NCVS Status and Progress Report
Volume 7/December 1994

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

Editorial and Distribution Information

Editor, Ingo Titze
Production Editors, Julie 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

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.
Andrea Beckmann, B.S.
Patricia Benjamin, B.S.
Bridget Berning
Florence Blager, Ph.D.
Diane Bless, Ph.D.
James Brandenburg, M.D.
Gail Brosovic
Geron Coale, M.A.
Stefanie Countryman, M.A.
Jeanne Delaney, B.A.
Heather Dove, M.A.
Christopher Dromey, M.A.
Wendy Edwards, B.A.
Penny Farrell
Charles Ford, M.D.
Steven Gray, M.D.
Chwen-Geng Guo, M.S.
James Guo, Ph.D.
Susan Hensley, M.M.
Marilyn Hetzel, Ph.D.
Margaret Hoehn, M.D.
Henry Hoffman, M.D.
Yoshiyuki Horii, Ph.D.
Heather Hughes
Kathy Ives
Bruce Jafek, M.D.
Antonia Johnson, B.S.
Joel Kahane, Ph.D.
Judith King, Ph.D.
Cynthia Kintigh, M.A.
David Kuehn, Ph.D.
Jon Lemke, Ph.D.
Julie Lemke
Christopher Linker
Russel Long, M.S.
Erich Luschei, Ph.D.
Sharon Maclay, M.A.
Kathryn Maes, Ph.D.
Amy Mast, B.S.
Brian McCabe, M.D.
Arlene 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.
Dawn Salyards, B.S.
Robin Samlan, M.S.
Jackie Schalow, B.S.
Ronald Scherer, Ph.D.
Elaine Smith, Ph.D.
Marshall Smith, M.D.
Nancy Pearl Solomon, 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
Mitch Wolfe, 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.
Brad Story, B.S.
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

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

Part II. Tutorial reports

Forward

The Seventh Status and Progress Report features a larger number of clinical studies than the previous volumes. This is understandable because clinical studies take a longer time to bring to maturity. In this fifth year of NCVS investigations, it appears that a significant number of data sets on treatment of voice and speech disorders have entered a summary and conclusion stage.

The Dissemination and Continuing Education arms of the NCVS have now reached a level of stability and focus. A traveling exhibit of NCVS materials was taken to the ASHA Convention in New Orleans last November. The exhibit was designed and produced by Julie Ostrem, our C.E. coordinator. Julie was also at the Convention in person to distribute information and answer questions. One of the more popular items was the voice bookmark. Several conventioneers took large quantities and asked for more. Another popular item was our new training brochure, featuring warm (and at times "hot") looking photographs of our students and fellows in action.

We are anticipating success with a series of Public Service Announcements on the use and care of the human voice. These PSA's are being produced by Cynthia Kintigh, our Dissemination coordinator, in conjunction with the Denver Center Media. The brief announcements (15-20 seconds) will show common real-life settings in which good oral communication can make or break a human relationship. Be on the look-out.

Ingo R. Titze, Director
December, 1994

Part I

Research papers submitted for peer review in archival journals

NCVS Status and Progress Report - 7
December 1994, 1-17

Effect of Voice Disorders on Quality of Life

Elaine Smith, M.P.H., Ph.D.
Department of Preventive Medicine, The University of Iowa
Katherine Verdolini, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa
Steven Gray, M.D.
Division of Otolaryngology, The University of Utah
Sara Nichols, M.S.
Jon Lemke, Ph.D.
Department of Preventive Medicine, The University of Iowa
Julie Barkmeier, Ph.D.
Department of Otolaryngology, The University of Iowa
Heather Dove, M.S.
Division of Otolaryngology, The University of Utah
Henry Hoffman, M.D.
Department of Otolaryngology, The University of Iowa

Abstract

Between 1991-1993, adult-age patients (N=174) from two university hospital voice disorder clinics completed a questionnaire designed to elicit information about the effects of voice impairments on quality of life. Questions were asked regarding work, social, psychological, physical, and communication problems related to a voice disorder. The results indicated that voice disorders were perceived by the majority of patients as adversely affecting past (53%), current (47%), and future (76%) job functions. Almost 75% of respondents felt that social interactions were adversely affected by their voice problems with participation in social activities limited whenever possible. Psychological problems were reported by a large portion of patients as well, particularly depression (65%) and adverse professional self-esteem (61%). Physical discomfort as the result of a voice disorder was also noted by the majority of respondents, with phonatory effort being the most common complaint (67%). The most commonly reported communication problems involved conversations with background noise (65%), difficulty with phone conversations (58%), and the necessity to repeat statements because of being poorly understood (58%). The elderly were the most commonly affected by these quality of life impairments; however, age may have been confounded with diagnostic category in this study. The results suggest that quality of life may be adversely affected in a large proportion of persons with voice disorders and that more research is needed for improving quality of life for these persons.

Increasingly, quality of life (QOL) is considered a fundamental parameter for evaluating the severity of health-related conditions and treatment outcomes (see for example, Hollister & Weintraub, 1993). This issue has been the focus of numerous reports across a range of diverse medical conditions including cancer (Aaronson, 1988; Cella & Cherin, 1987; de Haes & van Knippenberg, 1985; Fayers & Jones, 1983; Flanagan, 1982; Frank-Stromborg, 1984; Haberman, Bush, Young, & Sullivan, 1993; Hollandsworth, 1988; Moinpour, Feigl, Metch, Hayden, Meyskens, & Crowley,

1989; Najman & Levine, 1981; Tope, Ahles, & Silberfarb, 1991), Parkinson's disease (Fitzsimmons & Bunting, 1993), oral and dental conditions (Hollister & Weintraub, 1993), HIV infections (example, Lang, 1993), hemodialysis (Ferrans & Powers, 1993), asthma (juniper, Guyatt, Willan, & Griffith, 1994), and rheumatoid arthritis (Bendtsen & Hörnquist, 1993). The emphasis on QOL is consistent with a view of health and disease as not only by physical but also mental and social states (World Health Organization, 1971).

Little is known about the impact of voice disorders on QOL, because systematic studies have not been conducted. Such disorders might be anticipated to produce negligible effects because they usually are not life-threatening and usually are associated with normal speech, language, cognitive, and ambulatory functioning. The purpose of the present study was to evaluate patient-perceived QOL ramifications of voice disorders and selected sociodemographic risk factors for QOL disturbances. This report describes the preliminary results of the findings.

QOL was operationalized in this study based on a consensual view of QOL as a multi-dimensional construct minimally involving four critical domains, summarized as: (a) activities or work functioning, (b) social functioning, (c) psychological functioning, and (d) physical symptomatology (see for example, Aaronson, 1988, 1990; Bendtsen & Hörnquist, 1993; Calman, 1987; Haberman et al, 1993; Hollister & Weintraub, 1993; Juniper et al, 1994; Nayfield, Gantz, Moinpour, Cella, & Hailey, 1992; Schipper, Clinch, McMurray, & Levitt, 1984; Schraub, Bransfield, Monpetit, & Fournier, 1987; Smart & Yates, 1987; Spitzer, Dobson, Hall, Chesterman, Levi, & Shepherd, 1981; Tope et al, 1993). In the present study, these aspects of QOL were evaluated from the patient's perspective, along with an additional domain, (e) communication, because of the particular relevance of communication functions for persons with voice disorders.

Methods

Participants

Between September 1991 and April 1993, 174 patients aged 18 or older were enrolled from the voice clinics in the Department of Otolaryngology, University of Iowa Hospitals and Clinics, and the Division of Otolaryngology - Head and Neck Surgery, University of Utah. All patients who sought evaluation or treatment for a voice disorder in the voice clinics during this time-period were approached for participation, with an invitation to complete a questionnaire regarding the effects of voice symptoms on their life. The speech pathologists (J.B., H.D.) reviewed the study with potential participants using an approved Human Subjects Consent form and information summary. Less than 5% of patients refused or were unable to complete the questionnaire due to physical or mental incapacitation.

The resulting patient sample included persons who lived in Iowa or Utah, or who were referred from other locations to these specialty voice clinics. A variety of conditions underlying voice disorders were diagnosed in the two clinic populations based on medical and speech pathology examinations. Diagnoses included nodules, polyps, laryngitis/edema, neuromotor conditions/vocal fold paralysis, spasmodic dysphonia, contact ulcers, bowing, musculoskeletal tension, laryngeal trauma, and a small number of other conditions.

Questionnaire

The questionnaire (Appendix A) elicited information about subjects' sociodemographic characteristics, voice symptoms, effects of a voice disorder on QOL domains, potential chemical/physical risk factor exposures, and familial history of voice disorders. In addition to these questions, which subjects answered, a medical history form was completed by an experimenter documenting the specific diagnosis and co-existing or prior medical conditions, medications, surgeries, or hearing loss. The separate subsections of the questionnaire are described next.

Sociodemographic and employment characteristics. Sociodemographic characteristics including age, gender, race, income, education, and occupation (Appendix A) were examined as risks of voice disorders. Occupational groups (including homemaker, retired, and student categories) were defined using Department of Transportation codes (year), representing each job reported with a three-digit code and subsequently dividing the responses into six occupational categories: professional/clerical, craftsperson/laborer, unemployed, student, retired, or homemaker. The age groups of respondents, shown in Table 1, were categorized on the basis of the present data into different phases in life related to schooling and work/career status. The youngest age (21) was still primarily in school and only a few had started working. Most of those working outside of the home were between the ages of 22-65. Those between ages 22-39 were considered to be in the early stages of their career, particularly because many had received advanced degrees. Individuals aged 40-65 were considered to be more established in their careers or job, whereas those >65 years included primarily retired individuals and homemakers. Note that because of the small sample size, no attempt was made in the present study to relate different occupations to the risk of specific diagnostic conditions.

Voice symptoms. Patients were asked to identify the presence of 10 specific voice symptoms listed in Table 2, and to describe any other symptoms existing at the time of diagnosis. Patients were also asked to describe the durational characteristics of symptoms over time (Appendix A).

Quality of life questions. Quality of life (QOL) questions comprised the bulk of the questionnaire (Appendix A). As already noted, the QOL questions were designed

Table 1.
Sociodemographic Characteristics

Characteristic	%	(N)
Gender		
Female	68.2	(118)
Male	31.6	(55)
Age[1]		
≤ 21	23.6	(41)
22-39	27.0	(47)
40-65	27.0	(47)
> 65	22.4	(39)
Race		
White	94.3	(164)
Hispanic	1.7	(3)
Black	1.7	(3)
American Indian	1.1	(2)
Other	1.1	(2)
Education[1]		
≤ 12	37.6	(65)
13-16	47.9	(83)
> 16	14.5	(25)
Income		
< $20,000	36.1	(61)
$20,000 < 40,000	23.1	(39)
≥ $40,000	28.4	(48)
Refused	3.6	(6)
Unknown	8.9	(15)
Occupation		
Prof/Clerical	27.5	(47)
Crafts/Laborer	11.7	(20)
Retired	19.3	(33)
Homemaker	12.3	(21)
Student	24.0	(41)
Unemployed	5.3	(9)

[1] In years

Table 2.
Frequency of Voice Symptoms

Voice Condition	%	(N)
Symptom		
Hoarseness	67.8	(118)
H-Notes Difficult	53.4	(93)
Effortful	52.3	(79)
Weak Voice	47.7	(83)
Tired Voice	44.3	(77)
L-Speaking Voice	40.8	(71)
Breathy	34.4	(60)
L-Notes Difficult	24.7	(43)
Voice Spasms	17.2	(30)
H-Speaking Voice	8.6	(15)

[1] % Based on "moderate-extreme" positive responses

to cover several aspects of daily functioning that are considered relevant for health-related QOL assessments in general, and that the authors had experienced as relevant in their clinical practice of persons with voice disorders. Questions included queries about work, social, psychological, physical, and communicative impairments related to a voice disorder. Each QOL question was posed using a Likert-type (ordinal) scale indicating a range of responses from "not at all [a problem]" to "[extreme problem]," with each question asked twice with slightly different wording and the numbering (1-5) of "no problem" to "extreme problem" reversed across repetitions in an initial version of the questionnaire. Analyses showed that many questions were redundant, and were deleted (See Statistical Analyses.) The final version is shown in Appendix A.

The purpose of the QOL questions was to determine whether and to what degree participants perceived QOL to be compromised as the result of a voice disorder, rather than to cover the range of issues that might be associated with QOL disturbances. As a result, the questions may not have addressed all major problems among this patient population; however, participants rarely responded to open-ended questions with QOL concerns other than those specifically listed. In the present study, only patients' perceptions about QOL effects were evaluated, without any attempt to validate responses with the observations of third parties.

The impact of a voice disorder on work functioning was evaluated with five different types of questions (Appendix A) regarding (1) past job/career decisions, 2) recent job/career decisions, 3) current job performance limitations, 4) job/career change needed, and 5) future limitations in job/career options. These questions were designed to elicit perceived effects on the patient's career or jobs overall rather than to elicit specific information in reference to a current job. About a year after enrollment began, an additional

question was added to the questionnaire inquiring about an actual or perceived need to change a specific current job due to a voice impairment. One hundred nine patients responded to this question.

The effects of a voice disorder on social interactions were evaluated by asking about (1) overall adverse interactions, (2) the felt need to interact differently, (3) the desire to limit participation in social events, (4) the degree to which social interactions were avoided, (5) the felt need to interact differently with family or friends, and (6) the frequency of annoyance expressed by family/friends, as the result of a voice problem (Appendix A).

Psychological consequences of a voice disorder were assessed by asking about (1) the overall effect of a voice disorder on self-esteem, (2) the effect on professional/work self-image, 3) depression, (4) emotional pain, and (5) embarrassment as the result of a voice disorder (Appendix A).

Physical discomfort was evaluated by asking about the presence and degree of (1) vocal effort, (2) vocal roughness, (3) uncomfortableness related to voicing, (4) voice fatigue, (5) voice scratchiness, and (6) voice aching (Appendix A).

The final set of QOL questions examined various perceived communication problems due to a voice disorder: (1) overall trouble being understood, (2) problems being understood on the phone, (3) the necessity of repeating statements that were not clearly understood, (4) inability to express oneself, and (5) inability to be understood in background noise (Appendix A).

Potential chemical/physical risk factor exposures. Physical, chemical, and traumatic risk factors were examined as possible etiologic agents in the development of specific conditions leading to a voice disorder, including prior surgeries, intubation, drug use, occupational exposures, smoking, and alcohol use (Appendix A). These factors will not be examined or further discussed in this study because of a limited number of subjects in each diagnostic category, and the differential relevance of the factors for specific diagnostic conditions.

Family history of a voice disorder. A series of questions about family history was also included a year subsequent to the initial enrollment to determine if patient symptoms/diagnoses were similar to those of family members, or if the presence of a voice problem in other family members was more frequent among those with a current voice disorder. The motivation for these questions was that some diagnostic conditions underlying voice disorders might have a familial basis. For these questions, patients were asked to specify the family member and his/her symptoms or diagnosis (Appendix A). Responses will not be examined or discussed further in this study because of the small number of subjects in each diagnostic category.

Medical questionnaire. The medical abstract, which was filled out by one of the experimenters based on the medical and speech/pathology evaluation and on the hospital chart, included questions about the voice-related diagnosis and other medical conditions that might affect the likelihood of developing a voice disorder, including diseases, injury, medications, surgeries, and hearing loss (Appendix A). These responses werenot evaluated in the present study because ofthe small number of patients in each diagnostic group, but will be described in a future report.

Statistical Analyses

Included in this analysis are the results for the initial 174 participants enrolled in the study. Questionnaire results regarding QOL were previously tested for reliability. That is, as noted previously, each question about QOL originally was asked twice using slightly different wording, with the valence of the responses reversed across repetitions. The similarity in the responses was evaluated using Kappa and McNemar's statistics (Nichols, unpublished data). Based on the statistical results, about half of the questions were deleted because they were found to be redundant.

Responses to each QOL question were evaluated for the subject group as a whole but not for specific diagnostic categories, because the assessments would not be statistically valid due to a small number of respondents in each category. Diagnostically-specific results will be described at a later date when additional patients have been enrolled.

To evaluate the frequency of an adverse effect of a voice disorder in general on QOL, ordinal responses to QOL questions were categorized so that responses of "not at all [a problem]/very little [problem]" was defined as "no," is not a problem, and responses of "moderately/very much/extremely [a problem]" were defined as "yes," is a problem.

The same response groupings were used in multivariate analyses, in which conditional logistic regression was used to examine the multivariate relationship between each QOL impairment and possible risk factors, including age, gender, education, and income (Voice symptoms, smoking, and alchol use were not included in the analyses because they were considered to be relevant for specific diagnostic categories, and the number of patients in each group was too small for valid statistical interpretation). Statistical interpretations were reported as odds ratios (OR) and 95% confidence intervals (CI) (Fleiss, 1981). Odds ratios designate the risk of disease or other outcome (for example "limitations in current job performance") in the "exposed" group (for example, "elderly") compared with the risk of disease or outcome ("limitations in current job performance") in the "unexposed" or comparison group (for example, "young"). TheOR associated with disease or outcome in the unexposed group is arbitrarily set at 1.0. An OR >1.0 refers to an increased risk of disease or other outcome in the exposed group whereas an OR <1.0 refers to a decreased risk of disease in the exposed (ie., a protective effect) compared to the unexposed group. Confidence intervals (CI) refer to the

probability of the OR falling between the range of OR values presented by the CI results. CI that do not include unity (1.0) are statistically significant, regardless of whether the range is greater or less than 1.0. CI also provide information about the power of a study to truly detect a difference based on the sample size in the study. That is, wider CI connote poorer power to detect a true effect.

Work-related QOL impairments were further evaluated in univariate analyses, as a function of employment category. The Mantel-Haenszel chi-square test was used to conduct these tests, because of the small number of subjects in each employment category.

Results

Before discussing the results of the association between voice disorders and QOL effects, background characteristics of the voice disorder clinic patients are described because these characteristics may influence the risk of perceived severity of QOL issues.

Sociodemographic and employment characteristics

Sociodemographic and employment characteristics of the subject sample are shown in Table 2. Two-thirds of these clinic patients were female with an even distribution of subjects across age groups. The retired group was almost all over age 65. Homemakers were equally divided between ages 22-65 and >65 whereas the student group was evenly divided in the younger two age groups. Most patients were Caucasian. A large proportion of the subject sample had at least a high school degree (84.4%), and 62.4% had additional college or graduate education with 68% ages 22-65, and 25% ages >65. There was a fairly even distribution of family incomes across the subjects with a portion of the sample, primarily those less than age 21, who did not know the family income.

Among the currently employed, the most common occupation represented in the voice clinics was teachers: 16.4% (10.4% primary/secondary, 6% university). No other specific occupational group was identified with a high frequency. Another 4.4% each were actors/entertainers (including singers), insurance/sales agents, office managers, secretaries, or waiters; 3% each were bill collectors, buyers, ministers, or craftspersons/machine operators. Most of the patients were employed in occupations that required using the voice on an ongoing basis in their daily work activities. A small percentage of patients was unemployed.

Voice symptoms

At the time they were seen in the voice clinics, 77% of the participants reported having a current, chronic voice problem (data not shown). Among these patients, for most, once their voice problem began it continued to persist (84.8%). About a quarter of the patients had had symptoms for nine months or less, with an average duration of two years. Disturbingly, another 25% had had problems for 12 years or more. Greater than 50% of the unemployed, retired, or homemaker groups had had symptoms for more than five years before seeking medical care at our voice clinic, as compared with 21% of the professional/clerical, 33% of the crafts/laborers, and 25% of students. These percentages included both those who sought treatment for the first time as well as those who previously had been seen by other health care practitioners but were either misdiagnosed or unsuccessfully treated.

The most commonly reported symptom at the time of diagnosis was hoarseness, followed at some distance by an effortful, weak, tired, or low speaking voice (Table 2). Another voice symptom noted was the increased inability to sing high, compared with low, notes. Among those >age 65, voice spasms and the inability to sing low notes was more frequent compared with the younger age groups ($p < .05$). A breathy voice was associated with younger age ($p < .05$), and a weak voice was more likely to be found in the working age groups (22-65, $p < .05$; data not shown).

Quality of life effects

Frequency of QOL impairments. A major focus of this study was to determine the frequency with which voice disorders adversely affected quality of life (QOL) in work, social, psychological, physical, and communicative domains, from the patient's perspective. The results for each domain are shown in Table 3.

Work impairments were noted by a large number of patients (Table 3). The results regarding past and potential future employment were based on individuals who had ever been employed, including those not currently working outside the home but who may have worked previously (i.e., unemployed, retired, housewife, see Table 1). Results regarding current work-related issues were evaluated based on individuals currently working outside the home (professional/clerical and craft/laborer groups). A large proportion of patients felt that either past or future job opportunities were or would be adversely affected by a voice disorder. Current job performance, decisions about changing a job, and job options were viewed as relatively less compromised by the disorder, with voice problems perceived as more likely to limit current job performance than to limit either current job decisions or the ability to change jobs.

Regarding social effects of a voice problem (Table 3), almost three-fourths of the patients found social interactions to be problematic because of the disorder, with over half preferring to avoid or limit them as much as possible. The majority of respondents felt a need to interact differently than they would have liked in general, and even with family/friends, because of their voice problem. About a third of the respondents felt that interactions were annoying to family members or friends. However, interactions with family/friends generally received fewer negative responses than with other persons or in other settings.

Table 3.
Frequency of Adverse Effects of Voice Disorders on Lifestyle and Communications

Effect[1]	%	(N)
Work		
Past job/career	53.1	(78)
Current: limit job decisions	41.1	(51)
Current: limit job performance	49.2	(63)
Current: change job/career	30.6	(38)
Future job/career options	75.8	(108)
Social		
Interactions: negative	73.5	(125)
Avoid social situations	57.9	(99)
Limit participation	65.3	(111)
Interact differently: general	57.3	(98)
Interact differently: family/friends	52.0	(89)
Family/Friends annoyed	32.7	(56)
Psychological		
Overall	51.7	(89)
Professional/job	61.2	(89)
Depressed	64.5	(111)
Emotional	52.0	(89)
Embarrassed	52.9	(91)
Physical		
Effortful	66.9	(75)
Rough	62.1	(69)
Uncomfortable	57.6	(64)
Tiring	56.3	(63)
Scratchy	52.8	(57)
Ache	48.2	(54)
Communication		
Trouble being understood	55.2	(95)
Phone conversation problems	57.9	(99)
Repeat statements	57.6	(99)
Can't express self	53.5	(91)
Conversations: background noise	65.3	(111)

[1] % based on "moderate-extreme" positive responses

The results from questions about psychological functioning related to a voice disorder (Table 3) indicated that over half of the patients felt generally psychologically devastated by their voice limitations. An adverse psychological impact related to career or job functioning was reported by over 60% of currently or formerly working patients. An even higher proportion reported depression, and over half described emotional pain and embarrassment as the result of a voice problem.

Physical discomfort from a voice disorder was described by the majority of patients. Effortfulness in voicing was the most common complaint, followed by complaints of a rough, uncomfortable, tiring, scratchy, or achy voice (Table 3).

Communication problems were also indicated by most patients (Table 3), including trouble being understood in general, problems with intelligibility in phone conversations, the need to repeat statements because of being poorly understood, difficulty expressing oneself as one would like, and difficulty being understood in background noise. Problems in loud environments (with background noise) were the most commonly reported communication impairment.

Risk factors for QOL impairments. Using multivariate analyses, QOL impairments were evaluated as a function of age, gender, education, and income, controlling for other independent effects, confounders, and interactions. Table 4 shows the results of conditional logistic regression based on the final model for QOL impairments, with only age and gender as variables that remained in the model. Table 5 shows the univariate relations between work-related QOL impairments and employment category, based on Mantel-Haenszel chi-square tests because of the small number of subjects in each employment category.

As shown in Table 4, age was significantly associated with QOL problems in this sample group. The elderly expressed the greatest concern about work-related effects on past work performance whereas the youngest expressed the greatest concern about future career effects. These results are not unexpected since the two age groups also were more likely either to have been employed in the past (elderly) or more likely to expect to be employed in the future (young). There were no differences by age in the effect that a voice disorder had on current job decisions, with over half of the patients in each age group reporting adverse effects (data not shown). Nor were there age differences among the currently employed associated with limiting current work performance or current ability to change profession/job although those aged 40-65 were less likely than those aged 22-39 to respond negatively about current ability to change jobs/career because of a voice problem.

With the exception of work-related issues, for which age effects were found in both directions, adverse effects of voice disorders were otherwise most prominent in the elderly. Examination of social effects showed that the elderly were more likely than other age groups to report that voice problems led to negative interactions, causing respondents to act differently than they would like as well as leading them to avoid social situations as much as possible. The elderly were significantly more likely to feel a need to interact differently with family and friends when communicating with them, although age was not a factor in the frequency of

Table 4.
Demographic Risks[1] of Voice Disorders on Lifestyle and Communication

Effect[2]	Age		Gender	
	OR	(95% CI)[3]	OR[4]	(95% CI)
Work				
Past job/career	1.02	(1.00-1.04)	–	–
Current: limit job decisions	–	–	–	–
Current: limit job performance	–	–	–	–
Current: change job career	–	–	–	–
Future job/career options	0.94	(0.92-0.97)	–	–
Social				
Interactions: negative	1.03	(1.01-1.05)	–	–
Interact differently: general	1.03	(1.01-1.05)	–	–
Limit participation	–	–	–	–
Avoid social situations	1.03	(1.01-1.05)	–	–
Interact differently: family/friends	1.02	(1.01-1.04)	2.42	(0.99-5.95)
Family/friends annoyed	–	–	–	–
Psychological				
Overall	1.02	(1.00-1.03)	–	–
Professional/job	–	–	–	–
Depressed	–	–	–	–
Emotional	–	–	–	–
Embarrassed	1.02	(1.01-1.04)	0.48	(0.23-0.98)
Communication				
Trouble being understood	–	–	–	–
Phone conversation problems	–	–	–	–
Repeat statements	1.02	(1.00-1.04)	–	–
Can't express self	1.03	(1.01-1.05)	–	–
Background noise	1.02	(1.00-1.04)	–	–

[1] controlling for age, gender, education, and income at diagnosis

[2] % based on "moderate-extreme" positive responses

[3] OR = odds ratio, CI = confidence intervals

[4] Female v male

[5] OR > 1.00 based on 3 decimals

responses regarding perceived annoyance of family/friends because of a voice problem. The elderly reported the heaviest psychological toll overall from having a voice problem and they were more likely to feel embarrassed by their condition than other age groups. Analyses assessing the relation between physical discomfort and age and gender were not conducted, because the discomfort variables were predicted to strongly vary with diagnostic condition and sample sizes for the different conditions were still too small for valid statistical evaluation. Finally, those aged >65 were

Table 5.
Frequency of Adverse Work Effects by Employment Status

Employment Status	Current[1]		Past		Future	
	%	(N)	%	(N)	%	(N)
Prof/Clerical	64	(30)	68	(30)	89	(41)
Unemployed	–	–	63	(5)	88	(7)
Crafts/Laborer	63	(12)	50	(8)	74	(14)
Student	31	(11)	45	(18)	77	(30)
Homemaker	30	(3)	38	(6)	56	(5)
Retired	22	(2)	41	(9)	36	(5)

[1] Limit job performance

more likely than other age groups to report the following communication problems: (1) the need to repeat statements because they were poorly understood, (2) an inability to express oneself as one would like because of a voice condition, and (3) difficulty being understood in a noisy environment (70% of the elderly reported difficulty being understood in a noisy environment, as compared with 23% of the youngest age group who gave that response).

The only other sociodemographic factor that influenced QOL because of a voice disorder was gender. Men were more likely than women to indicate a need to interact differently with family/friends, but less likely to state that they were embarrassed by their voice disorder.

Table 5 shows the association between specific employment category and past, current, and future adverse job-related effects due to a voice problem. Past and future adverse work responses are based on responses to the question regarding the effect that the voice problem had or would have on job/career limitations. Current adverse work effects were based on the question that received the highest frequency for adverse effects shown in Table 3: limitations on current job performance. Responses included those from persons who worked part as well as full-time. Those most likely to report current limitations on their work were non-students and persons working outside of the home, i.e. in the professional/clerical and craft/laborer groups. Although working individuals might be expected to indicate that their voice had in the past and would have in the future significant adverse effects on their career, it was unclear how those who were not currently working would respond. The professional/clerical, and unemployed groups, followed by the craft/laborer, retired, homemaker, and student groups in frequency, felt adverse effects on both past and future job opportunities. Professional/clerical and unemployed groups responded with similar frequency to the question regarding

adverse effects on future career options with a slightly lower frequency noted by students, craft/laborer groups, and homemakers (who might consider reentering the work force). These groups responded in a similar order to the question regarding past career options having been affected by a voice problem. Thus, the unemployed were highly likely to feel adverse work-related effects associated with their voice disorder. The craft, laborer, and unemployed groups also were more likely than other employment groups to state that their work-related self-esteem was impaired by the voice disorder (data not shown).

In the second year of the study we included a question regarding whether voice problems significantly affected the ability to work currently, asking about the need to change jobs, retire, or go on disability due to a voice problem. Among those who answered this question (n=109), 16.6% either had changed jobs entirely or were considering a job change because of a voice disorder. There was no apparent association between those who responded in the affirmative by job categories shown in Table 1. Thus, occupational groups were about equally likely to feel the need to change jobs. However, because of the small sample, this null result is extremely tentative, as seen by wide confidence intervals. The number of years of employment prior to job change due to a voice disorder ranged from less than one to 23 years.

Discussion

This is the first report to assess the effects of voice problems on quality of life in a large group of patients. Many health care providers may not consider voice disorders serious because they generally do not involve life threatening conditions, and as for many other communication disorders are an invisible handicap apparent only when the individual attempts to communicate. Despite these considerations, this investigation indicates that persons with voice disorders commonly perceive markedly adverse effects on quality of life (QOL). A moderate to severe negative impact is perceived by persons with a voice disorder in several critical functional domains, including work, social, psychological, physical, and communicative domains.

It is interesting to compare the QOL responses in this study with those from investigations of other illnesses. Somewhat surprisingly, patient reports of QOL impairments because of a voice disorder were in a similar range of severity - or worse - than impairments reported by patients with medical conditions generally considered more serious, including rheumatoid arthritis (Bendtsen & Hörnquist, 1993), hemodialysis treatments (Ferrans & Powers, 1993), previous bone marrow transplantation (Haberman et al, 1993), asthma (Juniper et al, 1994), repaired cleft lip or palate (Hollister & Weintraub, 1993), and oral conditions including dentures and temporomandibular joint syndrome (Hollister & Weintraub, 1993). And although there are important differences between discomfort and pain, it is interesting that the 70% or so of patients with a voice disorder who reported moderate to severe discomfort in phonation due to effortfulness is in the range of the 20-75% of patients anticipated to experience moderate to severe pain with cancer (Tope et al, 1993). Cross-study comparisons of this type are not straightforward due to differences in study methodologies and most likely, different internal scaling criteria by patients with different health-related conditions. Patients with voice disorders might retrospectively evaluate their current QOL impairments differently if they developed a life-threatening health condition. Further, QOL severity ratings might have been somewhat inflated in the present study relative to ratings for the population of persons with voice disorders in general, because the present data were based on patients evaluated in a tertiary care center with the possibility of a high rate of difficult and recalcitrant cases. The point is that, as they experience them, many persons with voice disorders perceive disorder-related functional impairments that are markedly distressing and functionally limiting.

From the employment and therefore societal perspective, a striking finding was the high rate of reporting a perceived limitation in current job performance because of a voice disorder (49%) and a perceived limitation in future job options because of the disorder (75%). The retreat from social situations and communication difficulties noted with a voice disorder would likely contribute to work limitations, leading to an actual job loss or a restriction of job opportunities in some cases. Seventeen percent of the later respondents to whom the question was posed (N = 109) indicated that they had in fact changed jobs or had thought about doing so because of a voice disorder. The high frequency with which unemployed groups responded adversely regarding work-related effects because of a voice problem was disturbing and leads to the speculation that unemployment may have been partly determined by voice functioning. Verifications are, of course, required for such speculations. However, if confirmed, the findings regarding employment functioning have important implications in considering the economic impact of voice disorders on the job force at large, particularly in light of the high rate of occurrence of voice disorders in the general population (about 3-7% of the adult population at any given moment in time; Brindle & Morris, 1979; Laguaite, 1972).

A case in point was the high rate of occurrence of voice problems in teachers. Among the currently employed persons with voice disorders, a high frequency of teachers was identified. About 10% of the patient group listed their occupation as teachers in grades K-12, with another 6% teaching at the university level. Teachers were the most common professional category to seek medical care for a voice disorder. In Iowa and Utah, teachers comprise only about 1% of the working population. The findings thus

suggest that teachers' work activities place them at particularly high risk for developing voice disorders especially because, in our experience, teachers are reticent to and sometimes admonished for taking time off work. Persons in other voice-intensive occupations, such as singers, actors, or salespersons were not nearly so disproportionately represented as teachers in the present treatment seeking sample (however, see Miller, 1994, for a high rate of cumulative voice dysfunctions reported by teachers of singing as compared with control subjects).

Work-related functioning would be expected to affect not only employment and economic status, but also psychological status. Work provides a positive self-image, with coworkers serving not only as a reference group but also as a group that provides friendship and help with daily coping of stressful life events (Cohen & Wills, 1985; Kaplan & Criqui, 1983). Thus, some of the psychological distress that patients noted because of a voice disorder could have been due to work-related limitations.

Various aspects of voice effects on quality of life appear age-dependent, as evidenced by more extreme perceived impairments in social, psychological, and communicative domains in elderly as compared with younger subjects. Problems associated with social isolation have been identified as a major issue in studies of the elderly (Holahan, Holahan, & Belk, 1984; Krause, 1986; Oxman & Berkman, 1990); based on the present study, voice problems could contribute to that isolation particularly through inadequate voicing to be understood (in addition to the contribution of other medical conditions leading to restricted mobility). However, it should be noted that age and diagnostic category may have been confounded as variables in this study: elderly persons may have had more debilitating diagnostic conditions than younger subjects, such as vocal fold paralysis, Parkinson's disease, or other degenerative neurological conditions. Formal analyses of diagnostic categories by age will be reported as the subject numbers increase, to assess this possible confound.

Another finding of interest was the gender skewing of the data: the ratio of women to men seeking treatment for a voice disorder was over 2:1. A follow-up evaluation of this finding revealed that the higher percentage of female patients was unrelated to current work status, and thus greater time available to seek medical care, because employment rates by gender were similar. We cannot rule out that the reason was due to the higher frequency of health-seeking behavior associated with women. However, also this explanation seems unlikely based on recent findings of a similar gender ratio (2:1) for self-perceived voice problems in a non-treatment-seeking sample of subjects, teachers of singing and controls (Miller, 1994). A more likely explanation is related to conclusions from previous studies that women are more at risk than men particularly for common diagnostic conditions affecting voice (e.g., nodules, polyps, Reinke's edema; Colton & Casper, 1990). In some cases this finding has been attributed to biophysiologic differences associated with gender, and specifically, that females are more likely to phonate with a persistent posterior glottal gap (Sodersten & Lindestad, 1987), which may increase the risk of nodules (Morrison, Rammage, Belisle, Pullan, & Nichol, 1983). It is not clear why women were more likely to report feeling embarrassed about a voice disorder than men, but this finding may be due to a greater willingness to report adverse emotional reactions (Cohen & Wills, 1985). This study does suggest that although certain QOL functions associated with age, gender, and employment are affected by voice disorders, neither income nor education had a significant effect in ameliorating their adverse effects on quality of life.

One unsettling finding was that most patients had their voice disorder for a minimum of two years, with 25% having their problem for 12 years or more. Voice disorders are often thought of as being temporary and treatable. This assumption is called into question by the present study, which suggests that many cases of voice disorders seen clinically may represent long-term, chronic conditions.

Further study is needed to verify our findings and to examine mechanisms for disease prevention or earlier diagnosis, with emphasis on potential high-risk groups. In particular, additional studies are needed to evaluate QOL issues in a larger sample size of affected patient groups to allow for the assessment of QOL issues in relation to specific diagnostic categories. Normative data from a control group should also be established (such data collection is underway), and an attempt should be made to verify patients' responses of QOL effects by third parties (data collection also underway). Further comparisons of QOL effects should be made to other illness conditions. Additional questions about QOL might be added to the existing questionnaire by individual clinicians, based on their perceptions of relevant issues. As it is, the questionnaire used in this study is currently being employed to evaluated pre- and post-treatment outcome from the patient's perspective. The information gained from these responses and future studies will be useful in developing public education programs, in clarifying misinformation that people have regarding the seriousness of voice disorders in their effects on quality of life, in identifying specific risk factors for these difficulties, and in developing prevention and rehabilitation programs.

Acknowledgements

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

References

Aaronson, N.K. (1988). Quality of life: What is it? How should it be measured? Oncology, 2, 69-74.

Bendtsen, P., & Hörnquist, J.O. (1993). Severity of rheumatoid arthritis, function, and quality of life: sub-group comparisons. Clinical and Experimental Rheumatology, 11, 495-502.

Brindle, B.R., & Morris, H.L. (1979). Prevalence of voice quality deviations in the normal adult population. Journal of Communication Disorders, 12, 439-445.

Calman, K.C. (1987). Definitions and dimensions of quality of life. In N.K. Aaronson & J. Beckman (Eds.), The quality of life of cancer patients (pp. 1-9). New York: Raven Press.

Cella, D.F., & Cherin, E.A. (1987). Measuring quality of life in patients with cancer. In Proceedings of the Fifth National Conference on Human Values and Cancer (pp. 23-31). New York: American Cancer Society.

Cohen, S., & Wills, T.A. (1985). Stress, social support, and the buffering hypothesis. Psychological Bulletin, 98, 310-357.

Colton, R.H., & Casper, J.K.(1990). Understanding voice problems. A physiological perspective for diagnosis and treatment. San Francisco: Williams & Wilkins.

de Haes, J.C.J.M., & van Knippenberg, F.C. (1987). Quality of life of cancer patients: Review of the literature. In N.K. Aaronson & J. Beckman (Eds.), The quality of life in cancer patients (pp. 167-182). New York: Raven Press.

Department of Transportation

Fayers, P.M., & Jones, D.R. (1983). Measuring and analyzing quality of life in cancer clinical trials: A review. Statistics in medicine, 2, 429-446.

Ferrans, C.A., & Powers, M.J. (1993). Quality of life of hemodialysis patients. Anna Journal, 20, 575-582.

Fitzsimmons, B., & Bunting, L.K. (1993). Parkinson's disease: Quality of life issues. Neuroscience Nursing, 28, 807-819.

Flanagan, J.C. (1982). Measurement of quality of life: Current state of the art. Archives of physical medicine and rehabilitation, 63, 56-59.

Fleiss, J.L. (1981). Statistical methods for rates and proportions. 2nd ed. New York: Wiley.

Frank-Stromborg, M. (1984). Selecting an instrument to measure quality of life. Oncology Nursing Forum, 11, 88-91.

Haberman, M., Bush, N., Young, K., & Sullivan, K.M. (1993). Quality of life of adult long-term survivors of bone marrow transplantation: A qualitative analysis of narrative data. Oncology Nursing Forum, 20, 1545-1553.

Holahan, C.K., Holahan, C.J., & Belk, S.S. (1984). Adjustment in aging: The roles of life stress, hassles, and self-efficacy. Health Psychology, 3, 81-89.

Hollandsworth, J.C. (1988). Evaluating the impact of medical care and technologies on the quality of life: A 5-year update. Social Science in Medicine, 26, 425-434.

Hollister, M.G., & Weintraub, J.A. (1993). The association of oral status with systemic health, quality of life, and economic productivity. Journal of Dental Education, 57, 901-912.

Juniper, E.F., Guyatt, G.H., Willan, A., & Griffith, L.E. (1994). Determining a minimal important change in a disease-specific quality of life questionnaire. Journal of Clinical Epidemiology, 47, 81-87.

Kaplan, R.M., & Criqui, M.H. (1983). Behavioral epidemiology and disease prevention. New York: Plenum Press.

Krause, N. (1986). Social support, stress, and well-being among older adults. Journal of Gerontology, 41, 512-519.

Laguaite, J.K. (1972). Adult voice screening. Journal of Speech and Hearing Disorders, 37, 147-151.

Lang, C. (1993). HIV and quality of life: A rehabilitation perspective. Health Visitor, 66, 445-446.

Miller, M.K. (1994). The frequency of voice disorders in teachers of singing. Unpublished master's thesis, The University of Iowa, Iowa City, Iowa.

Moinpour, C.M., Feigl, P., Metch, B., Hayden, K.A., Meyskens, F.L., & Crowley, J. (1989). Quality of life endpoints in cancer clinical trials: Review and recommendations. Journal of the National Cancer Institute, 81, 485-495.

Morrison, M.D., Rammage, L.A., Belisle, G., Pullan, B., & Nichol, H. (1983). Muscular tension dysphonia. Journal of Otolaryngology, 12, 302-307.

Najman, J.M., & Levine, S. (1981). Evaluating the impact of medical care and technologies on the quality of life: A review and critique. Social Science in Medicine, 15, 107-115.

Nayfield, S.G., Gantz, P.A., Moinpour, C.M., Cella, D.F., & Hailey, B.J. (1992). Report from a National Center Institute (USA) workshop on quality of life assessment in cancer clinical trials. Quality of Life Research, 1, 203-210.

Oxman, T.E., & Berkman, L.F. (1990). Assessment of social relationships in elderly patients. International Journal of Psychiatry in Medicine, 20, 65-84.

Schippper, H., Clinch, J., McMurray, A., & Levitt, M. (1984). Measuring the quality of life of cancer patients: The functional living index-cancer: Development and validation. Journal of Clinical Oncology, 2, 472-483.

Schraub, S., Bransfield, D., Monpetit, E., & Fournier, J. (1987). Comparison of three methods to measure quality of life. In N.K. Aaronson & J. Beckman (Eds.), The quality of life of cancer patients (pp. 83-92). New York: Raven Press.

Smart, C.R., & Yates, J.W. (1987). Quality of life. Cancer, 60, 620-622.

Sodersten, M., & Lindestad, P.-A. (1987). Vocal fold closure in young adult normal-speaking females. Phoniatric and Logopedic Progress Report, vol. 5 (pp. 12-19). Sweden: Karolinska Institute.

Spitzer, W.O., Dobson, A.J., Hall, J., Chesterman, E., Levi, J., & Shepherd, R. (1981). Measuring the quality of life of cancer patients: A concise QP-index for use by physicians. Journal of Chronic Diseases, 34, 585-597.

Tope, D.M., Ahles, T.A., & Silberfarb, P.M. (1993). Psycho-oncology: Psychological well-being as one component of quality of life. Psychotherapy and Psychosomatics, 60, 129-147.

World Health Organization (1971). The economics of health and disease. WHO Chronicle, 25, 20-24.

Appendix A

Name:_____
Study #:_____
Form: 0101
Date:__-__-__

Voice Care Patient Questionnaire

First we would like to ask you a few questions that will be used to describe the patients in our study.

Please CIRCLE the NUMBER next to your answer and fill in any numbers that apply.

1. How old are you?_____ AGE

2. What is your date of birth? __ __ - __ __ - __ __

3. What is your sex?_____ MALE _____ FEMALE

4. Including grade school, high school, college, business, vocational, professional, and postgraduate schooling, how many total years of schooling have you completed?

_____ YEARS

5. Which of these best describes your race or ethnic background (CIRCLE ONE):

1 White
2 White, Hispanic Background
3 Black
4 American Indian
5 Chinese
6 Other (SPECIFY)

6. Including income provided by you, (your spouse), and any other person in your household, which range of figures comes closest to your total household income before taxes for the past year?

1 less than $10,000
2 $10,000-$19,999
3 $20,000-$29,999
4 $30,000-$39,999
5 $40,000-$49,999
6 $50,000 or more
7 Refused
8 Unknown

7. Did you change your occupation/job because of your voice problem?

__ NO (Go to question 10)
__ NO, but may need to (Go to question 10)
__ YES

8. If YES, what occupation/job did you have when your voice problem developed?

_____

9. How many years did you work in that occupation/job before you had to leave it because of your voice problem?

____ # YEARS

10. What is your current occupation/job?

_____

11. How many years have you been working in this occupation/job?

____ # YEARS

12. What was your primary reason for seeking treatment?

_____
_____

13. Do you think that you have a chronic voice problem?

___ NO (Go to question 15)
___ YES

14. When did you first notice that you had a voice problem?

__ __ - __ __
MO. YR.

15. Have you always had a voice problem since then?

___ NO
___ YES (Go to question 16)

16. Have you ever had a chronic voice problem in the past?

___ NO
___ YES (Go to question 17)

17. What voice symptoms do/did you have (CIRCLE ALL THAT APPLY):

1 Hoarse voice
2 Breathy voice
3 Lower speaking voice than normal
4 Higher speaking voice than normal
5 Could not sing high notes
6 Could not sing low notes
7 Weak voice
8 Effortful voice
9 Tired voice
10 Voice spasms
11 Other (SPECIFY)_____

18. Do you have a family history of anyone who has a voice problem?

___ NO (Go to question 19)
___ YES

If yes, specify family relationship (mother, brother, cousin)._____

What kind of problem did s/he have with her/his voice?

_____

19. How much have voice problems affected your past job or career option?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely
6 Do not know

Now we would like to ask you a few questions about your smoking history.

20. Have you ever smoked or chewed any of the following for a year or longer?

Tobacco Product	Ever Used		Date First Used	Still Smoke:		Date Last Used	Average Number or Chews/day
Cigarettes	Yes	No	__-__	Yes	No	__-__	_____
Cigars	Yes	No	__-__	Yes	No	__-__	_____
Pipe	Yes	No	__-__	Yes	No	__-__	_____
Chewing Tobacco/Snuff	Yes	No	__-__	Yes	No	__-__	_____
Other (Specify) _____			__-__	Yes	No	__-__	_____

21. Over your lifetime, did you drink any of the following alcoholic beverages?

Beverage	Ever Used		Date First Used	Still Use:		Number Days/Wk Drank Beverage	Number Bottles, Glasses, Drinks/ Wk
Beer	Yes	No	__-__	Yes	No	__-__	_____
Wine	Yes	No	__-__	Yes	No	__-__	_____
Liquor	Yes	No	__-__	Yes	No	__-__	_____
Other (Specify) _____			__-__	Yes	No	__-__	_____

22. If your voice does not become healthier, how much would this limit your future career or job options?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely
6 Does not apply; do not work

23. How much do voice problems negatively affect the way you interact in various social situations?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

24. How scratchy does your voice feel?

1 Not at all
2 A little
3 Moderately
4 Very much
5 Extremely

25. How much do voice problems limit your participation in social events?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

26. To what extent do you interact with family/friends differently than you would like to, because of voice problems?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

27. How effortful is it for you to use your voice?

1 Not at all
2 A little
3 Moderately
4 Very much
5 Extremely

28. To what extent do people not understand you clearly on the telephone, because you have voice problems?

1 Not at all
2 Very little
3 Moderately
4 Very little
5 Extremely

29. To what extent do voice problems negatively affect your self-image, professionally or at work?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely
6 Does not apply; do not work

30. How rough does your voice feel?

1 Not at all
2 A little
3 Moderately
4 Very much
5 Extremely rough

31. How much does your voice ache when you use it?

1 Not at all
2 A little
3 Moderately
4 Very much
5 Extremely

32. How much do voice problems negatively affect your professional or work image of yourself?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely
6 Does not apply; do not work

33. How much do people have trouble understanding you on the telephone, because of your voice problems?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

34. How uncomfortable is your throat when you talk?

1 Not at all
2 A little
3 Moderately
4 Very much
5 Extremely

35. To what extent have you thought of changing jobs or career plans because of a voice disorder?

1 Not at all
2 Very little
3 Moderately
4 Very Much
5. Extremely
6 Does not apply; do not work

36. To what extent do family/friends seem to get annoyed because of your voice problems?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

37. To what extent do voice problems make you act differently than you would like to act in social situations?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

38. How much do you interact differently with people, because of your voice problems?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

39. To what extent do you avoid various social situations because you have voice problems?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

40. To what extent do you feel you cannot get your point across, because of voice problems?

1 Not at all
2 A little
3 Moderately
4 Very much
5 Extremely

41. To what extent do voice problems restrict your job performance?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely
6 Does not apply; do not work

42. To what extent do friends/family seem "put off" by your voice problems?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

43. How much have you thought of changing what you do professionally or at work, because of poor vocal health?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely
6 Does not apply; do not work

44. How much do voice problems limit the way you do your job?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely
6 Does not apply; do not work

45. How much do people ask you to repeat yourself, because your voice was not clear enough?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

46. How much do voice problems make you feel bad or depressed?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

47. To what extent do you feel embarrassed about a substandard voice?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

48. How much does it hurt to use your voice?

1 Not at all
2 A little
3 Moderately
4 Very much
5 Extremely

49. To what extent has poor vocal health limited your recent job/career decisions?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely
6 Does not apply; do not work

50. How much does your voice tire you out?

1 Not at all
2 A little
3 Moderately
4 Very much
5 Extremely

51. To what extent do voice problems make you think poorly of yourself?

1. Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

52. To what extent do you feel you cannot express your ideas the way you want to because of your voice problems?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

53. How much trouble do people have understanding you in a background noise, because of your voice problems?

1 None
2 A little
3 Moderately
4 Very much
5 Extremely

54. How much effort is it for you to use your voice?

1 None
2 A little
3 Moderately
4 Very much
5 Extremely

55. If your voice does not become healthier, to what extent would this restrict what you do professionally or at work in the future?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely
6 Does not apply; do not work

56. How much do others have trouble understanding you, because you have voice problems?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

57. To what extent do you restrict the way you participate in social activities, because of voice problems?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

58. How much emotional pain do voice problems cause you?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

59. How much does your voice fatigue you?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

60. To what extent are voice problems the source of emotional pain?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

61. How uncomfortable is it for you to use your voice?

1 Not at all
2 A little
3 Moderately
4 Very much
5 Extremely

62. How much does a poorly functioning voice undermine your self-esteem?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

63. To what extent do you find yourself repeating what you said, because your voice was not clear enough?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

64. How much does your voice embarrass you, because you have voice problems?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

65. How much have your recent job/career decisions been negatively affected by poor vocal health?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely
6 Does not apply; do not work

66. To what extent are you sad or depressed, because you have voice problems?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

67. To what extent do people have difficulty understanding you because you have voice problems?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

68. To what extent does your self-esteem suffer, because your voice is not working well?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

69. To what extent do people not seem to understand you in noisy environments, because you have voice problems?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

70. How much do voice problems undermine your self-esteem?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely

71. To what extent have voice problems affected your past job and/or career options?

1 Not at all
2 Very little
3 Moderately
4 Very much
5 Extremely
6 Do not know

Other

IS THERE ANYTHING ELSE THAT YOU WOULD LIKE TO TELL US ABOUT YOUR VOICE PROBLEMS?

NCVS Status and Progress Report - 7
December 1994, 19-22

Benign Pathological Responses of the Larynx

Steven D. Gray, M.D.
Division of Otolaryngology, The University of Utah
Elizabeth Hammond, M.D.
Department of Pathology, LDS Hospital
Darrin Hanson, B.S.
Division of Otolaryngology, The University of Utah

Abstract

Benign laryngeal lesions were examined for patterns of injury indicated by deposition of fibronectin and collagen type 4. An immunoperoxidase technique was used to compare thirty-three fresh and paraffin - embedded tissues with regard to their staining of monoclonal antibodies directed against fibronectin and collagen type 4. Two types of patterns were recognized: One pattern showed basement membrane zone injury, indicated by thick collagen type 4 bands; this finding was often coupled with intense fibronectin deposition of the superficial layer of the lamina propria. The other pattern often showed no basement membrane zone injury and very little fibronectin deposition. The first pattern correlated more with nodules, the second pattern more with Rheinke's edema and some polyps. A fuller understanding of the effects of excessive deposition of structural glycoproteins such as fibronectin and of abnormal proteoglycan deposition may lead to a better characterization of vocal fold pathology and, ultimately, etiology and treatment.

Introduction

Pathological responses to benign laryngeal injury have been traditionally described using hematoxylin and eosin (H&E) staining. Electron and light microscopic identification techniques have revealed that the laryngeal responses to injury are varied and cannot be adequately described using only H&E staining.

The vocal fold lamina propria consists of two important tissue components: the cellular and extracellular. Biomechanically, only the extracellular tissue is likely to have a determining role since the cellular tissue in the lamina propria is sparse. (Only the epithelium is predominantly cellular, and it probably plays a significant role in providing a resilient cover.)[1]

The extracellular tissue, called a matrix, embodies two families of macromolecules: those molecules that provide a fibrous scaffolding of the lamina propria, which are the collagens and elastins, and those molecules that provide the "filling" between the fibrous scaffolding.[2] This latter family contains two macromolecular groups, the proteoglycans and the structural glycoproteins.

Traditionally, attention has been given to those molecules of the fibrous network as being a key influencing factor in determining tissue vibration. Hirano has observed the elastin and collagen fibers histologically. Elastin and sparse collagen fibers are present in the superficial aspect of the lamina propria. The elastin fibers are prominent in the intermediate layer and then decrease in the deep layer. Collagen fiber are more numerous in the intermediate and deep layer.[3]

The anatomic distribution of the fibrous proteins has been correlated with biomechanical observations of the vocal folds and led to the cover-body theory of vibration.[4] As a result of the elegant histological description of the fibrous network, these extracellular matrix fibrous components of vocal fold architecture became the focus of attention in studying normal and impaired vibration. Vocal fold pathology has been correlated with disturbed fibrous network.[2,5] These exquisite studies and the tremendous advances they provided in vocal fold physiology have resulted in the "filler" proteins of the extracellular matrix being largely forgotten.

This paper describes one of the filler proteins of the extracellular matrix, fibronectin, a critical structural glycoprotein found in the lamina propria and its relationship to

basement membrane zone injury. We have chosen fibronectin because of its deposition as a result of tissue injury.[6] We have selected collagen type 4 due to its location only in the basement membrane zone. Collagen type 4 is a good indicator of basement membrane zone thickness and location.[7]

Methods

Polyps, nodules, and Rheinke's edema specimens were obtained from the University of Iowa from 1984 to 1989 and from the LDS Hospital in Salt Lake City from 1980 to 1993. Specimens were analyzed only if the epithelium, basement membrane zone, and recognizable components of the superficial layer of the lamina propria were present. Because of the limited surgical excision involved with this type of surgery, specimens were often excluded due to insufficient tissue. Thirty-three specimens were acceptable.

Initially our goal was to correlate the tissue diagnosis and clinical diagnosis with our findings, but we realized that often the H & E stains of nodules and polyps look alike. Consequently, pathologists called the specimen a nodule or a polyp if the surgeon clinically indicated it as one or the other. Furthermore, occasionally the lesions were not labeled consistently by the surgeon, leaving little clinical record to use in interpretation. Therefore, although the pathological processes are described in this paper, the clinical correlation is performed cautiously and should be considered preliminary. The specific effect of fibronectin and collagen type 4 on vibratory behavior has not yet been quantified.

Fresh tissue specimens obtained at post mortem examination from normal adult larynges were compared to the same tissue fixed in formalin, embedded in paraffin, and stained with H&E. For immunocytochemical studies, fresh and paraffin-embedded tissues were compared with regard to their staining of monoclonal antibodies directed against fibronectin and collagen type 4; an immunoperoxidase (peroxidase-antiperoxidase) technique was used. Collagen type 4 and fibronectin were detected by an immunocytochemical technique (peroxidase-antiperoxidase, using Diaminobenezidine as a substrate). Antibodies were obtained from Dako Corp, Carpenteria, Calif. This procedure assured us that we could identify the components of fibronectin and collagen type 4 protein which survived the fixation process. Controls of kidney and skin (fixed and unfixed with known staining distributions) were used with each staining process.

Results

We found that normal adult vocal fold lamina propria has fairly moderate fibronectin deposition. This fibronectin deposition is intense through the superficial layer, and it remains fairly intense through the intermediate layer. In the deep layer, the collagen fiber predominates, and fibronectin staining becomes less intense. Although we are currently working on a technique to quantify the amount of fibronectin in vocal folds, so far we can only report our observations in degrees of intensity.

