensive storage system or storage medium. In this respect, directly-digitized and archived EMG recordings have great advantages. The fact that we do not already have well-established norms for using EMG for diagnostic purposes is probably a result of the prodigious task of analyzing EMG with the instruments that have previously been available for storing and analyzing these records. The development of digital data acquisition systems, which are currently just in their formative stage, will change this situation dramatically.

Source and Recognition of Artifactual Signals (Noise)

The 60-cycle Beast

As most people know, the power supplied to homes, hospitals and labs by the Electric Company comes out of the wall (in the U.S.) as a 60 Hz sinewave at a root-mean-square voltage of about 115 volts. This power line signal often contaminates EMG records. In dealing with 60-cycle interference, there are three important steps to take:

1. Minimize the electrode impedance, including the impedance of the reference electrode.
2. Remove the sources of the interference.
3. Shield against the interference, and/or selectively filter it out of the amplification process.

Reducing Source Impedance: Other factors being the same, the amount of 60-cycle problem one encounters can be related to the "source" resistance of the electrodes, which in this case is the sum of the resistances from each electrode to the reference electrode. Consider the extreme: suppose you directly connect both input electrodes directly to the reference electrode with a piece of wire. You could put that "electrode" configuration in the middle of a generator and you wouldn't have a problem! At the opposite extreme, suppose the electrodes are "open," i.e. there is nothing but air between the electrodes and the reference electrode. The

source impedance is now the input impedance of the amplifier, which is extremely high. In this condition, very small 60-cycle fields will be able to create comparatively large voltages (but see comment below on "blocking" of amplifiers). There is really little that one can do about the resistance of the electrode contact with the body fluids. One has a small area of bared stainless steel wire on the end of each wire, and that it that. The thing that one CAN do something about is the resistance between the reference electrode and skin of the patient. Good skin preparation and/or use of large gel electrodes, such as those for ECG recording in Intensive Care Units, will be helpful. Appreciation of the effect of having a high source impedance will help you understand a frequent problem in EMG recording: the mysterious appearance of high level of 60-cycle noise where it was absent earlier in the recording, or where exactly the same conditions were used in a previous recording session. The best guess is that either a connection is "open" somewhere in the circuit, or you have an air bubble between the gel in the skin electrode and the metallic part of the electrode.

Eliminating the Source of the 60-cycle Field: Interference from 60-cycle sources comes from two types of "fields." The easy problem (usually) to solve is caused by inductive fields from transformers or wires carrying large AC currents that are physically close to the electrodes. The 60-cycle interference from inductive sources usually has a characteristic waveform, which does not contain "spikes." (Fig. 3). You cannot shield against this type of field, but physically moving the source of the magnetic field away from the recording site is usually sufficient to solve the problem. Moving just a few feet sometimes helps a great deal. The other type of 60-cycle interference comes from electrostatic influences. Any two conductors separated by an insulator form a capacitor. If a capacitor has a significant capacitance (a number that indicates how much charge can be stored at a given voltage difference between the conductors), then an AC voltage impressed upon one of the conductors can "couple" some fraction of that AC voltage to the other conductor. Large currents do not have to flow to cause this type of interference. Equipment which is plugged into the wall socket, but which is not turned on, can still cause problems. The power cord to the equipment, and the wires running internally to the power switch can form a conductor having peak voltages of +/- 160 volts with respect to ground. Any other conductor in the area, in particular the leads to the amplifier, can have a large AC voltage coupled to it if the resistance to ground is large (see above re source impedance). In this case, the insulator between the "plates" of this capacitor is the air in the room. One strategy for finding the source of a 60-cycle interference problem is to systematically unplug any unnecessary equipment in the vicinity of the recording site. If the equipment has to be plugged in, then sometimes moving the offending equipment to another location is helpful.

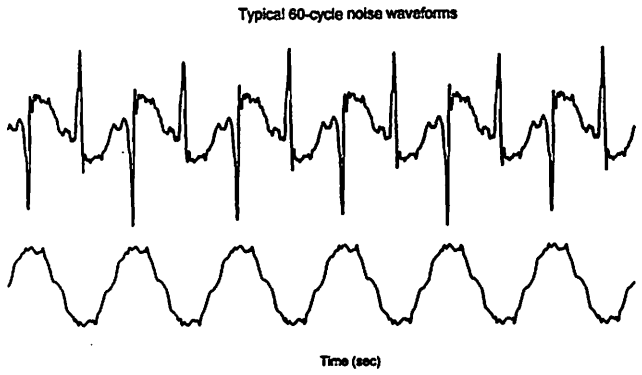


Figure 3. Typical waveforms of 60-cycle noise from florescent lights (top trace) and from power cord inductive field (botton trace). Traces are 0.1 seconds in duration , and therefore show exactly 6 cycles.

Florescent light can be a terrible problem. They all have a "ballast" in them, which increases the voltage across the florescent tube to about 750 volts. The waveform of this type of interference typically has large "spikes" in it (Fig. 3). The waveforms seen can vary a great deal, but they will always have a period of 16.667 msec. Keep in mind that a 60-cycle notch filter will not attenuate these "spikes" at all, so one should check the period of regular "spikes" of any type, even if there is no evidence of a sinusoidal waveform at a frequency of 60 Hz. Turning off florescent lights close to the recording site usually gets rid of the problem, but if one has to use high amplifier gains, then one may have to turn off all the florescent lights. Incandescent lights, which use only 115 volts, cause far less interference Sometimes florescent lights are not the source of the problem, however. You can answer this question by turning off the lights to see if the problem goes away.

One factor that seems to greatly increase 60-cycle interference is the presence of large ungrounded metal structures of any type close to the patient or amplifier. For example, we have found it very helpful, to eliminate 60-cycle noise problems in a minor surgery room used for EMG recording, to earth-ground the metal frame holding the patient's bed. It is important to note that the patient was not in contact with this frame, so this arrangement did not ground the patient.

Shielding and Filtering: After the source of the 60-cycle noise has been minimized, there are two additional strategies to reduce 60-cycle interference: shielding and filtering. If one is working in a normal environment, having an electrically-shielded room to do the recording is generally not practical. Some people shield the leads going from the subject to the amplifier, but this makes the leads rather stiff and much heavier than is desirable. A much more common strategy, that accomplishes the same thing, is to use

a small input stage on the amplifier (a first-stage amplifier in a small box) that is held close to the patient. Thus the leads can be kept fairly short. If these leads are only a few feet long, then putting them in a cable with a braided "shield" really doesn't help much. Shielded input leads can also be a major source of movement artifact (discussed later), so we recommend short unshielded leads to the amplifier. Although the primary attempt to control 60-cycle noise should be to get rid of the source, some environments, such as a surgical facility, make this very difficult. In this case, 60-cycle notch filters can be very helpful. They will help by reducing the 60-cycle component of the noise ten-fold, compared to what it would be without the filter. They will also produce corresponding amplitude and phase distortions of the biological signals in this frequency range, but that should not be a problem for diagnostic EMG recording.

Besides 60-cycle noise, one can encounter other forms of extraneous noise signals, such as high frequency fields from TV or computer monitors. Interference from a TV monitor will be seen on an oscilloscope as a broad band of "fuzz" around the recorded signal. You can't hear it on the audio monitor, because most audio amplifiers can't reproduce it and/or most of the "seniors" can't hear that high (about 16 kHz). Generally it isn't a serious problem, but if it is too large, it can "block" the input stage of the amplifier and thus affect amplification of the EMG signal. The easiest solution is to move the TV monitor as far away from the patient as possible. A few feet of movement can often totally eliminate the problem.

Movement Artifacts

Another type of artifact one has to recognize is related to movement of the electrode leads between the patient and the amplifier inputs. In most cases, these electrode lead movements cause large low-frequency excursions of the recording baseline. A "tap" or high frequency vibration of the leads may also cause artifactual signals whose frequency components are high enough to be in the pass-band of the amplifier. In fact, they sound like someone tapping on the bottom of a metal can when heard over the audio monitor. An important source of movement artifact is the electrochemical system at the interface of the electrode, which is a metallic conductor in a salt solution. There is usually a significant DC junction potential at each electrode. In effect, each electrode acts like a miniature "battery." The voltage from this battery charges the distributed capacitance of the electrode lead with respect to ground. Movements of the electrode leads cause changes in the distributed capacitances, and changes in these "capacitors" cause charging and discharging currents to flow back and forth across the electrode resistance, thus producing voltage changes. If the changes in the capacitance of both electrode leads are exactly the same, the voltage changes produced across each electrode will be identical, and thus will not be amplified by a differential amplifier, i.e. they will be a "common-mode signal." This conceptual model of movement artifact would explain why the use of bifiler wires, where both electrode leads have to move together, dramatically reduces problems with movement artifact.

Sometimes one sees a sinusoidal signal at the frequency of phonation when recording from laryngeal muscles. Don't suppose you are observing phase-locking of the EMG. This is another version of the movement artifact discussed above; the wires going to the amplifier are probably vibrating at the frequency of phonation. Another type of movement artifact may be observed when using hooked-wire electrodes: rather large "spikes" associated with the start of movements. They look somewhat like large single motor units, except for two things. They are usually monophasic (go in only one direction from the baseline), and typically have much longer durations than motor units responses. While they may occur several times in each burst, they don't exhibit a repetitive pattern at a respectable rate, i.e. something on the order of 10- 50 spikes/second. One possible cause of these large spikes would be intermittent contact of the two wires of the electrode within the muscle. Such contact would discharge the "standing" electrochemical junction potential (see above).

Even with bipolar hooked-wire electrodes having large tip exposures, it is fairly common to observe single motor unit activity in most muscles. One has to separate these potentials from artifacts that can produce spike-like potentials in the record. Real single motor unit potentials are distinguished by several features: 1) they are almost always biphasic, 2) they usually fire at least several times in a row at rather regular intervals, 3) they generally correlate with some aspect of motor activity, 4) they do not correspond to external events like someone switching a piece of electronic equipment on and off, and 5) they do not fire at a steady interval of 16.667 msec (the interval of 60-cycle interference). Another thing to look for, when recording from two or more muscles, are potentials occurring simultaneously in two or more channels. Such events are very unlikely to be biological in their origin. They can result from irregular powerline transients, sometimes caused by someone in another part of the building turning on or off large motors. We have seen "strange" potentials of this variety quite often when recording in a surgical suite.

"Blocking" of Amplifiers

If the lead to the reference electrode, or to either of the inputs electrodes is "open," i.e. has an extremely high resistance, the output of many (if not most) modern EMG amplifiers will show a perplexing behavior that can be quite confusing to the novice. The amplifier output may be very quiet for extended periods of time, as if there were absolutely no muscle activity, and even noise signals will be absent. Then one may observe brief periods of high amplitude 60-

cycle noise and/or large movements artifacts that gradually die out, leaving a quiet baseline once again. These periods of "noise" recording can usually be produced at will by moving the leads to the electrodes, or by any movements of the preamplifier or the patient/subject. This is the behavior of a "blocked" amplifier, and it can only be corrected by establishing a reasonable resistance from each electrode to the reference electrode, and making sure the reference electrode is connected to the reference input of the amplifier. "Blocking" in an amplifier is caused by the very small currents that have to flow from the inputs. These currents are, in modern amplifiers, in the picoampere range, and so are ordinarily of no consequence. However, if they have NO path back to the reference input of amplifier, they will create a relatively large DC differential signal at the inputs. This will cause the first-stage amplifier to "saturate" it's output at the plus or minus power supply voltage. Subsequent high-pass amplifier circuits, always present in an EMG amplifier system, will make this steady DC voltage into a steady zero-voltage output. Any movements of the wires or electrodes will momentarily allow this input current to redistribute itself, and thereby amplify the noise signals that one ordinarily sees when the electrodes have a very high resistance. Pragmatically, the main lesson is this: if you observe a VERY quiet baseline, interrupted by "funny" episodes of recording, check to see that you have intact electrode wires and good connections to the amplifier.

Analysis of EMG Records

Qualitative Analysis: Using the Calibrated Eyeball

Although it is natural, from a scientific standpoint, to gravitate toward numerical comparisons when using physical measurements for diagnostic purposes, the fact remains that probably one of the most sophisticated analyzers available to us is our own senses. Therefore it is very useful, in using EMG records to understand and/or diagnose dysarthrias, to be able to locate regions of interest in the record, and then observe them in detail. In doing this, however, certain simple transformations of the original waveforms may improve our ability to observe relationships in the recordings. The best way to make such observations is to use a computer program specifically designed for display and analysis of physiological recordings. Figure 4 illustrates an example of how this may be done, using a program such as Windaq (DATAQ Instruments Inc., Akron, OH). It is the result of an analysis of laryngeal EMG obtained from a normal subject during three repetitions of the phrase "Pop took his socks off". Keep in mind that this presentation is <u>not</u> intended as a definitive statement of how laryngeal muscles are used during speech, nor as a "sales pitch" for a particular commercial computer program. What <u>is</u> intended is to convey to the reader at least one approach to finding information in an EMG recording of some of the muscles involved in speech.

The recording to be presented below was obtained from a healthy normal adult female speaker. She repeated the phrase at normal pitch and somewhat increased loudness, using normal prosody and emphasis. Recordings were from made from the thyroarytenoid (TA) and cricothyroid (CT) muscles bilaterally using bipolar hooked-wire electrodes. Amplifier gain was 2000, and the bandpass of the amplifiers was 30Hz to 5 kHz. Before being digitized online at a sampling rate of 5 kHz, all channels were filtered at 2.5 kHz for purposes of anti-aliasing. Figure 4a illustrates a "compressed" portion of the recording. Compression of the record is accomplished by having the computer display a vertical line representing the maximum and minimum value of data found in each compressed "block" of data (in this case 80 data points) at each horizontal pixel of the screen. This algorithm is much preferable to other types of "compression," such as just displaying every 80th data point or displaying the average of the 80 data points; it preserves the occurrence of transient responses whose duration is shorter than the duration represented by each pixel, in particular single motor unit responses. The ability to display a compressed record is very useful in locating a region in which some activity of interest is occurring.

In a long record, where the subject perhaps speaks several phrases, one may wonder how you can be sure what the subject is saying at a particular point of the record. The computer program has the capability of outputting the data from a selected channel (in this case the microphone channel) to the digital-to-analog converter (DAC). The data to be output is selected using the cursor, and at a keystroke, they are played back and one hears the spoken phrase. Because of the relatively low sampling rate, such speech is not suitable for acoustic analysis, but it is clearly intelligible. We can thus be sure that this portion of the record is of interest.

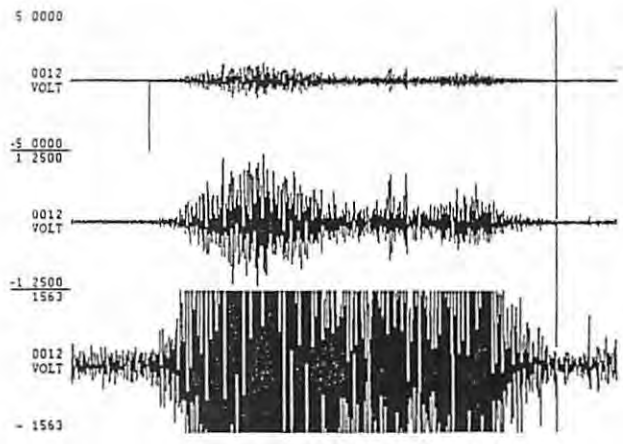


Figure 4. Three ways to view the same EMG record: a) compressed display; b) decompressed display (bottom three traces from 4a), plotting all data points, and c) rectified and smoothed (moving average). See text for discussion.

The details of the muscle activity are obscured by this degree of compression. This problem is easily solved by having the computer decrease the compression of the record corresponding to the location of the cursor. In effect, we want to "magnify" the record around the cursor. The decompressed record, displaying all data samples around the position of the cursor in Fig. 4a, is shown in Fig. 4b.

Even though the speaker is perfectly normal, an initial reaction to inspection of the EMG recordings of these laryngeal muscles is "What a bunch of junk!" There is clearly modulation of activity, but the pattern is uncertain. It might be supposed that the electrodes are not actually in the intended muscles, a possibility that always has to be considered when attempting to record from laryngeal muscles. However, certain transformation of the "raw" EMG records makes it easier to recognize patterns of modulation (Figure 4c). The "raw" EMG is first rectified (the sign bit of each data point is simply set to be positive). The record is then smoothed with a moving average. In this operation, the computer calculates, for each data point, the average of that data point and all the data points coming X points before and after the data point. In effect, a "window" containing 2X+1 samples is averaged , and the data point at the middle of this window is replaced with this average. Then the window is moved one data point to the right, etc. In Fig. 4c, the averaging window is 30 msec. The resulting record shows, with more clarity, the "trend" of EMG amplitude changes. In particular, there appears to be two peaks of activity in both TA muscles preceding the "his socks" portion of the phrase (immediately following the cursor), which correspond to two smaller peaks of activity in the CT that seem out of phase with the TA modulation. Obviously, appreciation of these features would benefit from being able to superimpose these repetitions, lining them up on some point that seems relatively invariant. This can be easily done by "clipping" 2.2 sec. portions of the record around each token, where the beginning of each of these subfiles begins exactly 1.1 sec before the peak of the "his" in the microphone record. The computer program allows these subfiles to be saved in a spreadsheet format. When these files are opened in a spreadsheet program, the three tokens may be aligned in time (superimposed), and the values normalized to the highest values found in each channel. The resulting data may then be graphed, resulting in Fig. 5. In this final form, the modulation of the laryngeal muscle activity, and its repeatability, is very clear.