It is curious that normal, non-injured vocal folds should have much fibronectin deposition. **In the normal superficial layer of the lamina propria, one of the major filler proteins is fibronectin.** In view of the sparse fibrous components in the superficial layer, the filler components may be the predominant determinant of tissue deformability.

In normal human vocal folds, the basement membrane zone is similar to human skin in construction and thickness (see center plate, Photo 1). This normally consists of a thin, highly organized structure securing the epithelium to the lamina propria. No increase in basement membrane zone thickness is seen. Collagen type 4 staining can be seen around arterioles and just below the basal cells as a thin brown line. Filler proteins of the lamina propria found in these normal specimens are fibronectin, hyaluronic acid, and other as yet unidentified mucopolysaccharides.[8]

We identified two general types of pathologic responses in the vocal fold lesions examined. One response is characterized by **intense fibronectin deposition (more than normal) in the superficial layer of the lamina propria (see center plate, Photo 2a-c), and is often associated with basement membrane zone injury (see center plate, Photo 3a-b).** The second type of response is noted for its lack of fibronectin (less than normal) in the superficial layer (see center plate, Photo 4a-b). Staining for fibronectin shows minimal or no fibronectin. This second type of pathologic injury usually shows no basement membrane zone injury and the pathology may be confined to the lamina propria. Some specimens occasionally showed variations of these findings.

Discussion

Etiology of vocal fold injury and the pathologic response to that injury have been debated for years. Vocal fold polyps have been attributed to vocal abuse that has resulted in a hemorrhagic event or increased vascular permeability.[9] Nodules have also been attributed to vocal abuse that has caused a thickening of the epithelium or lamina propria.

The results of this work indicate that one type of injury is due to lamina propria disruption usually associated with severe basement membrane zone injury. This injury leads to an aggravated healing response marked by fibronectin deposition in amounts exceeding that found in normal vocal folds. Since the basement membrane zone shows thickening and disorganization, this injury involves more than the lamina propria. The thickness of collagen type 4 band confirms the extent and aggravation of basement membrane zone injury sustained.

Normal basement membrane zones heal within days or weeks following injury. Basement membrane zone thickening and disruption to the degree seen in this vocal injury pattern are rarely seen anywhere else in the body. This is likely due to the repetitive and chronic nature of re-injury that the vocal folds frequently sustain. This inability to let the vocal folds rest enough to heal probably leads to the aberrant healing responses seen in these specimens.

Fibronectin is a structural glycoprotein deposited by fibroblasts in response to interstitial injury. Structural glycoproteins fill and permeate the areas between and around the fibrous structures. Fibronectin is a large macromolecule with a structure that promotes binding to several other macromolecules. Fibronectin and vitronectin (another structural protein found in extracellular matrix) bind to the collagens, other structural proteins, and proteoglycans.[10] Fibronectin is one of the early macromolecules deposited in wound healing and is a prominent protein in wound repair.[11]

Fibronectin is a precursor for collagen deposition and scar formation. However, in vocal fold benign laryngeal lesions, increased collagen or scar is not often seen. The increased fibronectin found in these lesions indicates that enough interstitial injury has occurred to cause a cellular fibroblastic response to repair the extracellular matrix. This response seems to be aborted before the phase of collagen type 1 or type 2 (scar) deposition is achieved.

These findings lead to formation of the following picture of etiology and pathologic response. The vocal folds sustain enough injury to lead to basement membrane zone disruption and superficial layer of the lamina propria injury. The injury, if repetitive, leads to aberrant healing and fibroblastic response with increased fibronectin deposition. The clinical implications are that in the acute phase, treatment must allow healing to occur without repetitive injury as fibronectin deposition is usually a permanent event.[6,11] Because of the biologic characteristics of this macromolecule, which are molecular binding, adhesion, and strength, its presence (in excess) probably does not contribute to easy tissue vibration.[2,10]

This type of injury was seen in nodules and, occasionally, in polyps. The deposition of fibronectin may explain why some nodules do not respond to voice therapy, even though voice habits to reduce abuse are strictly followed by the patient. Clinically, a patient with vocal nodules may show vastly improved phonatory habits while stroboscopy examination continues to show nodules or areas of stiffness. Fibronectin deposition would likely lead to either or both clinical findings. The disordered basement membrane and lamina propria may also predispose the patient for recurrence of the lesion since the pathological state is less likely to be as resilient to voice abuse as the original uninjured state. In cases of chronic injury, because these proteins are permanent, surgical excision may be required for improved voice.

The second type of pathologic finding shows few structural proteins such as collagen or elastin and little or no fibronectin.[12] Clinically, this absence of the structural glycoproteins and fibrous proteins would lead to a fold with excessive propensity for deformability (like a bag of jelly). The effect is then opposite to that of the first type of pathology described.

The pathogenesis of this second type of injury is not clear. The injury often seems to be confined to the lamina propria only and apparently does not consist of a fibroblastic response of tissue repair. This second type of pathologic response is seen more often in Rheinke's edema, polypoid corditis, and some polyps.

Conclusion

The extracellular matrix consists of many structural and filler proteins. The above descriptions represent only a few of them. Diseases such as some polyps, polypoid corditis, or Rheinke's edema may be disorders of abnormal proteoglycan deposition. Diseases leading to dampening of the mucosal wave or decreased tissue vibration may be due to increases in collagen fibers (scar), but also may be due to excessive deposition of structural glycoproteins such as fibronectin, as shown in this research. Further studies of normal and abnormal vocal folds will lead to a better characterization of fold pathology and, hopefully, etiology and treatment.

Acknowledgment

This work was supported by PHS Grant No. 5 K08 DC00036-06 Basement Membrane Zone in Vocal Injury.

References

1. Gray SD, Titze I, Lusk RP. Electron microscopy of hyperphonated vocal cords. J Voice 1987;1(1):109-115.

2. Labat-Robert J, Bihari-Varga M, Robert L. Extracellular matrix. Federation of European Biochemical Society 1990;268(2):386-393.

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

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

5. Hirano M, Kurita S, Sakaguchi S. Aging of the vibratory tissue of human vocal folds. Acta Otolaryngol 1989; 107:428-433.

6. Grinnell F. Fibronectin and wound healing. J Cell Biochem 1984;26:107-116.

7. Lauri WG, Leblong CP, Martin GR. Localization of type 4 collagen, laminin, heparin sulfate proteoglycan, and fibronectin to the basal lamina of basement membranes. J Cell Biol 1982;95:340-344.

8. Matsuo D, Watanabe S, Hirano M, et al. Acid mucopolysaccharide and glycoprotein in the vocal fold. Alterations with aging. Practica Otol Kyoto 1984;77:817-822.

9. Frenzel H. Fine structural and immunohistological studies on polyps of human vocal folds. In: Kirchner JA ed, Vocal fold histopathology, San Diego, CA: College-Hill Press, 1986:39-50.

10. Ruoslahti E. Fibronectin and its receptors. Annu Rev Biochem 1988;57:375-413.

11. Nagata H, Ueki H, Moriguchi T. Fibronectin. Arch Dermatol 1985;121:995-999.

12. Mossallam I, Kotby MN, Ghaly AF, Nassar AM, Barakah MA. Histopathological aspects of benign vocal fold lesions associated with dysphonia. In Kirchner JA ed, Vocal fold histopathology, San Diego, CA: College-Hill Press, 1986:65-80.

13. Gray SD. Basement membrane zone injury in vocal nodules. In: Gauffin J and Hammarberg B eds, Vocal fold physiology, San Diego, CA: Singular Publishing Group, Inc., 1991;21-27.

NCVS Status and Progress Report - 7
December 1994, 23-27

Magnetic Resonance Imaging (MRI) Assessment of Vocal Fold Medialization Surgery

Charles Ford, M.D.
June Unger, M.D.
Roger Zundel, M.D.
Diane Bless, Ph.D.
University of Wisconsin Clinical Science Center

Abstract

Historically, clinicians have used subjective assessment and perceptual judgments, supplemented with acoustic measures, aerodynamic studies, and videostroboscopy, to determine the effects of phonosurgery. When phonosurgical results are poor, MRI can be useful in determining how the surgical modifications contributed to the anatomical and functional status of the vocal folds.

We present here examples of MRI following vocal fold medialization by injection, thyroplasty, and arytenoid adduction. Findings reveal that the superior contrast resolution of MRI can precisely identify placement and persistence of injected implants, and is particularly helpful in showing effects of the size and shape of alloplastic prostheses on vocal fold displacement. Such information is useful in trouble-shooting suboptimal results, and in planning revision thyroplasty by defining modification in the design of prostheses and the placement of cartilaginous windows in medialization thyroplasty. MRI can also aid in confirming indications for and limitations of certain procedures.

Introduction

Vocal fold medialization can be achieved through a variety of surgical options. Treatment outcomes are based on resolution of dysphonia, aspiration, and other glottic insufficiency symptoms. In evaluating poor results, it is useful for the surgeon to appreciate the precise effects of surgery. Thorough analysis of the post-operative condition permits accurate diagnosis of the problem and facilitates planning subsequent modifications of treatment. It can also enhance our understanding of how surgery affects the vocal apparatus and lead to refinement of indications and surgical approaches.

Such analyses depend on optimal assessment techniques. Objective assessment of medialization phonosurgical procedures have focused primarily on voice results[1-4]. Laryngeal examination by indirect laryngoscopy, endoscopy techniques, and videostroboscopy provide indirect evidence as to the status of underlying tissues. Isshiki[5] used CT scanning to demonstrate the results of medialization thyroplasty, but this technique has not proven a valuable adjunct in phonosurgical assessment. Towler[6] and Hanafee and Ward[7] used magnetic resonance imaging (MRI) to elucidate laryngeal anatomy, and Curtin[8] found it particularly helpful in delineating soft tissue planes in the larynx. These characteristics have rendered MRI more useful than CT in defining the limits of soft tissue tumors in the larynx[9,10]. We believed that these same characteristics would make MRI a practical technique for studying results of vocal fold medialization procedures. We report here on the use of MRI to assess results after vocal-fold medialization by injection, thyroplasty, and arytenoid adduction.

Methods

Initially MRI studies were done in four patients with suboptimal medialization results. Subsequently we reviewed MRI studies on a variety of medialization procedures (20 scans from The Universities of Wisconsin and Iowa) and selected five that best demonstrated the soft tissue displacement characteristic of the specific procedure. All examinations were performed on either 1.5-T system with a cryogenic permanent magnet (Signa; GE Medical Systems, Milwaukee, WI), or a .4-T system with a hybrid magnet (Resonex, Resonex, Inc., Fremont, CA). An anterior neck coil was used, and T1-weighted images were acquired in the sagittal and coronal planes, followed by T1-weighted, proton density, and T2-weighted images which were obtained in the axial plane. The sagittal images were used to localize the supero-inferior and antero-posterior limits of the larynx.

The subsequent coronal images were then used to define both the larynx and paralaryngeal soft tissues within an area predetermined from the sagittal sections. The axial images were acquired from a supraglottic level corresponding to the mid-portion of the pyriform sinuses to an infraglottic plane which coincided with the inferior aspect of C-6, or the first tracheal cartilaginous ring. In general, a repetition time (TR) of 500 ms and echo time (TE) of 30 was used for the T1-weighted images, TR of 2000 ms and TE of 25 ms for the proton density images, and TR of 2000 and TE of 80 ms for the T2-weighted images. There was minimal variability between the scanners. A 16-cm field of view, 3-mm section thickness, 0.3-mm intersection gap, and 160 X 256 matrix were standard. Although additional T1-weighted images using fat suppression and intravenous Gd-DTPA (Magnevist®*), 0.1 m mole/kg were obtained on pilot study patients imaged following vocal fold medialization by injection, the technique failed to demonstrate expected enhancement of soft tissues around the implants and no paramagnetic contrast or fat suppression was used in any of the reported imaging studies. Voice production was assessed using the C-speech program for acoustic analysis. Morphologic and functional observations were made with indirect laryngoscopy and a 90° rigid telescope coupled to a Kay Elemetrics light source

Case Reports

Case 1. Analyzing Suboptimal Results: Revision Not Indicated: This 54-year-old female had undergone a medialization thyroplasty to correct glottic insufficiency resulting from vagal nerve injury during resection of a Schwanoma. Pre-operatively she had experienced severe breathiness during phonatory efforts and choked on some foods and liquids. Postoperatively her voice result was suboptimal and she was found to have a vocal-fold hematoma. At 4 months her voice was stronger and the hematoma resolved, but laryngeal videostroboscopy suggested that vocal fold closure was not on plane. It was unclear whether this was a product of poor implant placement or residua of the soft tissue hematoma. MRI revealed that there was soft tissue fullness on the superior surface of the true vocal fold (Photo 5A; see center plate) although the implant was adequately placed in the anterior-to-posterior and vertical axes (Photo 5B; see center plate). On the basis of this information, no further surgical intervention was attempted and the patient was treated expectantly. After one year her voice improved further and she was pleased. Voice function remained suboptimal, however, with a frequency range of 200-400 Hz, intensity range of 64-82 dB, and slight residual convexity in the mid-membranous vocal fold.

Case 2. Analyzing Suboptimal Results: Revision Indicated: This 33-year-old female had undergone multiple microlaryngeal procedures to remove nodules and polyps. She had bilateral bowed vocal folds, sulcus deformities, and stiffness, with resultant glottic insufficiency. After GAX collagen injections failed to effectively medialize the vocal folds, we performed bilateral thyroplasty. Since both vocal folds were mobile, the implants were placed anteriorly to avoid impeding arytenoid excursion. This resulted in improvement on the operating table but the voice was unstable. MRI demonstrated that the implants had been placed too far anteriorly (Photo 6; see center plate). Revision resulted in a stable and much improved voice at one year, with a frequency range of 187-400 Hz and intensity range of 70-90 dB.

Case 3. Modifying Procedure: Attention to Implant Design: This 68-year-old male with Parkinson's disease and presbylaryngis presented with breathy dysphonia that did not improve after conservative autogenous fat injections. We performed bilateral thyroplasties using a square interlocking shim design. The result was suboptimal; MRI showed failure of the medial edge of the implant to conform to the vocal fold deficit (Photo 7A; see center plate), although the vertical placement was appropriate (Photo 7B; see center plate). The implant was modified to decrease anterior displacement and rounded to compensate for the bowed vocal fold, which improved the voice result.

Case 4. Modifying Procedure: Attention to Cartilage Window Placement: In this 45-year-old female with right vocal fold paralysis, the result of medialization thyroplasty was satisfactory and the patient did not want revision. MRI, nevertheless, showed failure to adapt placement of the thyroid cartilage window to the shape of the implant (Photos 8A,B; see center plate). Failure to lower the window when the implant overlaps in the vertical axis can cause medialization of the ventricle and false vocal fold. Failure to place the window further posteriorly when the implant overlaps in the anterior-to-posterior axis can result in inappropriate medialization of the anterior third of the vocal fold. In this case, we compensated for anterior overlap by designing the implant with an extreme anterior taper.

Case 5. Defining Medialization Options: Arytenoid Adduction: At the age of 29, this young male had undergone resection of a vagus nerve neurilemmoma, which resulted in severe breathy dysphonia and chronic aspiration. Three years later, laryngoscopy revealed a large posterior glottic gap and barium penetration was noted in a swallowing study. Arytenoid adduction corrected the aspiration immediately, and resulted in a strong voice that was only a bit unstable for several months. MRI demonstrated how effectively arytenoid adduction closed the posterior glottis and interarytenoid area (Photo 9; see center plate). The coronal view showed that closure in the mid-membranous portion of the vocal fold was on plane.

Case 6. Defining Medialization Options: Thyroplasty: This 69-year-old patient had presbylaryngis with bilateral bowed atrophic vocal folds and a left vocal-

*Berlex Laboratories, Wayne, New Jersey

fold paralysis. The implant on the left fold was designed to compensate for the laterally placed vocal fold, and both implants were designed to be convex medially to compensate for vocal-fold bowing (Photo 10A,B; see center plate). This case demonstrates the versatility of medialization thyroplasty; by modifying the implant, it was possible to medialize the left vocal fold and also correct for bilateral vocal fold bowing.

Cases 7-9. Defining Medialization Options: Augmentation Injections: Case 7 was a 73-year-old male patient with left vocal-fold paralysis following carotid resection during extended neck dissection who was treated by autogenous fat injection. MRI demonstrated medialization of the membranous vocal fold (Photo 11A; see center plate), and also revealed the limitations of the injection procedure. In contrast with arytenoid adduction, injection provides for no appreciable closure of the posterior commissure. Unlike medialization thyroplasty, injection limits contour modifications and refinements by the bulk displacement characteristics of the material.

Attempts to demonstrate injected collagen and Teflon were less successful. In Case 8, the MRI was done only one week after Teflon injection and the vocal fold was clearly medialized although the Teflon could not be well demonstrated (Photo 11B; see center plate). GAX collagen was detected following vocal-fold injection but resonance characteristics were similar to the surrounding tissues. In Case 9 (Photo 11C; see center plate), MRI revealed GAX filling the vocal fold superior and inferior the plane of a sulcus vocalis defect.

Discussion

Critical factors in achieving good voice results in phonosurgical procedures are establishing accurate diagnosis and proper patient selection. These can best be assured through an appreciation of the options available for management, modifications and refinement of surgical techniques for optimal results, and critical analysis of poor results to guide subsequent management. MRI was chosen to help achieve these goals because the superior contrast resolution of MRI allows precise evaluation of silastic and fat implants. With MRI, loss of detail from laryngeal motion can be eliminated through averaging with multiple excitations and better spatial resolution in the coronal plane is available with MRI than with CT reconstructions.

In the examples presented, MRI was useful in trouble-shooting for patients with poor results, in modifying techniques in planning for future procedures (honing surgical skills), and in demonstrating the characteristics of selected medialization procedures. The two cases demonstrated the use of MRI in identifying the soft tissue changes responsible for poor results following medialization thyroplasty and in deciding if revision surgery might help. Observations in Cases 3 and 4 were useful as guides in planning subsequent medialization thyroplasties.

One of the main goals of any medialization procedure is to achieve glottic closure during phonation by simulating the plane, position, and contour of the normal adducted vocal fold. Case 3 was one example of an inappropriately shaped implant that failed to achieve these goals. The MRI served as a guide in planning implant design to address specific anatomical problems. It was central to the successful revision of that patient, and in the execution of subsequent procedures in patients with similar problems. For example, in Case 6, upon MRI analysis, the implant was designed with a convex contour to compensate for the bowed atrophic vocal fold and the anterior edge is more tapered.One of the most important aspects of medialization thyroplasty is the proper placement and fashioning of the thyroid cartilage window. Anatomic studies demonstrating the projection of the vocal fold on the thyroid cartilage skeleton have been used to suggest guidelines for cartilage incisions[11]. Isshiki[12] carefully described window placement such that the superior margin is in the plane of the midpoint of the anterior height of the thyroid cartilage and the anterior vertical margin is 5-7 mm from the midline. Koufman[13] proposed a specific formula based on the dimensions of the larynx by which he calculated placement of the window. The shape and orientation of prostheses employed varies greatly among surgeons. Netterville[14] described a three-sided interlocking silastic design; Montgomery[15] used a posterior displacement probe; Cummings[16] devised a vertically oriented hydroxyapatite ridge; and a variety ofsilastic designs has been generated by Koufman[11], Tucker[17] and others[18,19]. Most surgeons, nevertheless, adhere to the general guidelines proposed by Isshiki for placement and design of the window.

We continued to place the window in the same location as Isshiki without modification for the new implant shape[18]. We confirmed that the window was in the plane of the membranous vocal fold by passing needles through the window outlined in a cadaver larynx. Nevertheless, Case 5 was one example where window placement should have been modified to allow for overlap of the prosthesis. Failure to recognize anticipated overlap in a vertically-interlocking prosthesis can result in obliteration of the ventricle if the window is not placed at a lower level (Photo 12; see center plate). Similarly, anterior overlap of a prosthetic implant must be accommodated by more posterior window placement (Photo 13; see center plate). This concept can be appreciated by reviewing the Type I Thyroplasty as described by Isshiki[12]: the cartilage window is left attached to the inner perichondrium and displaced medially by a simple wedge of silastic that is the exact dimensions of the window (Photo 14; see center plate). Most techniques subsequently described removal of cartilage, different shapes of implants extending well beyond the dimensions of the windows, and, in some cases, removal of inner perichondrium. This study points out the need to accommodate the placement of the window to the prosthesis.

The final five cases highlight some of the distinguishing features of three approaches to vocal fold medialization. Case 5 shows the effective overcorrection of the posterior glottis that can be achieved with arytenoid adduction. This can be extremely useful in correcting aspiration problems, as in the patient presented. Case 6 demonstrates the scope and precision of correction possible with medialization thyroplasty. In this patient, simultaneous asymmetric bilateral thyroplasties were used to compensate for presbylaryngis and unilateral paralysis. The potential for intraoperative adjustments, postoperative revision, and reversibility make this a very useful approach for optimal voice results.

Case 7 is an example of medialization by fat injection and demonstrates the general anatomical changes in a successful injection augmentation. The convex displacement of the vocal fold is a function of the spherical distribution of the injectate, which can vary in degree. Subtle adjustments are difficult to achieve once the material is injected and precision is compromised by variable inflammatory response to Teflon and by variable resorption of bioimplants. Cases 8 and 9 show Teflon and collagen-injected vocal folds but it was difficult to distinguish the injectate from the adjacent tissues. This was thought to be possibly a result of migration, resorption, or failure of MRI to detect distinct resonance characteristics of the injectate. The latter seems most likely, since the study in Case 8 was performed just one week after injection, voice was improved, and the vocal fold was clearly medialized. Injected bovine GAX-collagen also did not show clearly on MRI. In Case 9 there was clearly residual medialization of vocal fold tissue in a patient who remained dysphonic due to a sulcus vocalis deformity that extended along the entire free margin of the vocal fold. Previous studies have shown collagen injection effective in treating the scarred vocal fold 20 but MRI demonstrates the distribution of collagen in this case and reveals the difficulty in getting appropriate dispersion of the injectate.

In summary, MRI is useful in assessing several aspects of vocal fold medialization procedures. When phonosurgical results are poor, MRI can be used to troubleshoot and can help determine if revision is indicated. MRI has been helpful in planning revision surgeries, defining modification in the design of prostheses and in the placement of cartilaginous windows in medialization thyroplasty. MRI can also help clinicians better appreciate the relative merits and limitations of various options in medialization procedures

Acknowledgements

The authors would like to acknowledge the contributions of the National Center for Voice and Speech (Grant No. P60 DC00976 from the National Institute on Deafness and Other Communication Disorders) and Michael G. Bayerl, M.D. and Henry T. Hoffman, M.D. for their assistance in providing material for case 7 from the University of Iowa. Thanks also to Joan Kozel for the timely completion of the illustrations.

Bibliography

[1] Watterson T, McFarlane S and Menicucci A: Vibratory characteristics of Teflon-injected and noninjected paralyzed vocal folds. JSHD, 55:61-66, 1990.

[2] Hammarberg B, Fritzell B and Schiratzki H: Teflon Injection in 16 Patients with Paralytic Dysphonia: Perceptual and acoustic evaluations. JSHD, 49:72-82, 1984.

[3] Sasaki CT, Leder SB, Petcu L and Friedman CD: Longitudinal voice quality changes following Isshiki Thyroplasty type I: the Yale experience. Laryngoscope, 100:849-852, 1990.

[4] Slavit DH and Maragos NE: Physiologic assessment of arytenoid adduction. Annals of ORL, 101:321-327, 1992.

[5] Isshiki N: Phonosurgery: theory and practice. Springer-Verlag, Tokyo, 1989.

[6] Towler CR and Young SW: Magnetic resonance imaging of the larynx. Magn Reson Q, 5:228-241, 1989.

[7] Hanafee WN and Ward PH: The Larynx. Radiology, Surgery, Pathology. Thieme Medical Publishers, Inc., New York, 1990.

[8] Curtin HD: Imaging of the larynx: current concepts. Radiology, 173:1-11, 1989.

[9] Hoover LA, Wortham DG and Hanafee WN: Magnetic resonance imaging of the larynx and tongue base: clinical applications. Otolaryngol Head Neck Surg, 97:245-256, 1987.

[10] Teresi LM, Lufkin RB and Hanafee WN: Magnetic resonance imaging of the larynx. Radiol Clin NA, 27:395-406, 1989.

[11] Meiteles LZ, Lin PT and Wenk EJ: An anatomical study of the external laryngeal framework with surgical implications. Otolaryngol Head Neck Surg, 106:235-240, 1992

[12] Isshiki N, Okamura H and Ishikawa T: Thyroplasty type I (lateral compression for dysphonia due to vocal cord paralysis or atrophy). Acta Oto-larlyngol, 80:465-473, 1975.

[13] Koufman JA: Laryngoplasty for vocal cord medialization: An alternative to Teflon. Laryngoscope, 96:726-731, 1986.

[14] Netterville JL, Stone RE, Luken ES, Civantos FJ and Ossoff RH: Silastic medialization and arytenoid adduction: the Vanderbilt experience. Ann Otol Rhinol Laryngol, 102:413-424, 1993.

[15] Montgomery WW, Montgomery SK and Warren MA: Thyroplasty simplified. Op Techn Otol-HNS, 4:223-231, 1993.

[16] Cummings CW, Purcell LL and Flint PW: Hydroxlapatite laryngeal implants formedialization: preliminary report. Ann Otol Rhinol Laryngol, 102:843-851, 1993.

[17] Tucker HM: Combined laryngeal framework medialization and reinnervation forunilateral vocal fold paralysis. Ann Otol Rhinol Laryngol, 99:778-781, 1990.
[18] Ford CN, Bless DM and Prehn RB: Thyroplasty as primary and adjunctive treatment of glottic insufficiency. J Voice, 6:277-285, 1992.
[19] Maves MD, McCabe BF and Gray S: Phonosurgery: indications and pitfalls. AnnOtol Rhinol Laryngol, 98:577-580, 1989.
[20] Ford CN and Bless DM: Collagen Injection in the Scarred Vocal Fold. J. Voice, 116-118, 1987.

NCVS Status and Progress Report - 7
December 1994, 29-32

Maternal Speech Rate and Childhood Stuttering: Is Slower Always Better?

Patricia Zebrowski, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa

Introduction

A number of contemporary theories of stuttering onset and development views the emergence and perpetuation of stuttering as the result of a complex and multidimensional interaction between the child and his or her environment (e.g., Andrews, Craig, Feyer, Hoddinot, Howie & Neilson, 1983; Bloodstein, 1970; Conture, 1990; Johnson & Associates, 1959; Starkweather & Gottwald, 1990). Although many factors contribute to this interplay, in recent years researchers and clinicians have specifically examined the verbal interactions between children who stutter and their parents, in the hope of delineating aspects of parent-child discourse that may be related to either the perpetuation or amelioration of stuttering. In particular, researchers have been interested in the relationship between "time pressure," which children who stutter may experience in conversation, and how such pressure may lead to increased stuttering for the child (e.g., Conture, 1990; Schulze, 1991; Starkweather, Gottwald & Halfond, 1990; Van Riper, 1973). Attempts to identify which conversational timing behaviors create perceptions of communicative time pressure for children who stutter have led to the examination of a number of parent and child discourse behaviors associated with the timing or tempo of both their utterances and turns. These behaviors include: (1) speaking rate, (2) response-time latency, also referred to as turn-switching or interspeaker pause duration, (3) number of turns, as well as (4) rate of turn-taking, (5) frequency of interruption, and (6) duration of co-conversationalist overlap or "simultalk" (Kelly & Conture, 1992) associated with interruption.

Of all the different turn-management timing behaviors used in conversation, parent and child speaking rate has received the most attention in studies of parent-stuttering child interaction. In particular, a number of studies of the speech rate of the mothers of children who stutter has recently been published. These studies have examined maternal speaking rate in naturally occurring mother-child verbal interaction, as well as within conversations in which mothers were instructed to reduce their rate of speech. The purpose of this paper is to present the results from descriptive studies of maternal speech rate in the spontaneous conversations of children who stutter and their mothers, as well as results from studies of maternal speech rate reduction in similar contexts. Following this review, some conclusions about the relationship between parent and child speech rate and the emergence and perpetuation of childhood stuttering will be offered.

Studies of Maternal Speech Rate in the Conversations of Children Who Stutter and Their Mothers

Since 1985, a relatively small number of studies has been published that reported the speaking rates of mothers of young children who stutter while they engaged in conversation with their children. In the first of these investigations, Meyers and Freeman (1985) observed that the mothers of 12 four to six-year-old boys who stutter spoke significantly faster to their sons than the mothers of a matched group of nonstuttering boys spoke to their children. In their conclusions, Meyers and Freeman speculated that "fast talking mothers contribute to the development or maintenance of stuttering in their children" (p. 441). Besides its provocative findings, an important aspect of this study is that it motivated several replication studies of maternal speech rate in the conversations of children who stutter and their mothers. In two of these subsequent investigations, the speech rates of mothers of preschool children who stutter were measured and compared to those of a control group. In each case, there was no significant difference in rate of speech between the two groups of mothers (Kelly & Conture, 1992; Schulze, 1991).

In a third and somewhat different investigation, Yaruss and Conture (1992) studied the relationship between mother and child speech rate in adjacent, perceptually fluent utterances. Ten pairs of utterances were selected from each mother-child dyad's videotaped conversational sample. As in the work of Kelly and Conture (1992) and Schulze (1991), results indicated that there were no significant differences in speech rate between the mothers of the stuttering and nonstuttering children. In addition, individual paired comparisons and multiple regression models for each of the mother-child pairs (N = 20) revealed no significant relationships between any of the mothers and their children for rate of speech.

Discrepancies between these more recent studies and the work of Meyers and Freeman are most likely due to a number of factors, including differences in the ages of the children used as subjects, as well as inter-subject and intra-subject disparities in stuttering severity, conversational samples obtained, and the methods used for making measures of speech rate (e.g., derived from video tape time-code as opposed to the acoustic waveform). With these possibilities in mind, however, it seems reasonable to conclude that as a group, the mothers of children who stutter do not differ from the mothers of nonstuttering children in the speech rate they use when talking with their child. Further, in individual mother-child dyads, there may exist a clinically significant difference or "mismatch" between the speaking rates of the mother and her child, and it is this "dyadic" or "relative rate" to which clinicians should attend (Kelly & Conture, 1992; Starkweather et al., 1990).

Studies of Maternal Speech Rate Reduction in the Conversations of Stuttering and Nonstuttering Children and Their Children

Although research has shown that the mothers of children who stutter do not tend to speak excessively fast when talking with their children, a common therapeutic strategy for working with young children who stutter is to counsel their parents to reduce their speaking rate when talking with the child. This advice is based on both the clinical assumption and clinical observation that a slower speech rate on the part of the parent results in more fluent speech for the child.

Since the late 1980s, there have been several studies which have empirically assessed the effects of maternal rate reduction training. These investigations have asked the questions, (1) Do mothers slow their speaking rate when specifically trained to do so? (2) If so, do their children slow their rate when talking with their mothers? (3) Is either maternal or child rate reduction related to changes in the child's disfluencies? These studies have examined the impact of maternal rate reduction in dyads comprised of stuttering children and their mothers (Guitar, Kopff, Donahue-Kilburg & Bond, 1992; Savelkoul, Zebrowski & Buizer, 1993; Stephenson-Opsal & Bernstein Ratner, 1988), as well as those consisting of nonstuttering children and their mothers (e.g., Bernstein Ratner, 1992; Marchinkoski & Guitar, 1992). Further, these studies reported analyses of both group and individual dyad and subject data.

A consistent finding across all of these studies was that while in conversation with their child, the mothers were able to produce and maintain a significantly reduced speech rate. However, analysis of the relationship between the mothers' slower rate and their children's speech rate and frequency of disfluency suggests that parental rate reduction as a clinical strategy should be used differentially. That is, while the majority of mothers in these studies successfully reduced their speech rate, their children did not typically follow suit, nor were dramatic, across-the-board decreases in frequency of speech disfluency reported. Those children who exhibited increased fluency when their mothers talked to them more slowly, tended to speak at faster, rather than slower, rates in response to their mother's reduced rate (Stephenson-Opsal & Bernstein Ratner, 1988), or were children who produced a high proportion of whole-word or sound/syllable repetitions (as opposed to prolonged sounds; Savelkoul et al., 1993) as well as a tendency to reduce "primary" as opposed to "secondary" stuttering behaviors as a result of their mother's slowed speech rate (Guitar et al., 1992).

Based on these findings, one can conclude that parental rate reduction is not fluency facilitating for all children who stutter. Or it may be the case that the overall tempo or perceived tempo of the conversation (i.e., the combined influence of rate, turn-switching pause duration, number and rate of turn exchanges) may be more important than rate alone in contributing to either the child's increased fluency or disfluency. As well, perhaps the overall tempo of the parent-child verbal interaction influences the child's perception of the parent's affective state. It may be these latter two conditions which account for those cases in which speaking more slowly to a child who stutters does result in increased fluency, even when the child himself does not reduce his rate.

Conclusions

Presently, there is a need for continued exploration into the temporal dynamics of the verbal interactions between children who stutter and both their mothers and fathers. Additional research may uncover specific parameters, or clusters of parameters which, if manipulated, can enhance a child's fluency within a particular conversational context. Based on current research findings, however, several factors should be considered when deciding which children who stutter might benefit from a parental model of

slowed speech. These factors include: the frequency of speech disfluency the child produces, as well as overall severity of the problem and the types of disfluencies the child most frequently produces (i.e., sound/syllable and whole-word repetitions versus sound prolongations; or "primary" as opposed to "secondary" stuttering), the age and general speech articulation, language, cognitive, and emotional abilities and characteristics of the child, and the difference or "gap" between the child's and parent's rate (dyadic or relative rate). Further, it is likely that differences exist between mother-stuttering child and father-same stuttering child conversational dyads, such that a child may benefit from either one or both of his parents using a reduced rate when talking with him (Kelly, 1994). In addition, it is important to keep in mind that parents will need continued guidance and monitoring of their "new" way of talking, so that it is acceptable to them and their child.

Finally, it is essential for clinicians and researchers alike to consider the complex characteristics of influence in parent-child interaction. Specifically: (1) Influence is bidirectional. That is, as Kelly (1993) notes, within the parent-child verbal interaction, the parent and the child exert mutual influence over each other's speech rate, turn-taking behaviors, and the like. (2) Influence may be indirect. It may be that the parent's speech rate, produced at one point in time, may not affect the same behavior exhibited by the child at that same point in time, but instead may influence a related parameter (such as turn-switching pause duration, for example), which in turn may be highly correlated with increased fluency. Or, the parent's rate may have a direct influence on the child's rate, but with no observable effect on the child's speech fluency at that point in time. The obvious explanation is that there is no relationship between the child's speaking rate and stuttering. An alternative explanation, however, is that the relationship between speaking rate and stuttering or fluency is cumulative; that is, the parent's and child's consistent use of as slowed speech rate over time eventually contributes to the child developing normally fluent speech at a later point in time. (3) Influence is variable. Considering all the sources of possible variation, a high degree of both inter- and intra-dyad variability with regard to any conversational timing behavior and its relationship to disfluency should be expected. As such, the need for well designed single-subject studies is clear. Finally, (4) Influence in parent-child verbal interaction is only one piece of the puzzle. There is consensus among clinicians and researchers alike that the problem of childhood stuttering is both multidimensional and complex. It is probably safe to say that all of the specific factors which contribute to the emergence and development of stuttering are not know, and it is likely that there are multiple etiologies. Therefore, the significance of environmental, and specifically, parental influence is only one "piece of the puzzle" (Kelly, 1993).

References

Andrews, G., Craig, A., Feyer, A., Hoddinott, A., Howie, P., & Neilson, M. (1983). Stuttering: A review of research findings and theories circa 1982. *Journal of Speech and Hearing Disorders*, 48, 226-246.

Bernstein Ratner, N. (1992). Measurable outcomes of instructions to modify normal parent-child verbal interactions: Implications for indirect stuttering therapy. *Journal of Speech and Hearing Research*, 35, 14-20.

Bernstein Ratner, N. (1993). Parents, children and stuttering. *Seminars in Speech and Language*, 14(3), 238-250.
Conture, E.G. (1990). *Stuttering*. Englewood Cliffs, NJ: Prentice-Hall.

Bloodstein, O. (1987). *A handbook on stuttering* (4th ed.). Chicago, IL: The National Easter Seal Society.

Guitar, B., Kopff, H., Donahue-Kilburg, G., & Bond, L. (1992). Parental interactions and speech rate: A case study in stuttering: *Journal of Speech and Hearing Research*, 35, 742-754.

Bloodstein, O. (1970). Stuttering and normal nonfluency - A continuity hypothesis. *British Journal of Disordered Communication*, 5, 30-39.

Johnson, W. and Associates (1959). *The onset of stuttering*. Minneapolis: University of Minnesota Press.

Kelly, E. (1993). Speech rates and turn-taking behaviors of children who stutter and their parents. *Seminars in Speech and Language*, 14(3), 203-214.

Kelly, E. (1994). *Speech rates and turn-taking behaviors of children who stutter and their fathers*. Unpublished manuscript, Purdue University.

Kelly, E.M., & Conture, E.G. (1992). Speaking rates, response time latencies, and interrupting behaviors of young stutterers, nonstutterers, and their mothers. *Journal of Speech and Hearing Research*, 35, 1256-1267.

Marchinkoski, L., & Guitar, B. (1994). *When mothers speak more slowly*. Unpublished manuscript, University of Vermont.

Meyers, S. & Freeman, F. (1985). Mother and child speech rates as a variable in stuttering and disfluency. *Journal of Speech and Hearing Research*, 28, 436-444.

Peters, T., & Guitar, B. (1991). *Stuttering: An integrated approach to its nature and treatment.* Baltimore: Williams & Wilkins.

Riley, G. (1980). *Stuttering Severity Instrument for Children and Adults.* Pro-Ed.

Savelkoul, E., Zebrowski, P., & Buizer, A. (1993). *Reduced maternal speech rate and childhood stuttering and speech rate.* A paper presented to the annual meeting of the American Speech-Language-Hearing Association. Anaheim, CA.

Schulze, H. (1991). Time pressure variables in the verbal parent-child interaction patterns of fathers and mothers of stuttering, phonologically disordered and normal preschool children. In H.F.M. Peters, W. Hulstijn & C.W. Starkweather (Eds.), *Speech motor control and stuttering* (pp.441-452). New York: Excerpta Medica.

Smith, A. (1990). Factors in the etiology of stuttering. In J. Cooper (Ed.), *Research needs in stuttering: Roadblocks and future directions* (p. 18). Rockville, MD: ASHA Reports.

Starkweather, C.W., & Gottwald, S.R. (1990). The demands and capacities model II: Clinical implications. *Journal of Fluency Disorders*, 15, 143-157.

Starkweather, C.W., Gottwald, S., & Halfond, M. (1990). *Stuttering prevention: A clinical method.* Englewood Cliffs, NJ: Prentice-Hall.

Stephenson-Opsal, D.,& Bernstein Ratner, N. (1988). Maternal speech rate modification and childhood stuttering. *Journal of Fluency Disorders*, 13, 49-56.

Van Riper (1973). *The treatment of stuttering.* Englewood Cliffs, NJ: Prentice-Hall.

Yaruss, J.S.,& Conture, E.G. (1992). *Relationships between mother-child speaking rates in adjacent fluent utterances.* Paper presented at the annual meeting of the American Speech-Language-Hearing Association. San Antonio, TX.

Zebrowski, P.M. (1994). Duration of sound prolongation and sound/syllable repetition in children who stutter: Preliminary observations. *Journal of Speech and Hearing Research*, 37, 254-263.

NCVS Status and Progress Report - 7
December 1994, 33-39

The Parkinson Larynx: Tremor & Videostroboscopic Findings

Kathe Perez, M.A.
Lorraine Olson Ramig, Ph.D.
Department of Communication Disorders and Speech Science, The University of Colorado - Boulder
Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts
Marshall Smith, M.D.
Department of Otolaryngology/Head and Neck Surgery, The University of Colorado Health Sciences Center
Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts
Christopher Dromey, M.A.
Department of Communication Disorders and Speech Science, The University of Colorado - Boulder
Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts

Abstract

Patients with Parkinson disease have a high incidence of speech, voice, and laryngeal abnormalities. To characterize laryngeal abnormalities, visual-perceptual ratings of endoscopic and stroboscopic examinations of twenty-two patients diagnosed with idiopathic Parkinson disease and seven patients with Parkinson plus syndromes were carried out by four trained viewers. Incidence of tremor, tremor location, phase closure, phase symmetry, amplitude and mucosal waveform were scored. Tremor was observed in one or more of these conditions, rest, normal pitch and loudness, or loud phonation for most of the twenty-nine patients. Fifty-five percent of the idiopathic Parkinson patients had tremor with the primary location being vertical laryngeal tremor. Sixty-four percent of the Parkinson plus patients had tremor with the arytenoid cartilages being the primary location. Laryngeal tremor was observed early in the disease in these Parkinson patients. The most striking stroboscopic findings for the idiopathic Parkinson disease patients were abnormal phase closure and phase asymmetry. Amplitude and mucosal waveform were essentially within normal limits in the majority of the idiopathic Parkinson patients.

Parkinson disease, a movement disorder associated with degeneration of the brainstem nuclei, particularly the substantia nigra (1) effects approximately 1.5 millon Americans (2). Common symptoms include resting limb tremor (hand & arm), limb hypokinesia, rigidity, and poor control of coordinated movements such as walking. There is a high incidence of speech, voice, and laryngeal abnormalities (3;4;5) in these patients with some speech and voice symptoms occurring at the onset of the disease. It is estimated that 89% of Parkinson disease patients will experience voice difficulties as the disease progresses (5). Characteristics of disordered voice include: reduced loudness, monotone, hoarseness, breathiness, and vocal tremor (6).

Vocal tremor (in conversation and reading) has been reported as a common feature in the voices of idiopathic Parkinson disease (IPD). However, visual observation of vocal tremor in relation to anatomical and physiological correlates has only recently been described. Hanson et al. (4) reported tremor involving the tongue and strap muscles in 31% of the Parkinson Disease patients they studied. During quiet breathing, 12% of these patients demonstrated tremulous movement of the arytenoid cartilages. Aside from Hanson's findings, the incidence of resting laryngeal tremor has not been reported. Resting limb tremor is often an initial symptom of IPD with an incidence as high as 50% (7). Kinetic limb tremor (occurring during goal directed movement) has not been reported in these patients (5), but postural tremor (when the limb is outstretched) is common (8). Considering the perceptual reports of voice tremor and the high incidence of tremor reported in the limbs in IPD patients, this paper sought to examine tremor characteristics of the larynx and surrounding structures for both incidence and location.

Tremor has also been reported in patients classified with Parkinson plus syndromes (PPS), which includes multi-system degeneration, Shy-Drager syndrome, basal ganglia degeneration, and progressive supranuclear palsy. In PPS patients, rapidly progressive deterioration of motor function and more severe speech and voice deficits have been observed (9;10). It is not unusual for PPS patients to go months or even years before proper diagnosis is made. Additional information about the severity and incidence of tremor could assist in diagnosing these patients, which would allow effective treatment choices to be implemented sooner. Based upon the reports of more severe speech and voice deficits, it was expected that the PPS patients would have a greater incidence of laryngeal tremor.

Only recently have the physical correlates for speech and voice disorders in Parkinson disease patients been investigated. In a paper describing the laryngeal characteristics of IPD patients, Hanson, Gerratt, and Ward (4) examined the larynges of thirty-two patients using cinegraphic film. They reported that 94% of their patient sample demonstrated vocal fold bowing. Using videostroboscopic methods, Smith et al. (11) found a 38% incidence of vocal fold bowing and 67% incidence of glottal incompetence in patients with IPD. Stroboscopy has become widely accepted as the method of choice for examining anatomical and function characteristics of vocal fold vibration (12,13,14,15). It was hypothesized that the vibratory patterns of the vocal folds would reflect the perceptual speech and voice characteristics of IPD patients. Reduced loudness observed in Parkinson disease patients would be expected to correlate with phase closure abnormalities since phase closure reflects the degree of medial compression of the vocal folds. Difficulty with coordinated movements is one of the hallmark characteristics of Parkinson disease and is seen as difficulties in voice and speech production because of the coordinated effort needed to couple the various systems of speech production. Respiratory and phonatory coordination, as well as vocal fold tissue integrity are necessary for normal phase symmetry and it was expected such coordination difficulty would present as abnormalities in phase symmetry in these patients. Phase symmetry abnormalities are often seen in both central and peripheral pathologic conditions (12). Amplitude of vibration is the degree of lateral or horizontal excursion of the vocal folds and abnormalities might exist in relation to increased stiffness of the vocal folds, vocal fold bowing or short glottal closure time (12). The mucosal wave occurs along the entire length of the membranous portion of the vocal folds and is affected by many factors. Like amplitude, mucosal waveform is closely related to glottal closure. Mucosal waveform abnormalities would be expected in the presence of vocal fold tissue damage, in relation to peripheral nerve damage, and with short phase closure durations. The extent of tissue loss related to bowing is not well understood and mucosal waveform abnormalities related to bowing may be possible in these patients (12).

This study was designed to compare tremor for IPD and PPS patients and to describe vocal fold vibratory movements in the IPD patient group. Comparisons of the incidence of tremor for years-since-onset of the disease and the location of the tremor were made in the IPD and PPS patient groups. Select characteristics of laryngeal function were investigated through stroboscopic analysis in the IPD patients.

Methods

Subjects

Twenty-two patients diagnosed with IPD and seven patients classified with PPS (e.g.multi-system degeneration, Shy-Drager syndrome, progressive supranuclear palsy, and basal ganglia degeneration) were studied. This sample was taken from a total population of 58 subjects who were used in a larger study to examine voice treatment efficacy in patients with Parkinson disease. Table 1 shows the patients demographic data at it relates to staging in each group of patients (IPD & PPS). Staging of the IPD (n=22) and five of the seven PPS patients was completed by a neurologist using the Hoehn & Yahr scale (16) and the Unified Parkinson's Rating Scale (17). The age range for the IPD patients was 49 to 81 years, the mean age was 65.5 years (SD 8.4 years) with a mean YSO of the disease was 6.9 years. The age range for the PPS patients was 57 to 77 years, the mean age was 65.4 years (SD 7.0 years) with a mean YSO of 4.57 years. Two of the PPS patients had been recently diagnosed when they entered the study. Their YSO was zero, which lowered the mean YSO for this group. These two newly diagnosed PPS had not received a staging of the disease at the time the study began. There were no Stage IV or V IPD patients and no Stage I, II or V PPS patients.

Table 1.
Patient Demographic Data

22 IPD & 7 PPS patients

	Mean	SD	YSO	Gender
IPD Patients				
Stage				
I n=1	51	0	3	1F
II n=7	65.4	16.4	5.1	5M - 2F
III n=14	66	6.8	8.1	12M - 2F
PPS Patients				
Stage				
III n=3	70.7	6.5	6.7	2M - 1F
IV n=2	58	1.4	6	2M - 0F

Instrumentation

The patients included in this study had usable pre-treatment videoendoscopic and stroboscopic imaging data for sustained phonation of /i/. Stroboscopic assessment was a part of the pre-treatment examination protocol and was performed according to well-documented techniques (12;13). An otolaryngologist performed the examinations. Endoscopic examination was conducted with an Olympus ENF-P3 flexible nasopharyngoscope and a Nagashima SFT-1 70 degree rigid telescope. The light sources included an Olympus halogen constant light and a Nagashima LS-3A laryngostrobe. Images were captured on VHS and SVHS videotape with a CCD camera, using a 35 mm lens for the flexible scope and a 60 mm lens for the rigid telescope. Each patient's videoendoscopic examination was recorded on a separate tape. A master tape of each patient's exam was created for the purpose of this study.

Examination Protocol

The nasal passage was topically anesthetized with 4% lidocaine spray. A standard examination protocol was used. The flexible fiberscope was used first under constant light during quiet respiration, followed by sustained phonation of /i/ at a comfortable pitch, comfortable loudness, and production of short phrases. During the stroboscopic portion of the exam, a variety of tasks were used, but only productions of /i/ during normal-pitch normal-loudness (NPNL) and loud phonation were analyzed in this study. After completion with the flexible scope, examination was continued with the rigid telescope.

Rating Procedures

A master tape was created from the twenty-nine individual patient tapes (22 IPD and 7 PPS patients). These views were dubbed onto a master videotape and were rated individually by four raters. The raters consisted of three speech-language pathologists, and an otolaryngologist each experienced in assessment measures of laryngeal function. The order of presentation of the views was randomly assigned and coded. The person who dubbed the master tape (the prompter) was not one of the raters and only the prompter knew the code. To minimize observer bias, the audio portion was deleted. The prompter identified the subject number and each task on the audio track. In addition to the master tape, a practice tape was developed for training purposes. During the training session, five views were examined and raters agreed upon the specific distinctions for the ratings of the variables described below. For comparison, samples of videostroboscopic examinations of normal subjects were available. There was no time limit to viewing and rating the video tapes. Each rater viewed the master tape individually and was allowed to examine each view as often as necessary.

Rating Scale

Adaptation of the Wisconsin scale (12) was used. There were 8 parameters of interest. They included:

1. tremor at rest
2. tremor during phonation of /i/ at NPNL
3. tremor during phonation of /i/ at loud phonation
4. tremor location in each condition (rest, NPNL, loud)
5. phase closure
6. phase symmetry
7. amplitude
8. mucosal waveform

Tremor was scored at-rest, during normal-pitch normal-loudness (NPNL) and during loud phonation. Tremor was analyzed in the flexible scope condition and was rated as either absent or present. In addition to rating the presence of tremor, the location of the tremor for each of these conditions was rated. Raters scored the location of the tremor from the following choices: true vocal folds, vertical laryngeal movement, arytenoid cartilages, pharyngeal walls and epiglottis.

Glottal closure characteristics were rated on a scale ranging from -4 to +4, with a score of 0 being normal; -4 being an extreme open phase and +4 being an extreme closed phase. Phase symmetry was rated on a five-point scale where a score of 0 indicated that symmetry was always regular, i.e., normal. A score of 1 = 25% irregular at the beginning and end of tasks; 2 = 50% irregular; 3 = 75% irregular; and 4 = 100% irregular or asymmetric. Amplitude was rated for each vocal fold on a scale where a score of 0 = normal amplitude; 1 = amplitude slightly decreased; 2 = moderately decreased; 3 = severely decreased; and 4 = no visible movement or amplitude. Mucosal waveform was rated for each vocal fold on a five-point scale. A score of 0 = normal Mucosal waveform; 1 = slightly decreased; 2 = moderately decreased; 3 = severely decreased; and 4 = no visible movement or mucosal waveform completely absent.

Tremor was rated under continuous light. Only the stroboscopic characteristics were rated under strobe condition.

Data analysis

All ratings were analyzed by frequency and magnitude (where applicable) and cross referenced with stage and year-since-onset (YSO). Frequency or incidence was a count of the number of times a score occurred for each variable. Magnitude (mean severity rating) was obtained for the strobe variables; phase closure, phase symmetry, amplitude and mucosal waveform. Tremor was scored as either absent or present and the location of the tremor was determined. A score was included if there was agreement in rating of at least 3 out of 4 judges.

Results

Tremor

The incidence of tremor data for each of the three conditions (rest, NPNL, loud) for the IPD and PPS patients are presented in Figure 1. Some of the endoscopic views unusable because images were too dark, trials were too short, or the views (which happened more frequently in the loud condition) were not obtained by the examiner. Overall, 55% of the IPD patients (n=22) had tremor (rest, NPNL, & loud conditions combined) with the primary location being vertical laryngeal tremor. Sixty-four percent of the Parkinson plus patients (n=7) had tremor (rest, NPNL, & loud conditions combined) with the arytenoid cartilages being the primary location.

In the IPD patients (Figure 1), resting tremor occurred 7 times out of 20 usable views (35%) and was observed most frequently as vertical laryngeal movement. Arytenoid cartilages and pharyngeal wall tremor were also observed, but with less frequency. Phonation tremor (NPNL) occurred 10 times out of 14 usable views (71%) and was observed most frequently as vertical laryngeal movement. Arytenoid cartilage and pharyngeal wall tremor were observed less often in the NPNL condition. Loud phonation tremor occurred 4 times out of 7 usable views (57%) and was only seen as vertical laryngeal movement. To examine trends, these cross sectional data for the IPD patients were compared with a subgroup of five cases of IPD patients. Each of the three conditions (rest, NPNL, and loud phonation) was examined for the same five patients and the incidence of tremor occurred with the same distribution. That is, phonation tremor (NPNL) occurred more frequently than loud tremor, and resting tremor occurred with a similar frequency of occurrence.

Overall, tremor was observed more frequently in the PPS patients, even though a Chi-square test was not statistically significant for the differences in incidence of tremor in the two patient groups (p=.09 for resting tremor; p=.35 for phonation tremor; and p=.57 for loud tremor). In the 7 PPS patients, tremor was observed 71% of the time at rest, 50% in the NPNL condition, and 71% in the loud phonation condition.

The incidence of tremor location with each view (the flexible and rigid scope) and each condition (rest, NPNL, loud) was combined and is shown in Figure 2. The IPD patients had a greater incidence of vertical laryngeal movement (VLM) tremor and the PPS patients had a greater incidence of arytenoid cartilage tremor. There were no instances where true vocal fold tremor was observed consistently by the judges in the IPD patient group. There were no instances were VLM tremor was observed in the PPS group.

The incidence of tremor for the age (YSO) of the disease in the IPD and PPS patients was compared. There was a trend for the PPS group to show a greater incidence of tremor (100%) early in the disease (1-5 YSO) compared to the IPD patients in the same 1-5 YSO group (48.3%).

In comparing tremor across stage of the disease for the IPD patient group, there was only one patient diagnosed in Stage I and no tremor was observed in this 51 year old male who was 3 years since diagnosis. The average age for patients diagnosed in Stage II was 65.4 years with and mean YSO of symptoms of 5.14 years. Incidence of resting tremor for these early Parkinson patients was 33%, and phonation (NPNL) tremor was 40%. There were no usable views for the loud phonation condition in the Stage II patients. Comparison to age revealed tremor gradually increasing across all conditions in the older age groups.

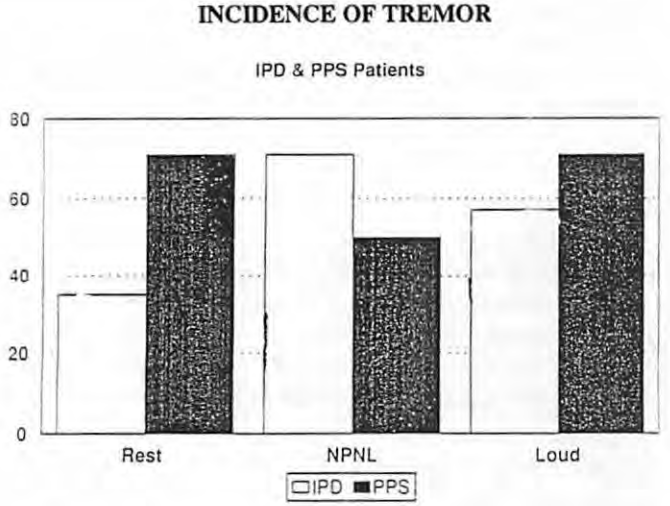


Figure 1. Incidence of tremor for IPD and PPS patient groups. Rest, NPNL and loud conditions for each scope (flexible and rigid) were combined. For IPS patients (n=22), rest tremor was 35%; NPNL tremor was 71%; and loud tremor was 57%. For PPS patients (n=7), rest tremor was 71%; NPNL tremor was 50%; and loud tremor was 71%.