Even knowing that the subject studied in Fig. 5 was a normal speaker, we cannot assert that this pattern of laryngeal muscle activity is "normal." It could in fact be a quite atypical: one would only know by applying the analysis to a large number of speakers. Had the health status of the speaker been unknown, however, the results illustrated in Fig. 5 would establish important facts:

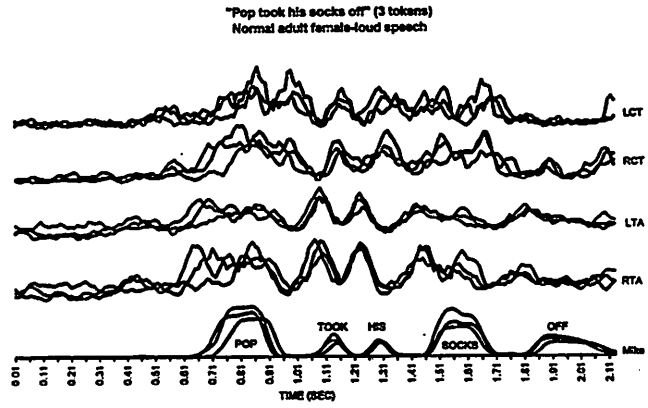


Figure 5. Superimposed traces of rectified and smoothed laryngeal EMG generated after importing data into a spreadsheet:. Nomal adult female speaker. All trace amplitudes have been normalized to the highest value occurring in the three tokens.

1. The nerves to the TA and CT muscle are intact bilaterally.
2. The motoneurons to these muscles are activated in a coordinated manner that seems related to speech, and thus corticopontine neural systems controlling laryngeal motoneurons are intact.
3. Absence of long duration or polyphasic single motor unit waveforms would be evidence against previous nerve injury followed by reinnervation.

With additional research, it may eventually be possible to make more detailed interpretations of these records. Asymmetric modulation of pairs of muscles involved in speech, or absence of modulation, may be characteristic of some diseases of the nervous system. For example, Fig. 6 shows an analysis of the recordings from the laryngeal muscles of an elderly male with Parkinson's Disease during three repetitions of the phrase "Pop took his socks off." Methods of analysis were identical those described for Fig. 5. The TA muscles in this patient are quite active during the phrase, but there is no modulation of their activity in the central part of the phrase. Cricothyroid muscles are, on the other hand, active and well-modulated. This observation could possibly be related to the rather weak and breathy voice of the patient. Warning! This comparison is VERY preliminary, and is only presented here to illustrate the potential use of EMG analysis.

Quantitative Analysis of the EMG

There are two basic features of the EMG signal that can be fairly easily reduced to a constrained set of numbers: its amplitude and its temporal properties (when it starts and stops). There are some surprisingly "sticky" problems about making these measurements, however, and our eventual ability to say whether a measure of EMG amplitude is

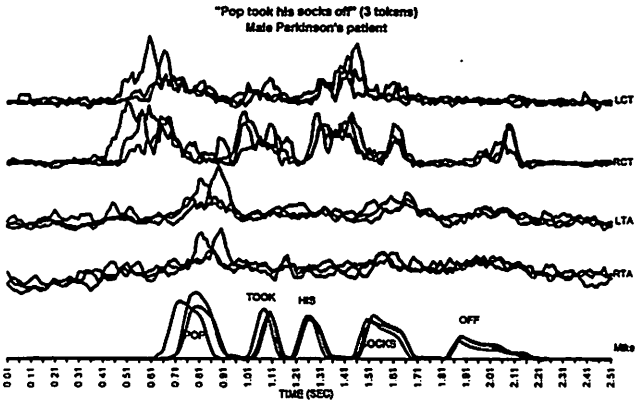


Figure 6. Analysis of the laryngeal EMG of a Parkinson's patient. Methods exactly as described for Fig. 5.

"normal" or "abnormal' will depend upon attention to these details. Let's first consider the problem of measuring the EMG of a muscle in a patient with weakness during a static maximal response. The amplitude of this response, measured in volts, could have diagnostic value. For example, if this measure in this patient were statistically the same as a normal subject, it would suggest that the contractile aspects of muscle, rather than muscle activation, were disordered. In arriving at this conclusion, what factors need to be considered?

Electrode Characteristics: First of all, the electrode type and/or physical configuration, such as interelectrode spacing, that is used to make the measurement has to be the same as that used to generate the normative data. For those interested in muscles of the articulators, intramuscular electrodes are usually required. Rigid needle electrodes, in which tip exposure and electrode spacing is fixed, would have advantages in this regard, but making maximal responses could be potentially painful, and thus might interfere with the patient's ability or willingness to "give it all they've got." Hooked wire electrodes are not painful, even for maximal responses. Historically, their tip exposure and spacing has been treated rather casually, but this could be easily improved, particularly by the use of bifiler wire. Let us suppose we make the measurement with hooked-wire electrodes manufactured with great attention to uniformity of their tip characteristics. We will still have to consider that some degree of variation will occur because of differences in the geometry of the electrode in the muscle or because of its location in a particular part of the muscle. A good example of this problem is illustrated in Fig 4a. All EMG channels were amplified at a gain of 2000. Note that the top trace (LCT) is displayed at twice the sensitivity as the other traces (2.5 rather than 5.0 volts full-scale). From this display one may see that a 2:1 difference in the general amplitude of EMG's may be expected in a normal subject, when observing bilateral pairs of muscles (which we are assuming are excited to the same degree). This problem, that of difference of EMG amplitude as a function of "whimsy", could be reduced by repeated measurements in a patient. If several electrode placements in the muscle all yielded low amplitudes, then much more significance could be attached to the observations.

Measures of EMG Amplitude: Periods of relatively steady EMG may be numerically summarized in two ways:

1. Rectification (taking the absolute value) and computing the mean amplitude over an established interval.
2. Computing the RMS (root-mean-square) value of the unrectified signal over an established interval.

Given the same record, these two methods give somewhat different values. The squaring operation in the RMS method gives greater "weight" in the final value to high-amplitude portions of the record, which are usually single-motor unit responses[204]. Root-mean-square values will obviously be larger than mean rectified EMG values, but how much larger is difficult to predict from a theoretical viewpoint. Figure 7 presents an empirical treatment of this problem. Six 200 msec periods representing low, medium, and high levels of EMG were chosen from the "pop took his socks off" record shown in Fig. 4a, and the ratio of the RMS value and mean rectified value was determined from each sample. These values are plotted as a function of the mean rectified amplitude of the LTA, the muscle in this record usually having the largest value. One may observe the ratio of RMS/mean rectified EMG ranges between one and two. For three of the muscles, the ratio does not change much as the EMG signal goes from small to large. Nevertheless, it worth noting that the LTA ratio does decrease as the signal

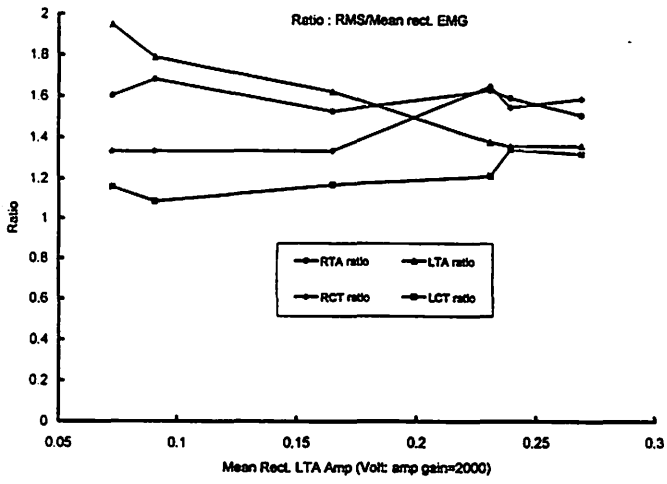


Figure 7. Ratio of values obtained for RMS vs. mean rectified EMG from the same EMG record as a function of the mean rectified EMG amplitude.

gets larger, and one may also observe that the ratio for LCT is quite a lot smaller than the other muscles. This is just an example, but it makes it clear that there is a complex relationship between RMS and mean rectified EMG "amplitudes." There is no "right" or "wrong" about the use of either measure, but the use of mean rectified EMG seems much more common in current publications than the use of RMS values, even though it requires the additional step of rectification.