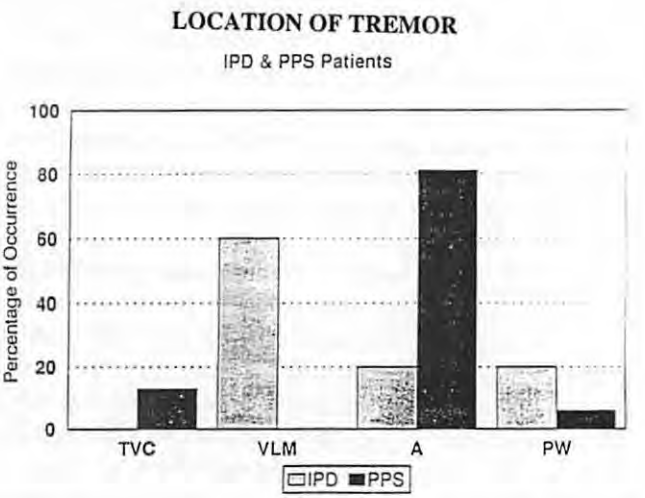


Figure 2. Location of tremor for IPD and PPS patient groups. Rest, NPNL and loud conditions for each scope (flexible and rigid) were combined. TVC=true vocal folds; VLM=vertical laryngeal movement; A=arytenoid cartilages; PW=pharyngeal walls. There was no TVC tremor for IPD patients and no VLM tremor for the PPS patients.

Phase Closure

Phase closure is reported here as the frequency of the presence of some abnormality (e.g.,incidence) and the severity of that abnormality (mild, moderate, etc.) in the IPD patient group only. Figure 3 shows the results of phase closure in the NPNL condition. Abnormal phase closure was observed 50% of the time. In rating severity, a scale ranging from -4 to +4, with a score of 0 being normal; -4 having the longest open phase; and +4 having the longest closed phase was used. The severity rating (Figure 4) for this variable was -1.99, which represents a moderately long open phase during production of /i/ at NPNL. There was a trend for the most abnormal phase closure to occur in Stage III patients.

Phase Symmetry

Asymmetry of the vibratory pattern was observed 46% (Figure 3) of the time in the NPNL condition. The severity rating (Figure 4) was 2.7, indicating that phase asymmetry was severely abnormal. A score of 0 indicated that symmetry was always regular. A score of 1 - 25% irregular at the beginning and end of tasks; 2 = 50% irregular; 3 = 75% irregular; and 4 = 100% irregular or asymmetric. The severity rating for these patients was indicative of irregular vibratory patterns during nearly 75% of the glottal cycle.

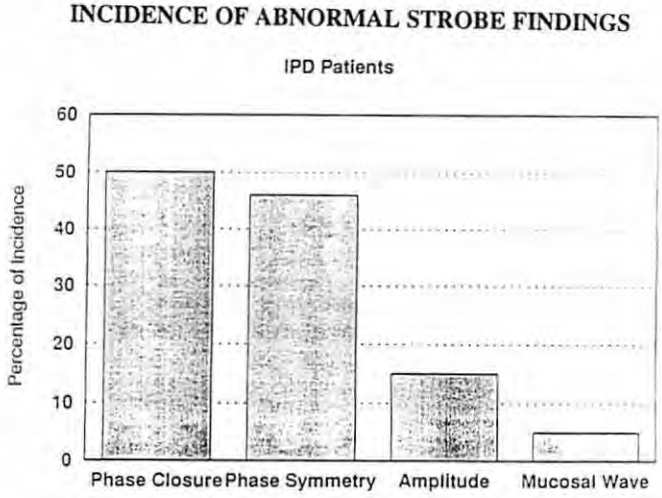
Amplitude

Amplitude was rated separately for each vocal fold, then the results were combined to report overall ratings. Although abnormal amplitude can result in an increase (12;13) or greater than normal finding, the ratings in this study are related to decreases in amplitude. The incidence of the number of abnormal findings is shown in Figure 3. Decreased amplitude of one or both vocal folds was observed (by at least 3 out of 4 judges) 15% of the time. The severity rating (Figure 4) for this abnormal finding was 1.35, indicating slightly decreased amplitude. Amplitude abnormalities were only found in Stage III patients.

Mucosal Waveform

Abnormal mucosal waveform (MW) was observed in only one patient. This was a 67 year old male, 8 years since onset in Stage II of the disease. Moderately decreased MW, a score of 2.0 on a 4 point scale was observed in this patient.

A subgroup of eight patients was compared to these cross sectional stroboscopic findings. Eight patients were selected who had complete data sets, that is, there were usable views in all conditions. These data were compared with the cross sectional data and are shown in Figure 5. The distribution reveals a similar pattern to the cross sectional data.



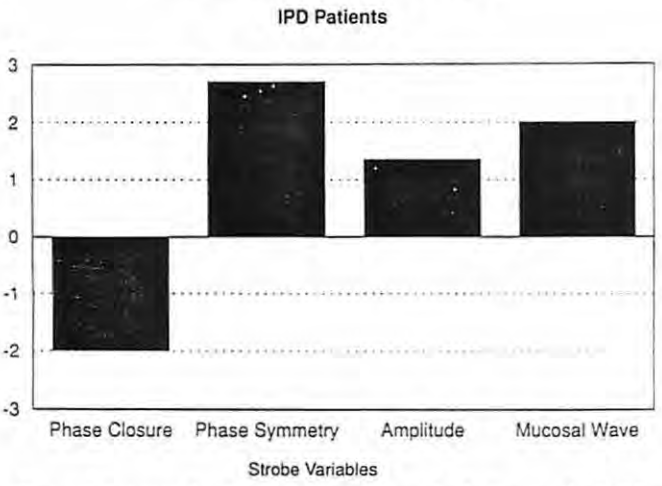


Figure 3 (upper). Incidence of abnormal strobe findings in the NPNL condition for the 22 IPD patients. Figure 4 (lower). Mean severity ratings of the strobe findings in the NPNL condition of the 22 IPD patients.

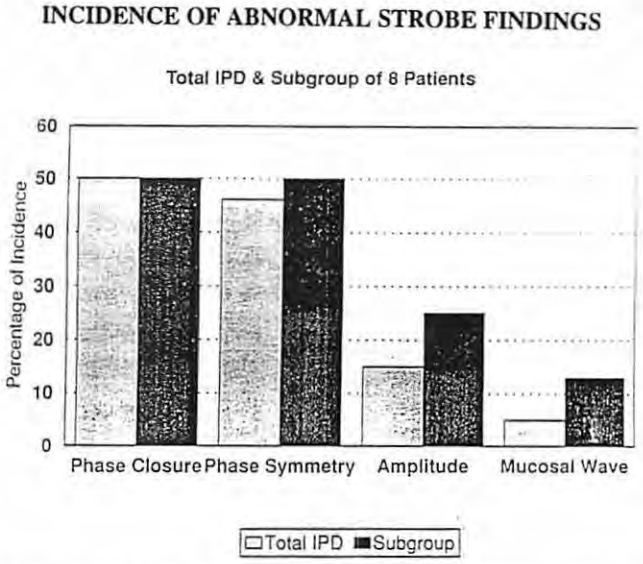


Figure 5. Incidence of abnormal strobe findings for a cross section of the 22 IPD patients and a subgroup of 8 IPD patients.

Discussion

Laryngeal tremor (incidence and location) and vocal fold vibratory characteristics were described in twenty-nine patients with Parkinson disease (22 IPD; 7 PPS) using visual-perceptual ratings of endoscopic and stroboscopic examinations. Comparisons of tremor for IPD and PPS patients were made and description of vocal fold vibratory movements in only the IPD patient group was completed. Limb tremor is widely observed in Parkinson disease patients; it was not surprising that 55% of the IPD patients and 64% of the PPS patients demonstrated laryngeal tremor. Resting limb tremor is a common early symptom in Parkinson patients with an incidence about 50% (7). In the IPD patient group, the incidence of resting laryngeal tremor was 35% and occurred most frequently as VLM. While both kinetic limb tremor and laryngeal tremor involve coordinated movements of structures and muscles to achieve particular outcomes, kinetic limb tremor has not been reported in Parkinson patients. However, there was a relatively high incidence of phonation tremor in a NPNL (71%) condition and a loud (57%) condition in the IPD patients. This is suggestive that kinetic tremor exists in the laryngeal structures of Parkinson patients. Postural laryngeal tremor was not investigated here, but given that postural limb tremor is a common finding in Parkinson patients, it would be of interest to define a postural laryngeal task and examine tremor in this condition.

Generally, the PPS patients showed more tremor than the IPD patients. In the 1-5 YSO group for the PPS patients, the mean YSO was lower (4.57 years) than the mean YSO group for the IPD patients (6.9 years). However, the incidence of tremor was greater in the first 5 years since onset of symptoms for the PPS patients. The speech and voice problems of PPS patients are more severe and deteriorate more quickly compared with IPD patients (9) and our tremor findings support this. There was a clear trend in differences in the location of tremor in these two groups. Tremor was primarily seen as vertical laryngeal movement in the IPD patients and in the PPS patients, tremor was seen primarily in the arytenoid cartilages. Often PPS patients are misdiagnosed early on as having IPD and it is only as the disease progresses that PPS syndrome is better classified (10). It may be that the presence of vocal tremor and the location of the tremor as an early symptom may help contribute to a more accurate diagnosis in PPS patients. Additional studies with larger patient numbers are necessary to substantiate this finding.

The phonatory abnormalities in Parkinson disease patients are typically related to reduced loudness, monotone, hoarseness, and breathiness. We found laryngeal characteristics consistent with expectations. A hypophonic profile was seen in the incidence and severity of the phase closure abnormalities. As reported elsewhere (4;11), vocal fold bowing and degree of glottal incompetence is relatively high in Parkinson patients. The predominantly open phase configuration observed here is consistent with those findings. Phase symmetry is a feature of the timing of the opening and closing of the vocal folds (entrainment). Changes in mass and tension may result in asymmetric phase (12;13). The high incidence of phase asymmetry in these patients was not surprising. The magnitude of phase asymmetry was the most severe of all the stroboscopic variables and may be related to phase closure abnormalities. While abnormalities in amplitude were observed, the incidence was considerably lower than phase closure or phase asymmetry. Vocal fold stiffness is often seen in the aged larynx. However, knowledge of the nature, incidence and severity of stiffness in the senile larynx is limited. Decreased amplitude was only a minor finding in the IPD patients in this study. Decreases in mucosal waveform are associated with peripheral or structural damage. Aside from the diagnosis of Parkinson disease, these patients had no other laryngeal abnormalities. Mucosal wave was within normal limits in all but one patient.

Stroboscopy as a laryngeal imaging method can provide information about the subtle changes in the vibratory motion of the vocal folds. In this study, stroboscopy was used to examine these characteristics in patients with Parkinson disease. We have demonstrated that a certain stroboscopic vibratory profile predominates in patients with IPD. Further study is needed in investigating the mechanisms of bowing, in examining stroboscopic abnormalities in a variety of tasks, in larger patient groups (IPD and PPS).

Acknowledgments

Grateful appreciation is acknowledged to Kathy Ives who spent numerous hours dubbing the master video tape and entering the data into the master spread sheet and to the patients who so generously contributed their time.

The research presented here was supported in part by NIH Grant no. RO1 DC01150 and The National Center for Voice and Speech Grant no. P60 DC00976.

References

1. Forno LS. Pathology of Parkinson's disease. In: Marsden CD & Fahn S (Eds.) *Movement Disorders.* London: Buttersworth, 1982.

2. The American Parkinson Disease Association, Inc. 60 Bay Street, Staten Island, NY 10301.

3. Ramig LO. Speech therapy for patients with Parkinson's disease. In: Koller W, & Paulson G (Eds.) *Therapy of Parkinson's disease.* Marcel Dekker, Inc., New York (in press).

4. Hanson DB, Gerratt BR, Ward PH. Cinegraphic observations of laryngeal function in Parkinson's disease. *Laryngoscope.* 1984;94:348-353.

5. Logemann JA, Fisher HB, Boshes B, Blonsky ER. Frequency and occurrences of vocal tract dysfunction in the speech of a large sample of Parkinson's disease patients. *Journal of Speech and Hearing Disorders.* 1978;42:47-57.

6. Darly FL, Aronson A, Brown J. Clusters of deviant speech dimensions in the dysarthrias. Journal of Speech and Hearing Research. 1969;12:462-496.

7. Koller WC, Huber SJ. Tremor disorders of aging: Diagnosis and management. *Geriatrics.* 1989;44(5):33-41.

8. Gresty MA, Findley LJ. Postural and resting tremor in Parkinson's disease. *Advances in Neurology.* 1984;40:361-364.

9. Quinn N. Multiple system atrophy-The nature of the beast.*Journal of Neurology, Neurosurgery, and Psychiatry.* 1989;52(suppl):78-89.

10. Countryman S, Ramig LO, Pawlas AA. Speech and voice deficits in Parkinson plus syndromes: Can they be treated? *Journal of Medical Speech-Language Pathology.* 1994;2(3):211-25.

11. Smith ME, Ramig LO, Dromey C, Perez KS, Samandari R. Intensive voice treatment in Parkinson's disease: Laryngoscopic findings. *Journal of Voice.* (in submission).

12. Hirano M, Bless D. *Videostroboscopic Exam of the Larynx.* San Diego: Singular Publishing Group, Inc., 1993.

13. Sataloff RT. Professional voice: the science and art of clinical care. New York: Raven Press, 1991.

14. Sataloff RT. The Human Voice. *Scientific American.* December, 1992.

15. Bless DM, Hirano M, Feder RJ. Videostroboscopic evaluation of the larynx. *Ear, Nose and Throat Journal,* 1987;66:363-8.

16. Hoehn M, Yahr M. Parkinsonism: Onset, progression and mortality. *Neurology,* 1967;17:427.

17. Fahn LS, Elton, & members of the Unified Parkinson Disease Rating Scale, 1987.

NCVS Status and Progress Report - 7
December 1994, 41-64

A Comparison of Two Forms of Intensive Speech Treatment for Parkinson Disease

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, Denver Center For The Performing Arts
Stefanie Countryman, M.A., CCC-SLP
The Wilbur James Gould Voice Research Center, Denver Center For The Performing Arts
Laetitia Thompson, Ph.D.
Neuropsychology Laboratory, The University of Colorado School of Medicine
Yoshiyuki Horii, Ph.D.
Department of Communication Disorders and Speech Science, The University of Colorado-Boulder
The Wilbur James Gould Voice Research Center, Denver Center For The Performing Arts

Abstract

This study investigated the effect of two forms of intensive speech treatment, respiration (R) and voice and respiration (Lee Silverman Voice Treatment [LSVT]), on the speech and voice deficits associated with Idiopathic Parkinson disease. Forty-five subjects with Idiopathic Parkinson disease completed extensive pretreatment neurological, otolaryngological, neuropsychological and speech assessments. All subjects completed 16 sessions of intensive speech treatment, 4 times a week for one month. Measures of intensity and maximum duration of sustained vowel phonation, and intensity, fundamental frequency and its variability, utterance duration and pause duration during reading of the "Rainbow Passage" and conversational monologue were made pre- and post-treatment. Perceptual measures of subject self-ratings and family ratings were completed pre- and post-treatment for the variables loudness, monotonicity, hoarseness, overall intelligibility and initiation of conversation. Significant improvements were observed following both forms of treatment. However, significant pre- to post-treatment improvements were observed for more variables and were of greater magnitude for the subjects who received the voice and respiration treatment (LSVT). Only subjects who received the LSVT rated a significant reduction in the impact of Parkinson disease on their communication post-treatment. Correlations between descriptive prognostic variables (i.e. stage of disease, speech severity rating, depression, and time since diagnosis) and magnitude of treatment-related change indicated these factors did not significantly influence treatment effectiveness. Age and cognitive ability were correlated with magnitude of treatment-related change in intensity in conversational monologue for the voice and respiration (LSVT) group. These findings suggest that intensive voice and respiration (LSVT) treatment, focusing on increased vocal fold adduction, is more effective than respiration (R) treatment for improving vocal intensity and decreasing the impact of Parkinson disease on communication.

Parkinson disease is a progressive, degenerative neurological disease resulting from nigrostriatal dopamine deficiency (Hornykiewicz, 1966; Hornykiewicz & Kish, 1986). At least seventy-five percent of the 1.5 million patients with Parkinson disease have a speech disorder which may limit their ability to function fully in society (Canter, 1965; Hoberman, 1958; Logeman, Fisher, Boshes, & Blonsky, 1978; Oxtoby, 1982; Streifler & Hofman, 1984). The characteristics of disordered speech and voice in Parkinson disease include reduced loudness, monotone, imprecise articulation and disordered rate (Darley, Aronson, & Brown, 1969 a, b; Logemann, et al., 1978). Traditionally, speech treatment for Parkinson disease has met with limited success (Allan, 1970; Aronson, 1985; Greene, 1980; Sarno, 1968; Weiner & Singer, 1989). Consequently, many Parkinson disease patients do not receive speech therapy (Mutch, Strudwick, Roy & Downie, 1986; Oxtoby, 1982) and their speech deteriorates as their disease progresses (Logemann, et al., 1978; Morley, 1955).

Table 1.
Logic for development and rationale for therapy goals and measurement variables for the Lee Silverman Voice Treatment (LSVT)

Perceptual characteristics of speech	Hypothesized laryngeal and/or respiratory physical pathology	Therapy goals and tasks	Acoustic, physiologic variables measured
Reduced loudness Breathy, weak voice	Bowed vocal folds Reduced glottal adduction Reduced respiratory support	• Increase vocal fold adduction via isometric effort (pushing, lifting) during phonation • Increase maximum duration vowel phonation at greater vocal intensity • "Think loud/shout"	• Videolaryngostroboscopic examination • Maximum phonation time • Sound pressure level
Hoarseness	Rigidity, hypokinesia, in laryngeal and/or respiratory muscles	• Improve vocal quality during sustained phonation and speaking	
Monotone	Rigid cricothyroid muscle	• Increase maximum fundamental frequency range • High/low pitch glides • Sustained phonation at highest and lowest pitches	• Variability of fundamental frequency in connected speech (STSD)

Recently Ramig and colleagues (Countryman & Ramig, 1993; Countryman, Ramig & Pawlas, 1994; Ramig, in press; Ramig, Bonitati, Lemke, & Horii, 1994; Ramig, Horii, & Bonitati, 1991) have reported positive findings following a voice treatment program for Parkinson disease (The Lee Silverman Voice Treatment [LSVT]) that focuses on increasing vocal loudness by increasing phonatory effort. Subjects have been trained to increase vocal fold adduction in maximum phonatory effort tasks and generalize the increased vocal effort to louder speech. Studies have documented significant increases in maximum duration sustained vowel phonation, maximum fundamental frequency range and fundamental frequency and fundamental frequency variability in speech following this intensive voice treatment (Ramig, et al., 1994). Perceptual reports supported corresponding post-treatment increases in vocal loudness and improvements in speech intelligibility and functional communication (Ramig, 1992). As summarized in Table 1, the logic for the LSVT is based upon perceptual characteristics of voice and the corresponding laryngeal and respiratory physical pathology that has been hypothesized or observed in patients with Parkinson disease. For example, the reduced loudness and the breathy, weak voice that is observed in Parkinson disease has been associated with glottal incompetence such as bowed vocal folds (Hanson, Gerratt, & Ward, 1984). This is the basis for the treatment goal of improving vocal fold adduction in order to increase vocal loudness. It has been suggested that the hoarse voice quality observed in Parkinson disease is a result of rigidity and/or hypokinesia in the laryngeal and respiratory system. This is the basis for the goal of increasing phonatory stability in order to improve vocal quality. The monotone voice observed in Parkinson disease may be associated with rigidity in the cricothyroid muscle (Aronson, 1985). This is the basis for the goal of increasing range of motion of the cricothyroid muscle which may contribute to improved intonation.

Initial studies (Countryman & Ramig, 1993; Ramig, 1992; Ramig, in press; Ramig, Mead, Scherer, Horii, Larson & Kohler, 1988; Ramig, et al., 1991; Ramig, et al., 1994) have identified phonatory effort and increased loudness as key elements in efficacious treatment for Parkinson disease patients. However, in order to maximize the usefulness of this treatment approach a number of issues should be clarified. Reduced vocal loudness observed in Parkinson disease has been associated with limited respiratory support (Canter, 1965; Critchley, 1981) as well as reduced vocal fold adduction (Hanson, et al., 1984; Smith, Ramig, Dromey, Perez, & Samandari, in press). Treatment directed toward increasing vocal loudness is dependent upon improving the function of one or both of these systems (Dromey, Ramig, & Johnson, in review; Ramig, Dromey & Scherer, 1993; Ramig, Fazoli, Scherer, & Bonitati, 1990; Titze & Sundberg, 1992). In order to identify the most efficient form of treatment, the present study was designed to evaluate the speech and voice characteristics of patients with Parkinson disease following one of two forms of intensive treatment: 1) treatment designed to increase respiratory support for speech (R); and 2) treatment designed to increase vocal fold adduction and respiratory support (LSVT). Clarification of the interactive roles of the respiratory and laryngeal systems in treatment-related change will contribute to enhancing the effectiveness of treatment.

Identification of patient suitability or prognostic factors for treatment success is another issue of importance in generalization of findings from studies of voice treatment

efficacy. Previous research has not considered characteristics of Parkinson disease patients such as age, sex, stage of disease, time since diagnosis, and severity of speech disorder in relation to the efficacy of intensive voice therapy. In addition, depression and dementia (cognitive ability or functioning), which are frequently observed in Parkinson disease (Duvoisin, 1984; Gauthier, L., Dalziel, & Gauthier, S., 1987; Mayeux, Williams, Stern, & Cote, 1986) and associated with performance limitations (Cummings, 1986; Beatty, Staton, Weir, et al., 1989), should be considered for their potential impact on treatment-related change. While preliminary work (Mead, Ramig, & Beck, 1989) suggested that moderate dementia is not a limiting factor for voice treatment success, neither cognitive functioning or depression has been considered systematically for its impact on the efficacy of voice treatment in Parkinson disease. Clarification of the relationships among treatment-related change and patient characteristics will contribute to maximizing treatment effectiveness.

This study was designed to document pre- to post-treatment changes in sound pressure level and duration characteristics in sustained phonation, and sound pressure level, fundamental frequency and its variability, and utterance and pause duration in reading and conversational speech in patients with Parkinson disease who received respiration (R) or voice and respiration (LSVT) treatment. Utterance and pause duration measurements were included as a secondary measurement of respiration and rate. It was assumed that, if breath patterns of each group were altered with treatment, utterance and pause duration characteristics would change as well. Treatment-related changes were considered in relation to subject characteristics including: age, sex, stage of disease, time since diagnosis, severity of speech/voice deficits, depression and dementia (cognitive functioning).

Method

Subjects

Forty-five patients with idiopathic Parkinson disease were the subjects in this study. These patients volunteered to participate in a study of the effects of speech treatment on Parkinson disease-related speech and voice deficits. All subjects were residents of the Denver, Colorado, area. All subjects participated in pretreatment neurological, neuropsychological and otolaryngological assessments. After stratification on the variables age, sex, stage of Parkinson disease (Hoehn & Yahr, 1967), score on the motor examination section (section III) of the Unified Parkinson's Disease Rating Scale (UPDRS) (Fahn, Elton, & members of the UPDRS Development Committee, 1987), time since diagnosis, magnitude of glottal incompetence and severity of speech disorder, subjects were randomly assigned to one of two treatment groups: respiration (R) or voice and respiration (LSVT).

An analysis of variance revealed no significant differences between the groups of subjects at baseline for the variables age, stage of disease, score on the motor examination section of the UPDRS, time since diagnosis, glottal incompetence and severity of speech disorder. Lack of subject compliance with pre- and post-treatment assessments and experimental and treatment protocol resulted in unequal group sizes and gender distribution. Difficulty recruiting female patients resulted in a limited female sample. Descriptive characteristics of the 45 subjects are summarized in Table 2 a, b.

In order to provide additional descriptive information on the subjects and to evaluate the relationship between magnitude of speech and voice disorder and severity of Parkinson disease, correlation coefficients were completed on these data for all 45 subjects. Severity of speech and

Table 2a.
Descriptive characteristics of subjects who received the Lee Silverman Voice Treatment (LSVT; voice and respiration)

Subject	Sex	Age	Years since diagnosis	Hoehn &Yahr Stage (1967)	UPDRS Motor Exam	Medication *†	Speech/Voice Deficits (severity rating)	Glottal Incompetence
1	F	73	10	4.0	48.0	a, c, f	reduced loudness imprecise articulation (5)	2.25
2	F	52	13	3.0	17.0	a, f, o	reduced loudness rapid rate (3)	1.50
3	F	49	9	2.0	4.0	a, d, g	hoarseness (1)	1.5
4	F	32	1	2.0	UA~	a, j	hoarseness (1)	
5	F	74	1	1.5	12.0	b, c	hoarseness tremor (2)	
Mean (F) (sd)		56.0 (17.7)	6.8 (5.5)	2.5 (1.0)	20.25 (19.3)		2.4 (1.67)	1.75 (0.4)

(CONTINUED...)

6	M	49	1	2.0	13.0	a, b, c	monotone reduced loudness (1)	1.33
7	M	76	1	2.0	20.0	a	hoarseness monotone (2)	
8	M	76	5	3.0	28.0	b, c, l	imprecise articulation reduced loudness (3)	1.75
9	M	62	3	2.0	26.0	a, c	hoarseness reduced loudness (3)	2.0
10	M	71	11	3.5	43.5	j, h	hoarseness reduced loudness (4)	1.75
11	M	50	18	2.0	19.5	a, c, d	reduced loudness monotone (2)	
12	M	68	6	3.0	40.5	a, k	reduced loudness hoarseness (3)	2.0
13	M	75	20	3.5	47.0	b, n	reduced loudness imprecise articulation (4)	1.25
14	M	69	1	3.0	32.0	none	imprecise articulation reduced loudness (2)	
15	M	61	1	2.0	16.0	a, c	reduced loudness imprecise articulation (2)	1.0
16	M	72	5	3.0	30.0	a, c, h	reduced loudness monotone (2)	1.0
17	M	67	8	3.0	16.0	a	reduced loudness imprecise articulation (3)	1.75
18	M	64	7	2.0	26.0	a,	monotone reduced loudness (2)	1.0
19	M	73	10	3.0	27.5	a, h	monotone reduced loudness (3)	
20	M	79	1	3.0	33.0	a, c	monotone imprecise articulation (3)	
21	M	62	4	3.5	35.0	k, n	aphonia hoarseness (4)	
22	M	52	18	3.0	34.0	a, c, d, e	rapid rate reduced loudness (3)	
23	M	75	7	2.0	32.0	a, c	reduced loudness hoarseness (3)	
24	M	61	4	3.0	22.5	a, c, d	imprecise articulation reduced loudness (4)	
25	M	54	15	3.5	45.5	b, i, k	reduced loudness monotone (4)	2.25
26	M	56	7	2.0	2.0	a, c, h	rapid rate reduced loudness (3)	1.0
Mean (M) (sd)		65.3 (9.3)	8.7 (10.1)	2.7 (0.6)	28.0 (11.3)		2.86 (0.85)	1.5 (0.5)
Group Mean (sd)		63.5 (11.5)	8.3 (9.3)	2.7 (0.7)	26.8 (12.7)		2.77 (1.03)	1.6 (0.45)

* Parkinson medications
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
+ Antidepressants
n=amitriptyline, o=tofranil, p=elavil -unavailable
Note. The Hoehn and Yahr Stages range from 1-5; 5 is most severe
Note. Higher scores on the UPDRS indicate greater disability. Scores range from 0 - 108.
Note. Severity ratings for speech and voice deficits (in parenthesis) are on a scale of 1-5, 1=mild, 2=mild/moderate, 3=moderate, 4=moderate/severe, 5=severe
Note. Glottal incompetence ratings were on a scale of 1-5. 1=slight, 3=moderate, 5=severe

Table 2b.
Descriptive characteristics of subjects who received respiratory treatment (R)

Subject	Sex	Age	Years since diagnosis	Hoehn &Yahr Stage (1967)	UPDRS Motor Exam	Medication * †	Speech/Voice Deficits (severity rating)	Glottal Incompetence
27	F	64	2	3.0	29.0	a, c, p	hoarseness reduced loudness (1)	1.55
28	F	71	3	1.0	UA-	a	hoarseness tremor (2)	
29	F	68	1	2.0	1.0	a, b, c	reduced loudness (1)	
30	F	68	4	2.0	12.0	a, c	reduced loudness monotone (3)	1.75
31	F	63	3	3.0	42.0	a, c	imprecise articulation slow rate (4)	2.5
32	F	51	4	1.0	16.0	none	reduced loudness (1)	1.75
33	F	83	7	3.0	33.0	a	strained/strangled rapid rate (4)	
Mean (F) (sd)		66.9 (9.6)	3.4 (1.9)	2.1 (0.9)	23.6 (14.3)		2.29 (1.38)	1.8 (0.2)
34	M	70	4	3.0	32.0	a, c	reduced loudness monotone (3)	1.0
35	M	71	4	3.0	43.0	a, n	reduced loudness monotone (4)	1.25
36	M	69	1	2.0	27.0	i	monotone reduced loudness (3)	
37	M	65	8	3.0	19.0	none	imprecise articulation rapid rate (2)	1.0
38	M	52	14	2.0	31.0	a, d, e, p	pallilalia rapid rate (3)	1.0
39	M	67	18	3.5	UA	b, c, j	imprecise articulation reduced loudness (4)	
40	M	51	6	1.5	3.0	a, c	rapid rate pallilalia (4)	
41	M	66	12	3.0	41.0	a, c, i	strained/strangled reduced loudness (3)	
42	M	81	2	2.0	24.5	a, c, n	reduced loudness (2)	1.0
43	M	69	6	2.0	36.0	a, k	reduced loudness monotone (3)	
44	M	64	10	3.0	43.0	a	imprecise articulation monotone (4)	1.5
45	M	54	3	1.0	3.0	b, c, h	reduced loudness monotone (1)	
Mean (M) (sd)		64.9 (8.7)	7.3 (5.2)	2.4 (0.8)	27.1 (15.0)		3.0 (0.95)	1.1 (0.2)

(CONTINUED...)

Group Mean (sd)	65.6 (8.9)	5.9 (4.7)	2.3 (0.8)	25.6 (14.4)	2.74 (1.15)	1.38 (0.4)

* Parkinson medications
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
+ Antidepressants
n=amitriptyline, o=tofranil, p=elavil
~unavailable
Note. The Hoehn and Yahr (1967) Stages range from 1-5; 5 is most severe
Note. Higher scores on the UPDRS indicate greater disability. Scores range from 0 - 108.
Note. Severity ratings for speech and voice deficits (in parenthesis) are on a scale of 1-5, 1=mild, 2=mild/moderate, 3=moderate, 4=moderate/severe, 5=severe
Note. Glottal incompetence ratings were on a scale of 1-5. 1=slight, 3=moderate, 5=severe

voice disorder rated on a 1-5 point scale at the time of clinical intake was correlated (Pearson product moment) with the variables time since diagnosis (r=.43, p=.003), rating on the motor examination section of the UPDRS (r=.63, p=.000), Parkinson disease stage (r=.62, p=.000), and depression rating (r=.45, p=.004). These findings suggest that subjects who had Parkinson disease longer and were more motorically impaired also had more severe speech and voice deficits.

Neurological Assessment

Routine neurological assessment was carried out on all subjects pretreatment. Stage of Parkinson disease, score on the motor examination section (section III) of the UPDRS, time since diagnosis, and anti-Parkinson medications are summarized in Table 2 a, b. Subjects did not change medications or dosage during the treatment period.

Neuropsychological Assessment

A battery of neuropsychological tests was administered within one month pretreatment to determine status of cognitive functioning in these Parkinson disease subjects. These measures focused on attention and concentration, learning and memory, vocabulary and auditory verbal comprehension, and visual spatial skills. Four subtests from the Wechsler Adult Intelligence Scale-Revised were administered (Vocabulary, Block Design, Object Assembly, Digit Symbol; Wechsler, 1981). In addition, parts A and B of the Trail Making Test (Reitan & Wolfson, 1985), the Digit Vigilance Test (Lewis & Kupke, 1977), Story and Figure Memory Test (Heaton, Chelune, & Lehman, 1978), and the Complex Ideational Material auditory comprehension subtest from the Boston Diagnostic Aphasia Exam (Goodglass & Kaplan, 1972) were included. These measures produced 13 scores, all of which were converted to T scores (mean of 50, standard deviation of 10) that correct for demographic variables of age, education and sex (Heaton, Grant, & Matthews, 1991). After completion of the cognitive evaluation, subjects were interviewed by the clinical neuropsychologist, who then rated depression using the Montgomery Asberg Depression Rating Scale (MADRS; Montgomery & Asberg, 1979). At this time subjects also completed a self-report depression scale, the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) as well as the Sickness Impact Profile (SIP; Bergner, Bobbit, Carter, & Gilson, 1981). The SIP instrument assesses the extent to which patients perceive their illness to effect 12 different areas of their lives. For the purposes of this paper, the individual scores for the impact of illness on communication and social interaction were included. Both the BDI and the SIP were completed pretreatment at the neuropsychologist's office and then posttreatment in the speech treatment room after completion of the last therapy session. Pretreatment neuropsychological data for each individual subject as well as group means are summarized in Table 3 a, b. Results of analysis of variance revealed no significant differences between treatment groups in cognitive functioning, level of depression, or sickness impact on communication or social interaction at baseline. However, a main effect for gender was revealed for the self-rating of depression [F(1,41)=7.368, p=.010] as well as clinical rating of depression [F(1,41)=9.971, p=.003], with the females (n=12) being more depressed than the males (n=33).

Otolaryngological Assessment

An otolaryngological history and videolaryngostroboscopy examination were obtained on all subjects prior to beginning therapy. Laryngeal imaging and videolaryngostroboscopy examination were conducted utilizing well described techniques (Bless, Hirano, & Feder, 1987). The nasal passage was topically anesthetized with 4% lidocaine spray. Endoscopic examination was conducted with both an Olympus ENF-P3 fiberscope and Nagashima SFT-1 70-degree rigid telescope. Light sources included both an Olympus halogen constant light source and a Nagashima LS-3A laryngoscope. The subjects were instructed to say /i/ at a comfortable pitch and loudness.

Table 3a.
Neurophysiological characteristics of subjects who received the Lee Silverman Voice Treatment (LSVT; voice and respiration)

Subject	Cognition (T score)	Beck Depression (self-rating)	Depression (clinician's rating)	SIP Communication	SIP Social Interaction
1	40.38	10	12	65.24	12.55
2	40.08	11	7	56.41	13.03
3	40.54	20	14	0.00	33.93
4	45.31	20	9	9.66	13.17
5	42.62	14	11	9.66	11.03
Mean (F)	41.79	15.0	10.6	26.19	16.74
(sd)	(2.21)	(7.28)	(2.70)	(30.21)	(9.65)
6	46.54	6	3	0.00	0.00
7	47.00	5	0	9.66	0.00
8	42.69	0	3	21.10	3.59
9	44.54	2	7	9.66	9.03
10	42.69	6	6	56.00	7.38
11	39.54	3	8	31.31	6.00
12	41.15	4	7	39.17	3.52
13	42.62	13	15	22.48	9.10
14	44.08	4	5	67.86	0.00
15	45.69	1	4	0.00	0.00
16	39.77	12	3	41.79	18.62
17	48.69	12	8	65.24	11.79
18	50.31	3	1	30.24	0.00
19	43.54	16	5	40.41	0.00
20	41.23	4	6	51.86	17.03
21	37.38	12	5	63.03	24.89
22	42.08	12	8	63.86	3.52
23	53.69	15	13	11.45	11.45
24	44.38	8	13	52.41	6.83
25	50.38	10	5	28.97	6.83
26	48.31	11	3	30.76	15.72
Mean (M)	44.59	7.57	6.10	35.11	7.40
(sd)	(4.08)	(4.87)	(3.86)	(21.75)	(7.11)
Group Mean	44.05	9.0	6.96	33.76	9.19
(sd)	(3.92)	(5.62)	(4.05)	(23.07)	(8.33)

Note. T scores below 40 are considered impaired. Scores above 9 on self-rating of depression suggest some depression. Scores of 6 or greater on clinician ratings of depression are suggestive of depression. Scores that are larger on the SIP indicate more impact of the sickness in that area. Scores of 0.00 on the SIP indicate no difficulty in that area.

Table 3b.
Neurophysiological characteristics of subjects who received respiratory treatment (R)

Subject	Cognition (T score)	Beck Depression (self-rating)	Depression (clinician's rating)	SIP Communication	SIP Social Interaction
27	53.31	10	9	0.00	0.00
28	39.31	7	7	39.17	6.07
29	48.85	12	0	0.00	0.00
30	50.00	3	7	18.48	6.48
31	40.54	25	19	73.10	55.17
32	52.77	13	12	0.00	26.76
33	40.54	12	11	55.03	9.10
Mean (F)	46.47	11.71	9.29	26.54	14.80
(sd)	(6.14)	(6.62)	(5.79)	(29.77)	(19.95)
34	49.38	11	1	0.00	3.52
35	34.46	10	3	32.55	17.59
36	50.69	11	1	9.66	3.52
37	43.85	0	3	40.41	3.52
38	43.54	22	6	52.41	31.38
39	40.31	10	3	66.21	20.41
40	46.08	3	6	9.66	11.45
41	50.08	13	11	63.86	18.59
42	49.23	5	5	9.66	20.48
43	37.54	12	2	31.31	9.59
44	33.91	8	13	52.41	7.10
45	51.92	1	5	0.00	0.00
Mean (M)	44.25	8.83	4.92	30.68	12.27
(sd)	(6.42)	(6.04)	(3.75)	(24.52)	(9.45)
Group Mean	45.07	9.90	8.53	29.15	13.20
(sd)	(6.24)	(6.31)	(4.95)	(25.83)	(13.74)

Note. T scores below 40 are considered impaired. Scores above 9 on self-rating of depression suggest some depression. Clinician ratings of depression scores of 6 or greater are suggestive of depression. Scores that are larger on the SIP indicate more impact of the sickness in that area. Scores of 0.00 on the SIP indicate no difficulty in that area.

Images were recorded with a CCD camera, using a 35-mm lens for the fiberscope and 60-mm lens for the rigid telescope, and a VHS or SVHS tape recorder. The magnitude of glottal incompetence pre-treatment is summarized in Table 2 a, b for 25 of the 45 subjects. Videostroboscopy ratings could not be obtained on 20 subjects due to their inability to tolerate the procedure, excessive involuntary movement or technical problems. Results of analysis of variance revealed no significant difference between the two treatment groups in glottal incompetence at baseline. Additional details of pretreatment (Perez, Ramig, Smith, & Dromey, in review) and pre- to post-treatment laryngological characteristics (Smith, et al., in press) have been previously reported.

Patients were excluded from participation in this study if they had a form of Parkinson disease other than idiopathic or a laryngeal pathology which would contraindicate phonatory effort treatment (e.g., vocal nodules, gastric reflux). Voice treatment efficacy data from patients with non-idiopathic Parkinsonism such as Shy-Drager syndrome, multiple system atrophy, progressive supranuclear palsy and bilateral thalamotomy are summarized elsewhere (Countryman & Ramig, 1993; Countryman, et al., 1994).

Speech Assessment

Standard speech and voice assessments were completed at the time of the first pretreatment data collection session. None of the subjects exhibited oral motor and/or speech and voice characteristics uncommon to Parkinson disease. The primary speech and/or voice characteristics and rating of severity of speech disorder for each subject determined by clinical observations are presented in Table 2 a, b.

Instrumentation

As part of a larger protocol, sound pressure level (B & K 2230 sound level meter) and microphone (AKG C410) signals were transduced and recorded onto a Sony PC-108M 8-channel digital audio tape recorder (DAT). Forced vital capacity was measured with a Collins wet spirometer (Model RS 2785).

Tasks

Experimental tasks were designed to sample a broad continuum of speech production to address both underlying mechanisms of treatment-related change as well as functional aspects of speech production. Data were collected while subjects performed the following tasks: forced vital capacity, maximum duration sustained vowel phonation, reading of the phonetically balanced "Rainbow Passage" (Fairbanks, 1960) and 25-30 seconds of conversational monologue. Because the sound level meter and microphone offer non-invasive signal transduction, measures of intensity, frequency and duration could be collected from all subjects.

In order to evaluate the subjects' impressions of treatment effectiveness and the impact of treatment on functional communication, perceptual rating scales were completed by all 45 subjects at the time of data collection. Family members of 20 of the 45 subjects completed the perceptual rating scales as well.

Procedures

Pretreatment experimental data were collected within the week before treatment was initiated. Post-treatment data were collected within the week following treatment. For assessment of intrasubject variability and reliability, sixty-one percent of the subjects repeated their pretreatment data collection and forty-eight percent of the subjects repeated their post-treatment data collection. For experimental data collection, subjects were seated in a medical examining chair located in an IAC sound-treated booth and transducers were positioned and calibrated. The sound level meter was positioned 50 cm from the lips and the microphone was positioned on a head mount 8 cm from the lips.

After two minutes of tidal breathing, forced vital capacity was measured. Subjects were asked to take their deepest breath and blow out "as hard and fast and long as you can." This task was repeated two or three times at the beginning and end of the recording session and the best performance was taken as maximum forced vital capacity. Because of motor and balance problems of some subjects, all subjects performed the task in a seated position. One subject could not perform the task.

To obtain maximum duration sustained vowel phonation, subjects were instructed to "take a deep breath and sustain 'ah' for as long as you can". A timer with a second hand was within view of the subject and he/she was encouraged to monitor his/her performance and sustain phonation for a longer duration each time. No instructions were given for loudness level. This task was repeated three or four times at the beginning of the recording session and two times at the end of each recording. Because of potential instabilities in measures of maximum performance associated with subjects' inconsistent generation of maximum effort (Kent, R., Kent, J., & Rosenbek, 1987) as well as the variability in the performance of Parkinson disease patients (Canter, 1965; King, Ramig, Lemke, & Horii, 1994), the experimenter, who collected all the data, was careful to elicit consistent maximum effort performance from each subject at each recording trial as determined by her clinical decision.

Samples of reading and conversational speech were obtained by asking subjects to read aloud the "Rainbow Passage" at a comfortable rate and loudness and generate a 25-30 second monologue on a topic of their choice.

A visual analog scale (Kempster, 1984; Schiffman, Reynolds, & Young, 1981) was used to obtain subject self-ratings as well as family member ratings of loudness, monotone, hoarseness, overall intelligibility, and the extent to which the subject initiated conversation.

Treatment

Both forms of treatment were designed to be intensive (16 sessions in one month) and high effort (subject was pushed to maximum performance levels throughout the entire duration of each session). Each type of treatment stimulated high effort tasks in repeated drills and then stimulated carryover of the drilled behavior to speech tasks. The respiratory treatment (R) trained subjects to increase respiratory effort through tasks such as maximum inhalation and exhalation, maximum /s/ and /f/ duration and exhalation with resistance (Iowa Oral Performance Instrument [IOPI]; Robin, Goel, Somodi, & Luschei, 1992). Subjects were stimulated to take frequent breaths during reading and speaking and were given visual feedback (NIMS Respigraph System PN SY03) of the excursion of their rib cage and abdomen during both sustaining of /s/ and /f/ and while reading and speaking. Subjects were encouraged to take deep breaths frequently during speaking tasks. In the respiratory treatment group, no attention was directed toward increasing vocal fold adduction or increasing vocal loudness.

The voice and respiratory treatment (LSVT) focused on increased vocal loudness through increasing vocal fold adduction. Tasks such as pushing and lifting while phonating were used to stimulate increased vocal fold adduction. Drills were carried out on tasks such as maximum duration "ah" phonation and maximum phonation frequency range. Subjects were encouraged to "think loud" during these tasks as well as during reading and speaking. Attention was directed toward the respiratory system only in the form of encouraging these subjects to take deep breaths frequently and speak "on top of the breath." However, because maximum effort phonatory tasks such as maximum duration of sustained vowel phonation also stimulate respiratory function, the treatment was considered to address both voice and respiration. Feedback for adequate loudness was given using a voice light and tape recorder. The treatment techniques and measures are summarized in Table 1. Training materials for the treatment approach are summarized elsewhere (Ramig, in press; Ramig, Countryman, & Pawlas, 1994).

The treatment intensiveness, daily homework, daily quantification of treatment variables and carryover were all stimulated equally in both treatment groups. No attention was directed toward improvement of intonation, articulation or rate in either treatment group. Two clinicians delivered the treatment to all the subjects, both clinicians administered both forms of treatment. The clinicians worked closely together to insure consistency and equivalent high effort and motivation across both forms of treatment. A typical treatment session consisted of twenty-five minutes of repeated and intensive maximum effort drills and twenty-five minutes of high effort speech production tasks moving through a hierarchy of single words, phrases, sentences, reading and conversation during the four-week period. Performance feedback was provided to subjects using a variety of instruments: tape recorder, stop watch, voice light, respigraph, IOPI, visipitch, and vocal demodulator (Winholtz & Ramig, 1992) depending upon the treatment group.

Data analysis

Sound pressure level (SPL) measures for sustained phonation, reading and the monologue were analyzed using a custom-built software program. The sound level meter signal (Bruel and Kjaer 2230) was preamplified and then digitized at 5000 samples per second into a VAX 4000/200 system computer through a 16-bit resolution DSC-200 A/D converter. The software program displayed the sound level meter signal in decibels (dB). To avoid the inclusion of non-voiced segments in the analysis, data points below 58 dB were eliminated from the signal. This was achieved by cursor marking the first and last negative-going peaks of the signal and then identifying the percentage below which non-voiced data points occurred. The percentage was input into the analysis program and the signal was displayed with the percentage cutoff line. This enabled the analyzer to determine if the correct percentage was chosen. After the correct percentage was determined, the program displayed the mean SPL of the signal.

The SPL was measured from four to six maximum duration sustained "ah" vowel phonations and one reading of the "Rainbow Passage" per recording session for all 45 subjects. Due to limited and variable utterance duration, the SPL for conversational monologue was obtained from only 29 of the 45 subjects. The average duration of these monologues was 25-30 seconds.

Duration of sustained vowel phonation was also calculated by the SPL program using the markings between the first and last negative going peaks. Duration and SPL data were obtained from the same four-six sustained vowel phonations for each subject.

To obtain measures of mean fundamental frequency and fundamental frequency variability during reading and monologue, the microphone signal was digitized at 5000 samples per second into the VAX system computer through a 16 bit resolution DSC-200 A/D converter. The files were downloaded onto a 486 computer and then analyzed using C-speech software (Milenkovic, 1987). The program calculated mean F_0 and standard deviation in Hertz. Using standard procedures, the Hertz standard deviation was then converted to express frequency variability in semitones (STSD). Two subjects in the R group had unanalyzable fundamental frequency and fundamental frequency variability data during reading due to severe dysphonia.

Utterance and pause duration data were obtained from both the "Rainbow Passage" and the 25-30-second monologue. Four subjects' data during reading of the "Rainbow Passage" was unanalyzable; thus 41 passages and 29 monologues were analyzed. The recorded voices were played back, fullwave rectified and smoothed by an RC lowpass filter (Dobkin, 1969). The rectifying and smoothing extracted the intensity envelope in real time. This intensity envelope was digitized by a 12-bit analog-to-digital converter at a rate of 1000 times per second and stored on a 386 microcomputer disk using CSRE (Canadian Speech Research Environment) software.

The digital intensity envelope was subsequently analyzed for utterance and pause durations by software developed by Horii (1983). Given a specification of the maximum amplitude threshold for pause, the minimum duration of pause, and the minimum duration of utterance, the program identified utterances and pauses. The program then printed out means and standard deviations of utterance and pause durations, number of utterances, and speaking time ratios. For the purpose of this paper, utterance and pause durations are reported. Standard procedures for analysis of visual analog scales (Boeckstyns & Backer, 1989) were used to obtain perceptual data from the subjects and families.

The data for pretreatment glottal incompetence ratings were prepared by creating a "master" study videotape that included segments from the laryngostroboscopic videotapes of each subject. The order of the subject videotapes was determined by a computer random number generator. Six subjects' recordings were repeated to assess intrajudge reliability. The audio signal from each subject tape was removed to eliminate auditory perceptual cues. Instead, a "prompter" was overdubbed onto the audio track information that identified the tasks performed during examination of each patient to allow raters to identify them. Glottal incompetence was rated during comfortable pitch, comfortable loudness /i/ as a part of a larger protocol. Four raters with experience in voice research assessed the recordings. Three were speech-language pathologists and one was an otolaryngologist. They independently and as a group viewed a practice tape, and met several times together to develop standard rating criteria. Ninety-five percent inter-rater agreement was achieved on the practice tape. Following the practice procedures, they independently viewed and rated the master tape. Twenty-four percent of the glottal incompetence ratings were reanalyzed to determine reliability. Interexaminer reliability ranged between .78 and .94 while intraexaminer reliability ranged between .87 and 1.0.

Results

Reliability

Correlation coefficients reflecting intrasubject reliability for all acoustic and perceptual variables are summarized in Table 4. Sixty-one percent of the subjects completed a second pretreatment voice recording and perceptual visual analog scale and 48% of the subjects completed a second post-treatment voice recording and visual analog scale. Reliability of perceptual ratings completed by the families are also included in Table 4. Sixty percent of the 20 families completed a second pretreatment visual analog scale and 20% completed a second post-treatment visual analog scale.

Table 4.
Intrasubject (test-retest) reliability for acoustic and perceptual variables expressed in Pearson product-moment correlation coefficients

Variables	Pre 1 to Pre 2 correlations	Post 1 to Post 2 correlations
Acoustic- (n=45)		
SPL		
sustained phonation	0.80	0.93
Rainbow Passage	0.78	0.97
Monologue	0.83	0.83
Maximum duration	0.79	0.92
STSD		
Rainbow Passage	0.95	0.91
Monologue	0.70	0.64
F_0		
Rainbow Passage	0.99	0.99
Monologue	0.96	0.96
FVC	0.98	0.98
Perceptual- Subject (n=45)		
Loudness	0.63	0.63
Monotonicity	0.52	0.71
Hoarseness	0.69	0.44
Intelligibility	0.63	0.73
Initiate Conversation	0.81	0.80
Perceptual- Family (n=20)		
Loudness	0.60	1.00
Monotonicity	0.89	1.00
Hoarseness	0.66	0.92
Intelligibility	0.59	0.99
Initiate Conversation	0.93	0.87

Note. These correlations are based upon the following sample sizes: 61% of the 45 subjects repeated the pre-treatment and 48% of the subjects repeated the post-treatment acoustic and perceptual measures and 60% of the 20 families repeated the pre-treatment and 20% of the 20 families repeated the post-treatment perceptual ratings.

Table 5.
Sample sizes (n) for groups and genders for the acoustic and perceptual variables

		LSVT			R	
Variable	Group	Males	Females	Group	Males	Females
SPL						
Vowel	26	21	5	19	12	7
Rainbow	26	21	5	19	12	7
Conversation	16	11	5	13	8	5
Maximum Duration	26	21	5	19	12	7
STSD/F_0						
Rainbow	26	21	5	17	12	5
Conversation	16	12	4	13	10	3
Rate Data						
Rainbow	24	19	5	17	11	6
Conversation	16	11	5	13	9	4
FVC	25	20	5	19	12	7
Perceptual						
Subjects	26	21	5	19	12	7
Families	12	8	4	8	5	3

Twenty percent of the data were reanalyzed to determine intra-examiner reliability. Correlation coefficients were 0.99 for all measures of SPL, STSD and F_0 of the "Rainbow Passage," F_0 during conversation, and maximum duration of sustained vowel phonations and 0.92 for STSD in conversational monologue. Twenty percent of the perceptual data from the subjects' self-ratings as well as the families' perceptual ratings were reanalyzed to determine intra-examiner reliability. Correlation coefficients were 0.99 for all categories of the visual analog scale.

Acoustic Variables

An analysis of variance carried out to determine pretreatment comparability between the two groups on acoustic and perceptual variables revealed no significant differences. In addition, there were no statistically significant differences on these acoustic and perceptual variables between the two groups of males or the two groups of females at baseline.

Table 5 summarizes the 'n' of subjects for each treatment for each variable. A three factor (treatment group X gender X time) repeated measures analysis of variance (ANOVA) was used to evaluate the statistical significance of pre- to post-treatment changes. F and p values for the acoustic variables are reported in Table 6. All means and standard deviations of each acoustic variable for each group and gender pre- to post-treatment are summarized in Table 7.

SPL- Vowels

Statistical analysis of SPL during sustained vowel phonation revealed a significant time (pre- to post-treatment) by treatment group (R or LSVT) by gender interaction [$F(1,41)=7.37, p=.01$] (Table 6). As is shown in Table 7 and Figure 1a, both the males and females who received the LSVT improved post-treatment. However, the magnitude of improvement of the males in the LSVT group was statistically significant. Overall, neither the males nor females who received the R treatment increased significantly in SPL in sustained vowel phonation from the pre- to post-treatment condition.

SPL- Reading and Conversation

Statistical analysis for SPL during reading revealed a significant time by treatment group by gender interaction [$F(1,41)=4.85, p=.033$]. Both the males and females who received the LSVT improved pre- to post-treatment (Table 6). However, as is shown in Table 7 and Figure 1b, the males in the LSVT group made the greatest improvement post-treatment. Although female subjects in both treatment groups (Figure 1b) increased intensity, no females who received the LSVT decreased SPL during the "Rainbow Passage" as was the case for two of the females who received R treatment (Figure 2a). The males who received R treat-

Table 6.
F and $_p$ values for the acoustic variables

Dependent variables	Between Subject Effect: Group		Gender		Group x Gender		Within Subject Effect: Time		Time x Group		Time x Gender		Time x Group x Gender	
	F	p	F	p	F	p	F	p	F	p	F	p	F	p
SPL														
Sustained phonation	15.55	.000	0.49	.486	0.66	.421	43.63	.000	53.72	.000	0.11	.738	7.37	.01
Rainbow Passage	4.88	.033	5.30	.027	4.62	.037	29.41	.000	4.70	.036	1.62	.210	4.85	.033
Monologue	0.39	.537	3.65	.067	1.14	.296	14.02	.001	3.27	.083	3.67	.067	1.90	.180
Maximum Duration	1.68	.202	1.49	.230	.85	.361	9.29	.004	.63	.431	.51	.480	.48	.493
F_0														
Rainbow Passage	6.49	.015	86.41	.000	1.35	.252	35.91	.000	1.38	.247	1.32	.258	1.57	.218
STSD														
Rainbow Passage	2.37	.131	1.36	.251	.93	.342	40.57	.000	1.50	.229	.01	.925	2.30	.138
F_0														
Monologue	.88	.358	85.92	.000	1.39	.250	3.62	.069	2.97	.097	.61	.443	.03	.856
STSD														
Monologue	.22	.645	.50	.484	.32	.578	1.02	.321	8.48	.007	.43	.516	.94	.340
Utterance Duration														
Rainbow Passage	.07	786	60	.442	.90	.350	19.30	.000	.16	.696	.03	.864	.03	.875
Monologue	1.36	.254	8.90	.006	4.47	.044	2.23	.147	1.84	.186	.45	.506	.07	.793
Pause Duration														
Rainbow Passage	2.97	.093	1.67	.204	4.83	.034	30.15	.000	6.61	.014	.85	.364	2.79	.103
Monologue	2.02	.167	.29	.593	1.06	.313	1.87	.183	3.02	.093	.65	.427	3.06	.091
FVC	.64	.429	16.98	.000	.20	.657	.26	.615	.03	.858	.01	.927	1.33	.255

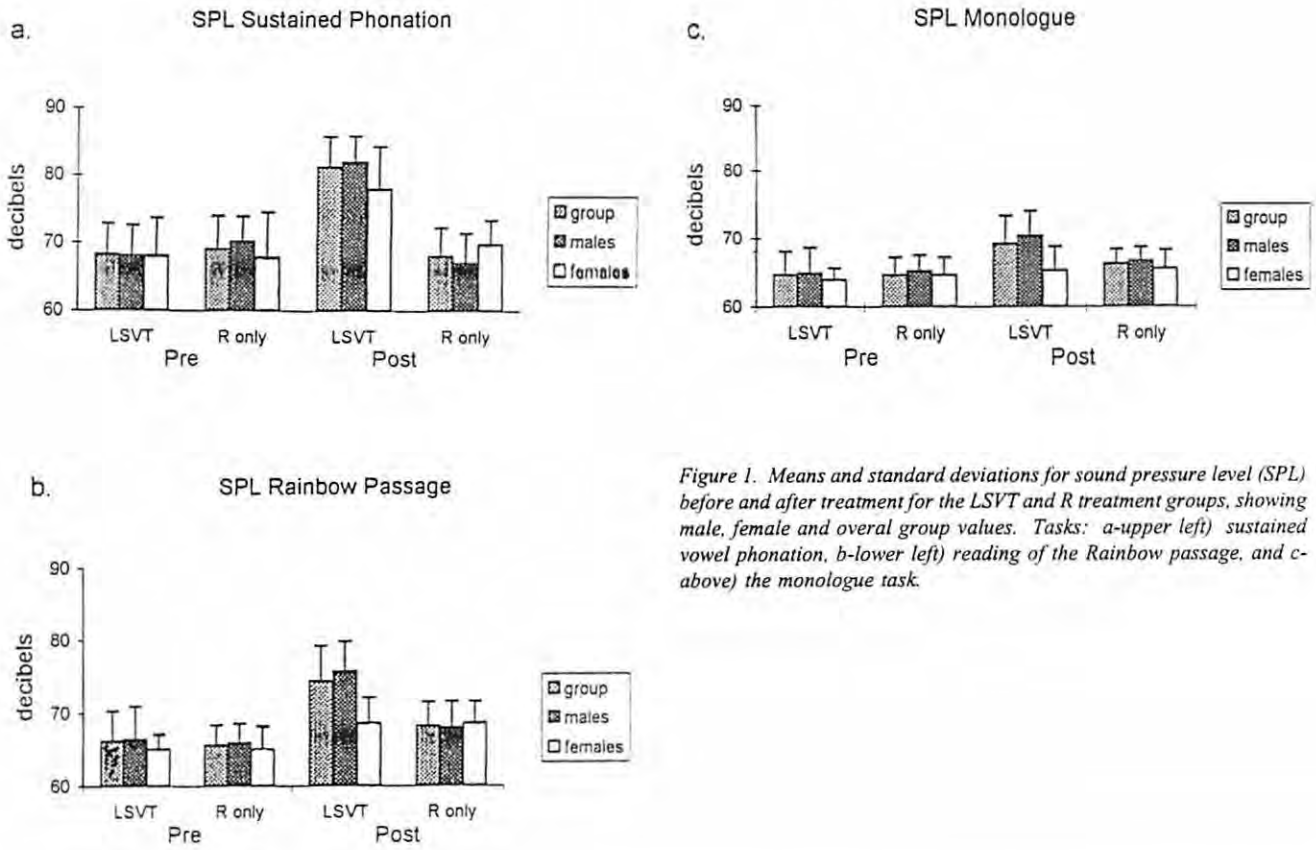


Figure 1. Means and standard deviations for sound pressure level (SPL) before and after treatment for the LSVT and R treatment groups, showing male, female and overal group values. Tasks: a-upper left) sustained vowel phonation, b-lower left) reading of the Rainbow passage, and c-above) the monologue task.

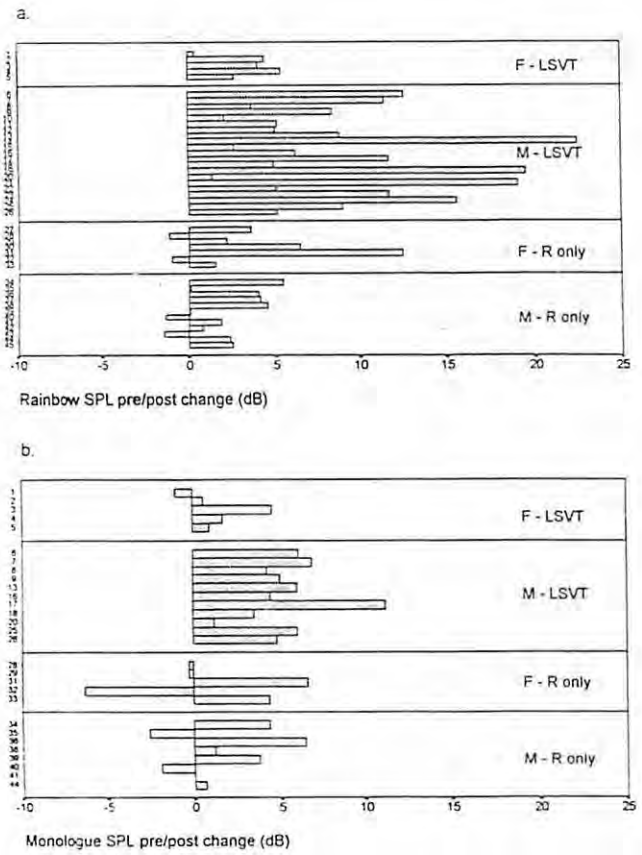


Figure 2. Individual subject values for the difference in sound pressure level (SPL) before and after treatment. A positive value represents an increase in SPL following treatment. Tasks: a) reading of the Rainbow passage and b) the monologue task. F = female, M = male. The y-axis shows the subject number.

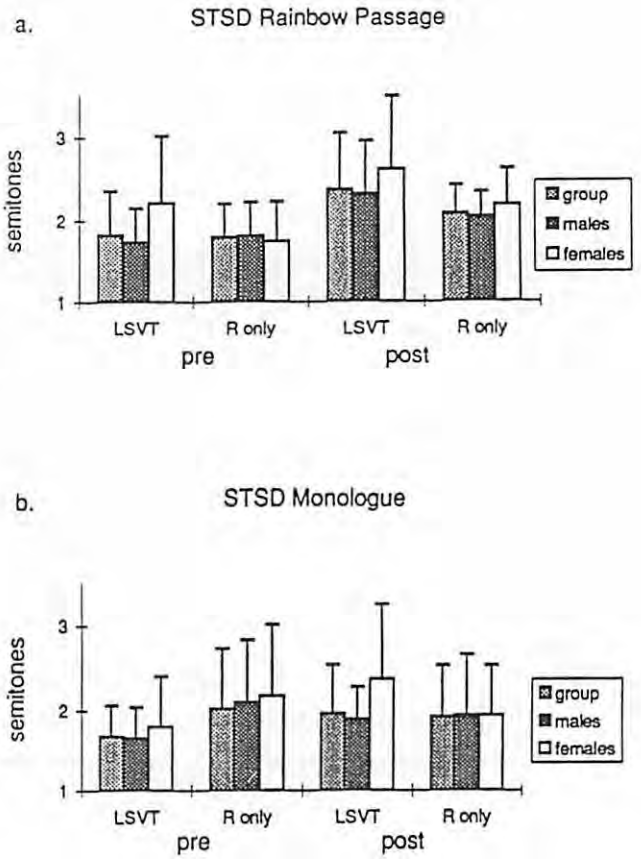


Figure 3. Means and standard deviations for semitone standard deviation (STSD) before and after treatment for the LSVT and R treatment groups, showing male, female and overall group values. Tasks: a) reading of the Rainbow passage and b) the monologue task.

ment (Figure 2a) exhibited smaller increases in SPL during the "Rainbow Passage" post-treatment than the males who received the LSVT.

Statistical analysis for SPL during conversational monologue revealed a significant main effect for time [$F(1,25)= 14.02$, $p=.001$]. There were no significant interactions between treatment group or gender (Table 6). Although subjects from both treatment groups increased intensity significantly during conversational monologue post-treatment, males who received the LSVT (Figure 1c and Figure 2b) had the greatest and most consistent increase in intensity.

Maximum duration of sustained vowel phonation

Statistical analysis of maximum duration of sustained vowel phonation revealed a statistically significant main effect for time [$F(1,41)=9.29$, $p= .004$]. There were no significant interactions between treatment group or gender (Table 6). As can be seen in Table 7, males and females from both treatment groups significantly increased maximum duration of sustained vowel phonation post-treatment. However, the females who received the LSVT had the greatest increase.

Fundamental frequency (Mean F_0) and variability (STSD)- Reading and Conversation

Statistical analysis of mean fundamental frequency while reading revealed a significant main effect for time [$F(1,39)=35.91$, $p=.000$]. There were no significant interactions (Table 6). Males and females from both treatment groups significantly increased their fundamental frequency post-treatment (Table 7). Statistical analysis of corresponding fundamental frequency variability (STSD) while reading the "Rainbow Passage" revealed a significant main effect for time [$F(1,39)=40.57$, $p=.000$]. Again, there were no significant interactions (Table 6). As can be seen in Figure 3a and Table 7, male and female subjects from both treatment groups significantly increased STSD post-treatment. However, Figure 4a demonstrates that the males who received the LSVT made the greatest and most consistent

Table 7.

Means and standard deviations (in parenthesis) for the acoustic variables collected pre to post-treatment for subjects who were in the LSVT or R treatment group. Means are listed for the entire group as well as separately for each gender.

Variables	LSVT		R	
	Pre-treatment	Post-treatment	Pre-treatment	Post-treatment
Intensity (dB)				
Sustained phonation				
Group	68.20 (4.70)	81.20 (4.65)	69.32 (5.05)	68.01 (4.31)
Males	68.03 (4.60)	81.99 (3.93)	70.23 (3.75)	66.99 (4.49)
Females	68.02 (5.70)	77.91 (6.40)	67.76 (6.80)	69.75 (3.64)
Rainbow Passage				
Group	66.30 (4.02)	74.33 (4.83)	65.68 (2.68)	68.16 (3.33)
Males	66.56 (4.37)	75.69 (4.06)	66.02 (2.53)	67.94 (3.62)
Females	65.19 (1.96)	68.58 (3.51)	65.11 (3.05)	68.56 (3.00)
Monologue				
Group	64.67 (3.37)	69.17 (4.10)	64.66 (2.58)	66.24 (2.11)
Males	64.94 (3.66)	70.35 (3.61)	65.19 (2.40)	66.70 (2.01)
Females	63.96 (1.66)	65.34 (3.46)	64.66 (2.58)	65.52 (2.77)
Max duration (sec)				
Sustained phonation				
Group	21.12 (7.63)	24.84 (10.35)	18.39 (9.65)	21.20 (8.78)
Males	21.62 (8.17)	24.67 (11.44)	20.67 (11.09)	23.47 (7.93)
Females	19.00 (4.77)	25.57 (3.79)	14.47 (5.10)	17.32 (9.37)
Fo (Hz)				
Rainbow Passage				
Males	124.44 (21.20)	140.78 (23.66)	120.13 (18.72)	125.57 (21.45)
Females	198.35 (12.13)	214.23 (13.56)	172.03 (10.91)	188.26 (17.58)
Fo variability (STSD)				
Rainbow Passage				
Group	1.83 (0.53)	2.38 (0.67)	1.80 (0.41)	2.09 (0.34)
Males	1.74 (0.42)	2.33 (0.63)	1.82 (0.41)	2.05 (0.30)
Females	2.22 (0.80)	2.62 (0.87)	1.75 (0.49)	2.19 (0.44)
F_O (Hz)				
Monologue				
Males	114.33 (13.19)	120.64 (16.81)	119.80 (21.85)	118.81 (18.62)
Females	190.52 (20.71)	201.40 (15.34)	179.15 (15.43)	180.99 (25.26)
Fo variability (STSD)				
Monologue				
Group	1.68 (0.39)	1.97 (0.58)	2.03 (0.71)	1.93 (0.61)
Males	1.66 (0.39)	1.90 (0.39)	2.11 (0.73)	1.95 (0.72)
Females	1.80 (0.61)	2.38 (0.88)	2.18 (0.84)	1.95 (0.59)
Utterance duration (msec)				
Rainbow Passage				
Group	2665.88 (666.03)	2186.24 (511.62)	2524.08 (530.33)	2132.55 (396.94)
Males	2671.59 (741.40)	2192.50 (512.17)	2409.05 (421.21)	2040.95 (361.57)
Females	2644.19 (276.33)	2162.44 (568.97)	2734.97 (680.22)	2300.48 (436.48)
Monologue				
Group	1943.19 (389.51)	1948.66 (672.05)	2165.12 (851.82)	1872.51 (569.48)
Males	1883.38 (346.72)	1915.60 (661.21)	1829.16 (582.93)	1604.47 (299.41)
Females	2098.69 (492.24)	2034.62 (771.47)	2921.04 (946.43)	2475.60 (598.07)
Pause duration (msec)				
Rainbow Passage				
Group	568.93 (175.95)	720.49 (198.89)	514.67 (138.15)	844.11 (342.80)
Males	605.95 (178.80)	771.64 (190.17)	536.54 (146.02)	783.43 (268.57)
Females	428.24 (53.83)	526.13 (71.22)	474.58 (124.09)	955.36 (456.97)
Monologue				
Group	683.64 (384.53)	746.83 (329.99)	714.79 (297.50)	941.33 (406.46)
Males	536.54 (146.02)	783.43 (268.57)	721.49 (242.41)	900.16 (332.84)
Females	705.50 (609.63)	463.83 (129.72)	699.72 (443.71)	1033.89 (590.71)
FVC (ml)				
Males	3500.00 (720.38)	3535.00 (672.99)	3458.30 (1015.80)	3391.70 (1014.90)
Females	2580.00 (601.66)	2520.00 (554.08)	2214.30 (533.63)	2228.60 (590.80)

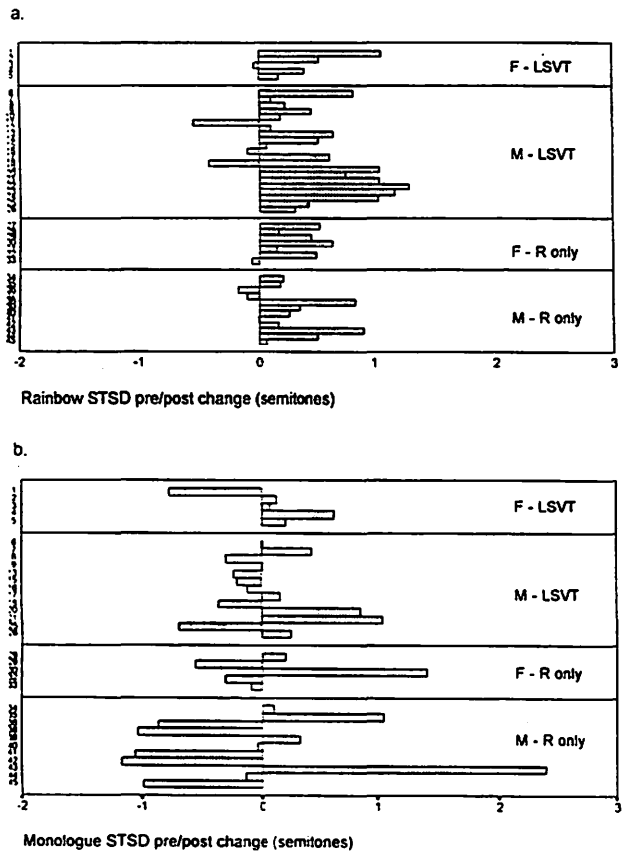


Figure 4. Individual subject values for the difference in semitone standard deviation (STSD) before and after treatment. A positive value represents an increase in STSD following treatment. Tasks: a) reading of the Rainbow passage and b) the monologue task. F = female, M = male. The y-axis shows the subject number.

gains while the females from both forms of treatment had approximately equal gains.

Statistical analysis of fundamental frequency during the conversational monologue revealed no significant changes pre- to post-treatment for either gender or treatment group. However, examination of the means in Table 7 reveals that both male and female subjects who received the LSVT increased somewhat in their fundamental frequency post-treatment, while male and female subjects who received the R treatment had no change. Statistical analysis of corresponding fundamental frequency variability (STSD) during the monologue revealed a significant time by treatment group interaction [F(1,25)=8.48, p=.007] (Table 6). As seen in Figure 3b, the subjects who received the LSVT significantly increased their STSD during monologue post-treatment while the subjects who received the R treatment decreased their STSD during monologue post-treatment. Despite the significant finding, Figure 4b demonstrates variability in the magnitude of change across treatment groups.

Rate

Statistical analysis for utterance duration during reading revealed a significant main effect for time [F(1,37)=19.30, p= .000]. As is shown in Table 7, both males and females in both treatment groups significantly reduced their utterance duration during reading relatively equally post-treatment. Statistical analysis for utterance duration during monologue revealed no significant main effect or interactions pre- to post-treatment. Although the results were not significant, both male and female subjects who received the R treatment decreased utterance duration during monologue (Table 7) post-treatment in contrast to subjects who received the LSVT. The females who received the R treatment had the greatest decrease on this measure.

Statistical analysis for pause duration during reading of the "Rainbow Passage" revealed a significant time by treatment group interaction [F(1,37)=6.61, p=.014]. As presented in Table 7, while both groups increased on this measure, only subjects who received the R treatment demonstrated a significant pre- to post-treatment increase. It should be pointed out that outlying data from one female subject in the LSVT group (n=4) has an effect on the mean for this variable. Statistical analysis for pause duration during the monologue revealed no significant main effect or interactions pre- to post-treatment (Table 6).

Forced Vital Capacity (FVC)

FVC data are summarized in Table 7. Statistical analysis revealed no significant main effect or interactions.

Perceptual Variables

F and p values for the analysis of variance carried out on the perceptual variables are presented in Table 8. Means and standard deviations for the perceptual variables for subject self-ratings are presented in Table 9. Statistical analysis for self-ratings of loudness revealed a significant time by treatment group by gender interaction [F(1,41)=4.17, p=.048]. Male and female subjects from both treatment groups rated an improvement in loudness post-treatment; however, the males who received the LSVT and the females who received the R treatment rated significant increases. Statistical analysis for the subjects' self-ratings of monotonicity revealed a significant main effect for time [F(1,41)=11.82, p=.001]. Both male and female subjects from both treatment groups rated significant improvement on this measure post-treatment (Table 9); the males who received the LSVT rated the greatest improvement. Statistical analysis for the subjects' self-ratings of hoarseness revealed a significant main effect for time [F(1,41)=11.64, p=.001]. Table 9 demonstrates that post-treatment improvement was comparable for both groups on this variable.

Statistical analysis of the subjects' self-rating of overall intelligibility revealed a significant main effect for time [F(1,41)=11.92, p=.001]. Both male and female sub-

Table 8.
F and p values for the perceptual variables

Dependent variables	Between Subject Effect: Group		Gender		Group x Gender		Within Subject Effect: Time		Time x Group		Time x Gender		Time x Group x Gender	
	F	p	F	p	F	p	F	p	F	p	F	p	F	p
Subject self-rating (n=45)														
Loudness	.03	.872	.06	.805	1.51	.226	19.26	.000	.10	.751	.42	.520	4.17	.048
Monotonicity	.01	.922	2.16	.149	2.94	.094	11.82	.001	.17	.686	1.87	.179	1.59	.215
Hoarseness	.22	.645	.51	.481	3.37	.073	11.64	.001	.39	.535	.00	.955	.04	.843
Intelligibility	.00	.978	.15	.697	.67	.416	11.92	.001	.02	.901	.21	.653	.67	.417
Initiate Conversation	.07	.791	4.62	.037	.72	.401	10.39	.002	.13	.718	.45	.506	3.98	.053
Family Ratings (n=20)														
Loudness	.50	.490	.28	.602	.53	.476	5.73	.029	.09	.773	.03	.867	.20	.660
Monotonicity	.56	.464	2.14	.163	.12	.735	1.50	.239	.15	.707	1.23	.283	.74	.403
Hoarseness	.01	.925	2.27	.151	1.81	.197	.44	.517	.28	.602	.60	.450	1.15	.299
Intelligibility	.40	.538	.87	.366	.62	.442	15.81	.001	6.74	.019	.10	.752	1.56	.230
Initiate Conversation	2.80	.114	8.66	.010	.22	.645	1.24	.282	2.62	.125	.07	.791	2.42	.139

Table 9.
Perceptual variables for the subjects' self-ratings by group and gender as well as the families' rating for both groups. Standard deviations are presented (in parenthesis).

Variables	LSVT: Pre-treatment		LSVT: Post-treatment		R: Pre-treatment		R: Post-treatment	
Loudness								
Group	46.85	(20.23)	64.15	(20.89)	51.37	(29.01)	65.42	(19.88)
Males	43.95	(18.21)	63.14	(18.59)	57.33	(27.83)	64.42	(23.33)
Females	59.00	(25.93)	68.40	(31.17)	41.14	(30.20)	67.14	(13.47)
Family (n=20)	45.50	(22.63)	60.50	(18.47)	49.73	(28.64)	63.12	(21.84)
Monotonicity								
Group	47.00	(22.54)	66.73	(16.28)	57.00	(26.13)	68.05	(20.67)
Males	41.38	(18.45)	64.76	(14.74)	57.42	(29.23)	68.75	(22.46)
Females	70.60	(24.69)	75.00	(21.56)	56.29	(21.91)	66.86	(19.44)
Family (n=20)	52.67	(27.49)	66.75	(19.34)	57.91	(27.03)	68.88	(21.92)
Hoarseness								
Group	55.65	(21.89)	70.11	(16.42)	64.58	(24.14)	75.11	(21.61)
Males	54.52	(23.13)	68.62	(17.66)	70.08	(23.03)	81.00	(23.30)
Females	60.40	(16.92)	76.40	(8.02)	55.14	(24.73)	65.00	(14.75)
Family (n=20)	61.58	(27.10)	66.58	(19.67)	70.45	(18.06)	70.38	(18.75)
Intelligibility								
Group	57.12	(20.30)	71.12	(16.77)	61.63	(20.20)	72.00	(19.99)
Males	54.95	(20.11)	70.52	(15.63)	63.00	(21.17)	72.50	(21.65)
Females	66.20	(20.62)	73.60	(22.96)	59.29	(19.81)	71.14	(18.39)
Family (n=20)	49.50	(17.93)	67.33	(19.59)	57.55	(22.81)	64.63	(17.94)
Initiate conversation								
Group	50.96	(22.37)	66.85	(16.39)	61.26	(22.06)	68.89	(20.83)
Males	45.90	(18.97)	64.90	(15.07)	59.83	(18.17)	64.50	(22.68)
Females	72.20	(25.10)	75.00	(21.00)	63.71	(29.03)	76.43	(15.91)
Family (n=20)	54.67	(20.70)	70.33	(16.70)	67.09	(28.55)	71.50	(18.41)

Note: Perceptual characteristics of loudness, hoarseness, intelligibility and initiation of conversation are on a scale of 0-100 percent with 0% the most severe and 100% the least severe.

jects from both treatment groups rated significant improvement in overall intelligibility post-treatment. Overall, the males who received the LSVT rated the greatest increase in intelligibility post-treatment. Statistical analysis for the subjects' self-ratings of initiating conversation revealed a significant main effect for time [F(1,41)=10.39, p=.002]. For this variable, subjects from both treatment groups rated significant improvement post-treatment. However, the males who received the LSVT rated the greatest improvement.

Family ratings (n=20) of pre- to post-treatment perceptual variables are presented in Table 9. Family ratings of loudness revealed a significant main effect for time [F(1,16)=5.73, p=.029]. As is seen in Table 9, families of subjects from both treatment groups equally rated significant improvement in loudness. Although families of subjects from both treatment groups rated improvement on monotone and families of the LSVT subjects rated improvement on hoarseness, there were no significant main effects or interactions pre- to post-treatment on these measures. Family ratings of overall intelligibility revealed a significant time by treatment group interaction [F(1,16)=6.74, p=.019]. Although families of subjects from both treatment groups rated improvement on overall intelligibility pre- to post-treatment, the families of the subjects who received the LSVT rated statistically significant improvement. For the families' ratings of initiating conversation, statistical analysis revealed no significant changes pre- to post-treatment. However, the families of the subjects who received the LSVT rated the greatest increase on this variable.

Neuropsychological Variables of Self-rated Depression and Sickness Impact

The Beck Depression Inventory (BDI: subject self-rating of depression) and the Sickness Impact Profile on communication and social interaction were completed post-treatment. The F and p values for the analysis of variance are presented in Table 10. Pre- and post-treatment means and sds are presented in Table 11. These means are based upon only the subjects who completed both a pre- and post-treatment BDI and SIP. Statistical analysis revealed no significant main effect or interactions for the BDI (n=35) pre- to post-treatment. Statistical analysis of the SIP on communication (n=35) revealed a significant time by treatment group interaction [F(1,31)=6.86, p=.014]. As is shown in Figure 5, only subjects who received the LSVT rated a significant reduction in the impact of their sickness on their communication post-treatment. Statistical analysis of the SIP on social interaction (n=35) revealed no significant main effect or interactions pre- to post-treatment.

Table 10.
F and $_p$ values for the neuropsychological variables.

Dependent variables	Between Subject Effect						Within Subject Effect							
	Group		Gender		Group x Gender		Time		Time x Group		Time x Gender		Time x Group x Gender	
	F	p	F	p	F	p	F	p	F	p	F	p	F	p
BDI (self-rating of depression)	.81	.374	9.17	.005	6.64	.015	.45	.508	1.04	.316	.30	.589	.65	.427
SIP Communication	.02	.886	2.19	.149	1.65	.209	10.95	.002	6.86	.014	.04	.839	.11	.741
SIP Social interaction	.14	.711	.08	.776	4.72	.038	2.36	.135	.23	.635	1.20	.282	.40	.533

Table 11.
Means and sd's (in parenthesis) for the neuropsychological variables completed by subjects both pre- and post-treatment.

Scale	LSVT		R	
	Pretreatment	Post-treatment	Pretreatment	Post-treatment
BDI (self-rating of depression)				
Group	8.3 (6.0)	7.4 (4.6)	8.7 (5.5)	8.9 (5.3)
Male	6.1 (4.4)	5.8 (3.3)	8.5 (6.3)	8.6 (6.2)
Female	16.3 (4.5)	13.3 (3.9)	9.0 (4.1)	9.6 (2.7)
SIP Communication				
Group	28.9 (21.8)	11.9 (11.7)	25.6 (25.1)	23.1 (23.6)
Male	29.2 (19.6)	12.4 (11.5)	32.7 (26.1)	29.1 (26.2)
Female	28.2 (30.2)	10.6 (13.8)	11.5 (17.4)	11.1 (12.0)
SIP Social Interaction				
Group	9.6 (8.1)	7.9 (4.4)	11.0 (10.4)	9.7 (11.7)
Male	7.2 (6.3)	6.8 (4.4)	12.6 (10.4)	11.8 (13.2)
Female	16.7 (9.6)	11.5 (2.4)	7.9 (11.0)	5.7 (7.5)

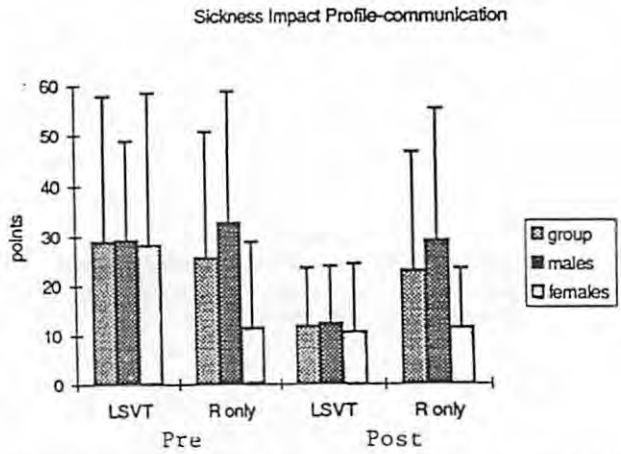


Figure 5. Means and standard deviations for the Sickness Impact Profile for Communication (SIP-C) before and after treatment for the LSVT and R treatment groups, showing male, female and overall group values.

Correlations to Treatment-related Change

The relationships between magnitude of treatment-related change and patient suitability or prognostic factors for treatment success were examined. Subject pre- and post-treatment change in SPL during the "Rainbow Passage" and conversational monologue were correlated (Pearson product moment) with pretreatment age, stage of disease, rating on the motor examination section of the UPDRS, time since diagnosis, severity of speech disorder, glottal incompetence, dementia (cognitive ability), and depression.

All 45 subjects

There were no significant correlations for the group as a whole between pre- to post-treatment change in SPL and age, stage of disease, rating on the motor examination section of the UPDRS, time since diagnosis, severity of speech disorder, glottal incompetence, cognitive ability, and depression variables or subject ratings of the impact of their disease on communication and social interaction.

All Males (n=33)

No significant correlations were observed for male subjects between prognostic variables and magnitude of treatment-related change.

All Females (n=12)

For the entire group of females there was a significant correlation between pretreatment self-rating of depression and magnitude of pre- to post-treatment change in SPL in monologue (r=.63, p=.039) and magnitude of pre- to post-treatment change in SPL "Rainbow Passage" (r=.56, p=.045). Female subjects who rated themselves more depressed pretreatment had greater post-treatment increases in SPL in monologue and the "Rainbow Passage" than subjects with less pretreatment self-rated depression. There were no other significant correlations between magnitude of treatment-related change and prognostic variables for the females.

LSVT Subjects (n=26)

There was a significant correlation (r=.56, p=.023) between the magnitude of pre- to post-treatment change in SPL in monologue and mean T score pretreatment. Subjects who had greater cognitive ability pretreatment had greater post-treatment increases in SPL in monologue than subjects with lower cognitive abilities. For the males only (n=21) in the LSVT group there were no significant correlations between magnitude of treatment-related change in SPL and prognostic variables. For the females only (n=5) there was a significant negative correlation (r=-.88, p=.05) between magnitude of pre- to post-treatment change in SPL on the "Rainbow Passage" and chronological age. Younger female subjects had greater increases in SPL than did older female subjects.

R Subjects (n=19)

There were no significant correlations for this group as a whole or for the males only (n=12) between magnitude of treatment-related change in SPL and prognostic variables. For the females only (n=7) there was a significant correlation (r=.99, p=.045) between magnitude of pre- to post-treatment SPL change in monologue and depression. Female R subjects who were more depressed pretreatment had greater increases in SPL than did the less depressed female subjects.

Discussion

This study was designed to compare the efficacy of two forms of intensive speech treatment (respiration [R] and voice and respiration [Lee Silverman Voice Treatment; LSVT]) on the speech and voice deficits in subjects with Idiopathic Parkinson disease. The study also examined the magnitude of treatment-related change relative to subject suitability or prognostic factors (age, stage of disease, time since diagnosis, severity of speech disorder, glottal incompetence, depression and cognitive ability) for treatment success.

Findings from this study support significant pre- to post-treatment changes for subjects from both treatment groups on the following variables: SPL during conversation, maximum duration of sustained vowel phonation, fundamental frequency and its variability during reading, utterance duration during reading, subject self-ratings of monotonicity, overall intelligibility, and initiation of conversation, and family ratings of loudness. Although subjects from both treatment groups improved on these variables, subjects who received the LSVT (particularly the males) (Table 7, Figures 1, 2, 3, 4) had a greater magnitude of treatment-related change than did the subjects who received the R treatment. The LSVT subjects alone had significant pre- to post-treatment improvements in SPL for sustained phonation and reading, fundamental frequency variability for conversation, subject self-ratings of loudness, family ratings of overall intelligibility and the Sickness Impact Profile for the impact of Parkinson disease on their communication. These pre- to post-treatment changes were most apparent in data from male subjects. The R subjects alone had significant pre- to post-treatment increases on the variable pause duration during reading and female self-rating of loudness. There were no significant pre-treatment differences between groups on any of the variables and test retest reliability for acoustic data collected from 61% of the subjects pretreatment and 48% of the subjects post-treatment was considered good. Therefore, the pre to post-treatment changes observed here may be attributed to the effects of the treatments administered and these effects may be considered reliable.

These findings are logical outcomes when the focus of treatment for each group is considered. The R treatment focused on increasing respiratory support for speech by training subjects to take deeper and more frequent breaths and to sustain exhalation for longer duration. The LSVT treatment focused on increasing vocal fold adduction, sustaining vowel phonation for longer duration and louder speech. Deeper and more frequent breaths were encouraged only as a method to maintain loudness more easily. However, because maximum effort phonatory tasks focusing on loudness also stimulate a deeper inhalation, the LSVT was considered to treat both the laryngeal and respiratory systems. Subjects from both treatment groups received the same frequency of treatment (16 sessions) with a "high effort" focus, as well as the same positive attention and support from the clinicians. Since conventional wisdom has suggested that treatment effects in patients with Parkinson disease may be associated with "positive clinical feelings and attention" rather than the efficacy of a technique, it is important to point out that positive attention was equally distributed across both treatment groups.

It is likely that the techniques taught and trained in the R treatment contributed to significant post-treatment increases in pause duration during reading. Subjects had practiced taking more frequent and deeper inhalations during speech tasks which would contribute to longer pauses between utterances. Both groups demonstrated a decrease post-treatment in utterance duration during reading as well as an increase in maximum duration of sustained vowel phonation. These changes are consistent with treatment-related increases in frequency and magnitude of inhalations. However, only with the addition of vocal fold adduction therapy (LSVT) were subjects able to significantly increase both maximum duration ($\bar{x}$=3.73 sec, sd=6.60 sec) and intensity ($\bar{x}$=13.18 dB, sd=5.28 dB) of sustained vowel phonation, as well as intensity during reading ($\bar{x}$=8.03 dB, sd=5.93 dB). Although subjects trained in the R treatment significantly increased maximum duration of sustained vowel phonation ($\bar{x}$=2.81 sec, sd=7.92 sec) post-treatment, the group slightly decreased intensity during the same task ($\bar{x}$= -1.31 dB, sd=5.29 dB). Furthermore, the pre- to post-treatment change in SPL during reading was not significant for the R group ($\bar{x}$=2.48 dB, sd=3.37 dB). While both treatment groups significantly increased SPL in monologue, the subjects who received the LSVT had a greater increase ($\bar{x}$=4.15 dB, sd=2.98) pre- to post-treatment than the subjects who received the R treatment ($\bar{x}$=1.26 dB, sd=3.73 dB). These objective acoustic measures were supported by the significant self-rated improvement post-treatment by the subjects who received the LSVT on the perceptual variable of loudness. Given the significant reduction in vocal loudness reported in patients with Parkinson disease (Aronson, 1985; Blonder, Gur, & Gur, 1989; Boshes, 1966; Critchley, 1981; Darley, Aronson, & Brown, 1969 a, b) and its association with reduced speech intelligibility (Ramig, 1992), these findings distinguish these two forms of speech treatment on an important variable.

Subjects from both treatment groups increased significantly on the variables fundamental frequency (F_0) and its variability (STSD) in reading post-treatment. However, only the subjects who received the LSVT improved significantly on the measure of fundamental frequency variability in monologue as well. Subjects who received the R treatment decreased fundamental frequency variability (STSD) during conversational monologue post-treatment. The changes observed in fundamental frequency and its variability in these subjects are interesting given the fact that these variables were not directly trained in therapy. These changes may have been the result of several factors. Adequate respiratory support may play an important role in fundamental frequency variability during speech. However, given the greater and more consistent improvements of the LSVT subjects, increased vocal fold adduction, which may increase overall vocal fold tension, may be the basis for the greater improvement in the LSVT group. This speculation may explain why the subjects who received the LSVT had the greatest increases in F_0 and STSD during monologue post-treatment. Another explanation for this finding may coincide with the significant improvement rated by the LSVT subjects on the neuropsychological scale Sickness Impact Profile. The subjects who received the LSVT rated a statistically significant reduction of the impact of their sickness on their communication skills post-treatment ($\bar{x}$ = -17.04 points, sd=18.66 points), while the subjects who received R treatment did not ($\bar{x}$ = -2.53 points, sd = 7.60 points). This rated improvement indicates that the LSVT subjects perceived that their communication skills were much less affected by their disease post-treatment. This finding suggests that improved communication can play an important role in the psychosocial well-being of individuals with Parkinson disease. Furthermore, it may be speculated that as the LSVT subjects improved their ability to communicate intelligibly, more effectively and with greater confidence, their attitude and affect improved. This could be reflected in increased pitch and intonation during speech. Higher pitch and increased pitch variability have been associated with descriptive characteristics of happiness, joyfulness and confidence (Fairbanks & Pronovost, 1939; Scherer, London, & Wolf, 1973). Improvements in attitude and affect are also consistent with our subjective observations of the subjects who received the LSVT immediately post-treatment. It is important to keep in mind that the assessment of fundamental frequency and its variability during speech may also be influenced by the topic of the monologue. For example, subjects who chose to talk about their grandchildren, for example, would likely have higher and more variable fundamental frequency than subjects who chose to talk about daily trips to the grocery store. This topic-related variability is likely associated with the lower intrasubject test retest reliability observed here for STSD measures during monologue. It seems important therefore to sample both more controlled speech (reading) as well as spontaneous (monologue) productions inorder to obtain representative measures of fundamental frequency and its variability during speech for comparative purposes.

The perceptual data presented here were considered as an initial probe of the impact of therapy-related changes on functional communication. While the ratings of patients and family members may be biased and less reliable (see Table 4), they represent one important level of assessment of the perceived impact of treatment on functional communication. Subjects from both forms of treatment rated significant improvement on self-ratings of monotonicity, overall intelligibility and initiation of conversation. Families of both treatment groups rated significant improvement on the perceptual variable of loudness. Male subjects who received the LSVT and female subjects who received R self-rated significant improvement on loudness. Only families of subjects who received the LSVT rated a significant improvement in ratings of overall intelligibility. Subjects who received the R treatment rated no significant improvement over the LSVT subjects on any perceptual variable. These findings suggest that both forms of treatment influenced subject and family perceptions of overall communication skills. However, they also suggest that subjects who received the LSVT perceived the greatest impact of the treatment on these perceptual variables.

Because of the heterogeneity of the Parkinson population, correlations between magnitude of treatment-related change and prognostic factors (i.e. age, stage of disease, time since diagnosis, severity of speech disorder, depression, and cognitive ability) were completed in order to evaluate the generalizability of treatment effects. Due to the significant differences between the two treatment groups in magnitude of pre- to post-treatment change, correlations were completed separately for each group. Correlations for subjects who received the LSVT indicate those who were less cognitively impaired had greater increases pre- to post-treatment in SPL in conversational monologue. This finding suggests that patients with less cognitive impairment are able to generalize the effect of increased SPL more easily than the more cognitively impaired patients. This is not surprising given the language and cognitive problems often observed in IPD patients (Brown, & Marsden, 1987; Cummings, 1988; Levin, Llabre, et al., 1989). However, it should be pointed out that subjects across the cognitive impairment range of mild to moderate achieved increases in SPL post-treatment following the LSVT. For the female LSVT subjects, it was observed that younger females had greater pre- to post-treatment increases in SPL in "Rainbow Passage" than the older females. This finding should be interpreted with caution given the small number of subjects. Subjects who received the R treatment did not have any significant correlations between magnitude of treatment-related change and prognostic factors. This was not unex-

pected given the fact that the increases observed post-treatment in these subjects were not as great or as consistent as the subjects who received the LSVT. For female subjects who received the R treatment, a correlation was observed between SPL in monologue and depression. For this variable, the R females who were more depressed had greater increases in SPL during monologue post-treatment than the less depressed females. This finding should also be interpreted with caution due to the small number of female subjects. Overall, these correlations suggest that while age and cognition may influence magnitude of treatment-related change for some subjects, stage of disease, time since diagnosis, severity of speech disorder, glottal incompetence, and depression measured in the subjects studied here did not influence treatment success in the LSVT or R programs.

The training tasks used in this treatment study were maximum effort tasks administered in a highly intensive program (16 sessions in one month). Subjects were encouraged to push themselves to new speech effort levels. It has been reported that IPD patients using increased effort are able to compensate for bradykinesia and improve task performance (England & Schwab, 1959; Hallet & Khosbin, 1980; McDowell, et al., 1986). We speculated that if increased effort alone were the underlying trigger for improved speech and voice performance in these subjects with Parkinson disease, then subjects from both treatment groups would have had equal improvement. However, the R treatment subjects did not have the same magnitude of improvement on overall speech and voice performance as the subjects who received the LSVT. We propose that the treatment tasks that stimulated increased vocal fold adduction in the LSVT subjects were essential to improve overall phonatory and speaking performance.

It can be postulated that the underlying physiological changes accompanying increased intensity in the LSVT subjects were improved vocal fold adduction as well as improved respiratory support. Data from videolaryngostroboscopic studies of these subjects have reported increased vocal fold adduction post-treatment in subjects who received the LSVT, while subjects who received the R treatment did not increase adduction and in some cases became more glottally incompetent (Smith, et al., in press). This finding is consistent with a report by Berke, Hanson, Gerratt, Trapp, et al., (1990) who concluded that increasing intensity (in an invivo canine model) by medial adductory compression was more efficient than by increasing airflow. They reported that increasing airflow alone produced a significantly greater open quotient and vocal fold vibratory excursion. Because the R group did not achieve as great an improvement in vocal intensity as the LSVT group and with no apparent increase in glottal closure (in fact in some cases increased glottal incompetence) it can be speculated that the LSVT approach, which directly trains vocal fold adduction, is more beneficial for improvement of intensity in IPD patients than R treatment. Furthermore, it appears that treatment to improve adduction and loudness stimulated adequate increases in respiratory support without a direct treatment focus on respiration (Ramig, Dromey & Scherer, 1993). This is an important finding since a therapy that is simple is optimal for IPD patients given their known cognitive disorders and difficulty with complex tasks (Beneckel, et al., 1986; Yanagisawa, et al., 1989).

Given these cognitive difficulties (Brooks, 1986; Berardelli, Dick, Rothwell, Day, & Marsden, 1986; it was not surprising that conversational monologue had the least pre- to post-treatment change of the three SPL conditions (sustained phonation, reading and monologue). The finding of a correlation between cognitive levels and changes in SPL in monologue suggests that it is difficult for the more cognitively impaired IPD patients to generalize treatment effects. In addition, increased SPL may be more easily achieved and maintained in a task that requires constant vocal fold adduction (e.g. sustained phonation). The reading and monologue tasks require more dynamic adductory gestures, as well as more attention and concentration, and might therefore have been more difficult for subjects to maintain at a high phonatory effort level. Nevertheless subjects in the LSVT group increased intensity in monologue on average 4.15 dB (sd=2.98).

It is well-established that Parkinson disease patients have problems scaling the amplitude of motor output (Brooks, 1986; Muller & Stelmach, 1991; Grill, Demirchi, McShane, & Hallet, personal communication, October 1994). For example, they have reduced amplitude of stride length in walking and of letter stroke in handwriting (Stelmach, 1991). Reduced loudness may be part of this reduction in the amplitude of motor output. We speculate that intensive voice treatment teaches the patients with Parkinson disease to rescale their phonatory effort to maximize speech output. The patient is "recalibrated" to the amount of phonatory effort it requires to produce a loudness level that is within normal limits. The suggestion by Berardelli et al. (1986) that in Idiopathic Parkinson disease there is a breakdown between "perceptual appreciation" of the task goal and "delivery of the appropriate instruction" to the motor cortex is consistent with frequent reports of patients with Idiopathic Parkinson disease that they feel that they are "shouting" when they increase their vocal loudness to within normal limits. It could be that the amplitude of phonatory effort that the LSVT subjects used to increase loudness to a level of intelligible conversational speech was similar to what they had previously (pre-Parkinson disease) used when talking loud or shouting. This "recalibration" or rescaling of the amplitude of output together with a focus on increased vocal fold adduction appear to be two critical elements of success-

ful speech treatment for patients with idiopathic Parkinson disease.

It is important to point out that the treatment goal of increasing vocal fold adduction in Parkinson disease is designed to maximize the efficiency of the phonatory source. Because of the glottal incompetence in patients with Parkinson disease, increased vocal fold adduction is most often essential to allow patients to generate adequate vocal intensity. It is never the goal of treatment to increase vocal fold adduction so the voice becomes pressed or hyperadducted. The goal is a voice with significant loudness generated with maximum phonatory efficiency.

While activities of respiration were the focus of the R treatment and activities of vocal fold adduction and respiration were the focus of the LSVT, it was apparent by clinical observation that post-treatment changes were observed in magnitude of articulatory gestures as well. These observations were most apparent in subjects who received the LSVT and are consistent with reports of Schulman (1989) who studied articulatory changes following loud speech in non-disordered speakers. These post-treatment changes in untreated articulation have been quantified in a patient with IPD following a course of the LSVT (Dromey, et al., in review). These findings are in agreement with those of Scott and Caird (1983) who reported that phonation in particular was a prerequisite for intelligible, effective speech in Parkinson disease and as voice production improved other aspects of speech also improved and Rubow and Swift (1985) who reported a positive relationship between improved loudness and articulatory skills in one patient with Parkinson disease. These findings suggest that multiple levels of speech production in patients with Parkinson disease can benefit from a single treatment focus of increased phonatory effort (Ramig, et al., 1994).

The gender differences observed in the pre to post-treatment data deserve further consideration. The smaller number of females in both treatment groups, the heterogeneity of the females in the LSVT group and the overall higher depression scores for females limit conclusive statements at this time. However, the findings indicate that further study of the effects of intensive speech treatment for female patients with idiopathic Parkinson disease is essential; that work is ongoing.

Treatment efficacy data in the field of speech-language pathology is lacking. Efficacious treatment programs for patients with progressive neurological disorders are virtually nonexistant (Yorkston, Beukelman, & Bell, 1988). This investigation documents short-term change in characteristics of speech and voice in 45 subjects with Idiopathic Parkinson disease following two forms of intensive speech treatment. The data of this study document that intensive speech treatment, particularly the LSVT, made significant short-term improvements in the voice production of the subjects studied. Given the observed perceptual and physiological speech and voice deficiencies in individuals with IPD, vocal fold adduction therapy (LSVT) appears to be the most beneficial of the two treatments studied here. Although the R treatment improved some aspects of speech, the subjects who received the LSVT made greater and more consistent changes in intensity and intonation as well as reducing the impact of Parkinson disease on oral communication.

Of the 45 subjects in this investigation, prognostic factors of stage of disease, time since diagnosis, severity of speech disorder, glottal incompetence and depression did not significantly influence post-treatment effects. Age and cognitive status must be studied more closely before conclusive statements are made. It should be noted, however, that most of the subjects studied here were in the mild to moderate range of their disease. Further research is needed which examines more severely involved subjects as well as very early patients to assess the generalizability of treatment across all stages of IPD. Given the heterogeneous nature of Parkinson disease, it is reasonable to assume that all subjects would not respond to treatment in the same way. This variability is clearly reflected in individual subject responses to treatment (Figures 2 and 4). It is important to keep in mind that subject heterogeneity may reduce statistically significant group effects despite positive findings for individual subjects. This is an important consideration in evaluating the efficacy of treatment for neurological disorders.

These data suggest that intensive speech treatment may improve select characteristics of speech production in patients with idiopathic Parkinson disease. This is in contrast to previous findings from Allan (1970), Aronson (1985), Greene (1980), Sarno (1968), Weiner & Singer (1989) who concluded that speech treatment was not effective for patients with Parkinson disease. The most reasonable explanation for the positive findings reported here is the focus and style of treatment. Previous approaches to treatment for Parkinson disease have focused on articulation and rate and were administered in the classic speech treatment style of once or twice a week. The treatment studied here focused on phonation, specifically increasing vocal fold adduction and vocal loudness and was administered four days a week for sixteen high effort sessions. We suggest that intensive treatment with the focus on phonation is a key to the the efficacious treatment for patients with Parkinson disease. Furthermore, these findings are consistent with Scott & Caird (1983) and Robertson & Thompson (1984) who also concluded that intensive voice treatment could be beneficial to patients with Parkinson disease.

Given the degenerative nature of Parkinson disease, the accompanying cognitive impairment, involvement at multiple levels of the speech mechanism, depression and general physical impairment, documentation of statistically significant pre to post-treatment changes for group data of the magnitude reported here is an important step in evaluat-

ing the efficacy of voice treatment for this population. The maintenance of these changes is being addressed in ongoing work. While the degenerative course of Parkinson disease cannot be altered at this time, improved oral communication may be an important component in developing the highest levels of functioning and independence for each individual.

Acknowledgments

The research presented here was supported by NIH grant No. RO1DC01150 and P60 DC00976. Our deepest gratitude is extended to the subjects who participated in this research study. The authors would also like to thank Drs. Marshall Smith, Margaret Hoehn, Christopher O'Brien, Ms. Annette Pawlas, and Mr. Christopher Dromey for their assistance during various parts of this research.

References

Allan, C. M. (1970). Treatment of non-fluent speech resulting from neurological disease: Treatment of dysarthria. British Journal of Disorders of Communication, 5, 3-5.

Aronson, A. E. (1985). Clinical voice disorders. New York: Thieme-Stratton.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. Archives of General Psychiatry, 4, 561-571.

Beatty, W. W., Stanton, R. D., Weir, W. S., et al. (1989). Cognitive disturbances in Parkinson disease. Journal of Geriatric Psychiatry and Neurology, 2, 22-33.

Beneckel, R., Rothwell, J. C., Dick, J. P, Day, B. L., & Marsden, C. D. (1986). Performance of simultaneous movements in patients with Parkinson's disease. Brain, 109, 739-757.

Berardelli, A., Dick, J. P., Rothwell, J. C., Day, B. L., & Marsden, C. D. (1986). Scaling of the size of the first agonist EMG burst during rapid wrist movements in patients with Parkinson's disease. Journal of Neurology, Neurosugery, and Psychiatry, 49 (11), 1273-1279.

Bergner, M., Bobbit, R. A., Carter, W. B., & Gilson, B. S. (1981). The sickness impact profile: Development and final revision of a health status measure. Medical Care, 19, 787-805.

Berke, G. S., Hanson, D. G., Gerratt, B. R., Trapp, J. K., Macagba, C., & Natividad, M. (1990). The effect of airflow and medial compression on vocal efficiency and glottal vibration. Otolaryngology, Head and Neck Surgery, 102, 212-280.

Bless, D. M., Hirano, J., & Feder, R. J. (1987). Videostroboscopic evaluation of the larynx. Ear, Nose, and Throat Journal, 66(7), 289-296.

Blonder, L. X., Gur, R. E., & Gur, R. C. (1989). The effects of right and left hemiparkinsonism on prosody. Brain and Language, 36, 193-207.

Boeckstyns, M. E., & Backer, M. (1989). Reliability and validity of the evaluation of pain in patients with total knee replacement. Pain, 38(1), 29-33.

Boshes, B. (1966). Voice changes in parkinsonism. Journal of Neurosurgery, 24, 286-288.

Brooks, V. B. (1986). The Neurological Basis of Motor Control. Oxford University Press, New York.

Brown, R. G., & Marsden, C. D. (1987). Neuropsychology and cognitive function in Parkinson's disease: An overview. In C. D. Marsden & S. Fahn (Eds.), Movement disorders (2nd ed.) (pp. 99-123).

Canter, G. J. (1965). Speech characteristics of patients with Parkinson's disease: II. Physiological support for speech. Journal of Speech and Hearing Disorders, 30, 44-49.

Countryman, S., & Ramig, L. O. (1993). Effects of intensive voice therapy on voice deficits associated with bilateral thalamotomy in Parkinson disease: A case study. Journal of Medical Speech-Language Pathology, 1, 233-249.

Countryman, S., Ramig, L. O., & Pawlas, A. (1994). Speech and voice deficits in Parkinsonian Plus syndromes: Can they be treated? Journal of Medical Speech-Language Pathology, 2(3), 211-225.

Critchley, E. M. R. (1981). Speech disorders of Parkinsonism: A review. Journal of Neurology, Neurosurgery, and Psychiatry, 44, 751-758.

Cummings, J. L. (1986). Subcortical dementia: Neuropsychology, neuropsychiatry, and pathophysiology. British Journal of Psychiatry, 149, 682-687.

Cummings, J. L. (1988). Intellectual impairment in Parkinson's disease: Clinical, pathologic, and biochemical correlates. Journal of Geriatric Psychiatry and Neurology, 1, 24-36.

Darley, F. L., Aronson, A., & Brown, J. (1969a). Differential diagnostic patterns of dysarthria. Journal of Speech and Hearing Research, 12, 246-269.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969b). Clusters of deviant speech dimensions in the dysarthrias. Journal of Speech and Hearing Research, 12, 462-469.

Dobkin, R. C. (1969). Precision AC/DC converters. National Semiconductor Corporation, LB8, 1-2.

Dromey, C., Ramig, L. O., & Johnson, A. (in review). Phonatory and articulatory changes associated with increased vocal intensity in Parkinson disease: A case study. Journal of Speech and Hearing Research.

Duvoisin, R. C. (1984). Parkinson's disease: A guide for patient and family (2nd ed.). New York: Raven Press.

England, A. C., & Schwab, R. S. (1959). The management of Parkinson's disease. AMA Archives of Internal Medicine, 104, 439-460.

Fairbanks, G. (1960). Voice and articulation drillbook. New York: Harper and Brothers.

Fairbanks, G., & Pronovost, W. (1939) An experimental study of the pitch characteristics of the voice during expression of emotion. Speech Monographs, 6, 87-104.

Fahn, S., Elton, R. L., & members of the UPDRS Development Committee (1987). Recent developments in Parkinson's disease. In Fahn, S., Marsden, C. D., Goldstein, M., & Calne, D. B. (Eds.). New York: Macmillan Press, 2: 153-163.

Gauthier, L., Dalziel, S., & Gauthier, S. (1987). The benefits of group occupational therapy for patients with Parkinson's disease. American Journal of Occupational Therapy, 41, 360-365.

Goodglass, H., & Kaplan, E. (1972). The assessment of aphasia and related disorders. Philadelphia: Lee and Leafebiger.

Greene, H. C. L. (1980). The voice and its disorders. London: Pitman Medical.

Grill, S., Demirchi, M., McShane, L. & Hallet, M. (1994) personal communication, October.

Hallet, M., & Khosbin, S. (1980). A psychological mechanism of bradykinesia. Brain, 103(2), 301-314.

Hanson, D. G., Gerratt, B. R., & Ward, P. H. (1984). Cinegraphic observations of laryngeal function in Parkinson's disease. Laryngoscope, 94, 348-353.

Heaton, R. K., Chelune, G., & Lehman, R. A. W. (1978). Using neuropsychology and personality tests to assess the likelihood of patient employment. Journal of Nervous and Mental Disease, 166, 408-416.

Heaton, R. K., Grant, I., & Mathews, C. G. (1991). Comprehensive norms for an expanded Halsted-Reitan Battery. Odessa, FL: Psychological Assessment Resources.

Hoehn, M., & Yahr, M. (1967). Parkinsonism: Onset, progression and mortality. Neurology, 17, 427.

Hoberman, S. G. (1958). Speech techniques in aphasia and Parkinsonism. Journal of the Michigan State Medical Society, 57, 1720-1723.

Horii, Y. (1983). An automatic analysis method of utterance and pause lengths and frequencies. Behavior Research Methods and Instrumentation, 15, 449-452.

Hornykiewicz, O. (1966). Metabolism of brain dopamine in human Parkinsonism: Neurochemical and clinical aspects. In E. Costa, L. Cote, & M. Yahr (Eds.), Biochemistry and pharmacology of the basal ganglia. New York: Raven Press.

Hornykiewicz, O., & Kish, S. J. (1986). Biochemical pathophysiology of Parkinson's disease. In M. D. Yahr & K. J. Bergman (Eds.), Advances in Neurology, 45, 19-34.

Kempster, G. B. (1984). A multidimensional analysis of vocal quality in two dysphonic groups. Unpublished doctoral dissertation, Northwestern University, Evanston, IL.

Kent, R. D., Kent, J. F., & Rosenbeck, J. C. (1987). Maximum performance tests of speech production. Journal of Speech and Hearing Disorders, 52, 367-387.

King, J. B., Ramig, L. O., Lemke, J., & Horii, Y. (1994). Parkinson's disease: Longitudinal variability in acoustic parameters of phonation. Journal of Medical Speech-Language Pathology, 2, 29-42.

Levin, B. E., Llabre, M. M., et al. (1989). Cognitive impairments associated with early Parkinson's disease. Neurology, 39, 557-561.

Lewis, R., & Kupke, T. (1977). The Lafayette Clinic repeatable neuropsychological test battery: Its development and research applications. Paper presented at the annual meeting of the Southeastern Psychological Association, Hollywood, FL.

Logemann, J. A., Fisher, H. B., Boshes, B., & Blonsky, E. R. (1978). Frequency and concurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson patients. Journal of Speech and Hearing Disorders, 42, 47-57.

Mayeux, R., Williams, J. B. W., Stern, Y., & Cote, L. (1986). Depression and Parkinson's disease. In M. D. Yahr & K. J. Bergmann (Eds.), Advances in Neurology (pp. 241-250). New York: Raven Press.

McDowell, F. H., Lee, J. E., & Sweet, R. D. (1986). Extrapyramidal disease. In A. B. Baker & R. J. Joynt (Eds.), Clinical Neurology, 38, 24-26.