Another detail to be considered when attempting to quantify EMG amplitude is the sampling rate used to digitize the signal. Whatever rate is chosen, the signal has to be low-pass filtered at a cut-off frequency of half the sampling rate to avoid the creation of false (aliased) components. Such filters are often called "anti-alias filters." Thus all high frequency components of the EMG signal above the cut-off frequency of the anti-alias filter will be lost. If the EMG signal contains large high frequency components, then the mean rectified or RMS value of this signal will obviously be spuriously low if these high frequencies are removed before they are digitized. One might avoid this problem by using a very high sampling rate, but if this is not necessary, then one would be wasting a great deal of storage space and processing time to work with the resulting large files. One might digitize the same EMG signal, recorded on a tape recorder, with a wide range of sampling rates to empirically determine the lowest sampling rate that would produce an accurate value of EMG amplitude. Spectral analysis of a typical example of EMG signal to be analyzed can also provide this information, however. Figure 8 compares the spectra of EMG signals recorded from the wrist extensor muscles of one of the authors using skin electrodes and also an intramuscular bipolar hooked-wire electrode. Signals from both electrodes pairs were anti-alias filtered at 5 kHz and digitized at 10 kHz. The skin electrodes produced an EMG signal whose high frequency components (above 300 Hz) are attenuated by 20 dB or more. The dominant frequency components are around 100 Hz. Therefore one can conclude that digitizing EMG from large skin electrodes at rates of 400Hz to 500Hz would provide accurate measures of EMG amplitude. The situation would be very different for the intramuscular recording, however. The spectrum of this signal does not fall to an attenuation level of 20 dB until frequencies of 2 kHz are reached. Therefore, we routinely use 5 kHz sampling for our studies of intramuscular EMG.

Measuring the Temporal Characteristics of EMG Activity: The duration and timing of EMG activity of muscles used in speech or swallowing have potential diagnostic value. In most respects, making these measures is merely a matter of placing the cursor at the beginning, peak, or end of periods of activity (a "burst" of EMG) and having the computer record the time (really the sample number in the record) of the cursor position. While this sounds simple, anyone who has actually done this task has

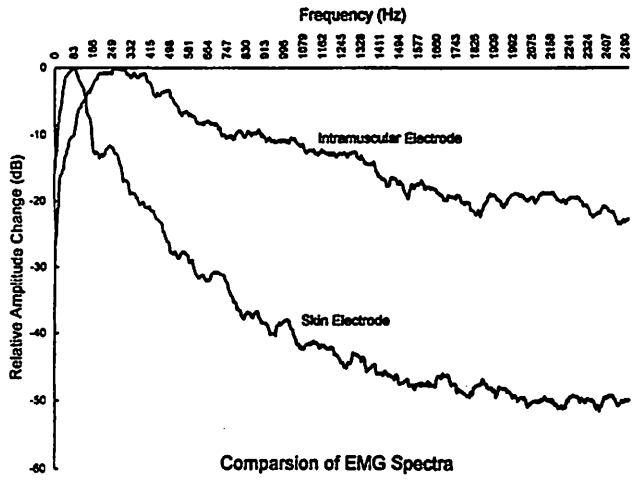


Figure 8. Comparison of the spectra of EMG obtained using intramuscular (hooked-wire) vs. skin (surface) electrodes.

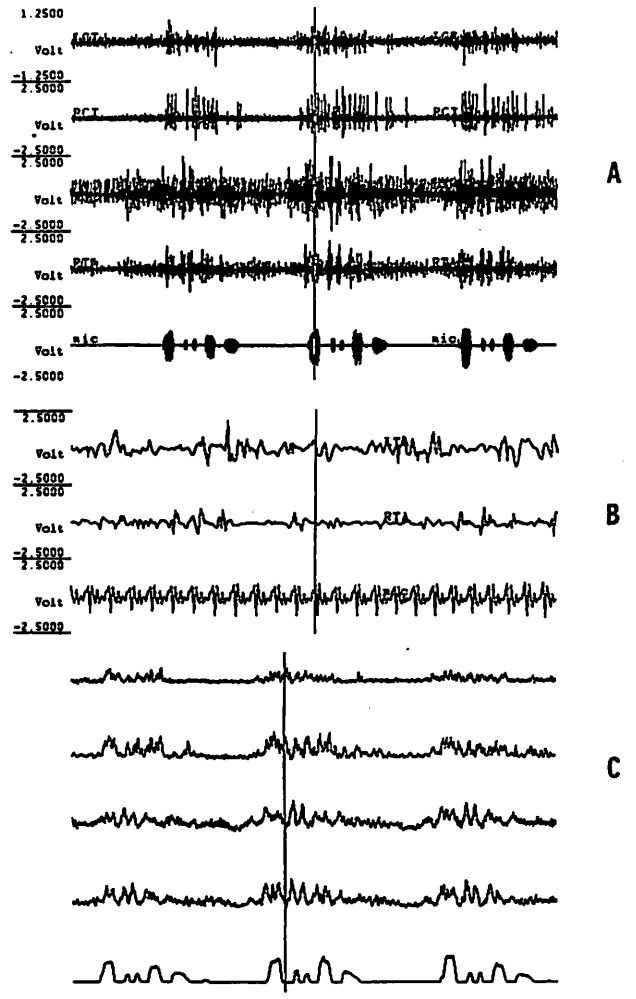


Figure 9. Subjective identification of onset of EMG activity can be influenced by "gain" of EMG display.

quickly come to discover that the "beginning" or "end" of EMG activity usually is somewhat arbitrary. The usual solution is to try to adopt some systematic rule, such as "the beginning is the first signal clearly larger than the baseline activity that persists and becomes larger." In making these judgments, however, it should be realized that very different values can be obtained by honest and diligent observers if they use the EMG trace displayed at different "gains." If the main body of the EMG burst is kept on the scale of the display, then very often the actual baseline activity cannot be seen with enough detail to know when it actually changes. This situation is illustrated in Fig. 9, which shows exactly the same burst of submental EMG activity, associated with a swallow, at three different display gains. The bottom trace, which displays the EMG at a very high gain, allows one to detect the true beginning the change in the baseline activity. The cursor position marking the beginning of the burst, as determined by the bottom trace, intersects the top trace at a point on the trace that appears to come about 200 msec before the beginning of the burst, as seen in this low gain display. In the end, people do their best, and check the reliability within and between observers. It is well to keep in mind, however, that very different measurements of EMG burst duration might result from the manner in which the signals to be measured were displayed on the computer.

Acknowledgments

We would like to acknowledge the support of the National Center for Voice and Speech (grant no. P60 DC00976 from the National Institutes of Deafness and Other Communication Disorders) and extend our thanks to Drs. Lorraine Ramig and Kristin Larson for permission to use the recordings used in figures 4-7. A special expression of thanks goes to Dunc for singing his heart out under adverse circumstances.

References

1. Piper, H.: Elektrophysiologie menschlicher Muskeln. Berlin, Springer, 1912.

2. Kimura, J.: Electrodiagnosis in diseases of nerve and muscle: principles and practice. 2nd ed. Philadelphia, F.A Davis Co., 1989

3. Faaborg-Andersen K: Electromyographic investigation of intrinsic laryngeal muscles in humans. Acta Physiol Scand 1957; 41:9-148.

4. Hiroto I, Hirano M,Tomita H: Electromyographic investigation of human vocal cord paralysis. Ann. Otol. Rhinol. Laryngol. 1968; 77:296-304.

5. Blair RL, Berry H,Briant TDR: Laryngeal electromyography:Technique and application. Otolaryngologic Clinics of North America 1978; 11:325-346.

6. Haglund S, Knutsson E,Martensson A: An electromyographic investigation of intrinsic laryngeal muscles in humans. Acta Oto Laryngologica 1972; 74:265-270.

7. Thumfart W.: Electromyography of the larynx, in Samii M ,Gannetta PJ, (eds): The cranial nerves. Berlin, Springer-Verlag; 1981, pp 597-606.

8. Parnes SM,Satya MS: Predictive value of laryngeal electromyography in patients with vocal cord paralysis of neurogenic origin. Laryngoscope 1985; 95:1323-6.