Mead, C., Ramig, L. O., & Beck, J. (1989). Parkinson's disease with dementia: Effectiveness of intensive voice therapy. ASHA, 31(10), 118.

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

Montgomery, S. A., & Asberg, M. (1979). A new depression scale designed to be sensitive to change. British Journal of Psychiatry, 134, 382-389.

Morley, D. E. (1955). The rehabilitation of adults with dysarthric speech. Journal of Speech and Hearing Disorders, 20, 58-64.

Muller, F., & Stelmach, G. E. (1991). Scaling problems in Parkinson's disease. Requin, J. and Stelmach, G. E. (Eds) Tutorials in Motor Neuroscience, Netherlands, Kluwer Academic Publishers, 161-174.

Mutch, W. J., Strudwick, A., Roy, S. K. & Downie, A. W. (1986) Parkinson's disease: disability, review and management. British Medical Journal 293, 675-677.

Oxtoby, M. (1982). Parkinson's disease patients and their social needs. London: Parkinson's Disease Society.

Perez, K., Ramig, L. O., Smith, M., & Dromey, C. (in review). The Parkinson larynx: Tremor and videostroboscopic findings. Journal of Voice.

Ramig, L. O. (1992). The role of phonation in speech intelligibility: A review and preliminary data from patients with Parkinson's disease. In R. D. Kent (Ed.), Intelligibility in speech disorders: Theory, measurement and management (pp. 119-156). Amsterdam: John Benjamin.

Ramig, L. O. (in press) Speech therapy for Parkinson's disease. In Koller, W. and Paulson, G. (Eds) Therapy of Parkinson's Disease Marcel Dekker, Inc.: New York.

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, 191-209.

Ramig, L. O., Countryman, S., & Pawlas, A. (in preparation). The Lee Silverman Voice Treatment workbook. Iowa City, IA: National Center for Voice and Speech.

Ramig, L. O., Dromey, C., Johnson, A., & Scherer, R. (1994, April). The effects of phonatory, respiratory and articulatory effect treatment on speech and voice in Parkinson's disease. Paper presented to the Motor Speech Conference, Sedona, AZ.

Ramig, L. O., Dromey, C. and Scherer, R. (1993) Aerodynamic, acoustic and kinemeatic characteristics of voice in Parkinson's disease following intensive treatment. ASHA 35, 10.

Ramig, L. O., Fazoli, K., Scherer, R., & Bonitati, C. (1990, January). Changes in phonation following voice therapy in patients with Parkinson's disease. Paper presented to the Clinical Dysarthria Conference, San Antonio, TX.

Ramig, L., Horii Y., & Bonitati, C. (1991). The effectiveness of speech therapy for patients with Parkinson's disease. National Center for Voice and Speech Research Report, 1, 61-86.

Ramig, L. O., Mead, C., Scherer, R., Horii, Y., Larson, K. & Kohler, D. (1988) Voice therapy and Parkinson's disease: a longitudinal study of efficacy. A paper presented at the Clinical Dysarthria Conference, San Diego, February.

Reitan, R. M., & Wolfson, D. (1985). The Halsted-Reitan neuropsychological test battery. Tuscon, AZ: Neuropsychology Press.

Robertson, S. & Thompson, F. (1984) Speech therapy in Parkinson's disease: A study of efficacy and long-term effect in intensive treatment. British Journal of Disorders of Communication 19, 213-224.

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.

Rubow, R., & Swift, E. (1985). A microcomputer-based wearable biofeedback device to improve transfer of treatment in Parkinsonian dysarthria. Journal of Speech and Hearing Disorders, 50, 178-185.

Sarno, M. T. (1968). Speech impairment in Parkinson's disease. Archives of Physical Medicine and Rehabilitation, 49, 269-275.

Scherer, K. R., London, H., & Wolf, J. (1973) The voice of confidence: Paralinguistic cues and audience evaluation. Journal of Research Personality, 7, 31-44.

Schiffman, S., Reynolds, M L., & Young, F. W. (1981). Introduction to multidimensional scaling: Theory, methods and applications. New York: Academic Press.

Schulman, R. (1989) Articulatory dynamics of loud and normal speech. Journal of the Acoustical Society of America 85, 295-312.

Scott, S., & Caird, F. L. (1983). Speech therapy for Parkinson's disease. Journal of Neurology, Neurosurgery and Psychiatry, 46, 140-144.

Smith, M., Ramig, L. O., Dromey, C., Perez, K., & Samandari, R. (in press). Intensive voice treatment in Parkinson's disease: Laryngostroboscopic findings. Journal of Voice.

Stelmach, G. E. (1991) Basal ganglia impairment and force control. In Requin, J. and Stelmach, G. E. (Eds) Tutorials in Motor Neuroscience Kluwer Academic Publishers. (Netherlands) 137-148.

Streifler, M., & Hofman, S. (1984). Disorders of verbal expression in Parkinsonism. In R. G. Hassler & J. F. Christ (Eds.), Advances in Neurolology, 40, 385-393.

Titze, I. R., & Sundberg, J. (1992). Vocal intensity in speakers and singers. Journal of the Acoustical Society of America, 91, 2936-2946.

Wechsler, D. (1981). WAIS-R manual. New York: The Psychological Corporation.

Weiner, W. J., & Singer, C. (1989). Parkinson's disease and nonpharmacologic treatment programs. Journal of the American Geriatrics Society, 37, 359-363.

Winholtz, W. S., & Ramig, L. (1992). Vocal tremor analysis with the vocal demodulator. Journal Speech and Hearing Research, 35, 562-573.

Yanagisawa, N., Fujimato, S., & Tamaru, F. (1989). Bradykinesia in Parkinson's disease: Disorders of onset and execution of fast movement. European Neurology, 29 (Suppl.), 19-28.

Yorkston, K. M., Beukelman, D. R. & Bell, K. R. (1988) Clinical Management of Dysarthric Speakers. Boston: Little, Brown and Company.

NCVS Status and Progress Report - 7
December 1994, 65-75

Sources of Error in Estimation of Laryngeal Airway Resistance in Patients with Spasmodic Dysphonia

Eileen Finnegan, M.A.
Erich Luschei, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa
and
Department of Otolaryngology, The University of Iowa Hospitals and Clinics
Julie Barkmeier, Ph.D.
Voice and Speech Section, National Institute on Deafness and Other Communication Disorders
Henry Hoffman, M.D.
Department of Otolaryngology, The University of Iowa Hospitals and Clinics

Abstract

Estimation of laryngeal airway resistance is a noninvasive method that has proven useful in the study of people with normal and some types of disordered voices. We were interested in examining more closely the application of this method to patients with spasmodic dysphonia (SD). Specifically, our purpose was to determine if SD patients were able to maintain steady airflow during the syllable repetition task used for estimation of laryngeal airway resistance. Oral pressure and airflow were collected from 10 subjects with SD and 10 control subjects during repetition of /pi/. The coefficient of variation (COV) of airflow during vowel production was calculated to quantify stability of airflow. The results indicated that although some SD subjects were able to produce steady flows during the syllable repetition task, others exhibited substantially varying flows. Inability on the part of the patient to attain steady flows could compromise the usefulness of a midpoint measure of airflow and/or estimation of subglottal pressure, resulting in sources of error in estimation of laryngeal airway resistance. As a result, of the 10 subjects with SD in this study, we were unable to determine if laryngeal airway resistance was within normal limits in the six SD subjects with unsteady flows. Laryngeal airway resistance was estimated in four SD subjects who were able to maintain steady airflow. Two of these subjects exhibited high laryngeal airway resistance with mean airflows at the low end of the normal range; the others exhibited normal laryngeal airway resistance with high subglottal pressures.

Introduction

Laryngeal airway resistance (LAR) can be calculated from simultaneous measures of translaryngeal pressure and translaryngeal airflow, based on Ohm's law (resistance = pressure/flow). Translaryngeal pressure is the difference between subglottal pressure and pharyngeal pressure. During vowel production, translaryngeal pressure is essentially equal to subglottal pressure since the mouth is open to atmospheric pressure, so under normal circumstances, pharyngeal pressure is zero. Tranglottal airflow is essentially the same as flow from the mouth and nose during vowel production when the velopharyngeal port is closed and the articulators are in a static position. Thus, measures of subglottal pressure and airflow from the mouth can be used to calculate LAR (Smitheran and Hixon, 1981). Airflow from the mouth is easily accessible and can be measured directly. It is possible to measure subglottal pressure in several ways. Continuous direct measurement of subglottal pressure can be made by placing a pressure transducer below the glottis either via a tracheal puncture or by passing the transducer through the glottis. Continuous indirect measurement of subglottal pressure can be made by having the subject swallow an esophageal balloon that is connected to an external pressure transducer via a catheter. Pressure changes in the lungs result in changing esophageal pressures due to compliance of the posterior tracheal wall. Simultaneous recording of lung volume with a body plethesmograph are needed for conversion of esophageal pressures to subglottal pressure (Bouhuys,

1966; Hixon, 1972; Smitheran and Hixon, 1981; Lofquist, Carlborg, and Kitzing, 1982). In an effort to avoid these invasive procedures and still obtain measures of subglottal pressure, Rothenberg (1973) proposed subglottal pressure during a vowel could be estimated from intraoral pressure during preceding and subsequent plosive consonants, as in the syllable /bæp/.

Several years later, Holmberg (1980) suggested laryngeal airway resistance could be estimated from a ratio of estimated subglottal pressure divided by airflow. She examined how this estimate changed as a function of intensity, fundamental frequency, and phonation type, and reported LAR was highest for pressed voice and lowest for breathy voice regardless of intensity or fundamental frequency. Based on these findings, she concluded that the estimated LAR appeared to be "a good index of laryngeal resistance". In 1981, Smitheran and Hixon provided a complete description of a method for estimating laryngeal airway resistance and the basic principles underlying the method. In addition, they collected LAR data from 15 normal adult males and reported that the estimates of LAR they obtained via this method were comparable to values obtained by other investigators using direct measurement (rather than estimation) of subglottal pressure.

Estimation of laryngeal airway resistance has been widely used by other investigators to obtain data from normal subjects of various age/gender groups (Leeper and Graves, 1984; Melcon, Hoit, and Hixon, 1989; Holmberg, Hillman, and Perkell, 1988; Higgins and Saxman, 1991; Hoit and Hixon, 1992; Stahopoulos and Sapienza, 1993; Netsell, Lotz, Peters, and Schulte, 1994) and from people with some voice disorders (McGuirt, Blalock, Koufman, and Feehs, 1992; Leeper, Gagne, Parnes, and Vidas, 1993; Lewis, Andreassen, Leeper, Macrae, and Thomas, 1993). Simultaneous measures of pressure and airflow during vowel production are needed to calculate LAR. In the method described by Smitheran and Hixon (1981) oral pressure and airflow data are collected during repetition of /pi/, as illustrated in Figure 1. Oral pressure (top tracing) rapidly increases during the closed phase of the /p/; attains a steady "peak" pressure when pressure throughout the vocal tract equalizes; rapidly decreases to zero when the plosive is released; and maintains a stable zero pressure, or a non-zero pressure representing pharyngeal pressure, during the vowel when the mouth is open to atmospheric pressure. Airflow (bottom tracing) peaks during release of the plosive, is steady during the vowel, and then falls to zero when the lips are sealed preventing flow from the mouth. The usual method for estimating laryngeal airway resistance is a four step procedure. First, airflow is measured at the midpoint of the vowel. Second, subglottal pressure at that same moment is estimated from the midpoint of a line drawn between adjacent peaks in oral pressure. Third, the pressure at the midpoint between peaks is subtracted from the estimated subglottal pressure. Finally, the estimated laryngeal airway resistance is calculated by dividing estimated subglottal pressure by airflow. Laryngeal airway resistance is calculated for the three central tokens of a series of seven repetitions of /pi/ and a mean laryngeal airway resistance is obtained.

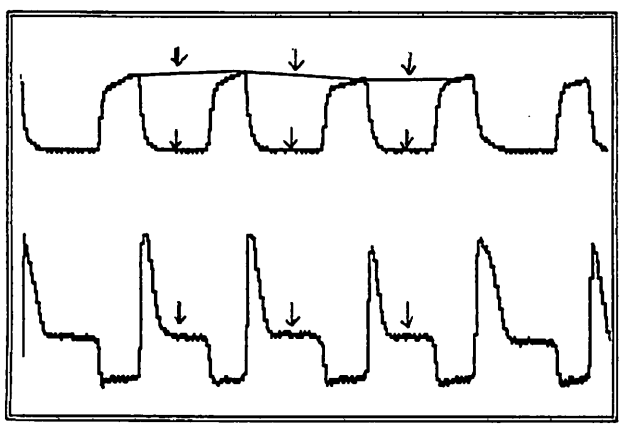


Figure #1: Noninvasive method for estimation of laryngeal airway resistance: measure airflow (bottom trace) at the midpoint of the vowel and estimate subglottal pressure (top trace) from oral pressure.

Adductor spasmodic dysphonia is thought to result from hyperadduction of the vocal folds, so there is an inherent attraction in the use of a measure such as laryngeal airway resistance, which could reflect an aspect of the underlying pathophysiology of the disorder. Recently, estimation of LAR was proposed as a means for "documenting results of botulinum toxin injection for treatment of focal laryngeal dystonia" (Witsell et al, 1994). However, Davis, Boone, Carroll, Darveniza & Harrison (1988) reported "marked variability" in airflow during sustained phonation in the majority of 20 subjects with SD, and identified three patterns of airflow: steady, oscillating, and irregular. The presence of unstable airflow during the vowel portion of the syllable repetition task could effect the estimation of laryngeal airway resistance. Initial review of the aerodynamic data from our SD patients indicated that although some of the patients were able to produce airflow tracings similar to examples published by previous investigators (Smitheran and Hixon, 1981), other SD patients were unable to achieve steady airflow during phonation. We were interested in examining this observation in a systematic manner to determine if SD patients were able to maintain steady airflow during the syllable repetition task used for estimation of laryngeal airway resistance.

Methods

Subjects: Ten subjects with diagnosis of spasmodic dysphonia (2 male, 8 females, mean age: 59.8 years) were included in this study. They were evaluated at the University of Iowa Hospital voice clinic and a diagnosis of spasmodic dysphonia (9 Adductor and 1 Abductor) was

made based on review of videoendoscopic, acoustic, aerodynamic, perceptual, and medical and social history information by two speech pathologists and an otolaryngologist. Since we were uncertain any of the spasmodic dysphonia subjects would produce phonation with steady flows, we included ten control subjects with no history of voice disorder (1 male, 9 females, mean age: 45.7 years), with the expectation that they would provide data with steady flows for comparison.

Data collection: Simultaneous recordings of intraoral pressure and airflow were obtained using the procedures delineated by Smitheran and Hixon, 1981. Airflow measurements were made by placing an anesthesia mask over the subject's mouth and nose to direct airflow through a pneumotachometer (Hans Rudolph model 4719) connected to a differential pressure transducer (Honeywell Microswitch 162PC01D). The airflow signal was amplified and recorded on a digital instrumentation taperecorder (Sony PC108-M). Oral pressure was transduced by a seven inch polyethylene tube with a 1.67 mm inner diameter (Intramedic PE#240) which was inserted through a small hole with a rubber dam in the front wall of the anesthesia mask. The tube was positioned in the subject's oral cavity to sample intraoral pressure without interfering with lip and/or tongue movement. The oral pressure signal from the pressure transducer (Honeywell Microswitch 162PC01D) was amplified and recorded on another channel on the digital recorder.

Calibration of the airflow and pressure instrumentation was performed on each day of data collection. Two DC flow levels were generated and monitored (0 and 1000 cc/s) using an airtank and a rotameter. Air pressure was calibrated at two levels (0 and 10 cm H_2O) using a syringe as a pressure source and a manometer. Signals from the pneumotachometer and differential pressure transducer were adjusted so that 1000 cc/s of DC airflow equaled 1 Volt. These signals were recorded onto the digital recorder and used to calibrate experimental signal measures.

Subjects were instructed to repeat the syllable /pi/ in a continuous manner, at normal pitch and loudness, as modeled by the investigator at 1.5 syllables/second. Subjects were directed to place the mask firmly against their face at the start of each trial and to maintain that placement during production of all tokens in that trial. Subjects were provided with practice and recordings were repeated until there was a plateau in the performance elicited from the subject. The data were low-pass filtered at 30 Hz, then digitized at 500 Hz per channel onto the software system used for data analysis (CODAS, Dataq Inc, Akron, Oh).

Measures: The stability of the flow (i.e. the "flatness" of the airflow tracing) during phonation was quantified by calculating the coefficient of variation of airflow during the vowel, as illustrated in Figure 2. This measure was calculated for tokens 3-5 for each of 3 trials, resulting in 9 measures of flow stability for each subject. These nine COVs were averaged to obtain the subject's mean COV.

Results

Control Subjects: Airflow tracings from the third, fourth, and fifth tokens of the second trial obtained from the ten control subjects are displayed in Figure 3 (see following page). The mean COVs are listed to the left side of each data set. The data are organized (top to bottom, left to right) in order of increasing mean COV. In general, the control subjects maintained steady airflow during vowel production, although some subjects exhibited mild instability in airflow on individual tokens. In a number of cases (57F, 55Fb, 55F, 64F) airflow was characterized by a slight increase at the end of vowel production. Generally this increase was associated with initiation of lip closure for /p/, and was not included in the measure of COV since the oral pressure started to increase at the same time (recall the airflow measure was made over the time that oral pressure was at baseline).

The airflow data collected from the control subjects are summarized in Table 1. Mean COVs from these subjects ranged from 6.9% to 21.3%, with individual tokens ranging from 3.7% to 32.7%. Control subjects with low mean airflows tended to exhibit greater instability of flow. Mean airflows ranged from .064 to .252 liters per second (L/s), and were comparable to airflow data collected from normal subjects by other investigators using this protocol (Leeper and Graves, 1984; Netsell et al 1994; Holmes et al 1994), as summarized in Appendix A.

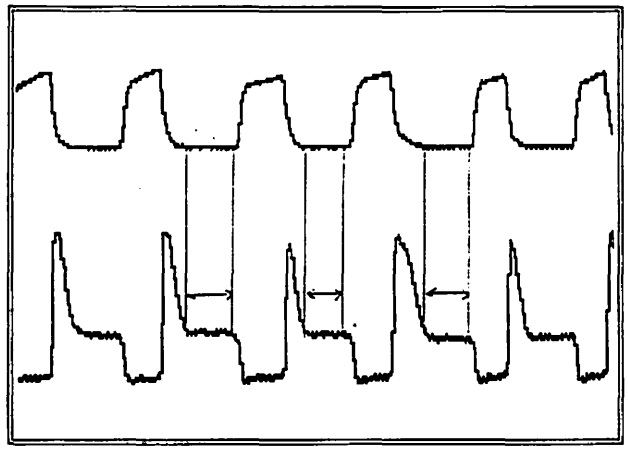


Figure #2: Method for measuring stability of airflow: 1) Onset and offset of the vowel were manually selected by marking the period during which the oral pressure trace returned to baseline (i.e., from release of the plosive to lip closure for the /p/); 2) mean and standard deviation of airflow during the vowel were calculated by CODAS; 3) the coefficient of variation (COV) for each token was calculated by dividing standard deviation by mean flow.

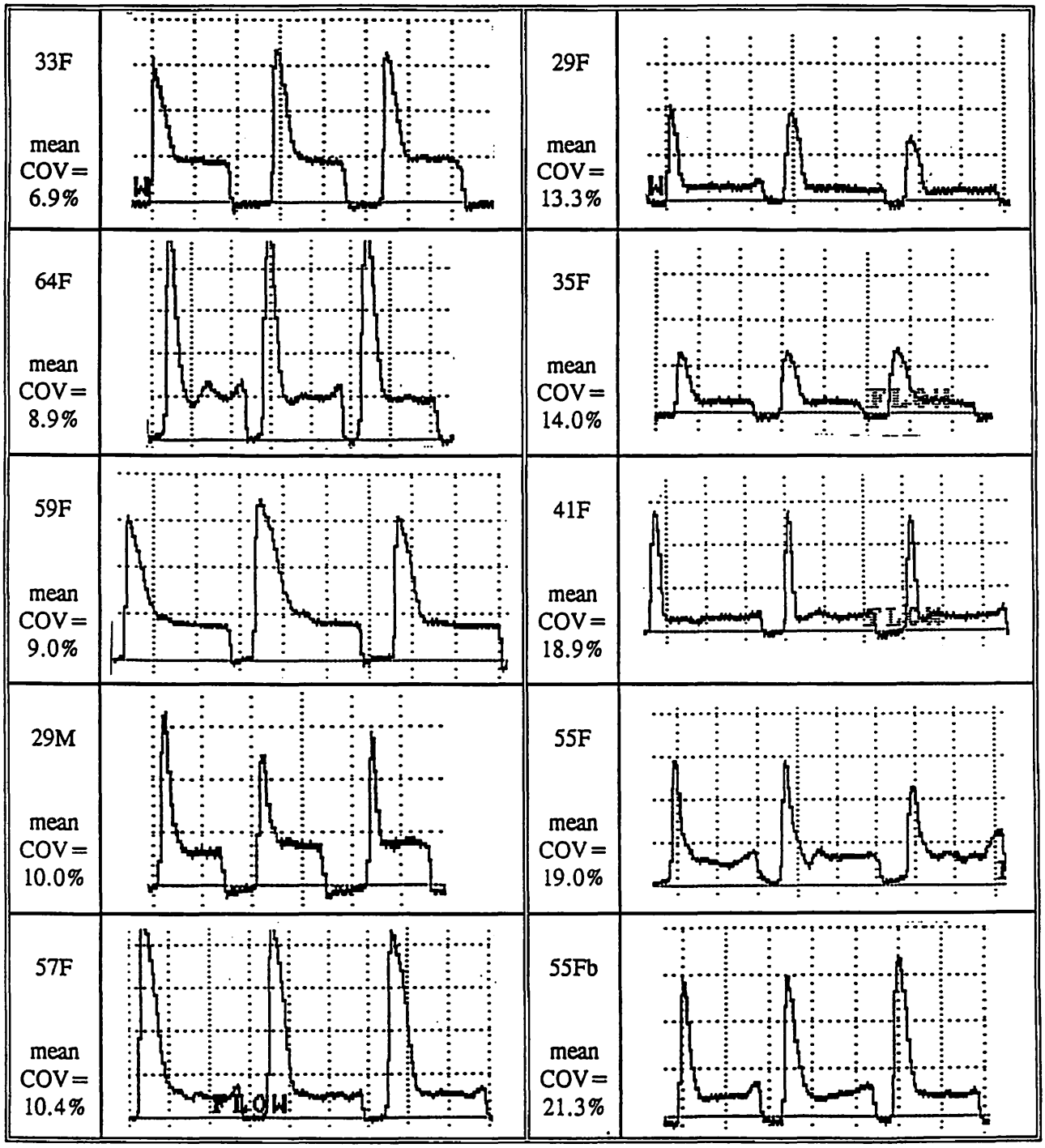


Figure #3: Airflow tracings from 10 control subjects. The mean COV for each subject was calculated from the three tokens pictured plus 6 additional tokens from the first and third trials. Time is displayed on the horizontal axis (200 msec per grid mark); amplitude of airflow is on the vertical axis (.25 LPS per grid mark).

Spasmodic Dysphonia: Airflow tracings from the third, fourth, and fifth tokens of the second trial obtained from the ten SD subjects are displayed in Figure 4, with mean COV listed to the left side of each data set. The data are organized (top to bottom, left to right) in order of increasing mean COV. Higher COVs indicate increased variability around the mean flow, presumably reflecting increased instability of the phonatory system. The first four data sets (32F, 65F, 33F, 59F) in Figure 4 are similar to the airflow data produced by the control subjects, displayed in Figure 3,

in that these four SD subjects demonstrated steady flows. Of the remaining six subjects, three exhibited oscillating flows (81F, 64M, 74F) and three had irregular flows (81Fb, 38M, 71F). The three types of flows (steady, oscillating, and irregular), which have been previously observed during sustained /a/ (Davis, Boone, Carroll, Darveniza, & Harrison, 1988), could also be identified during syllable repetition.

The mean COV values for airflow from the control and SD subjects are displayed graphically in Figure 5. Four SD subjects had COVs in the same range as the control subjects. Six SD subjects had COVs which were substantially larger. Table 2 summarizes airflow data collected from the SD subjects. The four SD subjects with COVs in the same range as the control subjects (shaded area of Table 2) had mean COVs ranging from 8.9% to 22.7%, and individual tokens ranging from 4.8% to 27.8%. Six SD subjects with high COVs had mean COVs ranging from 39.3% to 127.0%, and individual tokens ranging from 14.7% to 179.2%. Although some SD subjects were able to produce single tokens with low COVs, they were unable to maintain these low COVs of airflow across tokens.

The mean airflow values from the control and SD subjects are graphically displayed in Figure 6. Mean airflow for SD subjects ranged from .052 to .281 liters per second. Although similar to the control group in range, the SD subjects means tended to be distributed toward the extremes, rather than the center, of the range. The two SD subjects (32F and 65F) with mean airflow near the center of the range, were also the subjects with the lowest COVs, i.e. the most stable flows.

Table 1.
Summary of Airflow measures for 10 control subjects. Listed in same order as Figure 3. Means and standard deviations are based on nine measures, three tokens from each of three trials.

Subject	COV of Airflow		Mean Airflow (L/S)
	Mean +/- SD	Range	
33F	6.9% +/- 3.6%	3.7% - 13.5%	.252
64F	8.9% +/- 2.8%	4.9% - 14.0%	.205
59F	9.0% +/- 3.1%	5.4% - 15.8%	.197
29M	10.0% +/- 4.8%	5.9% - 20.1%	.170
57F	10.4% +/- 4.0%	6.7% - 16.7%	.140
29F	13.3% +/- 2.3%	10.6% - 15.6%	.064
35F	14.0% +/- 7.1%	4.9% - 22.5%	.093
41F	18.9% +/- 5.1%	13.7% - 28.9%	.085
55F	19.0% +/- 5.2%	11.2% - 24.3%	.169
55Fb	21.3% +/- 7.3%	6.8% - 32.7%	.098

Discussion

The results indicate that some subjects with spasmodic dysphonia (SD) were able to maintain steady flows during the syllable repetition task used for estimation of laryngeal airway resistance. However, other subjects with SD exhibited airflow that varied over a large range. This instability of airflow has been noted by previous investigators who examined airflow during sustained phonation in SD subjects (Davis et al, 1988). The presence of variable airflow could be an important aspect of the data to consider when describing laryngeal airway resistance in these patients. For example, in the case of a SD subject with steady flow (33F in Figure 4), a single measure of flow from the midpoint of the vowel appears to provide a good indication of flow during the entire vowel phonation. However, this same measure in a subject with unstable flow (81F in Figure 4) does not seem to provide an adequate description of airflow during phonation. Use of a single midpoint measure of airflow could compromise the estimation of laryngeal airway resistance. If the airflow was at a peak at the midpoint of the vowel, laryngeal airway resistance would be underestimated. If airflow was low at the midpoint, the estimate would be inflated. If mean airflow was used, laryngeal airway resistance could appear normal when it might be fluctuating widely from quite low to quite high, rarely within normal range.

We cannot assume varying airflow reflects varying laryngeal resistance, since it might reflect widely varying subglottal pressure. Estimation of subglottal pressure from oral pressure is based on the assumption that subglottal pressure remains steady during vowel production. The validity of this assumption has been investigated in normal subjects (Netsell, 1969; Shipp, 1973; Lofqvist, Carlborg, & Kitzing, 1982) and in some patients with laryngeal pathologies (Kitajima and Fujita, 1990), but not in patients with spasmodic dysphonia. Although we had no direct measure of subglottal pressure, we reasoned that unstable control of respiratory drive during the vowel might also occur during the production of oral pressure for the plosive. With this in mind, we examined the oral pressure data. Samples of oral pressure data collected from the SD subjects during tokens 2-6 of the second trial are presented in Figure 7. A stable maximum pressure, i.e. a flat top on the pressure tracing, would indicate equalization of pressure throughout the vocal tract and would validate the assumption that oral pressure is equal to subglottal pressure (Smitheran and Hixon, 1981; Holmberg, Perkell, and Hillman, 1984). We observed flat or slightly increasing oral pressure peaks during the /p/, and stable baseline pressures during the vowel in SD subjects with steady flows (32F, 65F, 33F, 59F), similar to oral pressure tracings from the control subjects. However, the SD subjects who had high COVs of airflow also exhibited pointed, decreasing, or erratic slopes during the /p/ and some

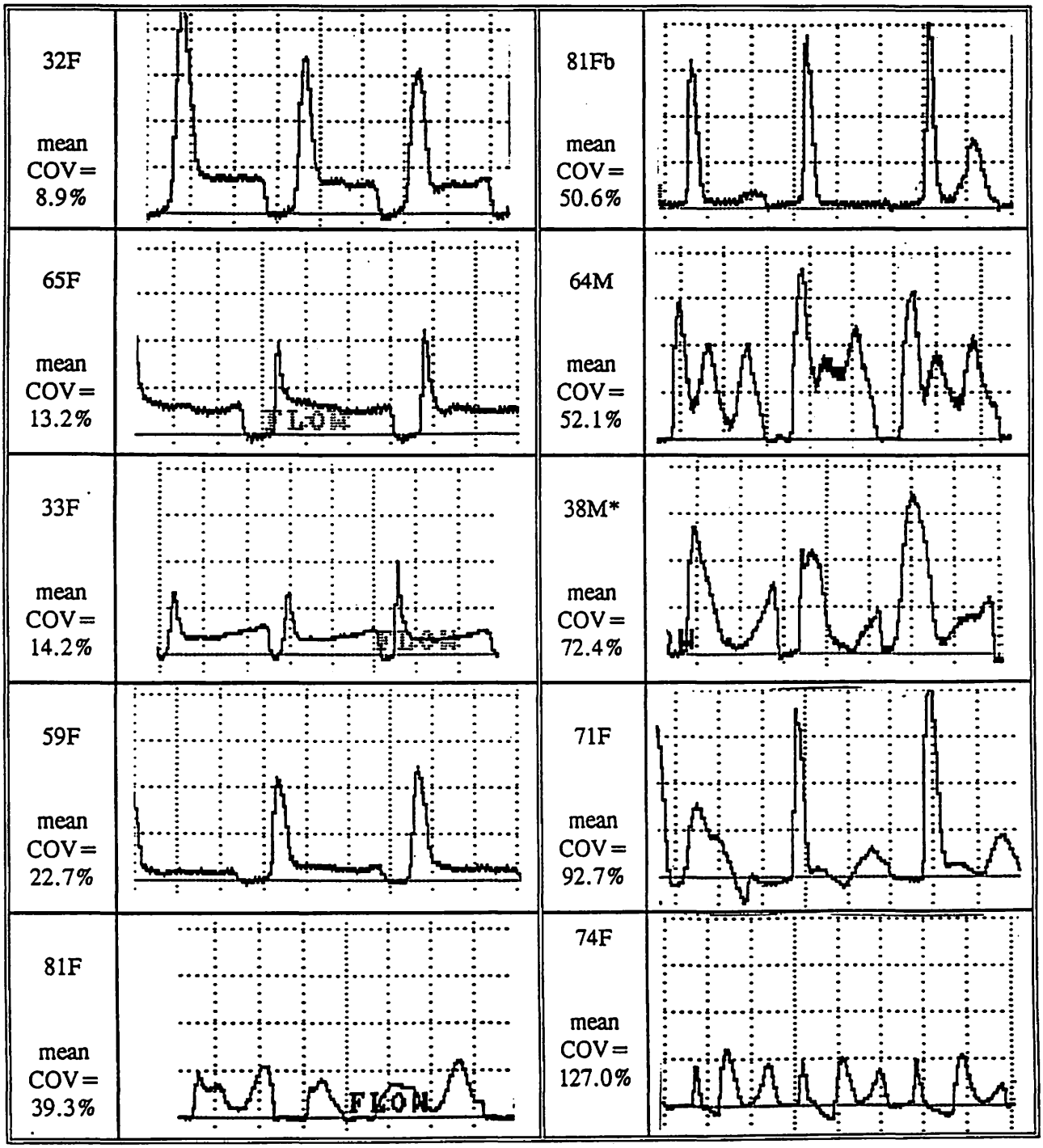


*Figure #4: Airflow tracings from 10 SD subjects. The mean coefficient of variation of airflow (COV) for each subject was calculated from the three tokens pictured plus 6 additional tokens from the first and third trials. Time is displayed on the horizontal axis (200 msec per grid mark); amplitude of airflow is on the vertical axis (.25 LPS per grid mark). [Note: *ABSD patient]*

(81F, 71F) demonstrated unstable baseline oral pressure during phonation. Irregular peaks and/or sudden negative or positive pressures in the oral cavity during the vowel could be due to instabilities in respiratory drive and/or changes in laryngeal configuration.

If stable subglottal pressure cannot be assumed when airflow is unsteady, estimation of subglottal pressure from oral pressure is called into question. Direct measurement of subglottal pressure in patients with SD has been reported. Shipp, Izdebski, Schutte, & Morrissey (1988)

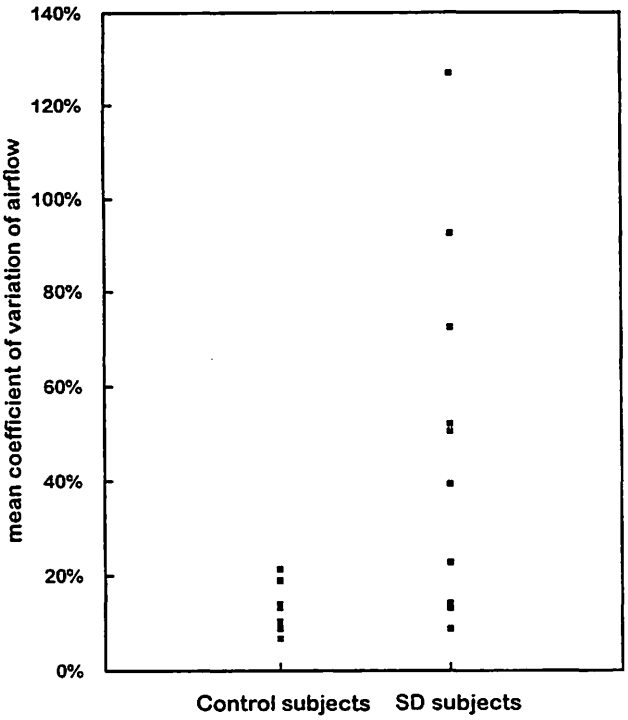


Figure #5: Coefficient of variation of airflow for 10 control subjects and 10 subjects with SD.

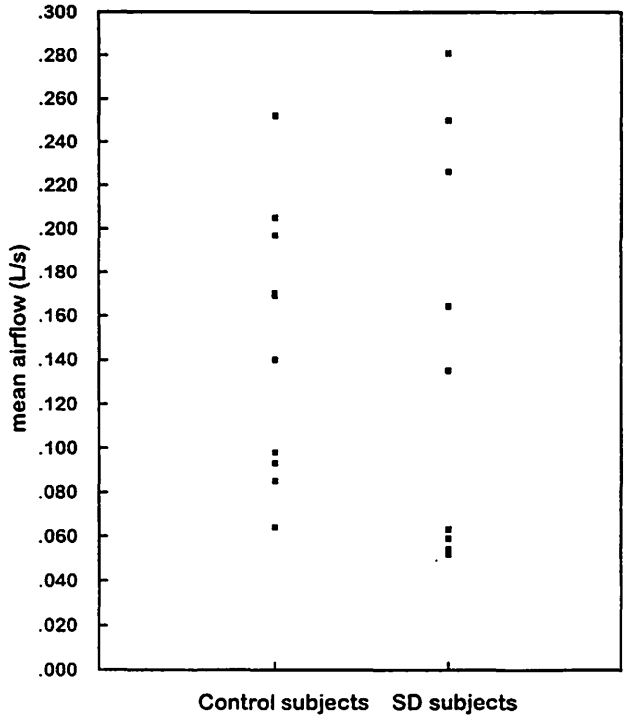


Figure #6: Mean airflow fro 10 control subjects and 10 subjects with SD.

Table 2.
Summary of Airflow measures from 10 subjects with SD. Listed in same order as Figure 4. Means based on nine measures, three central tokens from each of three trials. Shaded area highlights data from subjects with COVs similar to normal subjects.

Subject	COV of Airflow		Mean Airflow (LPS)
	Mean +/- SD	Range	
32F	8.9% +/- 2.5%	4.8% - 12.3%	.164
65F	13.2% +/- 4.3%	9.8% - 20.1%	.135
33F	14.2% +/- 3.9%	7.3% - 18.7%	.054
59F	22.7% +/- 7.9%	16.6% - 27.8%	.052
81F	39.3% +/- 13.9%	14.7% - 53.2%	.281
81Fb	50.6% +/- 32.1%	24.1% - 124.6%	.059
64M	52.1% +/- 18.1%	21.9% - 75.3%	.250
38M*	72.4% +/- 36.4%	21.3% - 127.1%	.226
71F	92.7% +/- 29.7%	32.4% - 127.3%	.063
74F	127% +/- 32.1%	74.3% - 179.2%	.063

measured subglottal pressure during reading of an all voiced sentence in two subjects with SD, using an esophageal balloon in one case and a tracheal puncture in the other. In addition to measuring average subglottal pressure they counted "the number, extent, and duration of short-term perturbations in the subglottal pressure trace". One subject showed a decrease in the number of perturbations during falsetto speech and during modal speech following RLN block compared with modal speech pre-block. Further direct investigation of subglottal pressure in these patients is needed.

Since it did not seem reasonable to estimate subglottal pressure in patients with unsteady airflow for the reasons delineated above, we did not calculate laryngeal airway resistance in those SD subjects. However, estimates of laryngeal resistance were determined for the four SD subjects with relatively steady flows. Table 3 lists mean estimated subglottal pressure, airflow, and laryngeal resistance for these subjects. Two SD subjects with steady flows (33F and 59F) exhibited higher than normal laryngeal airway resistance (Hoit and Hixon, 1992), and airflow at the low end of the normal range (Leeper and Graves, 1984; Holmes, Leeper, and Nicholson, 1994). The other two SD subjects (32F and 65F) exhibited laryngeal airway resistances that were high but were within two standard deviations above the mean for subjects with normal voices for their respective age groups (Hoit and Hixon, 1992; Holmes, Leeper, and Nicholson, 1994). The subjects with LAR within normal limits both exhibited higher than normal subglottal pressure. The finding of normal LAR in subjects with SD is in agreement with other laryngeal airway resistance data collected from SD patients. Witsell et al (1994) reported group mean laryngeal airway resistance of 61.0 +/- 17.5 cm H_2O/LPS in men and 151.7 +/- 25.9 cm H_2O/LPS in women with

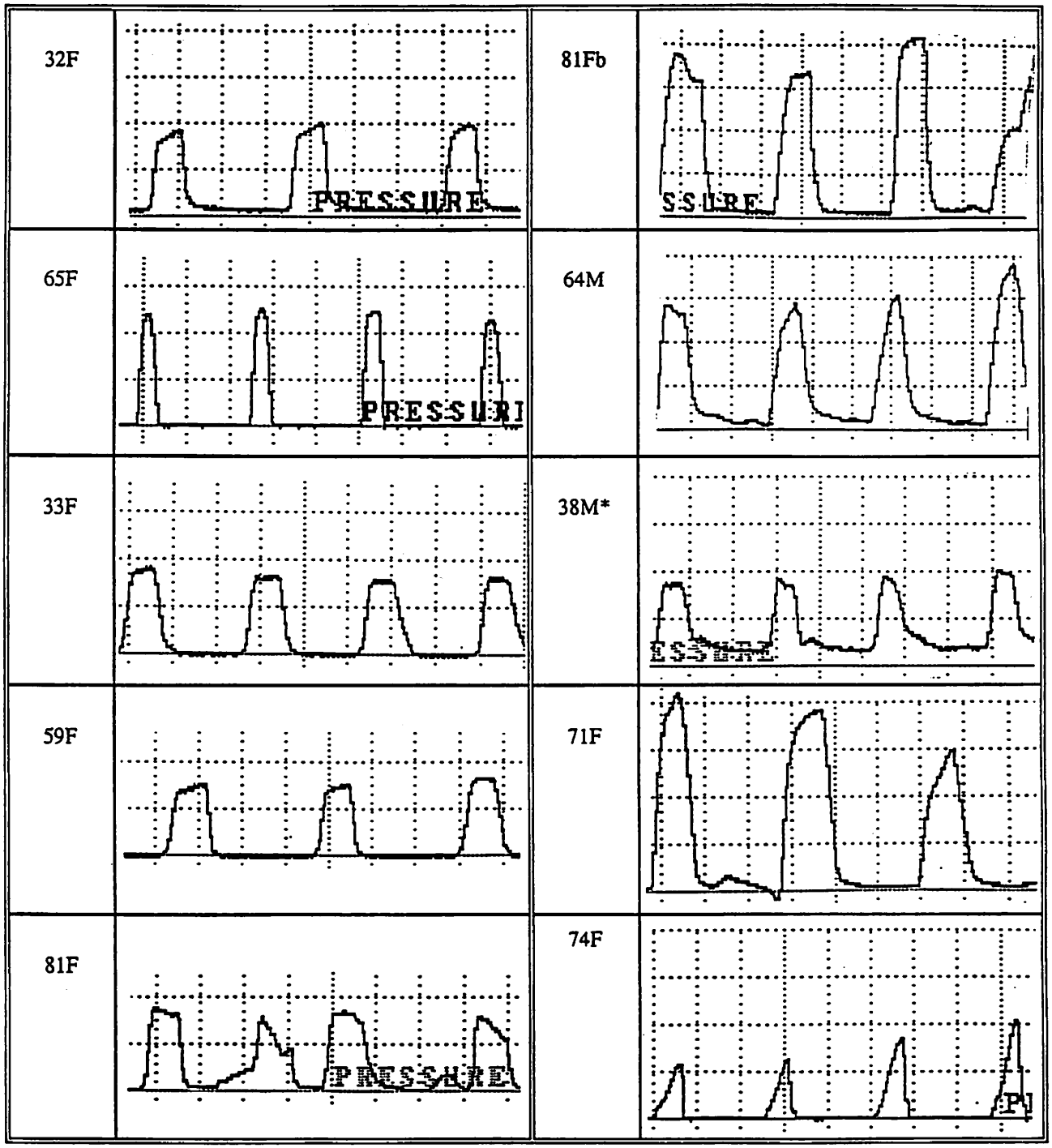


*Figure #7: Oral pressure tracings from 10 SD subjects. Listed in same order as Figure 4. Time displayed on the horizontal axis (200 msec per grid mark); amplitude of oral pressure is on the vertical axis (5.0 cm H_2O per grid mark). [Note: *ABSD patient]*

spasmodic dysphonia. Although individual data was not reported in the study by Witsell et al, the large variability suggests some of their male subjects may have demonstrated laryngeal airway resistances within two standard deviations above the mean for data from subjects with normal voices (Netsell et al, 1994; Holmes et al, 1994). See Appendix B for a summary of the LAR data collected from normal speakers.

The observation that some of the SD subjects exhibited steady airflow and normal laryngeal airway resis-

Table 3:
Mean estimated subglottal pressure, airflow, and laryngeal airway resistance of SD subjects with steady airflow. Each mean is based on measures from nine tokens. Shaded areas highlight data from patients with LAR estimate within normal range.

Subject	Estimated Ps (cm H_2O)		Airflow (liters/sec)		Laryngeal Airway Resistance (cm H_2O/LPS)	
	subject's mean	Norms	subject's mean	Norms	subject's mean	Norms
32F	8.1	6.43 +/- 1.87^	.163	.200 +/- .066^	52 +/- 11	38.33 +/-9.25^ 51 +/- 16.03+
33F	5.0	6.43 +/- 1.87^	.054	.200 +/- .066^	97 +/- 17	38.33 +/-9.25^ 51 +/- 16.03+
59F	7.7	3.6 +/- .5 *	.052	158 +/- .067*	163 +/- 62	27 +/- 11.1 * 55 +/- 20.72+
65F	9.6	4.0 +/- 1.2*	.135	125 +/- .056*	72 +/- 9	41 +/- 24.4 * 58 +/- 18.12+

^ Leeper and Graves (1984)
* Holmes, Leeper and Nicholson (1994)
+ Hixon and Hoit (1992)

tance during the syllable repetition task could indicate: (1) samples of strained phonation were not elicited during this task; or (2) some subjects with SD do not exhibit increased laryngeal airway resistance. The indirect estimation method relies on a specific utterance design so it is not possible to use this method with different speech tasks to elicit strained phonation. Thus, there are at least two problems that limit our ability to estimate subglottal pressure and/or laryngeal airway resistance data in patients with SD: 1) the presence of varying airflow and 2) the restriction of utterance type (i.e. the necessity of using the syllable repetition task, and the inability to collect data during other speech tasks). However, direct measurement of subglottal pressure is possible with tracheal puncture. This procedure would allow continuous collection of subglottal pressure data simultaneously with collection of airflow data using the standard methodology, resulting in calculation of time-varying laryngeal airway resistance. This is an invasive procedure that is generally justified as a research protocol but would be inappropriate for tracking change in voicing during the course of treatment. However, if variable resistance is suspected it may be informative to obtain a continuous rather than single measure of laryngeal airway resistance, at least during initial diagnostic procedures.

Estimation of laryngeal airway resistance is a noninvasive method that has proven useful in the study of normal and some voice disordered persons. However, clinical judgement should be exercised in the application of this method to patients with unstable airflow. Inability on the part of the patient to attain steady flows could compromise the usefulness of a midpoint measure of airflow and/or estimation of subglottal pressure, resulting in sources of error in estimation of laryngeal airway resistance. In this study of 10 subjects with SD, we were unable to determine laryngeal airway resistance in six SD subjects with unsteady flows. Laryngeal airway resistance was estimated in four SD subjects who were able to maintain steady airflow. Two of these subjects exhibited high laryngeal airway resistance with mean airflows at the low end of the normal range; the others exhibited normal laryngeal airway resistance with high subglottal pressure.

Conclusion

Acknowledgment

This research was supported by grant no. P60 DC00976-02 from the National Institutes of Deafness and Other Communication Disorders. We gratefully acknowledge the contribution of Dr. Emily Lin for assistance with data collection.

Bibliography

Bouhuys, A., Proctor, D., & Mead, J., (1966) "Kinetic aspects of singing" Journal of Appl. Physio. 21:483-496

Davis, P. J., Boone, D. R., Carroll, R. L., Darveniza, P., & Harrison, G. A. (1988). Adductor spastic dysphonia: heterogeneity of physiologic and phonatory characteristics. Annals of Otology, Rhinology & Laryngology 97:179-185

Higgins, M.B., & Saxman, J.H., (1991) "A comparison of selected phonatory behaviors of healthy Aged & Young adults" Journal of Speech & Hearing Research 34, 1000-1010

Hixon, T., (1972) "Some new techniques for measuring the biomechanical events of speech production: One laboratory's experience," ASHA Rep No. 7, 68-103

Hoit JD. & Hixon TJ. (1992) Age and laryngeal airway resistance during vowel production in women. Journal of Speech & Hearing Research 35(2):309-13

Holmberg, E. (1980) Laryngeal airway resistance as a function of phonation type, [Abstract] Journal of Acoustic Society of America Suppl 68, S101

Holmberg, E.B., Perkell, J.S., & Hillman, R.E., (1984) Methods for using a noninvasive technique for estimating glottal functions from oral measures. Working Papers Vol V: 47-58

Holmberg, Hillman, & Perkell (1988) Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice J. Acoust. Soc. Am. 84 (2): 511-528

Holmes, L.C., Leeper, H.A., & Nicholson, I.R., (1994) Laryngeal airway resistance of older men and women as a function of vocal sound pressure level Journal of Speech and Hearing Research 37 (4); 789-799

Katajima, K., & Fujita, F., (1990) Estimation of subglottal pressure with intraoral pressure Acta Otoloaryngol (Stockh) 109:473-478

Leeper HA. Gagne JP. Parnes LS. & Vidas S. (1993) Aerodynamic assessment of the speech of adults undergoing multichannel cochlear implantation. Annals of Otology, Rhinology & Laryngology, 102(4 Pt 1):294-302

Leeper HA Jr. & Graves DK. (1984) Consistency of laryngeal airway resistance in adult women. Journal of Communication Disorders, 17(3):153-63

Lewis JR. Andreassen ML. Leeper HA. Macrae DL. & Thomas J. (1993) Vocal characteristics of children with cleft lip/palate and associated velopharyngeal incompetence. Journal of Otolaryngology, 22(2):113-7

Lofqvist, A., Carlborg, B., & Kitzing, P., (1982) Initial validation of an indirect measure of subglottal pressure during vowels Journal of Acoustical Society of America, 72(2):633-35

McGuirt WF. Blalock D. Koufman JA. & Feehs RS. (1992) Voice analysis of patients with endoscopically treated early laryngeal carcinoma. Annals of Otology, Rhinology & Laryngology, 101(2 Pt 1):142-6

Melcon, M.C., Hoit, J.D., & Hixon, T.J., (1989) Age and laryngeal airway resistance during vowel production. Journal of Speech and Hearing Disorders, 54 (2): 282-286

Netsell R. Lotz WK. Peters JE. & Schulte L. (1994) Developmental patterns of laryngeal and respiratory function for speech production. Journal of Voice, 8(2):123-31

Netsell, R., (1969) Subglottal and intraoral air pressures during the intervocalic contrast of /t/ and /d/. Phonetica, 20:68-73

Rothenberg, M. (1973). A new inverse filtering technique for deriving the glottal airflow waveform during voicing. Journal of Acoustical Society of America, 53: 1632-1645

Shipp, T., Izdebski, K., Schutte, H., & Morrissey, P., (1988) Subglottal air pressure in spastic dysphonia speech, Folia Phoniat. 40: 105-110

Shipp T., (1973) Intraoral air pressure and lip occlusion in midvocalic stop consonant production Journal of Phonetics 1, 167-170

Smitheran, J. R., & Hixon, T. J., (1981) A clinical method for estimating laryngeal airway resistance during vowel production, Journal of Speech and Hearing Disorders 46: 138-146

Stathopoulos, E.T., & Sapienza C., (1993) "Respiratory and laryngeal function of women and men during vocal intensity variation" Journal of Speech and Hearing Research, 36(1), 64-75

Witsell, D.L., Weissler M.C., Donovan M.K., Howard, J.F., & Martinkosky, S.J., (1994) Measurement of Laryngeal Resistance in the Evaluation of Botulinum Toxin Injection for Treatment of Focal Laryngeal Dystonia, Laryngoscope 104: 8-11

Appendix A:
Airflow data collected during estimation of laryngeal airway resistance in normal adult speakers (20 - 65 years), repeating /pi/ at a comfortable level.

Gender	Age group	Airflow (Liters/second)			Source
		mean	SD	range	
FEMALE	15 "young adult women"	.200	.066	.078-.350	Leeper and Graves (1984)
	10 females (26-31 yrs) at 1.5 syl/sec at 3.0 syl/sec	 .127 .142	 .029 .036		Netsell, Lotz, Peters, and Schulte (1994)
	10 females 55.0 - 64.11 yrs 65.0 - 74.11 yrs	 .158 .125	 .067 .056		Holmes, Leeper, and Nicholson (1994)
MALE	10 males (21-36 yrs) at 1.5 syl/sec at 3.0 syl/sec	 .188 .212	 .051 .062		Netsell, Lotz, Peters, and Schulte (1994)
	10 males 55.0 - 64.11 yrs 65.0 - 74.11 yrs	 .249 .200	 .140 .079		Holmes, Leeper, and Nicholson (1994)

Appendix B:
Laryngeal airway resistance collected from normal adult speakers (20 - 65 years), repeating /pi/ at a comfortable level.

Gender	Age group	Laryngeal airway resistance (cm H2O/LPS)			Source
		mean	SD	range	
FEMALE	15 "young adult women"	38.33	9.25	14.04 - 65.62	Leeper and Graves (1984)
	10 females (26-31 yrs) at 1.5 syl/sec at 3.0 syl/sec	 43 44	 10 15		Netsell, Lotz, Peters, and Schulte (1994)
	10 women 25 +/- 2 yrs 35 " 45 " 55 " 65 "	 54.88 51.32 38.00 55.60 58.37	 15.04 16.03 13.74 20.72 18.12		Hoit and Hixon (1992)
	10 females 55.0 - 64.11 yrs 65.0 - 74.11 yrs	 27.2 41.7	 11.1 24.4		Holmes, Leeper, and Nicholson (1994)
MALE	15 males (21-40 yrs)	35.7	3.3	30 - 43.1	Smithern and Hixon (1981)
	10 males (21-36 yrs) at 1.5 syl/sec at 3.0 syl/sec	 34 37	 9 17		Netsell, Lotz, Peters, and Schulte (1994)
	10 men 25 +/- 2 yrs 35 " 45 " 55 " 65 " mean = 37.6	 39.36 35.69 40.72 35.97 38.11 mean = 6.0	 6.41 5.31 4.59 7.13 7.13 mean = 6.0		Melcon, Hoit and Hixon (1989)
	10 males 55.0 - 64.11 yrs 65.0 - 74.11 yrs	 .249 .200	 .140 .079		Holmes, Leeper, and Nicholson (1994)

NCVS Status and Progress Report - 7
December 1994, 77-94

Phonatory and Articulatory Changes Associated with Increased Vocal Intensity in Parkinson Disease: A Case Study

Christopher Dromey M.A., CCC-SP
Lorraine Olson Ramig Ph.D., CCC-SP
Department of Communication Disorders and Speech Science, University of Colorado-Boulder
and
Wilbur James Gould Voice Research Center, The Denver Center For The Performing Arts
Antonia Johnson, M.A.
Wilbur James Gould Voice Research Center, The Denver Center For The Performing Arts

Abstract

This study examined changes in voice and speech production in a patient with Parkinson disease as he increased vocal intensity following one month of intensive voice treatment. Phonatory function and articulatory acoustic measures were made before and after treatment as well as six and twelve months later. Pre- to post-treatment increases were documented in sound pressure level in sustained phonation, syllable repetition, reading and monologue. Consistent with mechanisms of intensity change reported in normal speakers, corresponding improvements were measured in estimated subglottal pressure, maximum flow declination rate, laryngeal airway resistance, open quotient, EGGW-25, harmonic-spectral slope and maximum vowel duration. Measures of phonatory stability in sustained phonation and semitone standard deviation in reading and speaking showed changes accompanying increased vocal intensity. In addition, changes were measured in articulatory acoustic parameters (vowel duration, whole word duration, transition duration, transition extent, transition rate) in single word productions. These findings indicate that this patient increased his vocal intensity using phonatory mechanisms that have been associated with the non-disordered larynx. In addition, the increased vocal intensity led to changes in articulation which were not targeted in therapy.

Introduction

Vocal intensity is a key variable in the production of intelligible speech (Moore, 1946; Pickett, 1956; Ramig, 1992). Subglottal air pressure (Isshiki, 1964), vocal fold adduction (Scherer, 1991; Titze & Sundberg, 1992) and vocal tract shape (Gauffin & Sundberg, 1989; Scherer, 1991) have been associated with the control of vocal intensity. Maximum flow declination rate (MFDR), which reflects the interaction of subglottal air pressure and vocal fold adduction, and is an index of the speed of glottal flow "shut-off," correlates highly with vocal intensity (Titze & Sundberg, 1992). The mechanism of intensity control has been described in normal (Holmberg, Hillman, & Perkell, 1988; Stathopoulos & Sapienza, 1993a, 1993b) and hyperfunctional voices (Hillman, Holmberg, & Perkell, 1989; Sapienza & Stathopoulos, in press), but has not been studied following treatment designed to improve intensity in patients with vocal hypofunction.