9. Hirano M, Nozoe I, Shin T,Maeyama T. Electromyography for laryngeal paralysis, in Hirano M, Kirchner J , Bless D (eds) Neurolaryngology: recent advances. Boston, College-Hill; 1987, pp :232-248.

10. Min YB, Finnegan EM, Hoffman HT, Luschei ES,McCulloch TM: A preliminary study of the prognostic role of electromyography in laryngeal paralysis. Otolaryngology-Head and Neck Surgery 1994; 111:70-75.

11. Berry H,Blair RL: Isolated vagus nerve palsy and vagal mononeuritis. Archives of Otolaryngology 1980;106: 333-8.

12. Crumley RL: Laryngeal synkinesis: its significance to the laryngologist. Annals of Otology, Rhinology & Laryngology 1989; 98:87-92.

13. Hoffman HT, Brunberg JA, Winter P, Sullivan MJ,Kileny PR: Arytenoid subluxation: diagnosis and treatment. Annals of Otology, Rhinology & Laryngology 1991; 100:1-9.

14. Lewis WS, Crumley RL, Blanks RH,Pitcock JK: Does intralaryngeal motor nerve sprouting occur following unilateral recurrent laryngeal nerve paralysis? Laryngoscope 1991; 101:1259-1263.

15. Mu L,Yang S: An experimental study on the laryngeal electromyography and visual observations in varying types of surgical injuries to the unilateral recurrent laryngeal nerve in the neck. Laryngoscope 1991; 101:699-708.

16. Kokesh J, Flint PW, Robinson LR,Cummings CW: Correlation between stroboscopy and electromyography in laryngeal paralysis. Annals of Otology, Rhinology & Laryngology 1993; 102:852-7.

17. Sutherland DH, Larsen LJ,Mann R: Rectus femoris release in selected patients with cerebral palsy: a preliminary report. Developmental Medicine & Child Neurology 1975; 17:26-34.

18. Keenan MA: Surgical decision making for residual limb deformities following traumatic brain injury. [Review]. Orthopaedic Review 1988; 17:1185-92.

19. Cahan LD, Adams JM, Perry J,Beeler LM: Instrumented gait analysis after selective dorsal rhizotomy. Developmental Medicine & Child Neurology 1990; 32:1037-43.

20. Scott AB: Botulinum Toxin injection into extraocular muscles as an alternative to strabismus surgery. Journal of Pediatric Ophthalmology and Strabismus 1980; 17:21-25.

21. Mauriello JA: Blepharospasm, Meige's syndrome, and hemifacial spasm: treatment with botulinum toxin. Neurology 1985; 35:1499-1500.

22. Tsui J, Eisen A, Mak E, Carruthers J, Scott A,Calne D: A pilot study on the use of botulimun toxin in spasmodic torticollis. Can J Neurol Sci 1985; 12:314-316.

23. Jankovic J,Orman J: Botulinum A toxin for cranial-cervical dystonia: A double-blind, placebo-controlled study. Neurology 1987; 37:616-623.

24. Brin MF, Fahn S, Moskowitz C, Friedman A, Shale HM, Greene PE, Blitzer A, List T, Lange D, Lovelace RE,McMahon D: Localized injections of botulinum toxin for the treatment of focal dystonia and hemifacial spasm. Movement Disorders 1987; 2:237-54.

25. Cohen SR,Thompson JW: Use of botulinum toxin to lateralize true vocal cords: a biochemical method to relieve bilateral abductor vocal cord paralysis. [Review]. Annals of Otology, Rhinology & Laryngology 1987; 96:534-41.

26. Cohen LG, Hallett M, Geller BD,Hochberg F: Treatment of focal dystonias of the hand with botulinum toxin injections. Journal of Neurology, Neurosurgery & Psychiatry 1989; 52:355-63.

27. Stager S,Ludlow C. Responses of strutterers and vocal tremor patients to treatment with botulinum toxin, in Janovic J , Hallett M (eds) Therapy with Botulinum Toxin. New York, Mercel Dekker, Inc, 1994.

28. Miller RH, Woodson GE,Jankovic J: Botulinum toxin injection of the vocal fold for spasmodic dysphonia. A preliminary report. Archives of Otolaryngology Head & Neck Surgery 1987; 113:603-5.

29. Blitzer A, Brin MF, Fahn S,Lovelace RE: Localized injections of botulinum toxin for the treatment of focal laryngeal dystonia (spastic dysphonia).Laryngoscope 1988; 98:193-7.

30. Ludlow CL, Naunton RF, Sedory SE, Schulz GM,Hallett M: Effects of botulinum toxin injections on speech in adductor spasmodic dysphonia. Neurology 1988; 38: 1220-5.

31. Ludlow C, Hallett M, Sedory S, Fujita M,Naunton R: The pathophysiology of spasmodic dysphonia and its modification by botulinum toxin, in International Medical Society of Motor Disturbances. Motor Disturbances II. San Diego, Academic Press, Ltd.; 1990, pp 273-288.

32. Ludlow CL, Naunton RF, Terada S,Anderson BJ: Successful treatment of selected cases of abductor spasmodic dysphonia using botulinum toxin injection. Otolaryngology Head & Neck Surgery 1991; 104:849-55.

33. Min YB, Luschei, ES, Finnegan EM, McCullough TM,Hoffman HT: Portable telemetry system for electromyography. Otolaryngology-Head and Neck Surgery 1994; 111:849-852.

34. De Luca C: Physiology and mathematics of myoelectric signals.IEEE Trabsactions on Biomedical Engineering 1979; BME-26:313-325.

35. Loeb, G.E. and Gans, C: Electromyography for Experimentalists, Chicago, University of Chicago Press, 1986.

36. Cole K, Konopacki R,Abbs J: A minerature electrode for surface electromyography during speech. J. Acoust. Soc. Am. 1983; 74:1362.

37. Smith A, McFarland DH, Weber CM,Moore CA: Spatial organization of human perioral reflexes. Exp. Neurol. 1987; 98:233-248.

38. Smith A: Neural Drive to muscles in stuttering.J. Speech Hear. Res. 1989; 32:252-264.

39. McClean MD: Lip muscle EMG responses to oral pressure stimulation. J. Speech Hear. Res. 1991; 34:248-251.

40. Blair C,Smith A: EMG recording in human lip muscles: can single muscles be isolated? J. Speech Hear. Res. 1986; 29:256-66.

41. Moore CA, Smith A,Ringel RL: Task-specific organization of jaw muscles. J. Speech Hear. Res. 1988; 31:670-680.

42. Basmajian JV,Stecko G: A new bipolar electrode for electromyography. J. Appl. Physiol. 1962; 17:849.

43. Basmajian JV, De Luca, C.J: Muscles alive: Their functions revealed by electromyography. Baltimore,Wiliams &Wilkins, 1985.

44. Hirano M,Ohala J: Use of hooked-wire electrodes for electromyography of the intrinsic laryngeal muscles. J. Speech Hear. Res. 1969; 12:362-73.

45. Hirose H: Laryngeal articulatory adjustments in terms of EMG, in Hirano M, Kirchner J, Bless D (eds): Neurolaryngology: Recent Advances Boston, College-Hill Pub. 1987, pp 200-208.

46. Freeman FJ,Ushijima T: Laryngeal muscle activity during stuttering. J. Speech Hear. Res. 1978; 21:538-62.

47. Thumfart WF: From larynx to vocal ability. New electrophysiological data. Acta Oto Laryngologica 1988; 105: 425-31.

48. Woo P,Arandia H: Intraoperative laryngeal electromyographic assessment of patients with immobile vocal fold. Annals of Otology, Rhinology & Laryngology 1992; 101:799-806.

49. Koda J,Ludlow CL: An evaluation of laryngeal muscle activation in patients with voice tremor. Otolaryngology Head & Neck Surgery 1992; 107:684-96.

50. Gartlan MG, Peterson KL, Luschei ES, Hoffman HT,Smith RJ: Bipolar hooked-wire electromyographic technique in the evaluation of pediatric vocal cord paralysis. Annals of Otology, Rhinology & Laryngology 1993; 102:695-700.

51. Hrycyshyn AW,Basmajian JV: Electromyography of the oral stage of swallowing in man. Amer. J. Anat. 1972; 133:333-340.

52. Sauerland EK,Mitchell SP: Electromyographic activity of intrinsic and extrinsic muscles of the human tongue. Texas Reports on Biology and Medicine 1975; 33:445-455.

53. Borden GJ,Gay T: On the production of low tongue tip /s/: a case report. Journal of Communication Disorders 1978; 11:425-31.

54. Mowrey RA,McKay IRA: Phonological primitives: electromyographic speech error evidence. J. Acoust. Soc. Am. 1990; 88:1299-1312.

55. Palmer JB, Rudin NJ, Lara G,Crompton AW: Coordination of mastication and swallowing. Dysphagia 1992; 7: 187-200.

56. Basmajian JV,Dutta CR: Electromyography of the pharyngeal constrictors and levator palati in man. Anat. Rec 1961; 139:561-563.

57. Fritzell B: The velopharyngeal muscles in speech: An electromyographic and cineradiographic study. Acta Oto-Laryng. (Suppl 250) 1969; 1-81.

58. Lubker JF, Fritzell B,Lindqvist J: Velopharyngeal function: An eletromyographic study. STL-QPSR 1970; 20: 9-20.

59. Ushijima T,Hirose H: Electromyographic study of the velum during speech. J. Phonetics 1974; 2:315-326.

60. Bell-Berti F: An electromyographic study of velopharyngeal function in speech. J. Speech Hear. Res. 1976; 19:225-240.

61. Benguerel AP, Hirose H, Sawashima M,Ushijima T: Electromyographic study of the velum in French. Ann. Bull. Research Institute of Logopedics and Phoniatrics, Tokyo 1975; 9:79-90.

62. Fritzell B,Kotby MN: Observations on thryroarytenoid and palatal levator activation for speech. Folia Phoniatrica 1976; 28:1-7.

63. Kiritani S, Hirose H,Sawashima M: Simultaneous X-ray microbeam and EMG study of velum movement for Japanese nasal sounds. Ann. Bull. Research Institute of Logopedics and Phoniatrics 1980; 14:91-100.

64. Seaver EJ,Kuehn DP: A cineradiographic and electromyographic investigation of velar positioning in non-nasal speech. Cleft Palate Journal 1980; 17:216-226.

65. Hairston LE,Sauerland EK: Electromyography of the human palate: Discharge patterns of the levator and tensor veli palatini. Electromyogr. Clin. Neurophysiol. 1981; 21:299-306.

66. Kuehn DP, Folkins JW,Cutting CB: Relationships between muscle activity and velar position. Cleft Palate Journal 1982; 19:25-35.

67. Cooper DS,Folkins J: Comparison of electromyographic signals from different electrode placements in the palatoglossus muscle. J. Acoust. Soc Am. 1985; 78: 1530-1540.

68. Kuehn DP, Folkins JW,Linville RN: An electromyographic study of the musculus uvulae. Cleft Palate Journal 1988; 25:348-55.

69. Hairston LE,Sauerland EK: Electromyography of the human pharynx: discharge patterns of the superior pharyngeal constrictor during respiration. Electromyography & Clinical Neurophysiology 1981; 21:299-306.

70. Rowe LD, Miller AJ, Chierici G,Clendenning D: Adaptation in the function of pharyngeal constrictor muscles. Otolaryngology Head & Neck Surgery 1984; 92:392-401.

71. Palmer JB, Tanaka E,Siebens AA: Electromyography of the pharyngeal musculature: technical considerations. Archives of Physical Medicine & Rehabilitation 1989; 70: 283-7.

72. Perlman AL, Luschei ES,DuMond CE: Electrical activity from the superior pharyngeal constrictor during reflexive and nonreflexive tasks. J. Speech Hear. Res. 1989; 32: 749-754.

73. O'Dwyer NJ, Neilson PD, Guitar BE, Quinn PT,Andrews G: Control of upper airway structures during nonspeech tasks in normal and cerebral-palsied subjects: EMG findings. J. Speech Hear. Res. 1983; 26:162-170.

74. Neilson PD,O'Dwyer NJ: Reproducibility and variability of speech muscle activity in athetoid dysarthria of cerebral palsy. J. Speech Hear. Res. 1984; 27:502-517.

75. Hunker CJ,Abbs JH: Uniform frequency of parkinsonian resting tremor in the lips, jaw, tongue, and index finger. Movement Disorders 1990; 5:71-7.

76. Palmer JB, Tippett DC,Wolf JS: Synchronous positive and negative myoclonus due to pontine hemorrhage.Muscle & Nerve 1991; 14:124-32.

77. Schaefer SD, Roark RM, Watson BC, Kondraske GV, Freeman FJ, Butsch RW,Pohl J: Multichannel electromyographic observations in spasmodic dysphonia patients and normal control subjects. Annals of Otology, Rhinology & Laryngology 1992; 101:67-75.

NCVS Status and Progress Report - 8
July 1995, 175-176

Training Update

Erich Luschei/Patricia Zebrowski, Training Coordinators
Department of Speech Pathology and Audiology, The University of Iowa

Predoctoral Trainees

Phyllis Palmer, M.A., chose to seek a doctoral degree at The University of Iowa following several years of clinical work, most recently as Director of Speech Pathology at The University of New Mexico Hospital. She had previously earned a bachelor's degree and master's degree in Communication Disorders from State University of New York-New Paltz, and Emerson College, respectively. Her primary areas of interest are voice and swallowing. More specifically, she is interested in studying clients with tracheostomies and the effects of premature birth on voice and swallowing skills. She works closely with Drs. Erich Luschei and McCulloch (the latter in the Department of Otolaryngology and Head and Neck Surgery at the University of Iowa).

Eileen Finnegan, M.A. earned a bachelor's degree in Communicative Disorders in 1980 and a master's degree in Speech Pathology and Audiology from the University of Iowa in 1982. For nearly ten years, she worked as a speech-language pathologist, focusing on voice as her major area of interest. She returned to the University of Iowa as a doctoral student in January 1992. Her pre-dissertation research project involved inherent difficulties of using indirect behavioral methods of estimating subglottic pressure during phonation to calculate the laryngeal resistance of individuals with spasmodic dysphonia. She passed her Comprehensive Exam in Summer 1995, and is developing her thesis research in an area of neurolaryngology under the direction of Dr. Erich Luschei.

Elease White, M.A. began doctoral study at Iowa in August 1993. Before coming to Iowa, she was an Assistant Professor of Communication Studies at Tougaloo College in Mississippi. Elease had earlier earned a bachelor's degree from Jackson State University and a master's degree from the University of Illinois. Her previous experience has also included an internship with the Mississippi Health Department, studying rural minority elderly clients and their communication with health professionals and the design of workshops to review the verbal and vocal strategies of African American orators. Future research goals include the study of gospel singers and ministers and incidence of vocal pathology as a consequence of singing and preaching.

Brad Story, Ph.D., former predoctoral trainee, completed his thesis in May of 1995. The title of his thesis is "Physiologically-Based Speech Simulation Using an Enhanced Wave-Reflection Model of the Vocal Tract." This research was done under the direction of Dr. Ingo R. Titze. Brad continues to work as an Assistant Research Scientist with Dr. Titze on quantitative models of the vocal tract.

Postdoctoral Trainees

Kristin Larson, Ph.D. continues her studies of the voice of patients with Parkinson disease under the sponsorship of Dr. Lorraine Ramig. During the last year, she has worked at the DCPA with Drs. Ramig, Luschei and an otolaryngologist, Dr. Marshall Smith, to record laryngeal EMG in normal elderly patients and patients with Parkinson's disease.

Robert Lange, Ph.D. earned a bachelor's degree in physics from Kenyon College, graduating with distinction for his senior thesis, an analysis of musical sounds. He worked as a research intern for the ultrasonics laboratory in the Physics Department at Ohio University and later as technician in the area of high-performance ceramics research at NASA Lewis Research Center. In 1988, he enrolled in the graduate program in acoustics at Penn State University. In 1994, Dr. Lange completed his doctoral degree in acoustics with a thesis entitled "A Wavelet Model

for Vocalic Speech Coarticulation". He began his postdoctoral fellowship with the NCVS at the Denver site in September 1994. Research interests include speech acoustics, articulatory recognition, wavelet signal processing, speech time-compression, and acoustic hall simulation.

David Berry, Ph.D. received both his bachelor and doctoral degrees in physics at Brigham Young University. His major area of interest in graduate work was the acoustics of speech, with a minor in digital signal processing. His dissertation, "Isolation and Characterization of Microevents in Speech," focused on time-frequency issues for enhancing speech analysis. In 1991, he began a three-year postdoctoral fellowship at the NCVS under the direction of Dr. Ingo Titze. At the NCVS, he is augmenting his studies of speech by gaining further experience with speech simulation and modeling. More specifically, he has been working on finite element modeling of the vocal folds, theoretical analysis of stress in vocal fold tissues, sources of irregularity and chaos in vocal fold vibrations, and high-speech vocal fold imaging.