Increases in vocal intensity also can affect supraglottal articulator movements. Reorganization of articulatory movements and timing relationships of the lips and jaw and changes in acoustic segmental duration have been identified when normal subjects have produced loud speech (Schulman, 1989). Specifically, loud speech (90 dB SPL at 50 cm) elicited jaw openings greater than for normal speech, lip movement increases, and more complete lip closure. Movement changes were accompanied by shorter acoustic durations of intervocalic bilabial stops and longer vowels, resulting in a change in the relative timing of acoustic segments. The increased articulatory movement accompanying loud speech was not always a simple linear amplification of normal articulation, but reflected a more

complex goal-oriented reorganization of specific movements for the maintenance of phonetic features. For example, even as vowel duration increased, the phonological distinction between duration of long and short vowels and distinctions in vowel height was maintained. This goal-oriented reorganization of articulatory movements and timing relationships is consistent with descriptions in the speech production literature of functional units (e.g. coordinative structures or synergistic actions) governing the spatiotemporal interactions among glottal and supraglottal articulators (e.g., Browman & Goldstein, 1990; Fowler, 1980; Gracco, 1988, 1994; Kelso & Tuller, 1981; Kelso, Tuller, Bateson, & Fowler, 1984). Additionally, this cooperativity among articulators to meet phonetic task demands is consistent with clinical observations that treatment effects seem to generalize across the speech mechanism (Hardy, 1967; Netsell, 1986). Although segmental durations and timing relationships have been studied in a limited manner when normal subjects have produced comfortable and loud speech, they have not been studied following treatment designed to increase vocal intensity.

Reduced vocal intensity and disordered articulation contribute to the impaired intelligibility of many patients with Parkinson disease (Ramig, 1992). Speech treatment has focused on improving articulation and rate, but such approaches have met with limited success (Allan, 1970; Greene, 1980; Sarno, 1968; Weiner & Singer, 1989). Reduced loudness in Parkinson disease has been attributed to glottal

Table 1.
Rationale for the Lee Silverman Voice Treatment. The therapy is intensive with a focus on phonation and immediate carryover into functional communication.

Perceptual characteristics of speech	Hypothesized laryngeal and/or respiratory pathophysiology	Therapy goals and tasks	Acoustic, physiologic variables measured
Reduced loudness Breathy, weak voice	Bowed vocal folds Reduced glottal adduction	• Increase vocal fold adduction via isometric effort (pushing, lifting) during phonation	• Videolaryngostroboscopic examination • Open Quotient • EGGW-25
	Rigidity, hypokinesia, in laryngeal and/or respiratory muscles	• Increase maximum duration vowel phonation at greater vocal intensity • Think "shout/loud" • Be able to speak over ambient noise	• Maximum phonation time • Sound pressure level • Subglottal pressure • MFDR
	Reduced inspiratory and expiratory volumes	• Increase respiratory support via — Posture — Deep breath before speaking — Frequent breaths	• Respiratory excursions
	Rigid cricothyroid muscle	• Increase maximum fundamental frequency range • High/low pitch glides • Sustained phonation at highest and lowest pitches	• Variability of fundamental frequency in connected speech (STSD)

incompetence (Hansen, Gerratt, & Ward, 1984; Perez, Ramig, Smith, & Dromey, 1994; Smith, Ramig, Dromey, Perez, & Samandari, in press) and reduced respiratory support (Critchley, 1981) associated with respiratory and laryngeal muscle rigidity and hypokinesia (Hirose & Joshita, 1987). Treatments addressing vocal intensity in these patients have included efforts at behavioral therapy (Robertson & Thompson, 1984; Scott & Caird, 1983), electronic amplification, and thyroplasty (Paul Flint, M.D., Charles Ford, M.D., James Kaufman, M.D., personal communication). Only recently has a behavioral treatment program been developed for Parkinson patients which directly targets the mechanism underlying the reduced intensity by increasing phonatory effort and vocal fold adduction. The rationale and key treatment elements of the Lee Silverman Voice Treatment (LSVT) for Parkinson disease are summarized in Table 1 and have been described in detail elsewhere (Ramig, 1994; Ramig, Bonitati, & Horii, 1991; Ramig, Bonitati, Lemke, & Horii, 1994).

While previous investigations have reported increases in the perceptual variable of loudness (Ramig, et al., 1991; Ramig, et al., 1994; Ramig, Fazoli, Scherer, & Bonitati, 1990), and its acoustic correlate of vocal intensity (Smith et al., in press) following the LSVT in patients with Parkinson disease, no study has evaluated these post-treatment changes relative to the mechanisms of intensity control which have been observed in normal speakers. An evaluation of the phonatory mechanisms associated with post-treatment changes in intensity would offer insights into the capacity of the Parkinson disease patient to voluntarily regulate intensity as well as reveal details of compensatory vocal function.

Furthermore, given the movement displacement and velocity changes accompanying loud speech in normal speakers (Schulman, 1989), an articulatory acoustic analysis of the Parkinson patient's speech would offer insight into modifications in articulatory patterns concomitant with increased vocal intensity. Because of increased jaw displacement, and the associated vocalic durational changes, second formant trajectory differences in transition duration and extent would be expected. The predicted increases in these parameters are anticipated because the articulators have more time to reach their target positions. Velocity increases in lip and jaw during loud speech (Schulman, 1989) are expected to be accompanied by increased velocity of tongue movement as reflected in transition rate. This is anticipated because of research by Munhall, Ostry and Parush (1985) which indicated a uniform basis for the temporal coordination of speech articulators. Knowledge of the phonatory and articulatory patterns associated with intensity change in Parkinson disease following treatment may contribute useful information to theories of voice and speech production as well as enhance the efficacy of treatment.

Method

Subject

The subject for this study was selected from a larger group of patients with Parkinson disease who were participants in an investigation of voice treatment efficacy. At the time the study began, he was forty-nine years old and employed as a family physician. He had been diagnosed with idiopathic Parkinson disease for two years and was in Stage II on the Hoehn and Yahr Scale (1967). His Parkinson disease medications included sinemet and eldepryl. There was no change in medication during the course of the study and neurological testing showed no progression of Parkinson symptoms. Neuropsychological testing revealed some possible mild attentional difficulties at the start of the investigation with no progression of symptoms during the year of the study.

Initial clinical speech examination revealed a normal oral peripheral mechanism both in structure and function; hearing was within normals limits. At the initial speech evaluation, the patient reported that his voice had become softer during the previous year and was raspy at times. He had associated his raspy voice with frequent upper respiratory infections, but it was his impression that people could understand him most of the time.

This subject was selected for participation in this case study because he was representative of early stage Parkinson disease and his livelihood and daily living were dependent upon his oral communication. In addition, the validity, reliability and scope of his multi-channel experimental data were considered excellent.

Equipment

Multi-channel recordings were made of the subject's speech production as he performed a series of tasks. The data were stored on a Sony PC-108M 8 channel digital audio tape (DAT) recorder. In addition, microphone and EGG signals were recorded on a Panasonic SV 3700 2 channel DAT recorder which allowed higher bandwidth storage for subsequent acoustic analysis. Digitization was performed using a 16 bit Digital Sound Corporation A/D converter to a VAX 4000/200 computer and a Data Translation DT 2821 to a 486 PC.

A sound level meter (Bruel and Kjaer Type 2230) was positioned 50 cm from the subject's lips and a head-mounted microphone (AKG C410) was located 8 cm from the lips to record sound pressure level and acoustic signals respectively. A Synchrovoice Research Electroglottograph (EGG) was used to obtain the electroglottographic signal. A Rothenberg mask (Glottal Enterprises MS 100-A2) was held on the subject's face to collect the oral air flow signal from which the glottal flow waveform was derived. An intraoral pressure tube rested in the center of the oral cavity to allow the estimation of subglottal air pressure during /p/ closure. A

pair of strain-gauge belt pneumographs (Respigraph-NIMS PN SY03) were used to measure rib cage and abdomen circumferential movement (Murdoch, Chenery, Bowler, & Ingram, 1989). A Collins wet spirometer (Model RS 2785) was used to measure forced vital capacity. Endoscopic examination was conducted with both an Olympus ENF-P3 fiberscope and Nagashima SFT-1 70 rigid telescope.

Tasks

Data were collected while the subject performed the following tasks: tidal volume, forced vital capacity, maximum duration sustained vowel phonation, maximum fundamental frequency range, a series of /pae/ syllables (Rothenberg, 1973; Smitheran & Hixon, 1981), reading the "Rainbow Passage" (Fairbanks, 1960), reading 70 individual words (Kent, Kent, Weismer, Martin, Sufit, Brooks, & Rosenbek, 1989; Weismer, Martin, Kent and Kent, 1992), and a 30 second monologue.

Procedure

Pre-treatment experimental data were collected twice within the week preceding treatment and twice during the week following treatment, and once six and twelve months later. The following procedure was carried out for each data collection session. After the respigraph bands were taped in position, the subject was seated in a medical examining chair in an IAC sound-treated booth. To limit extraneous movement, the subject's arms and legs were secured to arm and foot rests using three inch wide Velcro bands. After two minutes of tidal breathing, forced vital capacity (FVC) was measured. The subject was asked to take his deepest breath and blow out "as hard and fast and long as you can". This task was repeated three times at the beginning of the session and twice at the end. The best performance was taken as the FVC. The subject was asked to "inhale and exhale" using a 700 mL inspirese bag. Care was taken to ensure that the subject's lips were tightly sealed on the mouthpiece for both the FVC and inspirese tasks; to prevent any nasal air flow, nose clips were used for both procedures.

To determine the maximum duration of sustained vowel phonation, the subject was instructed to "take a deep breath and sustain /a/ for as long as you can." A timer with a second hand was within view of the subject and he was encouraged to monitor his performance and sustain phonation maximally for each vowel. No instructions were given regarding loudness level. This task was repeated four times at the beginning of the recording session and twice at the end of the session. For analysis of respiratory excursions associated with vowel prolongations, the interval between vowels included a sigh and three to five tidal breaths. Because of potential instabilities in measures of maximum performance associated with lack of consistently high effort on the part of the subject (Kent, Kent, & Rosenbek, 1987), as well as the variability in the performance of Parkinson disease patients (Canter, 1965; King, Ramig, Lemke, & Horii, 1994), the experimenter was careful to elicit consistent high effort from the subject for each trial in each recording session as determined by her clinical judgment.

For the collection of air flow and intraoral air pressure data, the subject repeated a series of seven /pae/ syllables with the Rothenberg mask held in place by the experimenter and the air pressure tube in the middle of the oral cavity. Syllables were produced at normal pitch and loudness and flat intonation at a rate of 1.5 syllables per second, as modeled by the experimenter. The flow and pressure signals were monitored on an oscilloscope to ensure correct task performance and the subject repeated the task series three times. Samples of reading and spontaneous speech were obtained by asking the subject to read aloud the phonetically balanced "Rainbow Passage" at a comfortable rate and loudness and to generate a 30 second monologue on a topic of his choice.

To gather data for the analysis of articulatory acoustics, the subject was asked to read individual words "as clearly as possible" from a series presented individually by videotape at approximately five-second intervals. This style of presentation was used in order to control for changes in speaking rate. The 70 words in the test were all monosyllabic and of the form CV, CVC, CCVC, CVCC, CCVCC and VC; data from 12 of the words will be described here. These words are "wax, sigh, sip, ship, sew, hold, row, cash, hail, ate, shoot and blend". These words were selected for acoustic analysis because they represent various patterns of F2 trajectories (e.g., gradually rising, sharply falling, etc.) that have been previously used in studies investigating the relationship between speech intelligibility and the slope or transition rate (TR) of the second formant in dysarthric patients (Ansel & Kent, 1992; Kent et al., 1989; Weismer et al., 1992). These words were also chosen because they represent a variety of vocalic nucleus types, including monophthong and diphthong, but were not intended to sample parametrically the range of vocalic nucleus types in American English. Trajectory data for 11 of the words (the exception being "hold") spoken by normal geriatric subjects and ALS speakers, have been reported previously (Weismer et al., 1992).

Laryngeal imaging and videolaryngostroboscopic examination were conducted using well-documented techniques (Bless, Hirano, & Feder, 1987). These data were collected within the week of the experimental voice recordings, but not on the same day.

Treatment

The subject, who had received no previous speech therapy, participated in 16 sessions of the Lee Silverman Voice Treatment within a four week period. The treatment was designed to increase vocal intensity by increasing phonatory effort, vocal fold adduction and respiratory support. Therapy sessions included maximum phonatory effort

tasks such as drills on maximum duration // phonation and maximum phonation frequency range for half of the session. The subject was encouraged to use this increased phonatory effort in louder speech production during the other half of the session. The details of this treatment approach have been summarized elsewhere (Ramig, 1994). It is important to point out that no effort in treatment was directed toward improving the patient's articulation. In the course of practicing loud voice, he may have spontaneously readjusted articulation to preserve phonetic distinctions, but during treatment his attention was directed exclusively to increasing vocal loudness.

The clinician reported that the subject was a cooperative and motivated patient once he became convinced that a louder voice would help people understand him. He attended all sessions and followed through on homework assignments. After the initial 16 sessions, the subject received no additional voice treatment.

Data Analysis

The variables selected for measurement and analysis allow inferences regarding the changes following treatment designed to increase vocal intensity. Phonatory, respiratory and articulatory acoustic measures previously associated with the control of vocal intensity were targeted.

SPL: The signals from the sound level meter recorded during sustained phonation, reading and monologue were digitized at 1 KHz and analyzed by custom software to derive the mean and standard deviation of intensity during these tasks. In order to eliminate the contribution of any silent segments to the SPL analysis, a cursor system was employed to set a floor criterion corresponding to the lowest level of SPL during any phrase in the recording. The data points from the sound level meter which fell above this level were analyzed for central tendency and variation. Analyses were carried out on a VAX 4000/200 computer.

Aerodynamic analysis: Simultaneous recordings of air flow, intensity and intraoral air pressure were digitized at 20 KHz and analyzed with custom software which interpolated between the pressure peaks during /p/ closure to allow the estimation of mid-vowel pressure. Flow and SPL values for their respective time-aligned channels were also obtained in this way, so that values for these measures represented the temporal midpoint of the vowel in each syllable. All analyses were performed on a VAX 4000/200 computer.

Laryngeal airway resistance was calculated by dividing the estimated mid-vowel subglottal pressure by the mean mid-vowel air flow. To measure maximum flow declination rate, the air flow signal was transferred to a 486 PC and was inverse filtered with CSpeech 4.0. The maximum flow declination rate was measured as the magnitude of the downgoing peak from the derivative of the glottal flow signal for 10 successive cycles at the vowel midpoint.

Open quotient was measured using a custom software program (Stathopoulos & Sapienza, 1993a, 1993b) at a 20% AC flow criterion level. Between 30 and 50 consecutive cycles were measured from the glottal flow waveform at the vowel midpoint during the /pae/ task. The mean values from three vowels from each of three trials for each session were calculated.

EGG: The electroglottographic (EGG) signal was analyzed with an in-house software program on a VAX 4000/200 computer to derive a measure called EGGW-25. This measure is based on the relative width of the EGG duty cycle at 25% of its height, and has been found to correlate with other measures of glottal adduction (Scherer & Vail, 1988).

Respiratory volume: Measures of lung volume were derived from the sum of the signals from the abdominal and rib cage bands, which was digitized at 300 Hz. Calibration values were derived from the voltage of these signals during the 700 cc inspirese exchange task. Initiation and termination volumes were measured for the phonation tasks relative to resting expiratory level as established during tidal breathing before each task. These analyses were performed using CSpeech.

Sustained phonation: Measures of phonatory stability were calculated after digitizing the sustained vowel phonation at 20 KHz. These digitized signals were analyzed with GLIMPES software (Titze, 1984) on a VAX 4000/200 computer to obtain the following measures: fundamental frequency, jitter and shimmer (with and without linear trend removed), coefficient of variation for frequency, coefficient of variation for amplitude, harmonic spectral slope and harmonics-to-noise ratio. The duration of each sustained vowel phonation was measured as a part of the sound pressure level program.

F0 and STSD: Mean fundamental frequency and fundamental frequency variability during reading and monologue were obtained after digitizing the microphone recordings at 5 KHz and analyzing them with CSpeech on a 486 computer.

Articulatory acoustics: Measures of articulatory acoustic variables were made after digitizing the 12 selected test words at 20 KHz on a 486 computer. Wide-band spectrographic (300 Hz) and waveform displays of the digitized signals were generated by CSpeech. The intensity level of the spectrogram was adjustable with a constant 45 dB range. The floor was set at -68 dB which was determined to be just above the level of ambient noise to assure that low intensity portions of the speech signal were included. The CSpeech program sampled x-y (time-frequency) coordinates 10 times per second. The sampled data were stored in a spreadsheet program as columns of time-frequency values. These files served as basis for quantitative analyses of the formant trajectory characteristics of each word.

Temporal measures included whole word duration and the duration of the vocalic nucleus. Measures of the formant trajectory included transition extent (TE), transition duration (TD), and transition rate (TR) or slope. Words were segmented and measured according to conventional criteria (Klatt, 1975, Peterson & Lehiste, 1960). Measurement of whole word duration included pre-and post-aspiration in order to more accurately assess articulatory agility in opening and closing the glottis and vocal tract to air flow. Vowel durations (VD) were measured from the first to last large-amplitude, complex shape glottal pulse, following measurement criteria defined by Weismer, Kent, Hodge, and Martin (1988) and Weismer et al. (1992). Additional procedures for determination of VD included using audio playback to differentiate vowels from liquids or semivowels, and inspection of the spectrogram for the initial and final vertical striations (glottal pulses) observed in the first and second formants, which were used to verify and validate the time waveform display.

Transitional segments of a trajectory for F2 were identified through direct inspection of the time-frequency files as determined by a rule developed by Weismer et al., (1988). This rule operationally defines the onset of a transitional segment as the first instant where there is a frequency change in F2 of at least 20 Hz within a time increment of 20 msec and offset as a succeeding instant where a time increment of 20 msec is not accompanied by a frequency change of 20 Hz or more. In cases where more than one interval met the criterion of a transition segment (i.e. first instance) within a single trajectory, the steepest slope corresponding to that obtained for normal geriatrics (Weismer et al., 1992) was used in the analysis. This criterion was used because of the finding in PD subjects of unusual formant movements at the beginning of a vocalic nucleus. This phenomenon has previously been described for ALS subjects by Weismer et al. (1992).

Transition measurements followed conventional rules developed by Weismer et al. (1988). Transition extent (TE) was operationally defined as the frequency change along the transitional segment. Transition duration (TD) was operationally defined as the duration of the transitional segment. Transition rate (TR) or slope was defined as TE/TD. The transition rate can be either positive or negative depending on the direction of F2 movement.

Temporal data were analyzed for individual words and for groups of words classified as monophthong or diphthong. This classification was chosen because research by Forrest, Weismer and Turner (1989) has shown that when PD dysarthric speakers are compared to normal geriatric (NG) speakers, they tend to have reduced second formant durations and transition extents for complex vowels (which correspond to diphthongs), but not for monophthongal vowels. In this study, the categories of monophthong and diphthong were defined according to Shriberg and Kent (1982). Temporal and trajectory data are presented for these phonemic categories to allow a determination of commonalties in temporal and trajectory characteristics of vowellike sounds produced with a gradually changing articulation (diphthong) compared to relatively static vocal tract configuration (monophthong). Note that one of the diphthongs /aI/ (sigh) is phonemic and cannot be reduced to a monophthong. The other diphthongs, /eI/ (ate) and /oU/ (sew) are nonphonemic and can be reduced to monophthongs. 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/ usually are found in weakly stressed syllables and occur much less frequently than the diphthongal forms. Since all words in this study with nonphonemic diphthongs had one syllable, they have been classified as diphthongs.

This study also includes a qualitative level of acoustic analysis, since traditional acoustic measures of temporal and spectral characteristics may not reveal the inherently "important" aspects of disordered speech production (Liss & Weismer, 1992). Qualitative analysis is the visual inspection and interpretation of acoustic data from spectrograms for evidence of pre- to post-treatment differences in the emergence of new (or different) spatiotemporal patterns. Specifically for this study, pre- and post-treatment waveforms and spectrograms for the words beginning with voiceless fricatives /s, symbol 242 \f "Symbol" \s 12 ,h/ were printed and visually inspected to identify changes in spatiotemporal patterning associated with peak glottal opening (widening) for voiceless fricatives and the degree of coupling between successive articulatory gestures (e.g., a consonant followed by a vowel).

Videostroboscopy: The videostroboscopic images were rated for the type of glottal closure (e.g., normal, bowed,. posterior gap), as well as the degree of glottal incompetence, and supraglottal hyperfunction (i.e., false fold and anterior-posterior hyperadduction). The ratings were done as part of a larger study (Smith et al., in press) by four trained judges (three speech pathologists and an otolaryngologist). Randomized images were presented without an audio channel and were rated blindly by the judges.

Reliability

Thirty percent of all the phonatory function data on each measure were reanalyzed to assess measurement reliability. A paired t-test revealed no significant differences between original and repeated analyses. Pearson correlation coefficients ranged from 0.998 to 1.000 between the original and the reanalyzed data.

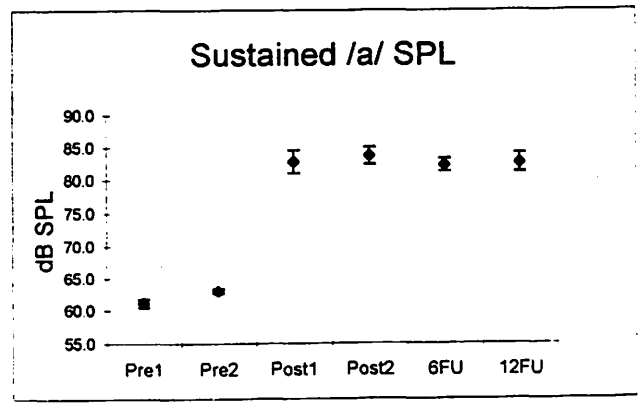
All segment duration and formant trajectory acoustic measurements were made by the third author. Twenty percent of the productions of each word were selected to assess intrajudge reliability. Spectrograms and waveforms were resegmented and remeasured. Reanalysis of the vowel segment durations yielded an average measurement differ-

ence of 3.7 ms (range 0-10 ms). Measurement differences tended to be greater for vowel durations of diphthongs and vowels flanked by the liquid /l/ (ranging from 4-10 ms), and least for vowels of short duration, and closed monophthongs (0 ms). Remeasurement for whole word durations yielded a measurement difference of 20 ms or less. The difference values do not exceed the differences of interest reported here.

Visual inspection of the data was performed to examine the main trends over time. The application of inferential statistical techniques to single-case studies is debatable (Barlow & Hersen, 1984; Kratochwill & Levin, 1992; McReynolds & Kearns, 1983). The present approach, which incorporated an irreversible treatment, makes the majority of these tests inappropriate, and a visual inspection of the data more suitable, particularly since our purpose is to examine and describe changes rather than to generalize treatment results to a population of patients.

Results

The goal of this study was to evaluate the mechanisms associated with changes in vocal intensity in a patient with Parkinson disease. The choice of variables and measurements for this study was guided by reports of the mechanisms of intensity change in non-disordered speakers. However, since the performance of many patients with movement disorders is frequently affected by extraneous movements (tremor, dyskinesia), cognitive limitations and fatigue, the application of conventional speech measurement techniques must be approached with caution. In many cases, simultaneous measurement of several apparently correlated variables can be essential in order to obtain an understanding of how an individual with a neurogenic disorder produces speech.



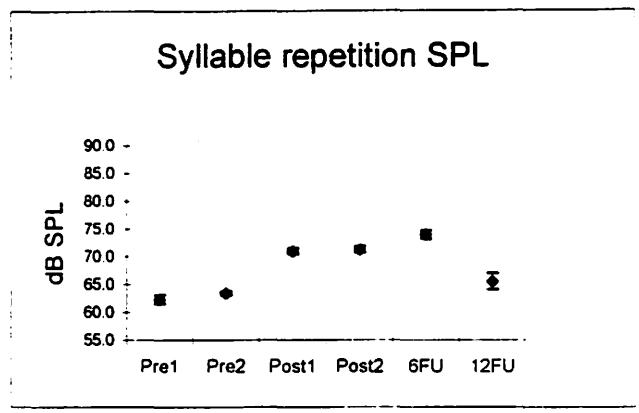


Figure 1. Means and standard deviations for sound pressure level (dB SPL at 50 cm) for vowels and /pae/ syllables.

Table 2.
Means and standard deviations for sound pressure level (dB SPL at 50 cm) for vowels, syllables, reading and monologue.

	Pre1		Pre2		Post1		Post2		6FU		12FU	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
sustained /a/	61.1	0.65	63.0	0.35	82.8	1.73	83.7	1.33	82.3	0.98	82.7	1.43
/pae/ syllables	62.3	0.80	63.4	0.36	70.9	0.54	71.2	0.57	73.8	0.73	65.5	1.46
reading passage	60.5	2.80	60.7	2.66	73.1	3.79	72.8	3.67	72.8	2.93	74.6	4.03
monologue	59.1	3.27	58.5	1.94	65.2	4.00	63.7	4.24	65.0	2.81	68.8	2.89

SPL

Sound pressure level (SPL) data for sustained vowel phonation, vowel midpoint for the /pae/ syllable series, reading and monologue are presented in Table 2 and two of the measures are plotted in Figure 1 (see previous page). The mean sound pressure level for all conditions increased following therapy and remained above pre-treatment levels through the 6 and 12 month follow-up recordings. Sound pressure level for vowels made the greatest pre- to post-treatment increase (20 dB), followed by reading (12 dB) and then monologue (5-6 dB).

Aerodynamic measures

The simultaneous aerodynamic and sound pressure level data for the /pae/ syllable repetition task are presented in Table 3 and four of the aerodynamic variables are plotted in Figure 2. Sound pressure level during the /pae/ task is lower than for the sustained vowels or reading task, and similar to the sound pressure level during the monologue. Estimated subglottal pressure increased 2-3 cm from the pre- to post-treatment condition and remained above pre-treatment levels through the 12 month follow-up recording. Maximum flow declination rate increased over 300 L/s/s from the pre- to the post-treatment condition and remained above pre-treatment levels through the 12 month recording.

Laryngeal airway resistance increased following treatment by 0.015 cmH20/cc/s and remained above pre-treatment levels through the 12 month follow-up recording. The mean mid-vowel air flow decreased somewhat following treatment, returning almost to the level of the first pre-treatment recording by the time 12 months had passed. Open quotient, as measured at the 20% AC flow criterion, decreased by approximately 0.1 following the intensive voice therapy and remained below pre-treatment levels through the 12 month follow-up. Forced vital capacity data are presented in Table 3 and show no pattern of difference throughout the course of the study.

Respiration

Respiratory kinematic data are presented in Table 4 and show that during maximum duration sustained vowel phonation, the subject consistently exhaled further below resting expiratory level immediately post-therapy when com-

Table 3.
Means and standard deviations of aerodynamic variables for /pae/ syllable repetition.

	Pre1		Pre2		Post1		Post2		6FU		12FU	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
dB	62.3	0.8	63.4	0.4	70.9	0.5	71.2	0.6	73.8	0.7	65.5	1.5
Psub	4.8	0.8	5.0	0.4	7.7	0.2	6.6	0.9	6.7	0.3	5.9	0.3
MFDR	240.6	11.7	253.3	12.6	540.0	17.1	580.0	14.8	614.4	36.6	416.2	13.1
Rlaw	0.023	0.003	0.022	0.002	0.037	0.005	0.037	0.005	0.040	0.006	0.030	0.003
Mean Flow	206.4	13.7	231.1	18.0	210.0	22.5	179.3	10.0	168.6	17.6	198.9	21.3
OQ-20	0.56	0.04	0.58	0.05	0.46	0.03	0.45	0.03	0.48	0.02	0.52	0.01
Max V.C.	4.8		4.5		4.8		4.6		4.7		4.7	

dB = midvowel sound pressure level (dB SPL at 50 cm)
Psub = estimated subglottal pressure (cmH2O)
MFDR = maximum flow declination rate (L/s/s)
Rlaw = laryngeal airway resistance (cmH2O/cc/sec)
Mean flow = mean midvowel airflow (cc/sec)
OQ-20 = open quotient using 20% AC flow criterion
Max V.C. = maximum vital capacity for session

pared to pre-therapy. Six months after treatment, the subject used both higher initiation lung volumes and lower terminations during this task. Twelve month respiratory data are not presented because of technical difficulties.

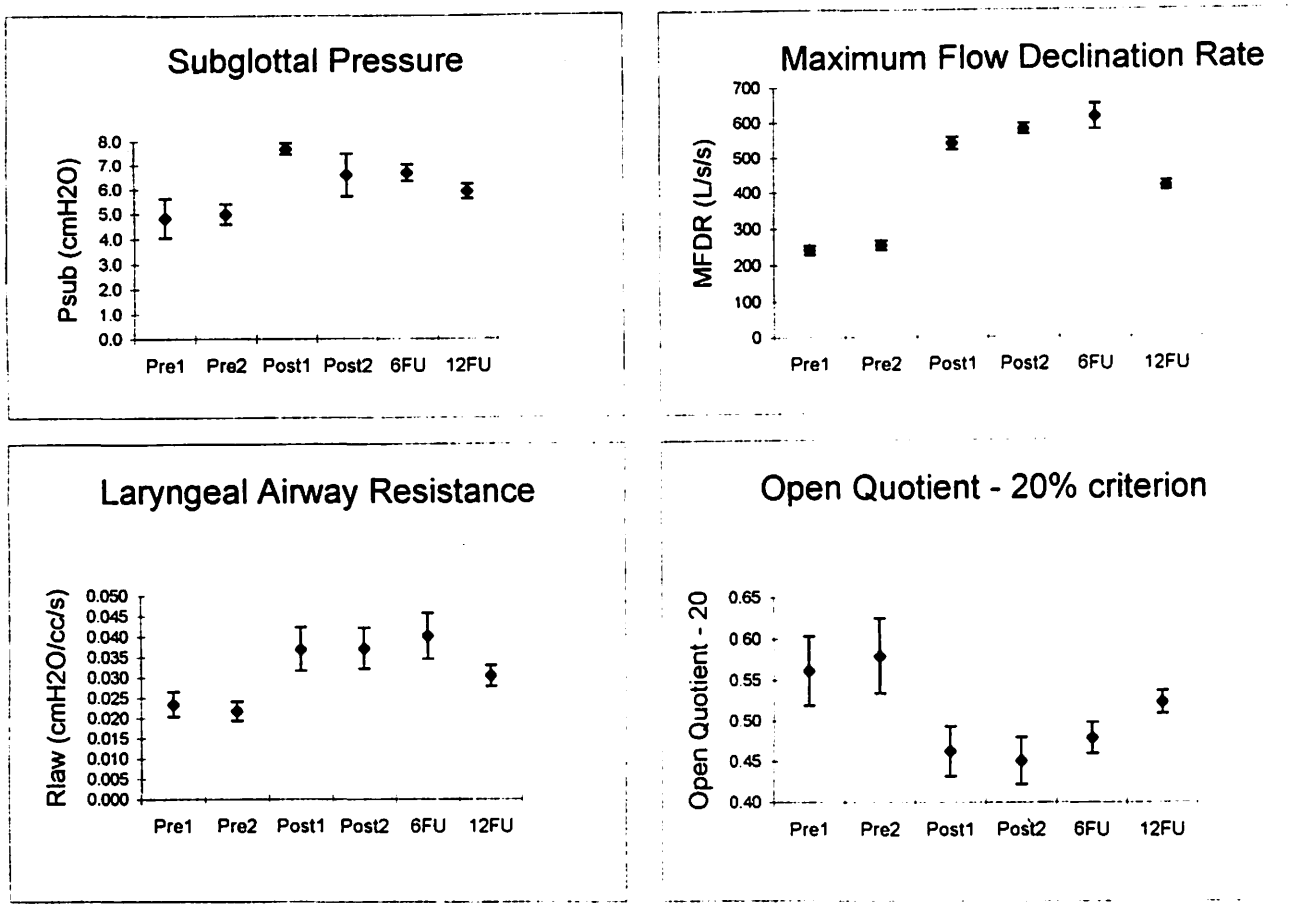


Figure 2. Means and standard deviations for estimated subglottal pressure, maximum flow declination rate, laryngeal airway resistance and open quotient.

Table 4.
Respiratory mean initiation and termination values for sustained vowels (liters relative to REL).

	Pre1		Post1		6FU	
	M	SD	M	SD	M	SD
Initiations	1.61	0.13	1.79	0.09	2.48	0.08
Terminations	-1.40	0.25	-2.22	0.02	-2.27	0.53

Notes:
Initiations = lung volume relative to REL at which vowel was initiated
Terminations = lung volume relative to REL at which vowel ended

Table 5.
Means and standard deviations for sustained phonation measures.

	Pre1		Pre2		Post1		Post2		6FU		12FU	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
MPT	24.3	2.5	23.3	2.9	38.6	2.2	41.8	2.2	40.3	2.7	35.0	2.0
SPL	61.08	0.648	62.96	0.347	82.83	1.733	83.71	1.329	82.31	0.977	82.74	1.429
EGGW-25	0.508	0.06	0.529	0.08	0.645	0.023	0.681	0.009	0.654	0.01	0.589	0.018
F0	119.3	2.4	126.0	0.8	176.8	11.0	173.0	7.2	152.3	7.1	147.8	3.5
HSS	8.36	0.95	9.22	0.45	6.39	1.33	4.88	0.98	2.59	0.81	4.53	0.82
HNR	19.95	2.66	19.98	2.14	24.57	1.02	25.11	1.53	23.01	1.19	22.76	1.64
CVA	9.96	3.69	8.88	2.69	3.60	1.05	2.03	0.67	3.89	1.54	4.78	1.46
Shimmer 1	2.57	0.59	2.46	0.53	1.28	0.60	0.95	0.21	1.32	0.18	1.58	0.61
Shimmer 2	1.74	0.51	1.90	0.53	0.97	0.54	0.75	0.19	1.02	0.16	1.20	0.54
CVF	0.61	0.13	0.56	0.02	0.41	0.14	0.50	0.10	0.58	0.12	0.49	0.08
Jitter 1	0.45	0.15	0.48	0.20	0.26	0.16	0.30	0.11	0.24	0.08	0.28	0.17
Jitter 2	0.40	0.16	0.43	0.21	0.23	0.15	0.27	0.11	0.20	0.07	0.25	0.17

MPT = maximum phonation time (seconds)
SPL = mean sound pressure level (dB SPL)
EGGW-25 - EGG pulse width adduction measure using 25% height criterion
F0 = mean fundamental frequency (Hz)
HSS = harmonic spectral slope (dB/octave)
HNR = harmonics to noise ratio (dB)
CVA = coefficient of variation for amplitude
Shimmer 1 = amplitude perturbation with linear trend
Shimmer 2 = amplitude perturbation - linear trend removed
CVF = coefficient of variation for frequency
Jitter 1 = frequency perturbation with linear trend
Jitter 2 = frequency perturbation - linear trend removed

Table 6.
Mean fundamental frequency (Hz) and semitone standard deviation for reading and monologue for six recording sessions.

	Pre1		Pre2		Post1		Post2		6FU		12FU	
	M (Hz)	STSD	M (Hz)	STSD	M (Hz)	STSD	M (Hz)	STSD	M (Hz)	STSD	M (Hz)	STSD
Reading	88.6	1.06	88.6	1.00	108.3	1.71	107.7	1.55	107.3	1.94	103.7	2.42
Monologue	91.7	1.72	85.6	1.31	98.5	1.86	105.4	2.43	92.1	1.18	93.5	1.77

EGGW

The data on EGGW25 are presented in Table 5. The values for this variable increased (reflecting more adduction) from pre-to post-treatment and remained above pre-treatment levels through the 12 month recording. This trend is consistent with that seen in open quotient, namely that the subject adducted his vocal folds more fully in the recordings following the treatment.

Sustained phonation, F0 and STSD

Acoustic measures from maximum sustained vowel phonation are presented in Table 5. Increases were measured in the duration (15-18 seconds) and sound pressure level (20 dB) of sustained phonation from the pre- to post-treatment condition which were maintained through the twelve month follow-up. Mean fundamental frequency increased 50 Hz pre- to post-treatment and at 12 months follow-up was 20 Hz above pre-treatment levels. Measures of amplitude perturbation (coefficient of variation of amplitude, and shimmer with and without linear trend) demonstrated increased amplitude stability pre- to post-treatment which was maintained through the 12 month follow-up. Measures of frequency perturbation (coefficient of variation of frequency, and jitter with and without linear trend) showed increased frequency stability pre- to post-treatment which was maintained through the 12 month follow-up. Spectral measures (harmonic spectral slope and harmonics to noise ratio) revealed increased high frequency spectral energy and reduced spectral noise pre- to post-treatment which was maintained through the 12 month follow-up.

Mean fundamental frequency and fundamental frequency variability (STSD) for reading and monologue are presented in Table 6. Both of these variables increased from pre- to post-treatment and, aside from the six month monologue sample, remained above pre-treatment levels through the twelve month follow-up.

Articulatory acoustic measures

Temporal parameters. Temporal data for word production are presented in Table 7 and the vowel and whole word duration data are plotted in Figure 3 (see following page). The data represent the mean segment durations and standard deviations for all words taken together. Also reported in Table 7 are mean vowel durations for the words with diphthongs (sigh, sew, row, hail, ate, hold) and mean vowel durations for words with monophthongs (sip, ship, wax, cash, blend, shoot).

The mean whole word duration for all words increased following therapy, but decreased below pre-treatment levels at the 6 and 12 month follow-up recordings. The average vowel duration for all words also increased following therapy and, though decreasing from post-treatment levels, remained above pre-treatment levels through the 6 and 12 month follow-up recordings. Thus, at the 6 and 12 month follow-up recordings, average vowel duration for all words was above pre-treatment levels while whole word duration was below pre-treatment levels. Although increases in vowel duration were greater for the more complex vocalic segments of diphthongs than for monophthongs, there was maintenance of the phonological distinction between durations of long and short vowels.

Qualitative acoustic analysis through visual inspection of spectrograms revealed that following therapy there was a decrease in duration of noisy frication for the initial glottal fricative /h/ accompanied by increased duration of vocalic nuclei. There was also an apparent decrease in duration of the noisy frication for initial voiceless fricatives (sip, sigh, sew, shoot, ship) along with an increase in vocalic

Table 7.
Mean and standard deviation segment durations of all words. Also mean and standard deviation segment durations for vowel monophthongs, ("wax, sip, ship, shoot, cash, blend") and diphthongs ("sew, hail, row, ate, sigh, hold").

	Pre1		Pre2		Post1		Post2		FU6		FU12	
ALL WORDS	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
VD (ms)	224	130	228	118	337	179	321	181	248	119	309	160
WWD (ms)	734	86	723	137	866	140	828	89	547	144	658	67
SELECTED WORDS												
VD-monoph (ms)	127	57	137	63	207	49	174	60	162	77	174	76
VD-diph (ms)	321	108	319	85	468	167	468	136	335	85	444	96

VD = vowel duration
WWD = whole word duration

nuclei duration immediately following therapy and in the 6 and 12 month follow-up recordings. Furthermore, visual inspection of waveforms and spectrograms indicated an apparent faster glottal opening (or widening) rise time for initial voiceless fricatives. Quantitative measurement of pre and post treatment frication duration and rise time were made using the operational definitions of Howell and Rosen (1983). Frication duration decreased post treatment when measurements were averaged across words and when individual words were compared. For this speaker, mean frication duration decreased from 133 to 82 ms for the glottal fricative /h/ and 933 ms to 801 ms for /s, ʃ/. Rise time decreased after treatment from a mean of 83 to 57 ms for /h/. Rise time decreased after treatment for three out of five words with initial /s, ʃ/ (sigh, sew, ship). Before treatment, mean rise time for all /s, ʃ/ fricatives was 509 ms (range = 174-771 ms); after treatment, mean rise time was 522 ms (range = 108-806 ms). For normal speakers, the mean rise time of isolated real-word fricatives in the initial position is reported to be 123 ms (Howell & Rosen, 1983).

Visual inspection of the spectrogram indicated a decreased gap between the initial voiceless fricative for "sigh" and the subsequent vowel following therapy and in the 6 and 12 month follow-up (see Figure 4 for representative spectrograms).

Trajectory Patterns. Measures associated with second formant trajectories are presented in Table 8 and selected measures are plotted in Figure 3. The mean transition duration (TD) for all words together increased over 30 ms from the pre- to the post-treatment condition and maintained that increase through the 12 month recording. Increases in transition duration following therapy were greater for the more complex vowel diphthongs (approximately 60 ms increase) than for monophthongs (with no overall increase). In the 12 month follow-up recording, although this difference between diphthongal and monophthongal change became smaller, the mean TD for words with diphthongs and monophthongs remained above pre-treatment levels. The greatest increase in TD post-treatment and the strongest maintenance of that increase were found for words with vocalic nuclei or off-glides in the high front position (e.g., ship, blend, sigh) and words reflecting lip rounding features (row and sew).

The mean transition extent (TE) of the second formant trajectory for all words taken together increased by 100 Hz following therapy and made further increases in the follow-up sessions. Mean TE values for diphthongs were practically identical to monophthongs before therapy, and

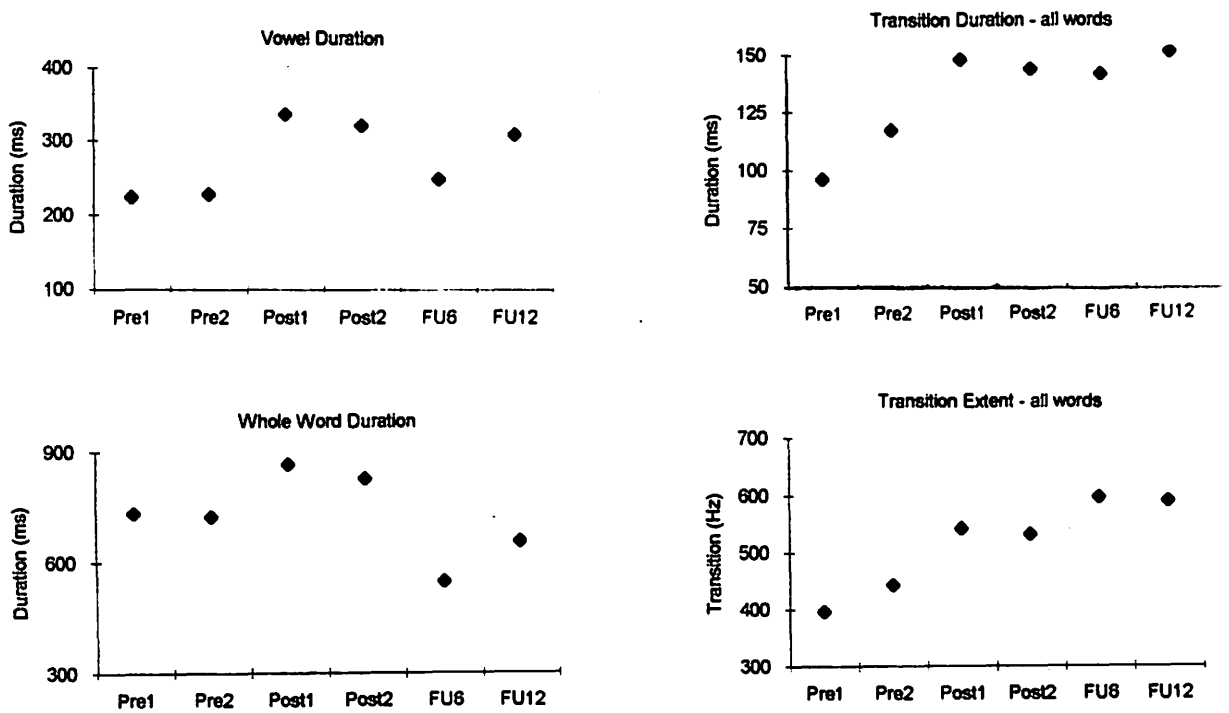


Figure 3. Mean values across all words for vowel duration, whole word duration, F2 transition duration, and transition extent.

both showed similar increases post therapy. However, in the follow-up conditions, monophthongs increased in mean TE by 60-140 Hz in contrast to diphthongs which only maintained their previous gains. In general, the greatest gains in TE for individual words were in words containing high front vowels and diphthongs with high front off-glides (e.g., sip, ship, blend, sigh).

Although the change in mean absolute transition rate (TR) was inconsistent for all 12 words immediately after therapy and in the follow-up recordings, in general there was

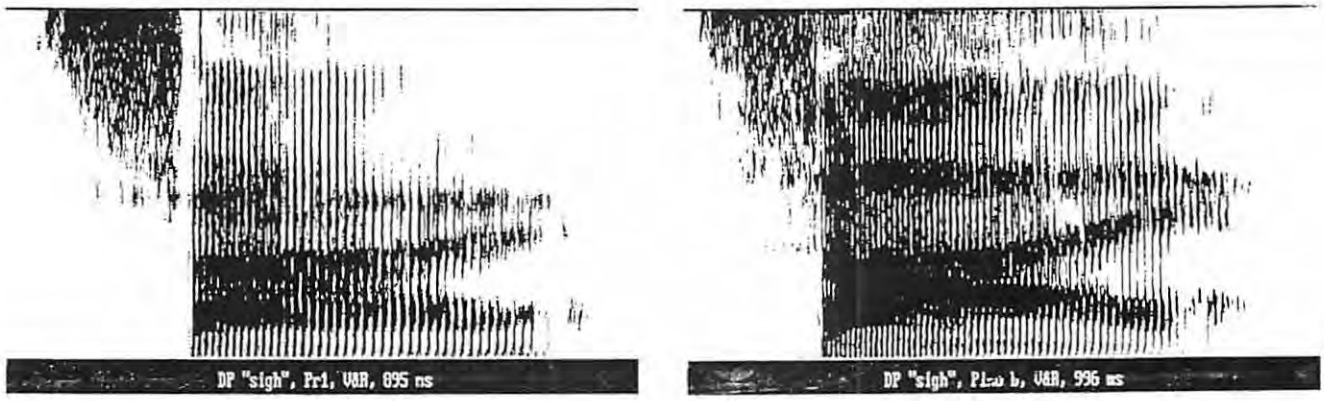


Figure 4. Wide band spectrograms of the word "sigh" before (left) and after treatment (right). Note the pre-treatment 'gap' between speech gestures for the fricative and the vowel.

Table 8.
Mean and standard deviation temporo-spectral measurements for selected second formant trajectories for all words, also for selected categories of words including monophthongs and diphthongs (see Table 6).

	Pre1		Pre2		Post1		Post2		FU6		FU12	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
ALL WORDS												
TD (ms)	96	72	117	43	148	85	143	85	141	72	151	77
TE (Hz)	396	329	443	300	542	286	532	335	598	344	591	357
TR (Hz/ms)	4.16	2.59	4.32	2.28	4.4	2.8	3.61	2.11	4.92	2.93	4.25	2.33
SELECTED WORDS												
TD-monoph (ms)	77	45	95	45	95	42	90	53	88	33	127	36
TD-diph (ms)	115	88	138	28	200	84	197	77	194	62	151	77
TE-monoph (Hz)	397	288	426	252	540	260	508	309	611	281	640	282
TE-diph (Hz)	396	365	460	340	544	309	557	358	584	396	541	414

TD = transition duration, TE = transition extent, TR = absolute transition rate

little overall TR change. This was primarily due to increases in TD being accompanied by increases in TE. However, comparison of individual words across conditions indicates a general trend for increases in TR primarily for words with front monophthongs and for diphthongs with high front offglides (e.g. sip, ship, blend, wax, sigh, ate). This was due primarily to increases in TE being considerably greater than increases in TD.

Videolaryngostroboscopy

Pre-treatment, the subject's vocal folds were judged to be bowed by all four judges; post-treatment they were judged to have a posterior glottal gap - a less severe deficit in adduction. On a five point scale (1 = normal, 5 = most severe), the degree of glottal incompetence was rated 1.33 pre-treatment and 1.00 post-treatment, for both soft and loud phonation. For follow-up recordings, this measure was 1.00 at six months and 1.25 at twelve months. Supraglottal hyperfunction was also reduced in the post-treatment condition; false-fold and anterior-posterior hyperfunction ratings both decreased from 1.50 to 1.00 pre- to post-treatment. At six months, anterior-posterior and false fold hyperfunction were both given ratings of 1.50, while at twelve months, the measures were 1.00 and 1.75 respectively.

Discussion

The purpose of this study was to document the phonatory mechanism associated with intensity change in a patient with Parkinson disease and investigate the concomitant effects on articulation. The sound pressure level data document increases in intensity in sustained vowels, reading, repetition of the /pae/ series and monologue that occurred after treatment and which were maintained up to twelve months without additional therapy. These data indicate that the patient was able to voluntarily increase his vocal intensity in spite of the physiologic impairment presented by his neurological condition. While intensity increases were measured in all speech tasks, the greater improvements for sustained phonation than for the other speech tasks are likely due to the fact that maximum vowel duration and high effort level were drilled extensively during therapy. Furthermore, maximum phonatory effort and increased SPL are more easily achieved in a task that requires constant vocal fold adduction. The reading and monologue tasks, on the other hand, require more dynamic adductory gestures, as well as more attention and concentration, and might therefore have been more difficult to maintain at a high effort level.

To examine the respiratory/phonatory mechanisms responsible for these SPL changes, we considered variables previously associated with vocal intensity: subglottal air pressure and vocal fold adduction. Pre- to post-treatment increases were observed in estimated subglottal air pressure which were maintained through the twelve month follow-up. Sound pressure levels measured during the /pae/ task are consistent with simultaneous estimates of subglottal air pressure. Two factors can be associated with the subject's increased subglottal air pressure. The first would be the increased effort expended by the subject in speech production. Higher lung volumes and greater expiratory force would allow generation of more subglottal air pressure. The respiratory excursion data are consistent with increased respiratory effort. A second factor involved in the maintenance of higher subglottal air pressure during prolonged phonation would be the improved valving efficiency of the larynx. If glottal adduction had not improved, the increased respiratory drive would have been partially wasted because of inefficiency at the glottis. However, the longer and louder sustained /a/ vowels produced by this subject after treatment suggest that this was not the case. Aerodynamic, videolaryngostroboscopic and electroglottographic data are indicative of post-treatment increases in vocal fold adduction that were maintained through the twelve month follow-up, suggesting that this was an important factor in the increased SPL.

Maximum flow declination rate (MFDR) increased substantially following treatment and remained at higher than pre-treatment levels through the twelve month follow-up. This variable has been found to correlate highly with sound pressure level in non-disordered speakers (Gauffin & Sundberg, 1989) and indicates that the subject was able to approximate the vocal folds with greater speed and shut off the flow more quickly at higher intensities. Although a more rapid approximation of the vocal folds might raise a concern that the treatment has the potential for leading to damaging vocal behavior, the videostroboscopic findings of improved vocal fold adduction and decreased supraglottal hyperfunction after treatment indicate that this was not the case. The subject was able to increase the excitation of the vocal tract through a more rapid flow shut-off, with improved true vocal fold approximation but without hyperfunctional behavior. These findings suggest that the patient was able to approach a more normal glottal configuration by overcoming the hypoadduction which had been a major contributor to his reduced intensity before treatment.

The open quotient, laryngeal resistance and harmonic-spectral slope data are consistent with post-treatment increases in vocal intensity and vocal fold adduction. Smaller open quotients are associated with longer closed phases and increased glottal efficiency (Colton, 1984). The main variable identified in lowering the open quotient is increased glottal adduction (Scherer, 1991). Laryngeal airway resistance was higher following treatment suggesting that the subject was able to adduct the vocal folds more fully and use the air supply more efficiently. The shallower harmonic spectral slope after treatment is consistent with louder pho-

nation. This observation documents the presence of increased high frequency energy, and when linked with greater power in the fundamental, led to stronger and more resonant phonation.

The acoustic perturbation data indicate that this subject's phonatory stability (in both amplitude and frequency domains) improved with increased sound pressure level. It is reasonable to consider that increased vocal intensity and adduction contributed to more stable vocal fold oscillation. Since the treatment tasks included 10 to 15 high effort maximum duration sustained vowel phonations daily, it could be speculated that as a result of treatment, the muscles of phonation became stronger and steadier in their function. This is consistent with the suggestion by Orlikoff and Kahane (1991) that when subjects phonate more loudly, there is recruitment of additional motor neurons and a higher frequency of firing in the active motor units, which could result in smoother muscle contraction and lower perturbation measures. Improved perturbation values with increased vocal fold adduction have been reported previously in Parkinson disease patients following treatment (Ramig et al. 1990), and appear to be additional benefits of a simple focus on louder phonation.

The increases in fundamental frequency variability for reading and monologue following therapy are consistent with previous observations (Ramig et al., 1994) and may be interpreted in at least two ways. Repeated maximum fundamental frequency range exercises were included as part of the daily phonatory effort treatment. It can be speculated that the cricothyroid muscle activity which was stimulated in these tasks generalized to improved intonation in connected speech. Another explanation is that as the subject's voice became stronger - and thus the effectiveness of his oral communication increased - his attitude and emotional outlook also improved and were reflected in the intonation of his speech. Both of these factors may have played a role in the greater fundamental frequency variability observed in this subject's speech following treatment. In either case, the patient's improved prosody in connected speech is further evidence of his capacity to compensate for the laryngeal and respiratory muscle rigidity associated with his condition.

The acoustic data on vowel characteristics and second formant trajectories provide additional evidence for a change in the valving efficiency of the larynx as well as for modifications to the relative timing of speech segments. Specifically, post-treatment and follow-up recordings in this study are consistent with evidence of articulatory reorganization found by Schulman (1989) in normal subjects producing loud speech. These included increased average vowel duration for all words accompanied by maintenance of phonological distinctions between durations of long and short vowels. Other changes in relative timing were observed in a reduction of the (silent or near-silent) gap between voiceless fricatives and the subsequent vowel, which reflects an increase in coarticulation or coupling between gestures. The ratio of vowel duration to whole word duration across the 12 words may also provide a rough ordinal index of relative timing relations among speech movement gestures. These ratios increased from .31 and .33 pre-treatment, to .39 post-treatment, and .45 and .46 in the 6 and 12 month follow-up recordings. Increases were primarily due to a decrease in frication duration and a decrease in post vocalic aspiration (which was included in the whole word duration measurement). These preliminary acoustic findings of increased coupling between gestures and increased ratio of vowel duration to whole word duration could be taken to reflect increased glottal coordination with supraglottal articulator movement.

Evidence of articulatory modifications accompanying louder phonation are apparent in second formant trajectory changes. As described previously, increases across all words for second formant transition duration, extent, and rate had been predicted because the expected increases in jaw displacement and vowel duration accompanying louder speech would allow more time for supraglottal articulator movement, including the tongue, to reach the target position. As predicted, there were increases in transition extent. However, there was considerably more gain in TE for words containing high front vowels and diphthongs with high front off-glides (e.g. sip, ship, blend, sigh, hail). Contrary to prediction, there was no overall mean transition duration increase for monophthongs following treatment. However, mean TD increased by approximately 60 ms for the more complex diphthongs. The difference between diphthongal and monophthongal change became smaller in the 12 month follow-up. Thus, even though there was more time "available" for supraglottal movement because vowel duration increased for all vowel types, in general, the transition duration stayed the same for vocalic nuclei produced with a relatively static vocal tract configuration (monophthong), but increased for vowel production with a gradually changing vocal tract configuration (diphthong).

Although there were corresponding changes in TR (computed as TE/TD), the change in mean absolute transition rate for all 12 words following therapy showed little overall change. This finding is a preliminary indication that the increased velocities of lips and jaw anticipated in louder speech are not always accompanied by increased velocity of the tongue or change in vocal tract configuration as measured by TR for this Parkinson patient. It is important to note, however, that general increases in TR were found primarily for words with high front vowels and diphthongs with high front off-glides, which was due to greater changes in TE than TD for these words. Acoustic findings summarized by Gracco (1988) seem to provide a reasonable explanation for the differences found for both TE and TR between vowel production involving primarily an anterior tongue position and those using a posterior tongue position. Specifically, it

has been found that jaw opening movement has little effect on posterior tongue position. This finding, along with the acoustic data for the subject in the present study, suggest that the biomechanical property of increased jaw opening for loud speech has a differential effect for anterior and posterior tongue position and movement.

These acoustic results are a preliminary indication that changes in segmental durations and timing relationships in a subject following treatment to increase vocal intensity may represent reorganization of articulatory movements and timing relationships which deserve further investigation with a larger sample of PD subjects. The post-treatment acoustic data on maintenance of phonetic distinctions between long and short vowels may be indicative of one or both of the following factors. First, the data may reflect goal-oriented reorganization to maintain phonetic distinctions. Second, the data may be indicative of the biomechanical properties accompanying increased jaw displacement combined the need for a gradually changing or static vocal tract configuration for diphthongs or monophthongs respectively.