Michael Bayerl, M.D. earned a bachelor's degree with distinction in biology and M.D. degree at The University of Michigan. In order to prepare for a career as an academic otolaryngologist, he chose to accept a fellowship through the NCVS. Dr. Bayerl resigned his traineeship at the end of Spring Semester to enter a residency program. During his training, his areas of investigation included the effect of local and systemic fluid status on phonation, kinematics of the cricoarytenoid joint complex, correlation of histology of the vocal fold with biomechanical properties and the clinical utility of phonatory function data in the initial diagnosis and disposition of voice patients.

NCVS Status and Progress Report - 8
July 1995, 177-179

Continuing Education Update

Julie Ostrem, Continuing Education Coordinator
Department of Speech Pathology and Audiology, The University of Iowa

Over the past five years, the Continuing Education project has continued to mature along with other areas of the NCVS. In the first granting period, investigators used written materials and oral presentations to convey their research findings and other information about voice and speech to practitioner groups. Members of the target audience are speech-language pathologists, otolaryngologists and other voice professionals: choral conductors, voice coaches and theatre trainers.

The experience of producing these materials and communicating frequently with clinicians has sharpened the focus of future goals. In the new granting period beginning September 1, 1995, NCVS investigators will scale back on the creation of written pieces, while concentrating heavily on the development of educational software. Specifically, a CD-ROM will be designed over the next four or five years to help practitioners better understand normal human voice production. Furthermore, a teaching and training schedule has been developed to further increase investigators' involvement in professional conferences and workshops. Also, an audio cassette tape will be made for clinicians containing numerous samples of disordered voices. Speech-language pathologists and otolaryngologists can listen repeatedly to this tape in their portable tapeplayers--perhaps as they exercise or while they are driving-- to better discern subtle differences between types of disorders. Finally, an innovative textbook is planned that introduces the new field of vocology.

National Meetings and Conferences

In the writing of the NCVS grant proposal last October, investigators tallied the total number of individuals reached through CE efforts. Representation of NCVS investigators at professional meetings was particularly impressive, with a combined audience of more than 50,000. It was especially gratifying to note the variety of meetings with NCVS representation. The full list of these conferences is presented in Table 1 (following page).

To encourage NCVS investigators to continue their strong tradition of "giving back" to their professions, a teaching schedule and travel support program has been established. These awards pay travel expenses for investigators to present information at professional meetings for which no honoraria nor travel support can be given. These funds also may be used to help sponsor a local conference to provide continuing education.

An example of CE-supported sponsorship was the *Voice and Resonance Disorders in Children Conference* organized by Dr. Marshall Smith in June 1995. The interactive forum allowed speech-language pathologists and pediatric otolaryngologists to discuss evaluation methods and behavioral, medical and surgical therapies for voice and resonance problems in children. Thomas Campbell, Ph.D., CCC-SLP, Associate Professor of Pediatric Neurology at the University of Pittsburgh School of Medicine, was guest faculty. Other faculty included Mary Kay Farina, Dr. Ronald Scherer, Dr. Smith and Dr. R.C.A. Weatherley-White.

Additionally, the successful series of Phonosurgery Conferences will be continued in 1996, 1998 and 2000 at The University of Wisconsin-Madison. Drs. Diane Bless and Charles Ford will continue to select both faculty and topics for inclusion. In response to feedback from past participants, Drs. Ford and Bless have decided to narrow the focus of topics, although the primary theme of the conference will remain unchanged: meeting multidisciplinary needs for education on topics related to phonosurgery. Conference coordinators also will videotape future meetings, so the proceedings will be available to those unable to attend the conference.

As an accompaniment to the oral presentations at meetings and conferences, a portable exhibit was created in 1994. In addition to providing information about ongoing research at the NCVS, the display provides an attractive backdrop from which NCVS printed materials can be distributed. The display made its debut at the November 1994 ASHA convention. Since that time, it has been also shown at

the Denver voice workshops, Care of the Professional Voice Symposium, Voice and Resonance Disorders in Children Conference and the National Association of Teachers of Singing annual convention.

Table 1.
NCVS Representation at National and International Meetings
October 1990-August 1995

A.B. Bell Conference for the Deaf
Acoustical Society meetings (ASA)
American Society of Pediatric Otolaryngology (ASPO)
American Academy of Pediatrics
American Cleft Palate Association (ACPA) conference
American Association of Pediatricians (AAP)
American Broncho-Esophageal Association (ABEA)
American Academy of Otolaryngology-Head and Neck Surgery conference (AAO-HNS)
American Otolaryngology Association
American Speech-Hearing-Language Association meeting (ASHA)
Association for Research in Otolaryngology (ARO)
California Speech-Language-Hearing Association
Canadian Association of Speech Language Pathologists and Audiologists
Care of the Professional Voice annual meeting
Clinical Aphasiology Conference
Clinical Conference for Otolaryngologists and Speech Pathologists (at Iowa)
Collegium Medicorum Theatri
Colloque International d'Audiophonologie conference
Conference on Laryngeal Cancer
Down Syndrome Association of Wisconsin
First World Congress on Fluency Disorders
Georgia Speech-Language-Hearing Association
Illinois Speech-Language-Hearing Association
International Congress on Cleft Palate and Craniofacial Anomalies meeting
International Congress of Voice Teachers
International Phonosurgery Association conference
Iowa Speech-Language-Hearing Association (ISHA)
London Voice Society
Mid-South Conference of Communicative Disorders
National Down Syndrome Congress
National Association of Teachers of Singing convention (NATS)
Oto-rhino-laryngologie et Audiophonologie Association (ORLAS) meeting
Pacific Voice Conference
Pediatric Voice and Resonance Disorders in Children Conference
Phonosurgery Symposia
Second Symposium on Treatment Research
Speech Motor Control Conferences
Stockholm Music Acoustics Conference (SMAC)
Texas Speech-Language-Hearing Association
Third Seminar on Speech Production
Triological Society meeting
Utah Speech-Language-Hearing Association
Vocal Fold Physiology meetings
Voice and Speech Trainers Association, Inc. conference (VASTA)
Voice Care Seminars (in Denver, CO)
Voice Symposium (Dublin, Ireland)
Wisconsin Society of Otolaryngologists meeting

Printed Materials

Nearly all publication goals for the first granting period now have been met. *The Vocologist's Guide to Voice Therapy and Training*, produced in 1994, has been distributed to more than 1,500 clinicians and students of speech pathology. A Spanish version of the guide was translated by Carmen Ana Ramos, a doctoral student at The University of Wisconsin-Madison. This version, entitled *Entrenamiento y terapia de la voz: Una guia para vocologistas* is available from the NCVS free of charge. Ms. Ramos has agreed to help distribute the guide to speech pathologists in her native Puerto Rico and other Central American countries.

The Lee Silverman Voice Treatment: A Practical Guide to Treating the Voice and Speech Disorders in Parkinson Disease is now available. Targeted specifically for speech-language pathologists, the guide fully explains how to administer this new and effective technique. Exercises, worksheets and other aids for delivering the LSVT are printed in the guide's comprehensive appendix. The clinician can easily photocopy any of these appendices for use in the treatment room. The LSVT manual is available from the NCVS offices at Iowa or Dr. Ramig at the DCPA. Profits from the workbook will be used to support further efficacy studies.

Other materials, co-developed with investigators of other projects, also accompany the display. An attractive brochure describes research training opportunities with the NCVS. "Workshop on Acoustic Voice Analysis: Summary Statement" was produced with investigators on the Technology Transfer project. It provides a synopsis of a workshop held in February 1994 between representatives of industry and academia. The "Tips to Keep you Talkin'..." bookmarks, as described in the Dissemination of Information Update, have been tremendously successful. Many speech pathologists have requested quantities of several hundred to give away during Better Speech & Hearing month, at inservices, or to display in their waiting rooms. Finally, copies of the *NCVS Status and Progress Report* are displayed, increasing the distribution list by about twenty percent.