It should be noted that in addition to the acoustic data reflecting increased valving efficiency of the glottis (glottal adduction), faster glottal opening rise time was measured following therapy for five out of seven initial voiceless fricatives. This change in fricative rise time may indicate a more rapid abductory valving action of the larynx, which is critical for generating early peak glottal opening in voiceless fricatives (Abramson, 1977; Kingston, 1990). It could be speculated that treatment variables aimed at promoting increased voice intensity and increased strength and stability of muscles of phonation not only improved the adductory but also the abductory valving action of the larynx.

Loud speech might be viewed as a naturally occurring scaling transformation, which modifies the activity of all muscles in the articulatory linkage, or coordinative structure, yet also preserves some consistent temporal relationships. Other naturally occurring scaling transformations include changes in speaking rate and degree of prosodic stress (Kelso, Saltzman, & Tuller, 1986). Change in speaking rate has been referred to as one of the major contributors to phonetic modification (Kelso et al., 1986). The present study provides a preliminary indication that scaling changes in intensity could be another cause of articulatory reorganization and resultant phonetic modification. Specifically, the data provide preliminary indications that treatment may be exploiting mechanical dynamics of the speech articulators which result from both a scaling transformation and increased phonatory functioning. These dynamics, in turn, influence movement reorganization, efficiency, and coordination.

The findings of this study contribute to our understanding of the mechanism of intensity change in one patient with Parkinson disease. The respiratory and laryngeal changes accompanying treatment designed to increase vocal intensity included increases in vocal fold adduction and subglottal air pressure. These mechanisms, which have previously been observed in normal speakers as they increase intensity, have been documented here in a patient with Parkinson disease. These findings suggest that compensation by behavioral means may modify, in part, the phonatory and articulatory acoustic effects of the physical pathology of this disease. The impact of this treatment on articulation, which was not targeted during therapy, suggests that a focus on improved phonation can have benefits beyond the larynx.

Acknowledgments

This research was supported in part by grants NIH-NIDCD #R01 DC01150 and #P60 DC00976. We wish to thank Bruce Smith, Gary Weismer, Donald Robin and an anonymous reviewer for their helpful comments on a previous version of this manuscript.

References

Abramson, A. S. (1977). Laryngeal timing in consonant distinctions. Phonetica, 34, 295-303.

Allan, C.M. (1970). Treatment of nonfluent speech resulting from neurological disease: Treatment of dysarthria. British Journal of Disorders of Communication, 5, 3-5.

Ansel, B. & Kent, R. (1992). Acoustic-phonetic contrasts and intelligibility in dysarthria associated with mixed cerebral palsy. Journal of Speech and Hearing Research, 35, 296-308.

Barlow, D.H. & Hersen, M. (1984). Single case experimental designs: Strategies for behavior change. New York: Pergamon Press.

Bless, D.M., Hirano, M., & Feder, R.J. (1987). Videostroboscopic evaluation of the larynx. Ear, Nose and Throat Journal, 66, 289-296.

Browman, C.P. & Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston & M.E. Beckman (Eds.), Papers in laboratory phonology I: Between the grammar and physics of speech. Cambridge: Cambridge University Press.

Canter, G.J. (1965). Speech characteristics of patients with Parkinson's disease II: Physiological support for speech. Journal of Speech and Hearing Disorders, 30, 44-49.

Colton, R.H. (1984). Glottal waveform variations associated with different vocal intensities. Transcript of the 13th Symposium on the Care of the Professional Voice. Voice Foundation, New York.

Critchley, E.M.R. (1981). Speech Disorders of Parkinsonism. A Review Journal of Neurology, Neurosurgery and Psychiatry, 44, 751-758.

Fairbanks, G. (1960). Voice and Articulation Drill Book. New York: Harper and Brothers.

Forrest, K., Weismer, G., & Turner, G.S. (1989). Kinematic, acoustic, and perceptual analyses of connected speech produced by Parkinsonian and normal geriatric adults. Journal of the Acoustical Society of America, 85 , 2608-2622.

Fowler, C. A. (1980). Coarticulation and theories of extrinsic timing. Journal of Phonetics, 8, 113-133.

Gauffin, J. & Sundberg, J. (1989). Spectral correlates of glottal voice source waveform characteristics. Journal of Speech and Hearing Research, 32, 556-565.

Gracco, V. L. (1988). Timing factors in the coordination of speech movements. The Journal of Neuroscience, 8, 4626-4639.

Gracco, V. L. (1994). Some organizational characteristics of speech movement control. Journal of Speech and Hearing Research, 37, 4-27.

Greene, H.C.L. (1980). The voice and its disorders. London: Pitman Medical.

Hanson, D.G., Gerratt, B.R., & Ward, P.H. (1984). Cinegraphic observations of laryngeal function in Parkinson's Disease. Laryngoscope, 94, 348-353.

Hardy, J. (1967). Suggestions for physiological research in dysarthria. Cortex, 3, 128-156.

Hillman, R. E., Holmberg, E. B., & Perkell, J. S. (1989). Objective assessment of vocal hyperfunction: An experimental framework and initial results. Journal of Speech and Hearing Research, 32, 373-380.

Hirose, H. & Joshita, Y. (1987). Laryngeal behavior in patients with disorders of the central nervous system. In Hirano, M., Kirchner, J., & Bless, D. (Eds.), Neurolaryngology. Boston: College Hill.

Hoehn, M., & Yahr, M. (1967). Parkinsonism: Onset, progression and mortality. Neurology, 19, 427.

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, 511-529.

Howell, P. & Rosen, S. (1983). Production and perception of rise time in the voiceless affricate/fricative distinction. Journal of the Acoustical Society of America, 73, 976-984.

Isshiki, N. (1964). Regulatory mechanisms of voice intensity variation. Journal of Speech and Hearing Research, 7, 17-29.

Kelso, J.A.S., Saltzman, E.L., & Tuller, B. (1986). The dynamical perspective on speech production: Data and theory. Journal of Phonetics, 14, 25-59.

Kelso, J.A.S. & Tuller, B. (1981). Toward a theory of apractic syndromes. Brain and Language, 12, 224-245.

Kelso, J.A.S., Tuller, B., Bateson, E.V., & Fowler, C.A. (1984). Functionally specific articulatory cooperation following jaw perturbations during speech: Evidence for coordinative structures. Journal of Experimental Psychology: Human Perception and Performance, 10, 812-832.

Kent, R.D., Kent, J.F., & Rosenbeck, J.C. (1987). Maximum performance tests of speech production. Journal of Speech and Hearing Disorders, 52, 367-387.

Kent, R.D., Kent, J.F., Weismer, G., Martin, R.E., Sufit, R.L., Brooks, B., & Rosenbek, J.C. (1989). Relationships between speech intelligibility and the slope of second-formant transitions in dysarthric subjects. Clinical Linguistics and Phonetics, 3, 347-358.

Kent, R.D. & Moll, K.L. (1972). Cinefluorographic analysis of selected lingual consonants. Journal of Speech and Hearing Research, 13, 453-473.

King, J.B., Ramig, L.O., Lemke, J.H., & Horii, Y. (1994). Variability in acoustic and perceptual parameters of phonation in patients with ParkinsonÕs disease. Journal of Medical Speech-Language Pathology, 2, 29-42.

Kingston, J. (1990). Articulatory binding. In J. Kingston & M. E. Beckman (Eds.), Papers in laboratory phonology I: Between the grammar and physics of speech. Cambridge: Cambridge University Press.

Klatt, D.H. (1975). Voice onset time, frication, and aspiration in word initial consonant clusters. Journal of Speech and Hearing Research, 18, 686-706.

Kratochwill, T. & Levin, J.R. (1992). Single-case research design and analysis: New directions for psychology and education. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Liss, J.M. & Weismer, G. (1992). Qualitative acoustic analysis in the study of motor speech disorders. Journal of Acoustical Society of America, 92 , 2984-2987

McReynolds, L.V. & Kearns, K.P. (1983). Single-subject experimental designs in communicative disorders. Baltimore: University Park Press.

Moore, P. (1946). Intelligibility related to routinized messages. Speech Monographs, 13, 46-49.

Munhall, K.G., Ostry, D.J., & Parush, A. (1985). Characteristics of velocity profiles of speech movements. Journal of Experimental Psychology: Human Perception and Performance, 11, 457-474.

Murdoch, B.E., Chenery, H.J., Bowler, S., & Ingram, J.C. (1989). Respiratory function in Parkinson's subjects exhibiting a perceptible speech deficit: A kinematic and spirometric analysis. Journal of Speech and Hearing Disorders, 54, 610-626.

Netsell, R. (1986). A Neurobiologic view of speech production and the dysarthrias. San Diego: College Hill Press.

Orlikoff, R.F. & Kahane, J.C. (1991). Influence of mean sound pressure level on jitter and shimmer measures. Journal of Voice, 5, 113-119.

Perez, K.S., Ramig, L.O., Smith, M.E., & Dromey, C.D. (1994). Videostroboscopic findings in patients with Parkinson disease. Paper presented at the Annual Symposium of the Voice Foundation, Philadelphia.

Peterson, G.E. & Lehiste, I. (1960). Duration of syllable nuclei in English. Journal of the Acoustical Society of America, 32, 693-703.

Pickett, J.M. (1956). Effects of vocal force on the intelligibility of speech sounds. Journal of the Acoustical Society of America, 28, 902-905.

Ramig, L.O. (1992). The role of phonation in speech intelligibility: A review and preliminary data from patients with Parkinson's disease. In R. Kent (Ed.), Intelligibility in speech disorders: Theory, measurement, and management. Amsterdam: John Benjamin.

Ramig, L.O. (1994). Speech therapy for Parkinson's disease. In W. Koller & G. Paulson (Eds.), Therapy of Parkinson's disease. New York: Marcel Dekker.

Ramig, L.O., Bonitati, C.M., Lemke, J.H., & Horii, Y. (1994). Voice therapy for patients with Parkinson's disease: Development of an approach and preliminary efficacy data. Journal of Medical Speech-Language Pathology, 2, 191-210.

Ramig, L.O., Bonitati, C.M., & Horii, Y. (1991). The efficacy of voice therapy for patients with Parkinson's disease. NCVS Status and Progress Report, 1, 61-86.

Ramig, L.O., Fazoli, K., Scherer, R.C., & Bonitati, C. (1990). Phonation in Parkinsons's disease. Paper presented at the Clinical Dysarthria Conference, San Antonio.

Robertson, S. & Thompson, F. (1984). Speech therapy in Parkinson's disease: A study of the efficacy and long-term effect of intensive treatment. British Journal of Disorders of Communication, 19, 213-224.

Rothenberg, M., (1973). A new inverse-filtering technique for deriving the glottal flow waveform during voicing. Journal of the Acoustical Society of America, 53, 1632-1645.

Sapienza, C.M. & Stathopoulos, E.T. (in press). Respiratory and laryngeal measures of children and women with bilateral vocal fold nodules. Journal of Speech and Hearing Research.

Sarno, M.T. (1968). Speech impairment in Parkinson's disease. Archives of Physical Medicine Rehabilitation, 49, 269-275.

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. New York: Raven Press.

Scherer, R.C. & Vail, V.J. (1988). Measures of laryngeal adduction. Journal of the Acoustical Society of America, suppl 84, (1) S81 (A).

Schulman, R. (1989). Articulatory dynamics of loud and normal speech. Journal of Acoustical Society of America, 85, 295-312.

Scott, S. & Caird, F.L. (1983). Speech therapy for Parkinson's disease. Journal of Neurology, Neurosurgery and Psychiatry, 46, 140-144.

Shriberg, L. D. & Kent, R. D. (1982). Clinical Phonetics. New York: Macmillan Publishing Co.

Smith, M.E., Ramig, L.O., Dromey, C.D., Perez, K.S., & Samandari, R., (in press). Videostroboscopic assesment of phonation in Parkinson's disease before and after intensive voice therapy. Journal of Voice.

Smitheran, J.R. & Hixon, T.J., (1981). A clinical method for estimating laryngeal airway resistance during vowel production. Journal of Speech and Hearing Disorders, 46, 138-146.

Stathopoulos, E.T. & Sapienza, C.M. (1993a). Respiratory and laryngeal function of women and men during vocal intensity variation. Journal of Speech and Hearing Research, 36, 64-75.

Stathopoulos, E.T. & Sapienza, C.M. (1993b). Respiratory and laryngeal measures of children during vocal intensity variation. Journal of the Acoustical Society of America, 94, 2531-2543.

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

Titze, I.R. & Sundberg, J. (1992). Vocal intensity in speakers and singers. Journal of the Acoustical Society of America, 91, 2936-2946.

Weiner, W.J. & Singer, C. (1989). Parkinson's disease and non-pharmacological treatmetn programs. Journal of the American Geriatric Society, 37, 359-363.

Weismer, G., Kent, R.D., Hodge, M., & Martin, R. (1988). The acoustic signature for intelligibility test words. Journal of the Acoustical Society of America, 84, 1281-1291.

Weismer, G., Martin, R., Kent, R.D., & Kent, J.F. (1992). Formant trajectory characteristics of males with amyotrophic lateral sclerosis. Journal of the Acoustical Society of America, 91, 1085-1098.

NCVS Status and Progress Report - 7
December 1994, 95-105

Acquisition of Speech by Children Who Have Prolonged Cochlear Implant Experience

Nancy Tye-Murray, Ph.D.
Linda Spencer, M.A.
Department of Otolaryngology - Head and Neck Surgery, The University of Iowa
George Woodworth, Ph.D.
Department of Statistics and Actuarial Science, The University of Iowa

Abstract

The four purposes of this investigation were to assess whether children acquire intelligible speech following prolonged cochlear-implant experience and examine their speech error patterns, to examine how age at implantation influences speech acquisition, to assess how speech production and speech perception skills relate, and to determine whether cochlear implant recipients who formerly used simultaneous communication (speech and manually coded English) begin to use speech without sign to communicate. Twenty-eight prelinguistically deafened children who use a Nucleus cochlear implant were assigned to one of three age groups, according to age at implantation: 2 to 5 yrs (N = 12), 5 to 8 yrs (N = 9), and 8 to 15 yrs (N = 7). All subjects had worn a cochlear implant for at least 24 mos, and an average of 36 mos. All subjects used simultaneous communication at the time of implantation. Subjects performed both imitative and structured spontaneous sampling speech tasks. The results permit the following conclusions: (a) children who have used a cochlear implant for at least two yrs acquire some intelligible speech; (b) children who receive a cochlear implant before the age of 5 yrs appear to show greater benefit in their speech production skills than children who are older, at least after a minimum of two yrs of use; (c) children who recognize more speech while wearing their cochlear implants are likely to speak more intelligibly; and, (d) signing does not disappear from a child's communication mode following implantation.

Some researchers and organizations have criticized cochlear implantation of prelinguistically deafened children (Lane, 1990; National Association for the Deaf, 1991). On the one hand, children who receive cochlear implants may not receive enough information to achieve a performance level that allows them to communicate and receive messages verbally, and to interact easily with members of the hearing community. On the other hand, because speech and listening skills receive so much attention, children may not acquire the necessary sign skills to associate with members of the Deaf community. Thus, young cochlear implant users may become culturally bereft, belonging to neither hearing nor Deaf groups (Evans, 1987).

A central role of research at this time is to document how children's speaking, listening, speechreading, language, and signing performance changes as a result of implantation. This documentation can either support or allay concerns about the efficacy of cochlear implantation, and can provide important information for counseling families and potential cochlear implant candidates and for developing habilitation programs. The present investigation had four purposes: 1) to determine the level of intelligibility that prelinguistically deafened children acquire following prolonged cochlear-implant experience, and to examine their speech error patterns; 2) to examine how age at implantation influences speech acquisition; 3) to investigate how speech production skills relate to speech recognition skills; and 4) to determine whether use of sign disappears.

Speech Intelligibility After Implantation

A number of investigators have demonstrated significant improvements in speech production following cochlear implantation (e.g., Osberger et al., 1991; Tobey et al., 1991; Tobey & Hasenstab, 1991; Tye-Murray & Kirk, 1993). However, overall intelligibility remains low, at least for the first 2 yrs. For instance, Osberger, Maso, and Sam (1993) reported that the mean intelligibility score derived from a set of 10 imitated sentences was only 18% words correct for a group of 12 prelinguistically deafened children

who use the Cochlear Corporation Nucleus multichannel cochlear implant.

In the present investigation, we evaluated the intelligibility of children who had worn their cochlear implants for a relatively long period of time. We performed fine-grained analyses to examine word intelligibility, phonemic errors (in both spontaneous and elicited speech tasks), and speech skills such as syllable stress production after children had acquired an average of 36 mos experience. This kind of examination not only indicates whether intelligibility changes, but also provides guidance for ordering speech therapy objectives. For example, phonemic contrasts that are shown to be produced relatively accurately might be targeted earlier in a speech therapy curriculum for new cochlear implant users; phonemic contrasts that are produced poorly might be targeted later.

Age at Implantation

Recent studies with prelinguistically deafened children suggest that children who receive a cochlear implant at a younger age might develop better speech skills than children who receive one at an older age. Osberger et al. (1993) reported that prelinguistically deafened children who receive a Nucleus cochlear implant before the age of 10 yrs achieve somewhat better speech production skills than children who receive a device later. In this investigation we examined this issue with a larger group of subjects, and three different age groups. We compared experienced users' speech performance with measures obtained pre-implant or measures obtained from children who do not use cochlear implants. We also examined the rate of progress demonstrated by younger versus older cochlear implant recipients. Such analyses and resultant data might aid family members of a particular cochlear implant candidate to develop realistic expectations about the benefits of implantation.

Relationship Between Speech Perception & Production

Children who have more residual hearing, as indexed by puretone thresholds, usually achieve greater success in speaking than children who have less residual hearing (Boothroyd, 1969; Gold, 1978;, Markides, 1970; Smith, 1975). Measures of speech recognition indicate how much of the speech signal children use in a meaningful fashion. Current research suggests that hearing-aid users with good speech recognition skills do not necessarily develop good speech production skills. However, children who speak relatively well are almost always good perceivers (Stark & Levitt, 1974; Osberger et al., 1993). In this investigation we related cochlear implant users' abilities to recognize speech to their abilities to perform a number of different speaking tasks. If a strong relationship exists, then we might assume that children who receive good audiological benefit from their devices begin to rely on auditory information for speech acquisition.

Use of Sign Language

Preliminary evidence does not suggest that cochlear implantation necessarily leads to a decrease in the use of sign if a child used sign beforehand (Firszt, Reeder, Zimmerman-Phillips, Tonokawa, & Proctor, 1991; Tye-Murray & Kirk, 1993). The final purpose of this investigation was to examine whether children who relied on simultaneous communication (i.e., speech and manually coded English) prior to receiving a cochlear implant continue to integrate sign into their communication mode following implantation, even after speech skills begin to improve. Both videotaped speech samples and questionnaire data from parents were collected to address this issue.

Method

Subjects

Twenty-eight children who had prelinguistic deafness (i.e., profound hearing losses incurred no later than 18 mos of age) and who use a Nucleus cochlear implant participated in this investigation. The Nucleus cochlear implant codes information about fundamental frequency, intensity, vowel formants, and aperiodic segments of the speech signals. It presents signal information to a 22-electrode array inserted into the child's cochlea.

Subjects' age at cochlear implant connection[1] ranged from 31 to 170 mos, and averaged 85 mos (SD = 47 mos). For purposes of data analysis, subjects were assigned to one of three groups according to age at implantation. The first age group (the 2 - 4 yrs group) included 12 subjects who were implanted between the ages of 31 and 58 mos, and an average age of 48 mos (SD = 8 mos). The second age group (the 5 - 8 yrs group) included 9 subjects implanted between the ages of 60 and 89 mos, and an average age of 69 mos (SD = 8 mos). The third age group (the 9 - 15 yrs group) included 7 children implanted between the ages of 113 and 184 mos, and an average age of 138 mos (SD = 35 mos). Subjects had to have attained at least 24 mos of experience with a cochlear implant to participate. On average, subjects had 36 mos of experience (SD = 13.5 mos). (Some subjects were also tested pre-implantation or on early dates following implantation, as indicated in the Results Section.) Subjects in the 9 to 15 yrs group had somewhat less experience on average than subjects in the 2 to 4 yrs and 5 to 8 yrs group, 29 versus 37 and 40 mos respectively. Data from an additional five subjects who did not have cochlear implant experience were used in one set of analyses, where post-implant performance was compared to pre-implant performance. The biographical characteristics of these children were similar to the subjects who received cochlear implants.

[1] "Cochlear implant connection" refers to the time that a child actually begins to wear the device. Connection typically occurs four to six wks following cochlear implant surgery.

All subjects lived at home and attended a public school at the time of testing. Parents reported that subjects used a simultaneous communication approach at home and at school. All but two of the subjects used Signed English (Bornstein, Saulnier, & Hamilton, 1983) or Signed Exact English (Gustason, Pfetzing & Zawolkow, 1980), either alone or in combination with American Sign Language (ASL). All but three subjects were mainstreamed in a classroom for normally hearing children for part of the day. One subject attended the Iowa State School for the Deaf.

In many ways, this group of subjects is representative of the population of children who have profound hearing losses in the United States. Most children who are deaf attend public schools, most live at home, and most use some form of signing (Moores, 1992; Kluwin, 1992).

Evaluation Measures of Speech Production

The test battery included the following tests: The Short Long Sentence Test, The Story-Retell Task, the Central Institute for the Deaf (CID) Speech Intelligibility Evaluation (SPINE, Monson, Moog & Geers, 1988),[2] Audiovisual Feature Test (Production Version) (Tyler, Fryauf-Bertschy & Kelsay, 1991), Fundamental Speech Skills Test (FSST, Levitt, Youdelman, & Head, 1990),[3] and a Parent Questionnaire (Tye-Murray & Kelsay, 1993). Because the Short-Long Sentence test, the story-retell task, the Audiovisual Feature Test, and the parent questionnaire were developed specifically for this investigation, they are described in greater detail below. Subjects were tested individually in a quiet room and productions were recorded via videotape with audio track. Subjects sat at a short table directly across from the examiner, who was a speech-language pathologist proficient in manually coded English. A Panasonic VHS Professional/ Industrial videocamera model AG-180 was placed across from the subject and was focused so as to record all hand and facial movements.

Short-Long Sentence Test. This test consisted of 14 sentence pairs, each pair containing a short and long item. The short version sentences contained between 1 and 3 words, and were ritual or simple phrases (e.g., How are you?; Please stop.). These sentences could be produced by very young children and permitted direct comparisons between younger and older children speaking the same sentences. The long version sentences were expansions of the short versions, (e.g., How are you going to get there?; Please stop making so much noise.). These items were more challenging to produce and matched the linguistic sophistication of older children. In this investigation, only the short versions were analyzed.

The speech-language pathologist presented each item with both speech and sign, and the child imitated her speech. The speech-language pathologist then transcribed the productions. The transcription was scored using a target transcription to compare with the actual production. The target transcription allowed for minor dialectal differences.[4] A word was considered correct if all phonemes in the word matched the target production. The phonetic transcriptions yielded a percent phonemes correct and percent words correct score.

The Story-Retell Task. This task was administered in order to obtain a spontaneous speech sample for articulation and intelligibility analysis. The speech-language pathologist described each of four pictures in a set, following a prepared script and using simultaneous communication. The child then retold the story using speech (and sign, if desired). This procedure was repeated until the child had retold all of six stories. If the child had difficulty in formulating a narrative, the speech-language pathologist was allowed to use a limited set of general prompts (e.g., Tell me more about this picture).

The story-retell speech samples were scored in three ways. First, the initial 100 spoken words were phonetically transcribed and the corresponding sign was orthographically transcribed. Only the initial 100 words were included in the analysis so as to control for unequal sample sizes. The accuracy of the spoken utterances was determined by referencing the phonetic transcriptions to the signed transcriptions, to yield a phoneme percent correct and word percent correct score (Tye-Murray & Kirk, 1993).

A second measure obtained from the story-retell samples was a subjective rating score. A group of ten listeners with normal hearing who were unfamiliar with the speech of deaf talkers heard the audio portions of the speech sample recordings. The listeners rated the overall intelligibility of two stories using a ten-point ascending scale, where 1 indicated that "the speech is completely unintelligible" and 10 indicated that "the speech is completely intelligible."[5]

A third measure derived from the story-retell speech samples was an identification score. This score reflected how well narrative information was conveyed verbally to unfamiliar persons. Ten new listeners (i.e., listeners who did not perform the rating task) with normal hearing and no experience listening to the speech of deaf talkers listened to the auditory signal of two stories told by each subject. As

[2] The SPINE requires a subject to name four sets of picture cards that are similar in phonemic context (e.g., bat/bag; bell/ball). The examiner records via pen and answer sheet which word she thought was said, to yield a percent words correct. The SPINE does not provide normative data with which to compare the present results.

[3] The FSST is comprised of a battery of subtests. It designed to assess speech production skills that include breath stream capacity, elementary articulation, pitch control, syllabification, stress, and intonation contour.

[4] Dialectical differences were allowed so that children were not penalized for using a regional dialect.

[5] Although rating scales do not provide specific information about articulation errors (Samar & Metz, 1988), they provide information about listeners' subjective impressions of intelligibility (Tye-Murray, Barkmeier & Folkins, 1991).

they listened, they viewed the four-picture sets that corresponded to each of the six stories. These picture sets were projected on a screen before the listeners with a slide projector. After each story presentation, the listeners indicated on a printed score sheet which of the six picture series corresponded to the story that they had just heard. Listeners must have understood more than a few words in order to have identified a sequence correctly, as the stories had similar actors (e.g., five of the six stories featured a boy) and actions occurring (e.g., four stories featured dressing or putting on clothes). An identification score for a particular child was computed by adding the number of stories correctly identified by the 10 listeners across the two stories, and then dividing by 20.

Audio-visual Feature Test (Production Version) (Tyler, Fryauf-Bertschy & Kelsay, 1991). This test was comprised of items that most 4- and 5-yr-old hard-of-hearing children can recognize: seven alphabet letters (e.g., B, T) and three common words (e.g., me), all with a consonant-/i/ structure. Before testing began, subjects were shown a picture of each stimulus item and asked to name it. If a child could not name an item after two attempts, the test was not administered.

During testing, subjects produced each stimulus three times in random order. Pictures of the items were used to elicit the productions. For analysis, a subject's responses were played to four listeners who had normal hearing and who were familiar with the speech of hard-of-hearing talkers. After each production was played, each of the four listeners circled the syllable that they heard from among the closed set. A percent correct score was computed by averaging the scores from the four listeners. The responses of the listeners were also compiled into confusion matrices so that the relative error rates for the initial consonants could be examined.

Parent Questionnaire (Tye-Murray & Kelsay, 1993). The Parent Questionnaire consisted of questions about a child's communicative skills. Each subject's parent, usually the mother, indicated agreement with each item using a 10-point scale, where 1 = Never, and 10 = Frequently. Seven items were selected from the questionnaire for analysis in this investigation, three that pertained to speech intelligibility and four that pertained to use of sign (Tye-Murray & Kelsay, 1993, p. 28).

Audiological Test Battery

The audiological test battery consisted of the 6-Choice Word Intelligibility by Picture Identification (WIPI, Ross & Lerman, 1971), the Audio-visual Feature Test (Perception Version) (Tyler, Fryauf-Bertschy & Kelsay, 1991) and the Children's Vowel Test, (Tyler, Opie, Fryauf-Bertschy, & Gantz, 1992). The Audio-visual Feature Test (Perception Version) (described above) was presented in an audition-only and an audition-plus-vision condition. The Children's Vowel Test, which consisted of five sets of four picture items (e.g., hit, hot, hat, hurt), and the WIPI were presented in an audition-only condition. Testing was performed live voice in a quiet room. The audiologist who administered the tests spoke at a normal conversational level.

Transcription Reliability

Phonetic and orthographic transcriptions were completed by two speech-language pathologists (SLP-1 and SLP-2) who were competent in Signed English and familiar with the speech of hard-of-hearing children. Each clinician transcribed approximately the same number of samples. To determine phonetic transcription reliability of each of the two clinicians, a total of 12 samples from the subjects were transcribed by 2 separate listeners who were also familiar with the speech of hard-of-hearing children and trained in phonetic transcription. A Lambda analysis (Hays, 1973) was performed. For this analysis, the percent of speech sound agreement was calculated by dividing the number of agreements about the occurrences of speech sounds by the total number of sounds produced. Average agreement between SLP-1 and listener/transcriber-1 was 74% (SD = 7) and average agreement between SLP-2 and listener/transcriber-2 was 73% (SD = 4). To determine reliability of the sign transcriptions, six stories from the Story-Retell task were chosen at random and transcribed orthographically by an audiologist competent in Signed English. A Lambda analysis (Hays, 1993) was performed. Average orthographic agreement was 90%.

Results

Speech Intelligibility after Implantation

In order to evaluate subjects' levels of speech intelligibility after an average of 36 mos of cochlear implant

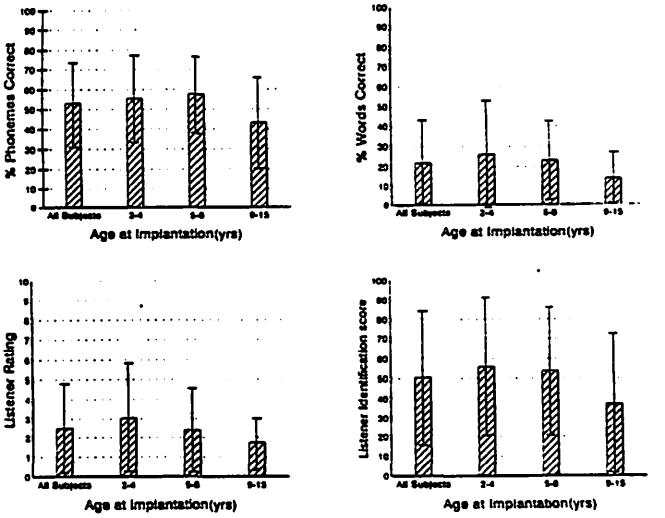


Figure 1. Means and standard deviations from the story-retell speech sample measures, including percent phonemes correct, percent words correct, listener ratings, and listener identification scores. Age ranges listed on axis represent age at implantation, not age at testing (which was at least 24 mos post-implantation).

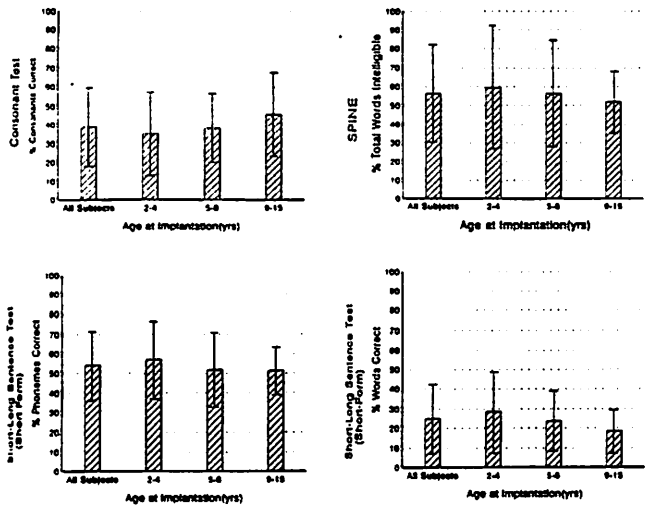


Figure 2. Means and standard deviations for the Audio-visual Feature Test (Production Version), the SPINE, and the short version of the Short-Long Sentence Test (both percent phonemes and percent words correct).

use, scores for each of the speech production tests were averaged. Scores for the entire group of subjects and also for each of the three age groups are presented in Figures 1-3.

Results from the story-retell speech samples appear in Figure 1. On average, the entire group of subjects produced 53% of the phonemes correctly and 22% of the words correctly. Performance within age groups varied widely, as indicated by the large standard deviation bars. Scores ranged from 14 to 92% phonemes correct and from 1 to 78% words correct. Ratings of intelligibility were low, corresponding with the relatively poor word production revealed by the transcription procedure. The mean rating of the 10 raters, averaged over all subjects, was 2.5, on a scale of 1 to 10. Scores ranged from 1 to 9. Listener identification scores indicated that subjects were sometimes able to convey narrative information successfully. The ten listeners who were assigned to the identification task identified which stories the children were describing in 51% of the samples (chance = 16%). Identification scores ranged from 5% to 100% correct.

A fine-grain analysis of the phoneme errors that occurred during the story-retell procedure was performed. Table 1 shows percent phonemes correct as a function of place of articulation. Vowels that are produced with a central place of articulation were somewhat more likely to be correct (approximately 70%) than front vowels (50%), back vowels (58%), or diphthongs (54%). Consonants that tend to be produced with visible articulatory movements were also more likely to be correct. The bilabials and labiodentals were produced more accurately (approximately 81% and 72%, respectively) than the linguadentals (47%), alveolars (47%), palatals (33%), or velars (40%). These trends were consistent across the three age groups.

The results from the syllable and sentence tests were examined next. Results from the Audio-visual Feature Test (Production Version), the SPINE, and the short version sentences of the Short-Long Sentence Test appear in Figure 2. On average, the group of subjects scored 39% consonants correct on the Audio-visual Feature Test (Production Version). The average score for the SPINE, which also indexes phoneme production, was 56% items correct. The percent phonemes correct score for the short version sentences of the Short-Long Sentence Test for all subjects was 54% and the percent words correct score was 25%, on average. The results from the three tests portrayed in Figure 2 suggest that the subjects produced about half of their phonemes correctly. These measurements correspond with the performance measured during the story-retell task (Figure 1).

The relative error rates for the phonemes produced by subjects during the Audio-visual Feature Test (Production Version) were examined. The response from all subjects were compiled into a single consonant confusion

Table 1.
Percent phonemes correct on average for each age group and standard deviations (in parentheses) for the story retell procedure, organized as a function of place of articulation.

Place of Articulation	Age Group (yrs)		
	2-4	5-8	9-15
Vowels			
Front (/i, æ, ɝ, ɚ, ɛ, ɪ/)	56 (24)	52 (21)	46 (17)
Back (/u, ʊ, ɔ/)	65 (25)	56 (36)	54 (27)
Central (/ə, ʌ, ɔ/)	72 (22)	67 (22)	69 (31)
Diphthongs (/oʊ, əɪ, əʊ, ɔɪ, ɪ/)	64 (31)	51 (34)	46 (27)
Consonants			
Bilabials (/p, b, m, w/)	83 (17)	87 (14)	72 (20)
Labiodentals (/f, v/)	76 (28)	83 (19)	57 (38)
Linguadentals (/θ, ð/)	36 (30)	59 (26)	46 (43)
Alveolars (/t, d, s, z, l, n/)	46 (29)	61 (31)	34 (22)
Palatals (/ʃ, ʒ, tʃ, dʒ, r, j/)	33 (35)	38 (19)	28 (34)
Velars (/k, g, ŋ/)	41 (37)	50 (31)	29 (30)

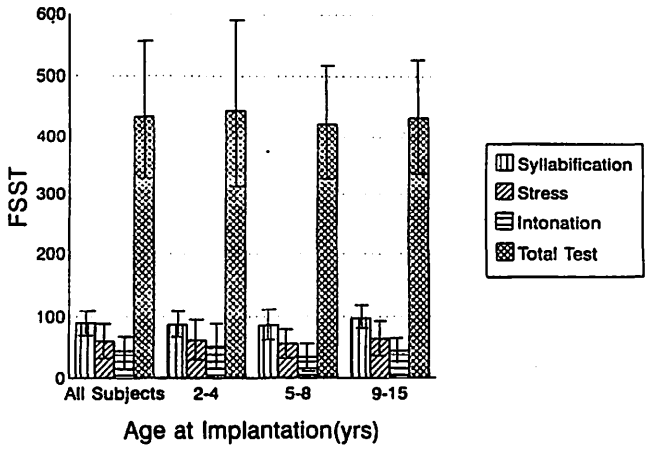


Figure 3. Means and standard deviations from the FSST for total score and suprasegmentals subtests including: number of syllable produced, stress, and intonation contours.

Table 2.
Consonant confusion matrice for the entire group of subjects, for the production version of the Children's Consonant Feature Test (Stimulus N = 276 for each phoneme).

Stimulus	Response										
	/p/	/t/	/k/	/s/	/b/	/d/	/v/	/z/	/n/	/m/	% Correct
/p/	130	3	4	4	102	8	8	3	5	9	47%
/t/	11	122	21	35	11	40	6	13	10	7	44%
/k/	11	48	92	20	12	25	7	13	27	21	33%
/s/	11	38	24	66	15	55	8	0	6	16	24%
/b/	46	4	4	1	181	10	8	0	6	16	66%
/d/	10	58	19	48	9	102	2	20	3	5	37%
/v/	45	23	6	14	54	26	78	10	7	13	28%
/z/	6	50	11	70	10	54	13	40	12	10	14%
/n/	6	16	20	0	29	52	9	6	93	45	34%
/m/	14	2	2	1	89	9	16	1	26	116	42%
Total	**290**	**364**	**203**	**259**	**512**	**381**	**155**	**139**	**202**	**255**	

matrix, which is presented in Table 2. The stop (/p, b, t, d, k/) and the nasal consonants (/m,n/) were produced with the most accuracy (45% and 38% correct, respectively) and the fricative consonants were produced with the least accuracy (22% correct). With the exception of /s/ and /v/, the most common errors produced for a particular target phoneme usually shared all but one feature with the target. These results suggest that children were achieving a close approximation of the intended targets. Voicing errors were especially common for the stop consonants. For instance, 70% of the errors produced for /p/ were /b/ substitutions; 48% of the errors produced for /b/ were /p/ substitutions; 26% of the errors produced for /t/ were /d/ substitutions; and 33% of the errors produced for /d/ were /t/ substitutions. Manner of articulation errors were common for the nasal consonants: the nasal /m/ was most commonly substituted by the non-nasal stop /b/ (56% of the errors) and the nasal /n/ was most commonly substituted by the non-nasal stop /d/ (28% of the errors). The phoneme /s/ was most often substituted by /d/ (26% of the errors), which differs from /s/ in both manner of articulation (a stop versus a fricative) and voicing (voiced versus unvoiced). The phoneme /v/ was most often substituted by /b/ (27% of the errors). The phonemes /p, b/ were most often produced correctly (66% and 47%, respectively) and /z/ was most often produced incorrectly (14%).

As shown in Figure 3, the final measure, the FSST, provided information about speech skills such as stress production and intonation contours. The total score averaged 435, and ranged from 221 to 594. Total scores are computed by adding together the percentages for the six subtests, so 600 is the highest score possible. The average total score was compared to the norms provided in the test. The total score for the present subjects fell between the 90 and 100 percentile for the test-normed age ranges of 6 to 9 yrs, 10 and 11 yrs and 12 to 14 yrs for subjects with hearing thresholds of 115+ dB. Individual subtest scores for Suprasegmentals for the entire group were 89% for correct number of syllables produced, 61% for correct production of stress, and 44% for correct production of intonation contours. All scores fall above the 70th percentile based on test norms for children with hearing thresholds of 115+ db.

Results from the parent questionnaire indicated that parents believed they recognized much of their child's speech, ranking their performance above that of family and strangers. On the following statements, parents provided the following ratings, on average: "I understand my child's speech", average rating = 7.4 (SD = 2.0); "Strangers understand my child's speech", average rating = 4.4 (SD = 2.0); and "Other members of the family understand my child's speech", average rating = 6.5 (SD = 2.0).

Age at Implantation

In order to examine the effects of age at implantation on speech production, four composite variables were created: word, phoneme, speech skills, and functional speech. The word composite variable reflects how accurately subjects spoke word stimuli, and was computed by averaging together the percent- words-correct scores from the story-retell speech samples and the short version sentences of the Short/Long Sentence Test. The phoneme variable reflects phoneme production, and was computed by averaging the percent-phoneme-correct scores from the story-retell speech samples, the short version sentences of the Short/Long Sentence Test, the Audio-visual Feature Test (Production Version), and the percent correct score for the SPINE. The speech skills composite score was the total score for the FSST, which encompasses scores from all of its subtests (see Footnote 3). A functional speech composite variable was computed by averaging the three measures

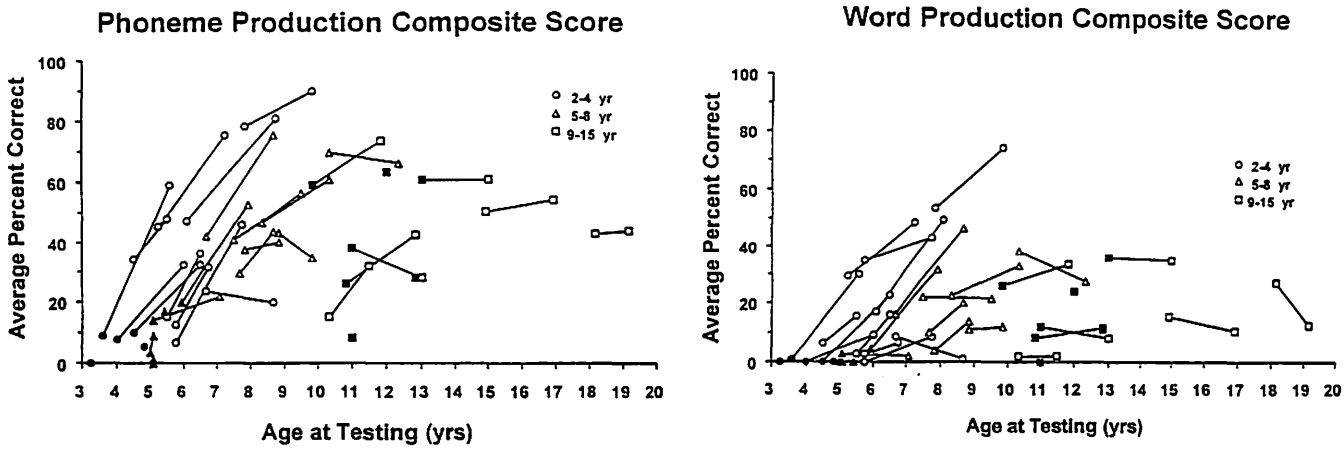


Figure 4 (left). Pre-implant and post-implant phoneme composite test scores plotted as a function of age at test administration. Filled shapes represent pre-implant test scores and unfilled shapes represent post-implant test interval scores. Lines connect scores from subjects who had data available pre-implant and post-implant, or data available from two post-implant test intervals. Figure 5 (right). Pre-implant and post-implant word composite test scores plotted as a function of age at test administration. Filled shapes represent pre-implant test scores and unfilled shapes represent post-implant test interval scores. Lines connect scores from subjects who had data available pre-implant and post-implant, or data available from two post-implant test intervals.

from the story-retell speech samples: percent words correct, percent stories identified, and the average rating score, multiplied by 10 (so it would encompass a 100-point range, like the other two story-retell measures).

The phoneme and word production composite scores are plotted in Figures 4 and 5 as a function of age at test administration. Those subjects who were tested on more than one occasion are denoted by symbols that are connected by lines. As noted in the Methods Section, some subjects were tested before implantation or shortly thereafter. Also included in Figures 4 and 5 are data points for the five subjects who did not receive cochlear implants. They are denoted by filled symbols without connecting lines.

Pearson correlations were performed between the composite scores and subjects' age at time of testing. The analysis revealed that older subjects tended to achieve higher scores than younger subjects prior to receiving a cochlear implant (word composite variable: $r = .786$, $p < .001$; phoneme composite variable: $r = .787$; functional speech variable: $r = .561$, $p < .01$, speech skills composite scores: $r = .436$, $p > .01$). However, after at least two yrs of experience with a cochlear implant, older children no longer performed better than the younger group. Age and post-implant performance measures were not significantly correlated (word composite scores: ($r = -.211$, $p > .01$, phoneme composite scores: ($r = -.020$, $p > .01$, speech skills composite variable: $r = -.033$, $p > .01$, functional speech composite scores: $r = -.228$, $p > .01$). One interpretation of these results is that younger children who use cochlear implants progress at a faster rate than older children.

In a second statistical analyses, we considered whether children who receive a cochlear implant perform better than children of the same age who do not receive a cochlear implant. Following Kirk and Hill-Brown (1985), we compared the post-implant composite scores of the younger group (ages 2 to 4 yrs) to the pre-implant scores of the middle age group (ages 5 to 8 yrs). After at least 2 yrs of cochlear implant experience, subjects in the younger group are about the same age on average as subjects in the middle age group were prior to receiving a cochlear implant. We also compared the post-implant composite scores of the middle age group to the pre-implant scores of the older group (ages 9 to 15 yrs).

T-tests revealed significant differences between the post-implant young age group scores and the pre-implant middle age group's scores for the word composite variable ($t = -3.99$, $p < .001$), the phoneme composite variable ($t = -6.12$, $p < .0001$), and the functional speech composite variable ($t = -2.49$, $p < .03$).[6] These results suggest that children who receive a cochlear implant at a young age surpass their age-matched peers who do not use a cochlear implant on measures of speech production, after at least 2 yrs of use. Similar analyses comparing the post-implant scores of the middle age group with the pre-implant scores of the older age group did not yield significant findings for any of the four composite variables ($p > .05$). That is, children in the middle age group who received a cochlear implant did not perform better than similarly aged children who were tested before receiving their cochlear implants.

[6]Because only two children in the younger group completed the FSST, we did not analyze the speech skills composite scores.

Table 3.
Correlations between the speech production composite scores and the audiological test scores.
(A = audition, V = Vision)

	Audiological Tests			
	Audio-Visual Feature Test		Children's Vowel Test	WIPI
Condition:	A	A+V	A	A
Speech Production Composite Scores				
Word	.898*	.839**	.569**	.739**
Phoneme	.807**	.730**	.464*	.683**
Speech Skills	.771**	.751**	.450*	.790**
Functional Speech	.935**	.899**	.589**	.789**

* $p = .01$
** $p = .001$

Relationship Between Speech Perception and Speech Production Performance

Most of the children recognized some speech when wearing their cochlear implants. Scores for the Audio-visual Feature Test (Perception Version) in an audition-only condition averaged 28% consonants correct (SD = 18%, chance = 10%), and in an audition-plus-vision condition, scores averaged 60% consonants correct (SD=24%). Scores for the Children's Vowel Test averaged 76% vowels correct (SD=24%, chance=25%) and scores for the WIPI averaged 51% words correct (SD = 23%, chance = 16%).

Scores from the four speech production composite variables were correlated with the four audiological test measures in order to examine the relationship between speech production and speech recognition skills. Pearson correlation coefficients are listed in Table 3 (next page). Subjects' abilities to produce speech corresponded well with their abilities to perceive it, as all relationships were significant.

Scores for the production version of the Audio-visual Feature Test were correlated with scores from the perception version. This analysis provided a direct comparison of subjects' abilities to produce and perceive the same stimulus items. Performance on the production version was significantly correlated with recognition performance when the test was presented in an audition-only condition ($r = .550$, $p < .01$) and also an audition-plus-vision condition ($r = .525$, $p < .01$). These results are presented in Figure 6.

Use of Sign Language

To ascertain whether children who have been implanted use primarily speech and little sign, we examined

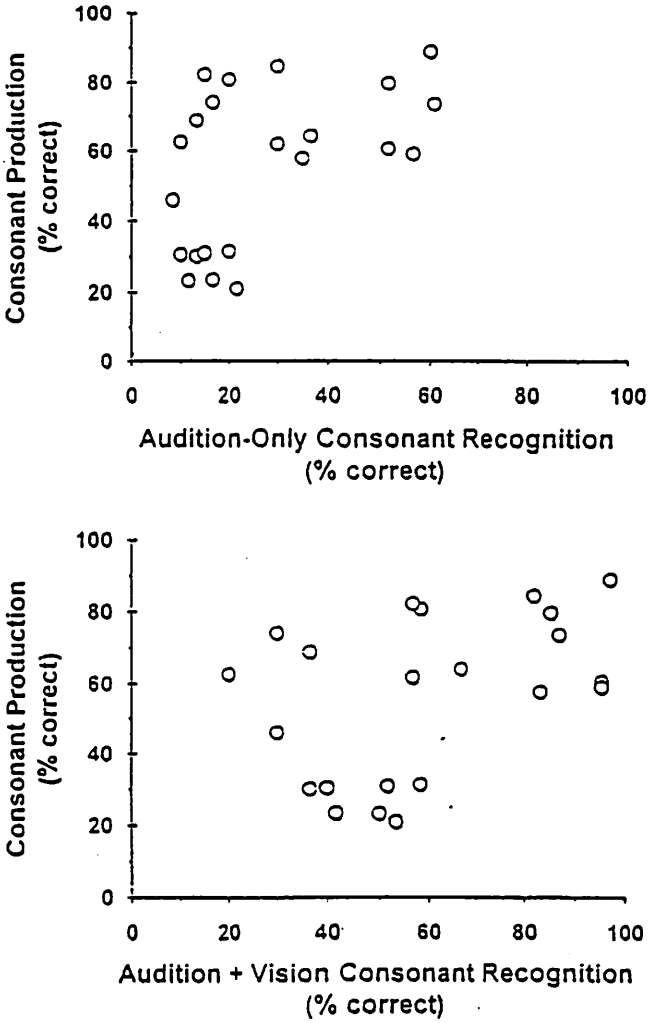


Figure 6. Scores for the Audio-visual Feature Test (Production Version) plotted as a function of performance on the perception version of the test, administered in an audition-only and an audition-plus-vision condition.

both the story-retell samples and the responses to the parental questionnaire. Performance on the story-retell task revealed that subjects clearly relied on both sign and speech to communicate with a speech-language pathologist, even after at least two yrs experience with a cochlear implant. Subjects used both sign and speech for an average of 83 words (SD = 29 words) within their 100-word samples. Subjects used speech-only for an average of 9 words (SD = 26 words) and sign without speech for an average of 7 words (SD = 16 words). Parents assigned an average agreement score of 4.8 (SD = 2.7) to the questionnaire statement, "My child communicates without sign, using speech only." Thus, the questionnaire results indicate a lower incidence of signing than in the story-retell procedure.

Discussion

This investigation included subjects who use simultaneous communication and who receive a public edu-

cation in their home communities. On these two counts, these subjects are representative of the majority of children who have profound hearing impairments in the United States. They may not be representative of children who attend private schools, nor children who receive an aural/oral education. The subjects had prelinguistic profound hearing impairments and wore a Nucleus cochlear implant for a minimum of 24 mos, and an average of 36 mos.

Speech Production After Implantation

Phoneme production accuracy for the present subjects exceeded reported performance levels of children with profound hearing loss who use hearing aids. The present subjects scored 53% phonemes (vowels and consonants combined) correct on the story-retell task and 54% phonemes correct on the Short-Long Sentence Test. Markides (1970) reports scores from a picture naming task of 44% vowels correct and 28% consonants correct for a group of profoundly hard-of-hearing hearing-aid users.

Overall intelligibility levels for the present subjects, however, were low, as indicated by the story-retell samples. Listeners assigned subjects low intelligibility ratings (2.5, on average, from a 10-point ascending scale), and only 22% of the words in the samples, on average, were produced without any phonemic errors. On the other hand, naive listeners were able to identify half of the stories from a closed set of six choices.

The relative error rates for the phonemes produced for the Audio-visual Feature Test (Production Version) suggest that subjects often approximate a target phoneme when they speak an error. Typically, the most common error made for a particular phoneme differed from the target by only one feature (i.e., voicing, manner of articulation, place of articulation). In some ways, their errors resemble those produced by children who are severely or profoundly hearing impaired and use hearing aids. For instance, Smith (1975) reported that her subjects frequently produced errors of voicing, which were also common in this investigation. Her subjects produced consonants that are visible on the mouth (/b, m, p/) relatively accurately and the fricative /z/ with many errors, as did the present subjects during both the story-retell task and the Audio-visual Feature Test (Production Version).

Most subjects performed well on the FSST, scoring above average for the normative data presented by the test developers. This good performance may indicate that the cochlear implant is especially beneficial for the development of suprasegmental aspects of speech, such as syllable stress and intonation. In addition or alternatively, changes in suprasegmental aspects of speech may be precursors to changes in segmental aspects. Ling's (1976) hierarchical stages of speech acquisition suggest that children who are hard-of-hearing may master syllable stress and intonation production before acquiring the skills to speak many speech sounds.

Age at Implantation

Children who are deaf and who receive a Nucleus cochlear implant before the age of 5 yrs may show greater benefit in terms of speech production measures than children who receive a cochlear implant after the age of 5 yrs. This tentative conclusion is based on two findings. First, younger subjects appeared to demonstrate a faster rate of improvement in their speaking skills than older subjects after receiving a cochlear implant, as indicated in Figures 4 and 5. The slopes of the lines connecting subject's pre-implant and post-implant phoneme and word composite scores in these figures were steeper for younger children than for older children, indicating that they were demonstrating an accelerated rate of improvement over time. The younger children may eventually surpass the older children with continued cochlear implant experience. It is possible, however, that, their improvement rate may begin to plateau with maturation and begin to resemble the improvement rate of older children.

The second indication that younger children may show greater benefit than older children comes from the analyses comparing the pre-implant measures of an older age group with the post-implant measures of a younger age group. The post-implant composite scores achieved by children in the 2 - 4 yrs age group were significantly better than the pre-implant scores achieved by children in the 5 - 8 yrs age group. A similar analyses comparing the post-implant scores of the 5 - 8 yrs age group to the pre-implant scores of the 9 - 15 yrs age group revealed no significant differences. Thus, younger children who used cochlear implants tended to perform better than similarly-aged children who used hearing aids. This was not true for older children.

Two caveats must be appended to this conclusion regarding age at implantation. First, it is quite possible that older children will also continue to improve their intelligibility with continued cochlear-implant experience. Progress may occur at a slower rate, or require longer lengths of experience. Secondly, even in the middle and older age groups, performance generally increased after subjects received a cochlear implant. This finding suggests that some older children may improve in their speaking behaviors. Future research must determine the proportion of older children who are likely to benefit from implantation, and identify those characteristics that predict gains.

Relationship Between Speech Perception and Production

Subjects' speech production skills were significantly correlated with their speech recognition skills. Chil-

dren with better speech production skills were more likely to have better speech recognition skills. Certainly the degree of speech recognition afforded by the cochlear implant appears to be an important factor in the variability that was noted in subjects' speech production performance.

These results underscore the importance of auditory information during speech acquisition. Tye-Murray (1992) suggests that auditory information plays at least five roles. First, auditory information potentiates the development of specific principles of articulatory organization. For instance, by listening to the speech of others in their community, children learn how to regulate their speech breathing, they learn how to flex and extend their tongue bodies, and they learn how to alternate rhythmically between open and closed postures of articulation. Secondly, by listening to others, children learn how to produce specific speech events. For instance, they learn to distinguish /p/ with a relatively rapid velocity opening gesture and /w/ with slow velocity. Thirdly, children develop a system of phonological performance (i.e., they learn the phonemes of their language community). Through listening, they also learn linguistic rules; for example, they learn which phonemes may occur in series (e.g., /tn/ does not occur in syllable-initial position). Fourthly, auditory feedback informs children about the consequences of their articulatory gestures, and how these consequences compare to sounds produced by other talkers. Finally, auditory feedback may provide information for monitoring ongoing speech production and for detecting errors.

The present results hint that a degraded signal (i.e., that provided via a cochlear implant) can perform some of the roles of an intact auditory system. For example, the significant correlations between scores for the production and perception versions of the children's Audio-visual Feature Test suggest that new listening skills are leading subjects to expand their systems of phonological performance. Their good performance on the FSST, compared to children who use hearing aids, suggest that access to auditory information is leading them to learn organizational principles, such as appropriate speech breathing behaviors.

Use of Sign Language

Results from the story-retell procedure suggest that children who used manual communication or simultaneous communication prior to receiving a cochlear implant do not discontinue signing after receiving a cochlear implant. Use of simultaneous communication increases. The results from the parent questionnaire indicate a somewhat lower incidence of signing. Scores from the questionnaire probably reflect the child's performance in the home setting, where listeners are familiar with the child's speech and perhaps not very fluent in sign language. We have noted clinically that some children have sign skills that surpass those of their parents. It is possible that they might attempt to communicate with only speech more often when interacting with their parents than when interacting with a clinician who knows sign.