Educational Software

Based on experiences in the first five years of providing CE, NCVS investigators have selected the CD-ROM as an ideal format for learning. First, a CD-ROM organizes and stores an enormous amount of information on one disk: roughly equivalent to 680,000 pages of written text. Second, mixed media can be combined into one product. Text, two-dimensional graphics, three-dimensional images, audio, quick-time movies, animations and combinations of these media are all possible strategies to explain concepts. With a full repertoire of these strategies, the

programmer is not limited by the boundaries of the media. Third, CD-ROMs can be programmed to give feedback to the learner. Educational strategies that emphasize interactive learning have long been recognized as superior to passive listening activities. Fourth, unlike short courses or conferences, CD-ROM allows learners to move at a rate suitable to their needs. Particularly complex concepts can be demonstrated an infinite number of times. Fifth, CE-ROMs are inexpensive. A CD-ROM can be purchased at a fraction of the cost of attending just one conference. Sixth, one CD-ROM can be shared by a number of individuals. It is anticipated that the CD-ROM developed by the NCVS will have as its primary consumers large clinics and departments of speech science. Individuals within these groups can share a single copy.

Assessment of Needs Survey

To determine the most useful content matter for the CD-ROM and other future CE products, a questionnaire was sent to the 215 programs of speech-language pathology in the U.S. The overall objective was simple: to determine the baseline understanding of voice production and voice disorders for graduates from master's programs in speech-language pathology.

Thirty-three percent of the programs responded to the survey. The results are shown in Table 2. Although nearly all programs offer a course in voice disorders, considerably fewer (less than ten percent) have course offerings in normal voice production. This leads to a concern that while clinicians possess acceptable technical skills acquired in a voice disorders class, they may lack a broad understanding of the vocal mechanism. This knowledge gap prevents the clinician from critically evaluating new therapies and keeping abreast of research findings.

Based on this information, the primary content of the CD-ROM will be normal voice production. Dr. Ingo Titze's textbook, *Principles of Voice Production,* will be used as a basis for much of the subject matter, although computer models, audio and video clips, and other embellishments will be gleaned from virtually every laboratory of the NCVS. It is anticipated that the project will be developed over a four- or five-year period.

As communications technology continues to expand, CE investigators have the ability to convey information faster, better and cheaper. It is anticipated that most of our products--including the NCVS Status and Progress Reports and the CD-ROM--will be distributed via World Wide Web in the near future. We welcome these technologic advances. They not only allow us to update information easily, but also provide another way to easily communicate with those we serve.

Table 2.
Needs Assessment for CE Programs
(Van Mersbergen, Ostrem, Titze, 1994)

Part I. Student interest

Percentage of graduate students with primary interest in voice.(criteria: opted for a thesis in voice, asked for extra clinical hours in voice, or stated preference): **18 percent**

Percentage of graduate students with a performance background in voice.(criteria: currently active in performance, college bachelors degree in music or theater): **4.4 percent**

Part II. Instructors/Clinical Supervisors

Number of SLP voice instructors in the department

1.6 = mean, 1 = mode, 0-6 = range

Programs with instructors strictly dedicated to voice

12 of the 71 schools surveyed (16.9 percent)

Number of clinical SLP voice supervisors employed in the department

1.7 = mean, 1 = mode 0-5 = range

Programs with voice supervisors strictly dedicated to voice

8 of the 71 schools surveyed (11.3 percent)

Part III. Coursework Required in Voice

Normal Voice Production

Academic semester hours required in normal voice production

2.25 = mean, 3 = mode, 0-5=range

Number of programs with a course wholly dedicated to normal voice production

7 of the 71 schools surveyed (9.9 percent)

Voice Disorders

Academic semester hours required in voice disorders

3.9 = mean, 3 = mode, 0-4=range

Number of programs with a class wholly dedicated to voice disorders

69 of the 71 schools surveyed (97.2 percent)

Part IV. Clinical Requirements

Clinical ASHA hours required in voice disorders

14 schools of 71 stated they had a requirement (19.7 percent)

Of those 14 schools, required hours:

15.3 = mean, 15 = mode, 5-25=range

Estimation of the average voice hours a graduate student receives (all programs)

14.3 = mean, 15 = mode, 0-100=range

NCVS Status and Progress Report - 8
July 1995, 181-182

Dissemination of Information Update

Cynthia Kintigh, Dissemination Coordinator
Wilbur James Gould Voice Research Center, The Denver for the Performing Arts

The goals of the dissemination project are to distribute information to the general public about care of the voice as well as prevention, detection, and treatment of voice and speech disorders. Methods for reaching the public include the use of media (electronic and print), educational presentations (workshops, seminars and lectures), and shared information with other professional organizations dedicated to voice care, speech and training.

This final year of the first NCVS grant saw placement of the finishing touches on projects begun in the first granting period with an eye on projects that will become the foundation for the campaign of the second granting period.

Voice Workshops

The public voice workshops continue to be one of the NCVS's most visible outreach efforts. In the initial five years of the grant, we have reached over 2500 members of the general public with an urgent need for information about the voice. Again, our three large public voice workshops were presented. In previous years, all of these large workshops had been presented in one of the theaters at the Denver Center for the Performing Arts. This year, as an additional community outreach effort, two of the three were presented at suburban arts centers, The Arvada Center for the Arts and Humanities and the Littleton Town Hall Arts Center.

The October voice workshop was targeted at choral singers; the February and May suburban workshops were focused on the speaking voice.

Highlights of other public outreach were a workshop presented in April through the Denver Center for the Performing Arts Native American Task Force and two workshops presented as part of the Denver Center for the Performing Arts Residencies in the Schools program which took place in October. The workshop for Native Americans focused on the effects of traditional Native American chanting and singing on the voice and offered applicable voice care techniques. The Residencies in the Schools program takes place annually through the DCPA's educational outreach program. Schools selected for the program typically have a student body comprised of underserved audiences. The residency is marketed to students who are now participating in school performing arts classes and programs and who may be considering careers in the performing arts.

Voices of People and Machines

Final editing continues on a thirty-minute video version of the arts-versus-science play, Voices of People and Machines. The award-winning Denver Center Media Department is producing the video of Voices of People and Machines. Copyright issues have delayed the distribution of the video but the video is in the final stages of production and editing and will be ready for distribution through the Acoustical Society of America in the fall of 1995. The finished product will have a running time of 30 minutes which makes it suitable for classroom use.

Public Service Announcement

A thirty-second television Public Service Announcement was produced for airing in ten select secondary markets. The goal of the PSA is to raise awareness in the general public. The PSA is a colorful, humorous, and clever play on expressions used to describe minor voice problems, such as frog in the throat and nagging hoarse voice. In the PSA, an aerobics instructor, a salesperson, a teacher, and some telemarketers are portrayed, however, the viewer does not hear their voices. Rather an animal sound, such as a horse neighing or a frog croaking is heard in place of the actors voice. In the voice over the viewer is given the following call to action:

"Talk is cheap, your voice isn't. If you have a nagging hoarse voice or a recurring frog in the throat, call your doctor. Or the National Center for Voice and Speech. And take care of your voice, before it goes to the dogs."

The phone number for the Denver site of the NCVS appears on the screen. Calls from outside the Denver area will be referred to an appropriate agency in the viewers area.

The markets selected for the first round of distribution were chosen based on proximity to a major voice research center (e.g.: Denver, Tucson) and/or a high number of voice users in the population (e.g.: Nashville, New Orleans).

Table 1.
Public Service Announcement Distribution

Market	Number of Television Households
Denver, CO	1,142,160
Cedar Rapids/Iowa City, IA	376,160
Salt Lake City, UT	616,700
Nashville, TN	749,060
Minneapolis/St. Paul, MN	1,410,630
Seattle, WA	1,468,730
New Orleans, LA	615,180
Atlanta, GA	1,567,300
Tucson, AZ	322,800
Boston, MA	2,104,900
Total potential exposure:	10,373,620

Source: May 1995 A. E. Nielson report.

Note: A television household is measurement used by the A. E. Nielson company in measuring ratings and is the standard by which television markets are ranked. The measurement represents the number of households within a city, not the number of occupants or televisions.

Tips to Keep You Talkin'...

A colorful bookmark, Tips to Keep You Talkin'... was produced. The bookmark is a handy guide to general voice care and is available in English, Spanish and French.

To date 8000 of these bookmarks have been distributed at events attended by the general public. Those events include the Denver Chamber of Commerce Business to Business Expo, the Boulder Creek Festival, the Capitol Hill Peoples Fair, the Denver Gay PrideFest, and the Denver Black Arts Festival. The bookmark has also received wide distribution at professional voice conferences including ASHA and the Voice Foundation.

Pavarobotti

Work continues on the second generation of Pavarobotti the singing robot. The long-range plan for Pavarobotti is for a self-contained, interactive exhibit which would teach young people about the voice, and would tour to science and children's museums across the country. The redesign of Pavarobotti allows for his expansion into that role. This new design will integrate audio and video using new multimedia technology. The robot will be upgraded to be more robust and dexterous.