Summary

The present study suggests that cochlear implantation leads to enhanced speech development, particularly if children are able to recognize some speech auditorally with their devices. Receipt of a cochlear implant may be more beneficial for younger children than older children, although additional data are needed to understand the effects of age more completely. Overall, children's intelligibility remains low after an average of three yrs of device use, although some children speak remarkably well. Cochlear implantation and subsequent improvements in speech skills appear not to eliminate children's use of sign to communicate.

Acknowledgements

This work was supported by the National Institutes of Health Grants DC00242 and DC00976-01, Grant RR59 from the General Clinical Research Centers Program, Division of Research Resources NIH, and a grant from the Lions Clubs of Iowa. The audiological data reported in this investigation were collected in an experimental protocol supervised by Dr. Richard S. Tyler. We thank Karen Iler Kirk for her assistance in the early stages of this project. We also thank Elizabeth Gilbert, Stephanie McCarthy, Barb Mullin, Shelley Witt, and Melissa Figland for their help with data analysis.

References

Bornstein, H., Saulnier, K., & Hamilton, L., (Eds.). (1983). The Comprehensive Signed English Dictionary. Washington, DC: Gallaudet University Press.

Boothroyd, A. (1969). Distribution of hearing levels in the student population of Clarke School for the Deaf (SARP Report #3). Northampton, MA: Clarke School for the Deaf.

Firszt, J. B., Reeder, R. M., Zimmerman-Phillips, S., Tonokawa, L., & Proctor, L. A. (1991). Rehabilitative plans and resulting performance of children using cochlear implants. Paper presented at the Academy of Rehabilitative Audiology Summer Institutes XXV, Breckenridge, CO.

Fryauf-Bertschy, H., Tyler, R. S., Kelsay, D., & Gantz, B.J. (1992). Performance over time of congenitally deaf and postlingually deafened children using a multichannel cochlear implant. Journal of Speech and Hearing Research, 35, 913-920.

Gantz, B. J., Woodworth, G. G., Knutson, J. F., Abbas, P. J., & Tyler, R. S. (1993). Multivariate predictors of success with cochlear implants. Annals of Otology, Rhinology, and Laryngology, 102, 909-916.

Gold, T. (1978). Speech and hearing: A comparison between hard of hearing and deaf children. Unpublished doctoral dissertation. City University of New York Graduate Center, NY.

Gustason, G., Pfetzing, D., & Zawolkow, E. (1980). Signing Exact English. Los Alamitos, CA: Modern Sign Press.

Hays, W. L. (1973). A measure of predictive association. In W. L. Hays (Ed.), Statistics for the Social Sciences, Second Edition, (pp. 745 - 749). New York, NY: Holt, Rinehart and Winston, Inc.

Kirk, K. I., & Hill-Brown, C. (1985). Speech and language results in children with a cochlear implant. Ear and Hearing, 6, 365-47S.

Kluwin, T. N. (1992). What does "Local public school" program mean? In T. N. Kluwin, D. F. Moores, & M. G. Gawstad (Eds.), Toward effective public school programs for deaf students: context, process, and outcomes (pp. 7-29). New York, NY: Teachers College Press.

Lane, H. (1990). Cultural and infirmity models of deaf Americans. Journal of the Academy of Rehabilitative Audiology, 23, 11-26.

Levitt, H., Youdelman, K., & Head, J. (1990). Fundamental speech skills test. Englewood, CO: Research Point, Inc.

Ling, D. (1976). Speech and the hearing impaired child: Theory and practice. Washington, DC: Alexander Graham Bell Association for the Deaf.

Markides, A. (1970). The speech of deaf and partially hearing children with special references to factors affecting intelligibility. British Journal of Disorders of Communication, 5, 126-140.

Monsen, R., Moog, J. S., & Geers, A. E. (1988). CID Picture SPINE. St. Louis: Central Institute for the Deaf.

Moores, D. F. (1992). An historical perspective on school placement. In T. N. Kluwin, D. F. Moores, & M. G. Gaustad (Eds.), Toward effective public school programs for deaf students: Context, process and outcomes (pp. 7-29). New York, NY: Teachers College Press.

National Association of the Deaf. (1991). Cochlear implants in children. The NAD Broadcaster, 13. The National Association of the Deaf, Silver Spring, MD.

Osberger, M. J., Maso, M., & Sam, L. K. (1993). Speech intelligibility of children with cochlear implants, tactile aids, or hearing aids. Journal of Speech and Hearing Research, 36, 186-203.

Osberger, M. J., Robbins, A., Berry, S., Todd, S., Hesketh, L., & Sedy, A. (1991). Analysis of the spontaneous speech samples of children with cochlear implants or Tactile Aids. American Journal of Otology, 12(Suppl.), 151-164.

Smith, C. (1975). Residual hearing and speech production in deaf children. Journal of Speech and Hearing Research, 18, 795-811.

Stark, R., and Levitt, H. (1974). Prosodic feature reception and production in deaf children. Journal of the Acoustical Society of America, 55(Suppl.), S23.

Tobey, E. A., Angelette, S., Murchinson, C., Nicosia, J., Sprague, S., Staller, S., Brimacombe, J., & Beiter, A. L. (1991). Speech production performance in children with multichannel cochlear implants. American Journal of Otology, 12(Suppl.), 165-173.

Tobey, E. A., & Hasenstab, M. S. (1991). Effects of a Nucleus multichannel cochlear implant upon speech production in children. Ear and Hearing, 12(Suppl.), 48S-54S.

Tye-Murray, N. (1992). Articulatory organizational strategies and the roles of audition. Volta Review, 94, 243-260.

Tye-Murray, N., & Kelsay, D. (1993). A communication training program for parents of cochlear implant users. The Volta Review, 95, 21-31.

Tye-Murray, N., & Kirk, K. I. (1993). Vowel and diphthong production by young users of cochlear implants. Journal of Speech and Hearing Research, 36, 488-502.

Tyler, R. S., Fryauf-Bertschy, H., & Kelsay, D. (1991). Audiovisual feature test for young children. Iowa City: The University of Iowa Hospitals, Department of Otolaryngology -- Head and Neck Surgery.

Tyler, R. S., Opie, J. M., Fryauf-Bertschy, H., & Gantz, B. J. (1992). Future directions in cochlear implants. Journal of Speech-Language Pathology, 16(2), 151-164.

NCVS Status and Progress Report - 7
December 1994, 107-117

Required Number of Tokens to Establish Reliable Voice Perturbation Values

Ronald Scherer, Ph.D
Vernon Vail, B.S.
Chwen Guo, M.S.
Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts

Abstract

Acoustic perturbation analyses of prolonged vowels are used in determining phonatory stability characteristics. When a number of tokens are analyzed, the average perturbation values create a stability profile of the voice. How many tokens are needed to establish a reliable average perturbation value? In this study, five perturbation measures were considered, namely, jitter, shimmer, harmonics-to-noise ratio, coefficient of variation for amplitude, and coefficient of variation for frequency. Subject groups were chosen on the basis of the subjects' average perturbation values. Results indicate that, except for the harmonics-to-noise ratio, generally the less stable the voice, the greater is the number of tokens needed to obtain reliable averages. For highly stable voices, at least six tokens are suggested; for voices with normal to high levels of instability, at least fifteen tokens are recommended. Regardless of vocal stability, at least ten tokens are suggested for the harmonics-to-noise ratio measure.

Introduction

Phonation involves cyclic vibration of the vocal folds and the creation of a repetitive glottal volume flow signal. The vibratory and flow characteristics are not exactly the same from cycle to cycle, giving rise to mechanical and acoustic instabilities of the voice (Simon, 1927). Certain levels of instability are considered normal, but higher levels may give rise to abnormal voice qualities and may reflect significant laryngeal problems (Lieberman, 1963).

The topic of vocal instability involves issues of methodology in obtaining instability information, interpretation of that information relative to laryngeal physiology and disorders, and the application of the interpretations to clinical and training decisions. The areas of methodology, interpretation, and application are not well developed.

Regarding methodology, a number of questions can be asked, the answers to which will affect the validity, reliablity, and usefulness of the information obtained. For example, (1) What signals should be recorded and why? Should s/he be neural, neuromuscular, aerodynamic, kinematic, or acoustic? (2) What should the subject or patient be asked to do? Should they produce steady vowel utterances or more dynamic voice tasks? Which pitch and loudness levels should be used? Which vowel should be chosen? How many tokens or utterances should be produced? When should the recordings be made? (3) What are appropriate recording procedures? For acoustic recordings, which microphone should be used and at what mouth to microphone distance? What are the necessary room characteristics? What should the recorder and A/D converter specifications be? And (4) what analysis procedures should be used? Again using the acoustic analysis example, what methods should be used to determine periods and amplitudes? How many cycles should be analyzed? What algebraic expressions and classification schemes should be chosen for defining the perturbation measures? Optimal choices for these and other issues are unknown or nonstandard. Many of these issues are discussed in a number of publications (e.g., Askenfelt & Hammarberg, 1986; Baken, 1987; Hillenbrand, 1987; Ludlow, Bassich, Conner, Coulter, Coulter & Lee, 1987; Titze, Horii & Scherer, 1987; Doherty & Shipp, 1988; Deem, Manning, Knack & Matesich, 1989,1991; Higgins & Saxman, 1989; Pinto & Titze, 1990; Karnell, Scherer & Fischer, 1991; Titze, 1991; Titze & Liang, 1993; Titze & Winholtz, 1993; Titze, 1994).

This paper addresses one of the methodological questions relevant to the tasks of the patient or subject - the number of tokens required to establish a reliable average

acoustic perturbation value. The hypothesis is that a reliable value for a steady phonation[1] may require a number of tokens because it is likely that there will be variability of laryngeal stability from one token to the next. This in turn is based on the assumption that the biomechanical conditions of the larynx and respiratory system can not be identical from one utterance to the next. Would the analysis of one token be sufficient (Jafari, Till, Truesdell & Law-Till, 1993), or must a relatively large number of tokens be analyzed and averaged?

Several studies that have discussed variability of measures and number of tokens (Scherer, Gould, Titze, Meyers & Sataloff, 1988; Titze et al., 1987; Linville & Korabic, 1987; Linville, 1988) have suggested that the average of a number of tokens is important, but no explicit rule has been determined. Scherer et al. (1988) accepted 7 to 15 tokens to obtain average values after realizing (Scherer, Titze, Raphael, Wood, Ramig & Blager, 1987) that 3 tokens were insufficient. Titze et al. (1987) used as many as 32 tokens in one of their experiments to guarantee the establishment of reliable mean perturbation values. Glaze, Bless, Milenkovic, and Susser (1988) used 10 tokens to obtain average perturbation values for children. Linville (1988) suggests that younger female subjects require fewer tokens to establish "steadiest" phonations compared to older females.

The purpose of this paper was to establish the number of analyzed tokens necessary to obtain a reliable average perturbation value. It was suspected that a reliable average perturbation value would be obtained with fewer tokens for a more stable voice.

Method

Previously analyzed voice recordings held at the Wilbur James Gould Voice Research Center were examined for the average perturbation values obtained for sustained steady /a/ vowels produced at comfortable and constant pitch and loudness levels. The tokens for every individual were recorded in consecutive order during one sitting. Usually only a single breath was taken between each vowel prolongation.

Subjects were recorded in IAC sound treated booths in medical facilities or at The Denver Center for the Performing Arts. The microphone used at each recording site was an AKG 451EB with a CK 22 capsule, an AKG N-62E phantom power supply, and an ATI M-1000 Version-1 amplifier. At the time of the subjects' involvement, vowel prolongations were digitized and video encoded through a Sony PCM-601ES and recorded onto video tape. The utterances were subsequently digitized (Digital Sound Corporation, 16 bit A/D convertion) onto computer disk (Digital Equipment Corporation VAX 11/750).

The utterances were analyzed for perturbation values using custom software (Titze, 1984; Scherer et al., 1988; software name: GL41). A peak picking procedure with quadratic interpolation was used to obtain period and amplitude reference points. The perturbation measures used in this study (refer to the Appendix for definitions of these measures) were jitter (JIT), shimmer (SHM), coefficient of variation for amplitude (CVA), coefficient of variation for frequency (CVF), and harmonics-to-noise ratio (HNR). JIT and SHM are "short-term" (cycle to cycle) measures, and CVA and CVF are "long-term" (over many cycles) measures (Scherer, 1991). One hundred consecutive cycles, taken at least 200 msec after phonation onset, were analyzed for each token produced.

Slight modifications to the normative data given in Scherer et al. (1988) following a reanalysis of that data (Table 1) were used to categorize subject recordings by level of perturbation value.[2]

For all 5 acoustic perturbation measures studied, subject recordings were chosen and grouped into three categories according to the level of the average perturbation values taken over the set of tokens (8 to 15) produced (Figure 1). Group A consisted of those with low perturbation values (less than or equal to approximately one standard deviation below the normative average perturbation value). Group B consisted of subject recordings with normal values (near the normative average perturbation value). Group C consisted of subject recordings with high values (greater than one standard deviation above the normative mean). Some of the

1. Measures of phonatory stability often involve perturbation analyses of vowels uttered with steady loudness, pitch, and vocal tract shape. This task is reasonable because it is essentially nonlinguistic, typically easily understood by the patient or subject, and can be imitated. In addition, prolonged steady vowels do not involve other potential perturbation causes inherent in linguistic laryngeal articulation or airway shape changes. Thus, the task of prolonging a vowel in a steady manner may elicit a person's minimal phonatory instability.

2. There are a couple of issues of a technical nature that should be mentioned. The first issue is the method by which the normative data were obtained (ref. Table 1). The original data were obtained during the dissertation process of Lorraine Olson Ramig (ref. Ramig & Ringel, 1983). A reel-to-reel tape recorder was used, which now can be thought to potentially create unwanted noise or fluctuations that would raise the values of the perturbations (Doherty & Shipp, 1988). The amount of this change is not well known at present, and may not be a serious issue for the present study. That is, even if the normative values for the perturbation measures used should be slightly lower, the norms were only used to separate the subjects into low, medium and high perturbation categories. The primary results of the study would be unchanged if the normative data were slightly shifted. Second is the definition of the perturbation measures jitter and shimmer. As Lemke and Samawi (1993) point out, the distribution of jitter and shimmer values, where the absolute difference between adjacent cycle values is taken, is not gaussian but skewed. This implies that a form of median value as the measure of central tendency should be considered in the jitter and shimmer definition itself, as well as quartiles for distribution markers. This can be considered an important refinement to present mean and standard deviation methods. For the present study, this issue is of some, but probably minimal, concern, since it is unlikely that changing the jitter and shimmer measure definition (for simple adjacent differences) will contradict the basic result here that higher perturbation values require more tokens to yield a reliable average value.

Table 1.
Group and normative means and standard deviations (in parentheses). The normative values are slightly changed from the values given in Scherer et al. (1988) following reanalysis of the recordings used in that study.

Groups	A	B	C	Norms
JIT	0.25	0.52	1.90	0.52
	(0.03)	(0.03)	(1.30)	(0.23)
SHM	1.05	1.88	5.68	1.88
	(0.08)	(0.15)	(1.66)	(0.83)
HNR	24.77	20.33	16.84	20.26
	(0.79)	(0.43)	(0.72)	(2.81)
CVA	3.26	6.77	11.37	6.68
	(0.75)	(0.24)	(1.05)	(3.03)
CVF	0.47	1.07	2.06	1.05
	(0.04	(0.06)	(0.46)	(0.40

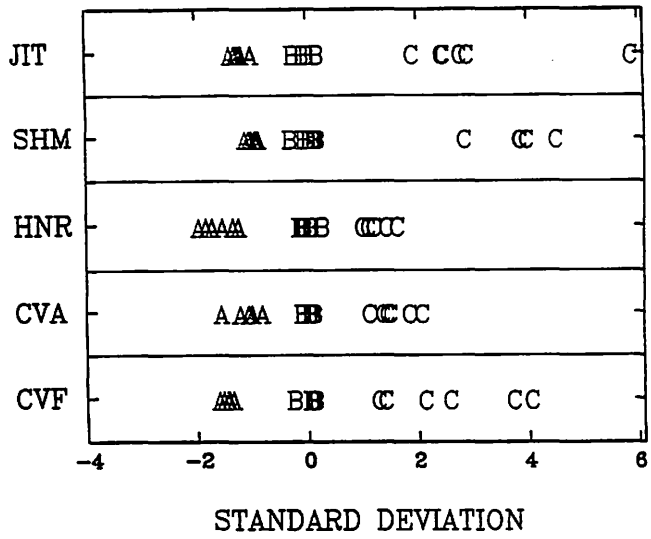


Figure 1. Distribution of mean perturbation values across perturbation measures and subject groups. JIT: jitter; SHM: shimmer; HNR: harmonics-to-noise ratio; CVA: coefficient of variation for amplitude; CVF: coefficient of variation for frequency; A: low mean perturbation subject group; B: average perturbation subject group; C: high mean perturbation subject group. The abscissa units are the standard deviations based on the normative values given in the right hand column of Table 1.

subjects were trained, some were untrained, and some had abnormal voices, but no decisions were made based on vocal training or normality per se. Figure 1 shows the distribution of mean pertubation values within the A, B, and C groups.

For JIT, the number of subjects in each subgroup A, B and C was 8. There were 6 subjects in each of the other subgroups of perturbation measures. For each subgroup, half of the subjects were men and half were women. Overall, the ages ranged from 28 to 81 years for the men, and 22 to 68 years for the women.

Some subjects were used for more than one perturbation measure. This was possible because the groupings were determined by the level of the perturbation values, not by diagnosis or other constraints. For example, the utterances of subject HSM (ref. Table A-1 in the Appendix) were used for a HNR and a CVA subgroup. Twenty four of the 51 subjects contributed only to one perturbation measure; 16 subjects to two measures; 7 to three measures; and 4 to four measures.

In the original recordings for all the subjects of prolonged, steady /a/ vowels, 8 to 15 tokens were produced (as mentioned above) and analyzed (average = 14.4 tokens). A *running average* of the token analyses was calculated; the tokens were taken as they were originally chronologically produced. That is, the average of a perturbation measure for tokens 1 and 2 was calculated, as well as the average of the analysis for tokens 1, 2, and 3. This technique continued up to the calculation of the average over all tokens produced by the subject.

The assumption was made that the more tokens used in the running average, the more asymptotic would be the average perturbation value. This required setting a window or range of values that defined the asymptotic criterion. Thus, the *minimal number of tokens* required to establish the reliable average value for a measure would be reached when adding more tokens to the running average kept the average within the criterion range. A narrow criterion range should require more tokens than a wide criterion range.

We used a set of criterion ranges based on different percentages of the normative standard deviations (ref. Table 1) for the perturbation measures. The criterion ranges were +/- 5% to +/- 50% of the normative standard deviations, in steps of 5%. For example, the normative mean value for JIT was 0.52, with a standard deviation value of 0.23. Thus an asymptotic criterion range of +/- 10% was 0.52 +/- 0.023, or a range of 0.497 to 0.543.

Results

If the perturbation values produced by a subject were to vary too much within a set of tokens, then the chronological cumulative average would not asymptote into a chosen criterion range. Figure 2 (next page) illustrates three cases for JIT at the +/- 25% criterion level (+/- 0.0575 JIT). For each case, successive data points indicate the cumulative average. The last data point on the right is the final cumulative average, around which the criterion range is indicated by the dotted lines. The cumulative average is within the criterion range even from the first token for Case 1. Case 2 illustrates asymptotic behavior by the eighth token. Case 3 does not asymptote over the available 15 tokens.

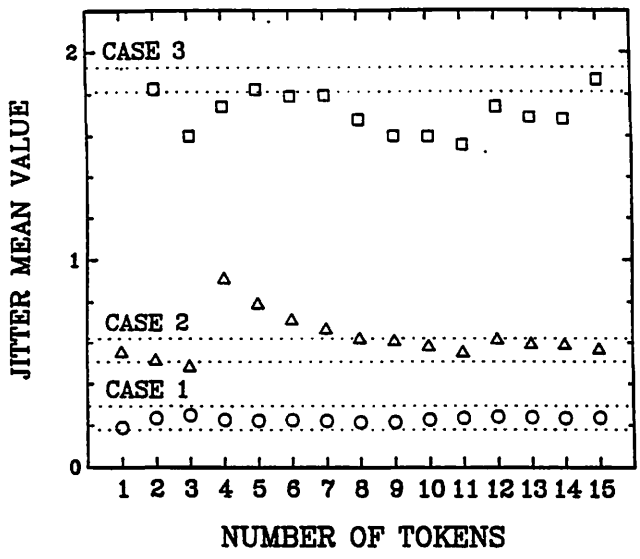


Figure 2. Chronological display of the cumulative average jitter values for three cases. Dashed lines refer to the criterion range of +/- 25% of one normative standard deviation for jitter. See text for explanation. The value for the first token for Case 3 was out of graphing range.

Table 2.
Percentage of cases for which the cumulative average perturbation value asymptoted within each criterion level within the subject's sequence of 15 tokens.

CRITERION LEVEL*	JIT	SHM	CVF	CVA	HNR
5	67%	89%	72%	78%	61%
10	75%	100%	88%	94%	89%
15	83%	100%	94%	94%	94%
20	87%	100%	94%	94%	100%
25	92%	100%	94%	94%	100%
30	92%	100%	94%	94%	100%
35	92%	100%	94%	94%	100%
40	96%	100%	94%	94%	100%
45	96%	100%	94%	100%	100%
50	96%	100%	94%	100%	100%

* +\- percent of one standard deviation

Table 2 gives the percentage of cases that showed asymptotic behavior within the criterion ranges. For example, for JIT and the +/- 5% level, 67% of the cases in the combined A, B, and C groups (16 out of 24) showed asymptotic behavior somewhere within the sequence of 15 tokens, and 33% (8 out of 24) did not. For those 8, more than 15 tokens would have been necessary to establish a reliable average within that criterion. Relatively more cases were asymptotic as the criterion range increased. By the +/- 25% level, more that 90% of all cases reached asymptote within the sequence of 15 tokens.

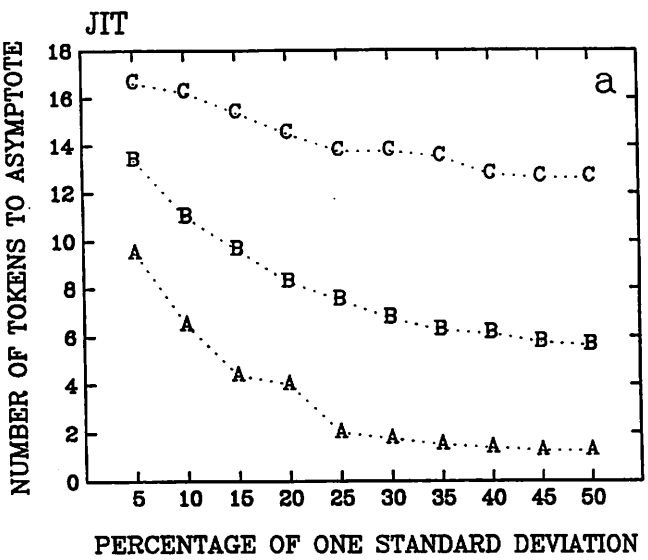
Figure 3 shows the main results of this study. The ordinate is the number of cumulatively averaged tokens that were necessary to satisfy the asymptotic criterion. The abscissa shows the asymptotic criterion in percent of the normative standard deviation for the given perturbation measure. The criterion ranged from +/- 5% to +/- 50%, increasing in steps of 5%. For each perturbation measure (Fig. 3a-3e), the data are given for each of the subgroups (A, B, and C). As the criterion range widened, fewer tokens were required to keep the cumulative average within the criterion range, and thus the values decrease left to right. For the JIT (Group C) cases without asymptotic values after all 15 tokens, 17 tokens was established as the approximate number of tokens it would have taken to reach an asymptotic average. This number did not seem unreasonable, and allowed greater ease in further analysis.

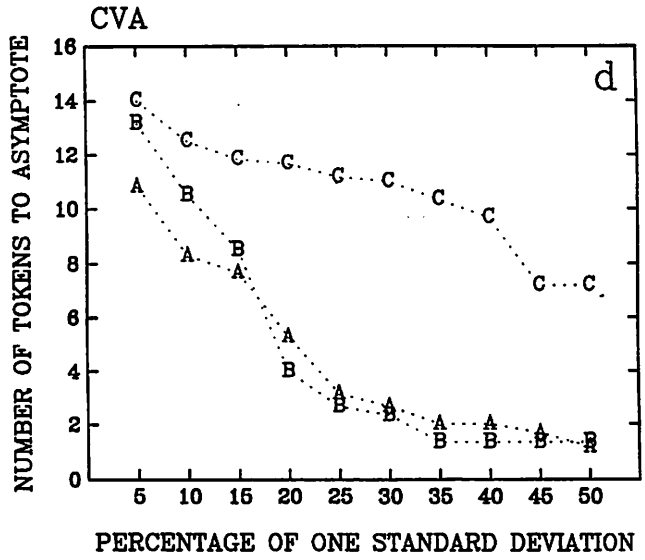
In Figure 3a, the asymptotic values for groups A, B, and C for JIT do not overlap. At any criterion level, the groups appear well separated. Table 3 gives statistical comparison results. T-tests were run using the SAS (1985) GL procedure (Bonferroni). This test was run at an alpha level of 0.05. All but one of the 30 comparisons for JIT were significantly different. This suggests, at least, that fewer tokens are needed to establish an asymptotic JIT average for a voice with low JIT (group A) than for a voice with high JIT (group C).

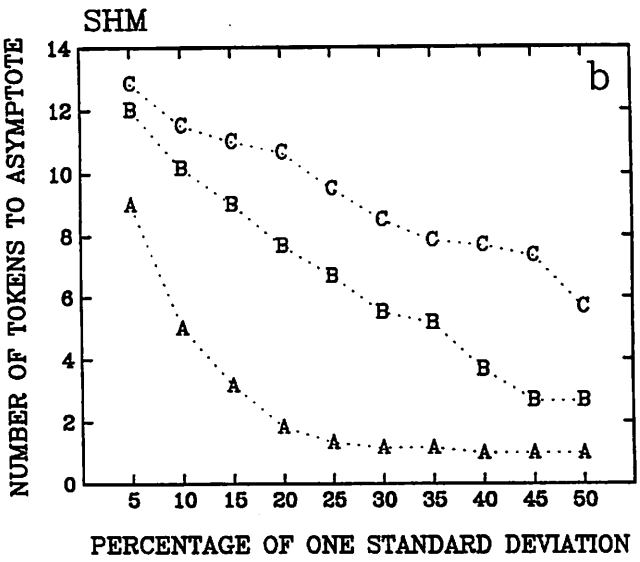
Figure 3 and Table 3 indicate that, of the perturbation measures examined, JIT appears to form the most consistent separation between groups A, B, and C. Other perturbation measures were less consistent. For all measures except HNR, a significant difference in the required number of tokens was found when comparing groups A and C for all criterion levels above 10%. HNR showed no significant differences between groups A, B, and C (Fig. 3e). That is, even when subjects were separated by relatively high and low HNR values, the number of tokens necessary to establish reliable HNR averages did not differ between the groups, but did decrease as the criterion level increased.

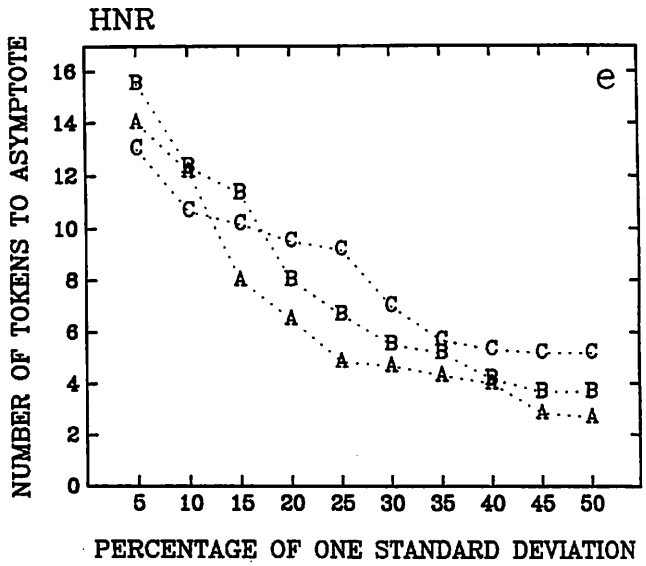
The results for the +/- 25% criterion level will now be shown in detail. The 25% level is chosen because it appears to be reasonable for practical considerations. We will base the final recommendations on this criterion level, but it is realized that the choice of criterion may be application specific.

Figure 4a gives the mean and standard deviation extents of the five perturbation measures for the +/- 25% criterion level. Thus, Figure 4a indicates how many chronologically produced tokens it took for the cumulative average to asymptote at +/- 25% of one standard deviation of the









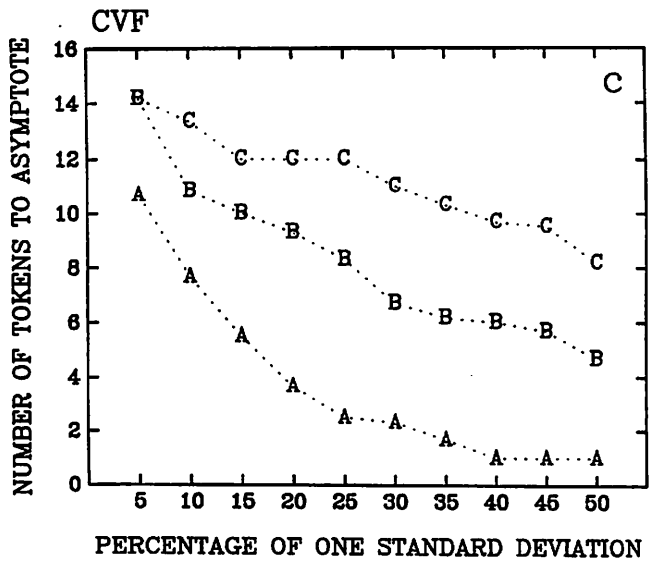


Figure 3. Number of tokens needed to reach asymptotic levels for cumulative average perturbation values. Each figure refers to a different perturbation measure (3a: JIT - upper left; 3b - middle left: SHM; 3c - lower left: CVF; 3d - top right: CVA; 3e - bottom right: HNR). The abscissa is the criterion range, and the ordinate is the required number of tokens to satisfy the criterion. Data are displayed relative to perturbation severity group (A,B, or C). See text for further explanation.

Table 3.
Paired comparisons T-tests for all criterion levels, subject (perturbation level) groups, and perturbation measures, to test for difference in number of tokens to asymptote, using the Bonferroni procedure to control the number of paired comparisons error. An "X" indicates that the comparison was significant at the 0.05 alpha level.

CRITERION LEVEL*	JIT	SHM	CVF	CVA	HNR
	A/A/B/ B C C	A/A/B/ B C C	A/A/B/ B C C	A/A/B/ B C C	A/A/B/ B C C
5	X X X				
10	X X X	X X		X	
15	X X X	X X	X	X	
20	X X	X X	X	X X	
25	X X X	X X	X X	X X	
30	X X X	X	X	X X	
35	X X X	X	X	X X	
40	X X X	X	X	X X	
45	X X X	X X	X	X X	
50	X X X	X	X	X X	

* +/- percent of one standard deviation. A comparison of the required number of tokens for the A group versus the B group is indicated by the letter "A" placed over the letter "B". Other comparisons are similarly indicated.

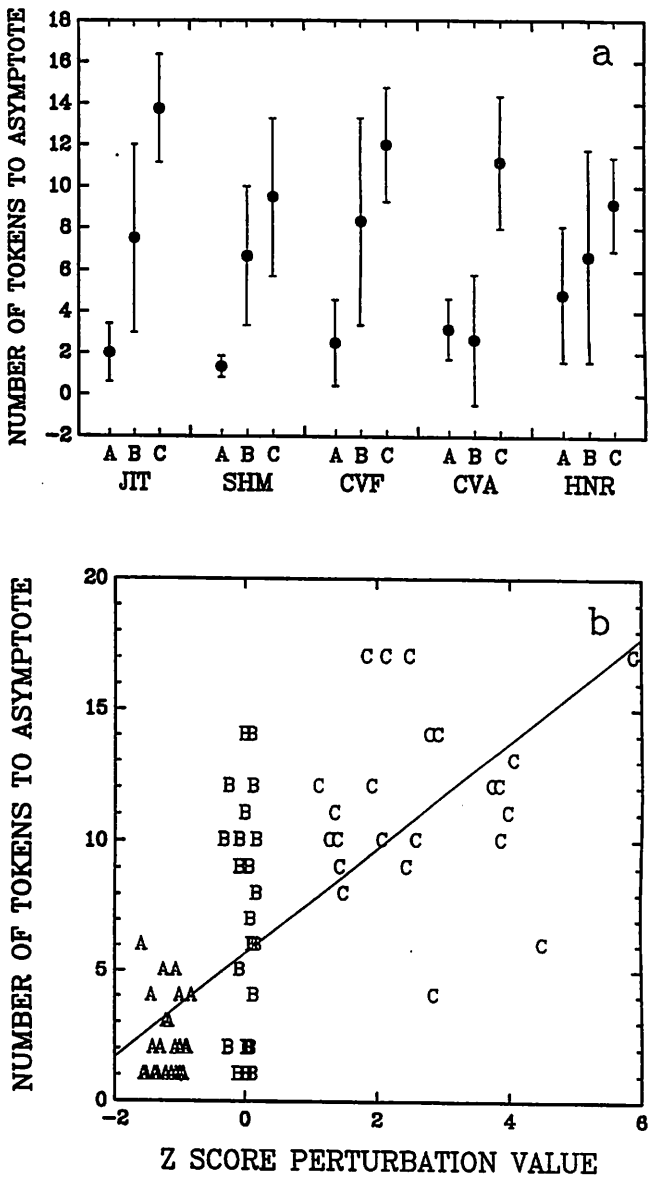


Figure 4. Results for the +/- 25% criterion range. Figure 4a (top) shows the mean and +/- 1 st dev of the asymptotic number of tokens for each of the perturbation measures. Figure 4b (bottom) shows the general trend across the z-score levels (ref. Figure 1), with the harmonics-to-noise ratio measure excluded. The equation for the line is y = 2.01x + 5.68.

normative data for each of the perturbation measures used. For JIT and the low perturbation subject group A, an average of only 2 tokens was needed to reach the asymptote; that is, the average JIT value for 2 or more consecutive tokens fell within the +/- 25% bounds. The standard deviation for this group was approximately one and one-half tokens. The average number of tokens needed to asymptote for the JIT subject group B was almost 8, with a relatively large standard deviation of about 4 1/2 tokens. The average number of consecutive tokens needed to asymptote for group C was about 14, with a standard deviation of almost 3. As indicated in Table 3, these three groups were statistically significantly different at the 0.05 level. Similar discussions can be offered for the other four perturbation measures, where the average values and significant differences vary.

The data shown in Figure 4a suggest, for example, why the HNR does not indicate significant differences (as shown in Table 3): The averages to reach criterion for the three groups A, B, and C have relatively large standard deviations that prevent a confident separation of the means, and therefore at the +/- 25% criterion level, groups A, B and C do not appear to differ in the number of tokens needed to obtain a reliable HNR value. However, Figure 4a does suggest that, in general (excluding HNR and noting the reversal for CVA groups A and B), as the perturbation of the

Table 4.
Mean of the standard deviations of the perturbation measures across subject groups A,B, and C, and results for the test for difference between variances of two independent samples (Bruning & Kintz, 1968) (at 0.05 level).

	Subject Groups			Significant Difference Between Variances		
	A	B	C	A vs B	A vs C	B vs C
JIT	.0752	.235	.889	Y	Y	Y
SHM	.166	.511	1.20	Y	Y	N
CVA	1.21	2.02	4.76	N	Y	N*
CVF	.147	.375	.648	N*	Y	N
HNR	2.11	2.19	2.34	N	N	N

* The only tests that differed in significance results compared to the significance tests for the +/- 25% criterion level given in Table 3.

voice increases, more tokens are needed to produce a reliable average perturbation value (a value that does not significantly change if more tokens are included into the average).

This trend is also indicated in Figure 4b, which groups all the data (excluding the HNR measure) by z-score (that is, the difference between the subjects' average perturbation scores and the normative averages for the measures, divided by the normative standard deviations; ref. Figure 1). Although there is a fair amount of scatter (R-square = 0.36; F = 42.07; sig. level = 0.0001), the trend is to increase the required number of tokens to reach reliable perturbation measure averages as the perturbation value for the subject is increased. The figure suggests that for the low perturbation group, the number of tokens is lower (max=6) than that required for the high perturbation group (max=17, the artificially specified value, see above). For the normal perturbation group near a z-score of zero, the scatter is relatively wide (1-14), suggesting that for normal voices, the choice of the number of tokens must be made with care: for example, a choice of 7 tokens may not be adequate.

Table 4 gives the mean of the standard deviation values for each subgroup for each perturbation measure. The approximately 15 token values for each subject for each perturbation measure were used to obtain a standard deviation value. Each cell in the left hand portion of Table 4 is the average of all the standard deviation values corresponding to the subject group - perturbation measure cell designation.

The right hand portion of Table 4 indicates the results of a significance test between variances (Bruning & Kintz, 1968) among the A, B and C subject groups. The "Y" represents a significant difference between two groups at the 0.05 level, and "N" indicates that this was not the case.

The results in Table 4 are consistent with the significance tests of Table 3 for the +/- 25% criterion level in that all but two of the 15 tests were the same in significance. That is, except for two cases, when the difference between two of the A, B and C groups was significant for the required number of tokens, the corresponding comparison between the variances of the perturbation measure values was also significant. For example, it took significantly more averaged tokens for group B than group A to reach asymptote for a reliable JIT measure (Table 3), and also the variance of the B JIT measures was significantly larger than for the A JIT measures (Table 4). Thus, in general, significant differences in the required number of tokens for a perturbation measure correspond to significantly different measure variances for that perturbation measure.

Discussion

Under the assumption that the acoustic analysis of prolonged vowels produced in a steady manner may reveal voice instabilities and other underlying physiological functions (e.g., Scherer et al., 1988), the practical matter of how many tokens to elicit from a patient or subject is important. It is important from the standpoints of both subject and analysis involvement. Essentially, the fewer the number of tokens that are required for a reliable measure, the less recording time and data storage that will be needed, and the lower the probability for patient/subject fatigue.

This study suggests that the variability of perturbation measure values for repeated tokens must be considered. The approach of this study was to average perturbation values obtained from consecutively prolonged (steady) /a/ vowels produced at comfortable pitch and loudness levels. The running average was found to change less as subsequent tokens were included. The number of tokens required for the running average to reach a relative constancy depended on the criterion range (Figure 3). For each perturbation measure, the criterion range was designated as a percentage of one normative standard deviation (Table 1). The study reported here found, as would be expected, that if the criterion range is wider, relatively fewer tokens are needed to establish constancy.

Not only did the required number of tokens vary according to the criterion range, but also according to the original perturbation values of the subjects. Excluding the HNR measure,subjects with greater acoustic perturbation values typically required more tokens to reach any given constancy (or satisfy an asymptotic) criterion. These results can be explained by the increased variability for a subject's

set of tokens for those measures with larger perturbation values (Table 4). That is, a greater variability of a perturbation measure across the produced tokens tends to correspond to a larger number of required tokens to reach an asymptotic constancy level. These generalizations hold best for the comparison between the low and the high perturbation groups, and do not hold for the HNR measure.

The reasons for the HNR measure to differ from the other measures also appear to lie in the variability of this measure. Note that in contrast to JIT and CVF, which involve min-picked frequency measurements in this study, and SHM and CVA, which involve peak-to-peak amplitude measurements in this study, HNR involves not only period and amplitude, but also waveshape measurements. Even though the HNR value may be relatively high or low, the various dimensions (period, amplitude, and waveshape) may interact to keep token variability (Table 4) and the number of required tokens (Figure 3e) consistent. Normal variation of the value of the HNR measure is small compared to the other perturbation measures. Table 1 indicates that for the normative data, the standard deviations are a relatively large portion of the normative mean values (JIT: 0.44; SHN: 0.44; CVA: 0.45; CVF: 0.38), but relatively small for HNR (0.14). This intrinsically smaller range of values may keep the variability relatively constant (Table 4) no matter which level of HNR if produced, thus having little effect on the cumulative averaging procedure used in this study. It is noted that an advantage of the HNR measure may be that the required number of tokens to reach an asymptotic level is independent of a subjects' perturbation level (A, B, C), since then the same number of tokens needed could be used for every subject.

For certain measures, such as JIT, the results of this study suggest that fewer tokens are needed if the voices already have known relatively low perturbation measures. This is problematic under most circumstances, since the objective usually is to measure perturbation for the first time. Eventually voice analysis may be conveniently automated so that an analysis could be performed each time the subject or patient produced a prolonged vowel. A running average of the perturbation values could then be obtained, with termination of the task of prolonging steady vowels when a criterion of little (but specified) change in the average was reached.

Until that time, a guideline related to *perceived voice quality* might be considered: Based on the criterion of +/- 25% of the normative standard deviations for the measures studied here, if the voice sounds normal or somewhat rough, or if the perception of the voice quality is not available, to guarantee a reliable average with reasonable confidence, 15 or more tokens (ref. Figure 4a,b) should be averaged. For a voice that sounds very clear, 6 or more tokens should be used for the JIT, SHM, CVA and CVF measures (and perhaps also for similar measures). For the HNR measure, 10 or more tokens should be considered regardless of voice quality.

There should be a trade-off relationship between the number of tokens required for a stable average of a perturbation measure and the number of consecutive cycles used in the analysis of each token. As more cycles are taken in the analysis of a token, an asymptotic value may be reached. Karnell (1991) found that, for subject groups separated by severity of hoarseness (no hoarseness to moderate hoarseness), jitter and shimmer appeared to asymptote at approximately 110 cylces. If the 25% criterion of our study is applied to the jitter ratio data of Jafari et al. (1993), asymptotic values developed at approximately 60 cycles for untrained normal voices. Deem et al. (1989) found that 40 cycles was sufficient for normal voices. It can be concluded that the 100 cycles used in the present study may have been near the required (asymptotic) number of cylces for most subjects.

Conclusion

This study examined the number of tokens required to determine a reliable average acoustic perturbation value for five voice perturbation measures. The number of required tokens typically varied with the asymptotic criterion level and the predetermined level of the voice perturbation.

Perhaps the best way to guarantee a reliable average value for a perturbation measure is to calculate the running average across token utterances while the person is producing them, and stop the process when the person's average falls within a reasonable criterion range. If this is not an option, the following should be considered.

1. The number of tokens required to establish a reliable average perturbation value depends on (a) the size of the asymptotic criterion range (the smaller the range, the more tokens required), (b) the degree of voice perturbation (the greater the voice perturbation of the subject, typically the more tokens needed), and (c) which perturbation measure is being used.

2. Based on the asymptotic criterion of +/- 25% of the normative standard deviation for the measures studied, low perturbation voices require 6 tokens, normal voices require 15, and high perturbation voices require 15 or more tokens to establish a reliable average with a high level of confidence. Some voices require fewer tokens, but these numbers can be considered as suggested guidelines.

3. For voices of unknown perturbation and quality characteristics, the results based on the criterion mentioned in #2 suggest that at least 15 tokens should be considered to obtain reliable average perturbation values.

4. The suggestions above may relate best to single dimension (time or amplitude) measures. The harmonics-to-noise measure (Yumoto, Gould, Baer, 1982), which com-

bines time, amplitude and waveshape factors, appears to be insensitive to the severity of the measure relative to the number of tokens required for asymptotic behavior. For the criterion given in #2, approximately 10 tokens were required regardless of measure severity.

5. The required number of tokens to reach a reliable average perturbation value may be affected by how many cycles in each token are analyzed. The trade-off relationship should be studied. Until then, using at least 100 consecutive cycles appears reasonable, and a cushion of a factor of 2 (200 cycles) most likely is near or beyond the required upper bound.

Acknowledgments

This work was supported by the National Institutes of Health grant number 8R01 DC00387. Appreciation is extended to R.T. Sataloff, A.D. Meyers, and the late W.J. Gould for use of selected patient data. Thanks also are given to Laura Goodwin for statistical advice and to Christopher Linker and Paula Saraceno for assistance. Special thanks are extended to Tom Doherty and Dale Metz for their helpful suggestions.

References

Askenfelt, A., & Hammarberg, B. (1986). Speech waveform perturbation analysis: A perceptual-acoustical comparison of seven measures. *Journal of Speech and Hearing Research*, 29, 50-64.

Baken, R.J. (1987). *Clinical measurement of speech and voice*. Boston: College-Hill Press.

Bruning, J.L., & Kintz, B.L. (1968). *Computational handbook of statistics*. Glenview, Illinois: Scott Foresman and Co.

Deem, J.F., Manning, W.H., Knack, J.V., & Matesich, J.S. (1989). The automatic extraction of pitch perturbation using microcomputers: Some methodological consideration. *Journal of Speech and Hearing Research*, 32, 689-697.

Deem, J.F., Manning, W.H., Knack, J.V., & Matesich, J.S.(1991). Comparison of Pitch Perturbation Extraction Procedures with Adult Male and Female Speakers. Folia Phoniatr, 43, 234-245.

Doherty, E.T., Shipp, T. (1988). Tape recorder effects on jitter and shimmer extraction. *Journal of Speech and Hearing Research, 31*, 485-490.

Glaze, L.E., Bless, D.M., Milenkovic, P., & Susser, R.D. (1988). Acoustic characteristics of children's voice. *Journal of Voice*, 2, 312-319.

Higgins, M.B., & Saxman, J.H. (1989). A comparison of intrasubject variation across sessions of three vocal frequency perturbation indices. *Journal of the Acoustical Society of America*, 86, 911-916.

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

Jafari, M., Till, J.A., Truesdell, L.F., & Law-Till, C.B. (1993). Time-shift, trial, and gender effects on vocal perturbation measures. *Journal of Voice*, 7, 326-336.

Karnell, M. (1991). Laryngeal perturbation analysis: Minimum length of analysis window. *Journal of Speech and Hearing Research*, 34, 544-548.

Karnell, M.P., Scherer, R.C., & Fischer, L.B. (1991). Comparison of acoustic voice perturbation measures among three independent voice laboratories. *Journal of Speech and Hearing Research*, 34, 781-790.

Lemke, J.H., & Samawi, H. (1993). The role of importance sampling in the establishment of normal ranges for speech characteristics. *Journal of the Acoustical Society of America*, 94, 1763(A).

Lieberman, P. (1963). Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. *Journal of the Acoustical Society of America*, 35, 344-353.

Linville, S.E. (1988). Intraspeaker variability in fundamental frequency stability: An age-related phenomenon? *Journal of the Acoustical Society of America*, 83, 741-745.

Linville, S.E., & Korabic, E.W. (1987). Fundamental frequency stability characteristics of elderly woman's voices. *Journal of the Acoustical Society of America*, 81, 1196-99.

Ludlow, C.L., Bassich, C.J., Conner N.P., Coulter, D.C., & Lee, Y.J. (1987). The validity of using phonatory jitter and shimmer to detect laryngeal pathology. In Baer T., Sasaki, C., & Harris, K. (Eds.), *Laryngeal function in phonation and respiration*. Boston: College-Hill Press.

Pinto, N.B., & Titze, I.R. (1990). Unification of perturbation measures in speech analysis. *Journal of the AcousticalSociety of America*, 87, 1278-1289.

Ramig, L.A., & Ringel, R.L. (1983). Effects of physiological aging on selected acoustic characteristics of voice. *Journal of Speech and Hearing Research,* 26, 22-30.

SAS User's Guide: Statistics, Version 5 Edition, Cary, NC: SAS Institute Inc.,1985, 956 pp.

Scherer, R.C. (1991). Aerodynamic assessment in voice production. In J.A. Cooper (Ed.), *Assessment of speech and voice production: Research and clinical applications.* NIDCD Monograph Volume 1-1991 (pp. 112-123). Bethesda: National Institute on Deafness and Other Communication Disorders, NIH.

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, 230-44.

Scherer, R.C., Titze, I.R., Raphael, B.N., Wood, R.P., Ramig, L.A., & Blager, F.B. (1987). Vocal fatigue in a trained and an untrained voice user. In T. Baer, C. Sasaki, & K.S. Harris (Eds.), *Laryngeal function in phonation and respiration*, San Diego: College-Hill Press.

Simon, C. (1927). The variability of consecutive wavelengths in vocal and instrumental sounds. *Psycological Monographs,* 36, 41-83.

Titze, I.R. (1984). Parameterization of the glottal area, glottal flow, and vocal fold contact area. *Journal Acoustic Society of America,* 75, 570-80.

Titze, I.R. (1991). A model for neurologic sources of aperiodicity in vocal fold vibration. *Journal of Speech and Hearing Research,* 34, 460-472.

Titze, I.R. (1994). Toward standards in acoustic analysis of voice. Journal of Voice, 8, 1-7.

Titze, I.R., Horii, Y., & Scherer, R.C. (1987). Some technical considerations in voice perturbation measurements. *Journal of Speech and Hearing Research, 30, 252-60.*

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). Effect of microphone type and placement on voice perturbation measurements. *Journal of Speech and Hearing Research,* 36, 1177-1190.

Yumoto, E., Gould, W.J., & Baer, T. (1982). Harmonics-to-Noise ratio as an index of the degree of hoarseness. *Journal Acoustic Society of America,* 71, 1544-50.

Appendix

The definitions of the perturbation measures used in this study are as follows:

Jitter and Shimmer:

$$\frac{100\sum_{i=2}^{N}|x_i - x_{i-1}|}{(N-1)\bar{x}}$$

where x_i is the frequency of the ith cycle for jitter and the amplitude of the ith cycle for shimmer. N is the number of consecutive cycles.

Coefficient of Variation for Frequency and for Amplitude:

$$\frac{100}{\bar{x}}\sqrt{\frac{\sum_{i=1}^{N}(x_i-\bar{x})^2}{N}}$$

where x_i is the frequency of the ith cycle for the coefficient of variation for frequency, and x_i is the amplitude of the ith cycle for the coefficient of variation for amplitude. N is the number of consecutive cycles.

Harmonics-to-Noise Ratio (Yumoto et al., 1982):

$$N\int_0^T f_A^2(t)\,dt \;/\; \sum_{i=1}^{N}\int_0^{T_i}[f_i(t)-f_A(t)]^2 dt$$

where

$$f_A(t) = \sum_{i=1}^{N}\frac{f_i(t)}{N}$$

where $f_A(t)$ is the average cycle waveform over N cycle waveforms $f_i(t)$, and T is the maximum period over the N periods T_i.

Table A-1.

Subject description and number of tokens required to asymptote at the 25% criterion level.
Dashes indicate that those perturbation analyses were not used for the particular severity group indicated.

					NUMBER OF TOKENS REQUIRED TO ASYMPTOTE					
SUB	GROUP	SEX	AGE	ANA#	JIT	SHM	HNR	CVA	CVF	CONDITION
HSM	C	M	81	15	--	--	11	11	--	HOARSE
SLS	C	F	42	15	9	4	--	12	10	CORDS ARE CLEAR
SFM	A	F	26	15	--	1	2	--	--	BILATERAL VOCAL NODULES
CB1	A	F	37	15	--	--	3	3	--	LEFT TVC NODULE (SMALL)
RMB	A	M	56	15	--	--	--	4	--	PARKINSON
VKS	B	F	32	15	--	--	5	1	--	SLIGHT EDEMA LVC, HOARSE
JHS	B	M	43	15	--	--	--	1	--	VENTRICULAR HYPERTHROPHY
GB1	B	F	31	15	14	--	--	--	--	BOWED RVC, HOARSENESS
CP1	B	F	42	18	--	--	--	--	12	VOCAL CORD PARALYSIS
CC1	B	M	59	15	11	--	--	9	--	STRIPPED VOCAL CORD
EW1	C	F	68	15	14	14	9	--	12	MILD EDEMA, HOARSE
FG1	B	M	49	15	9	--	--	--	--	NONE
GJS	B	M	49	15	--	7	--	--	--	BILATERAL HERNIA, HOARSE
CAR	C	F	44	15	14	--	11	--	--	DYSPHONIA, HOARSENESS
CAR	B				--	--	--	--	12	
MJF	B	M	46	15	8	6	--	--	--	NONE
MJF	A				--	--	5	--	--	
JLV	C	M	41	20	14	--	--	--	--	DYSPHONIA
JLV	B				--	--	14	--	--	
JAL	A	F	59	20	3	--	--	--	--	NONE
JAL	B				--	6	--	--	--	
KRA	C	F	31	15	13	11	--	--	13	HOARSENESS, DRYNESS
KRA	B				--	--	--	2	--	
GLR	B	M	43	15	10	--	--	--	--	POLYPOID DEGENERATION LVC
SFS	C	M	47	15	17	--	--	--	10	SPASTIC DYSFUNCTION
SFS	C				--	--	5	9	17	
GCK	B	M	37	15	--	--	11	--	--	MILD BOWING
BB1	B	F	29	15	5	1	--	--	2	SINGER
VL1	A	M	28	8	--	1	--	--	--	SINGER
CR1	A	M	30	15	5	--	--	--	1	SINGER
RC1	B	F	24	15	--	10	--	--	--	SINGER
RC1	A				--	--	--	--	1	
KK1	C	F	26	15	--	--	--	8	--	SINGER
THE	A	F	34	15	--	--	3	2	2	STIFFNESS OF VOCAL CORD
CDD	A	F	44	15	1	--	--	--	--	SPASMODIC DYSPHONIA,HOARSE
LRH	A	M	31	15	--	--	5	--	--	POLYPOID NODULE RVC
IK1	A	M	32	15	--	--	--	--	4	HOARSENESS, EDEMA
IG1	C	F	40	15	--	--	10	--	--	LVC POLYP
MB1	B	F	39	15	2	--	--	--	--	NONE
LNL	B	F	35	18	--	--	7	--	--	LVC NODULE
CW1	B	F	29	20	--	--	2	2	--	DECR. MUCOSAL MOTION LVC
MLL	C	M	38	15	--	6	9	10	--	TVC POLYP REMOVED 2 MONTHS
RJD	B	M	31	15	--	--	1	1	--	CYST ON LVC
LTH	A	F	35	14	2	1	--	--	--	NONE
LTH	A				2	2	--	--	--	
MCP	C	M	34	15	17	--	--	--	--	POSTERIOR VF GAP,HYPERFUNC
MCP	A				--	--	--	4	--	
MCP	B				--	--	--	--	14	
HK1	A	M	48	15	1	--	--	--	--	GRANULOMA R VOCAL PROCESS
HK1	B				--	10	--	--	--	
GLP	A	M	31	15	1	--	--	--	--	SMALL LARYNX
GLP	B				--	--	--	--	4	
DAH	C	M	52	15	--	--	--	--	10	SEVERE ABD. DYSPHONIA
GH1	C	M	58	15	--	12	--	--	--	MUSCULAR TENSION DYSPHONIA
GH1	B				--	--	--	--	6	
DH1	A	M	30	15	1	--	11	--	6	NORMAL LARYNX
MRE	C	F	68	15	--	--	--	17	--	FUNCTIONAL DYSPHONIA
JAK	C	M	53	17	--	10	--	--	--	LVF MOVES, LESS MOTION
PAJ	B	F	40	20	1	--	--	--	--	NONE
WLT	A	F	22	15	--	--	--	1	1	NONE
DG1	A	M	34	15	--	--	--	5	--	NONE
WES	C	M	71	15	12	--	--	--	--	POSSIBLE SUP NERVE DEFICIT
VF1	A	M	35	15	--	1	--	--	--	SINGER
MK1	A	M	50	12	--	2	--	--	--	SINGER

NCVS Status and Progress Report - 7
December 1994, 119-129

Pulsatile Airflow During Phonation: An Excised Larynx Model

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

Abstract

Pulsatile airflow in the excised larynx was investigated with simultaneous recordings of air velocity, subglottal pressure, volume flow, and the electroglottograph signal (EGG) for various conditions of the larynx. Canine larynges were mounted on a bench with sutures attached to cartilages to mimic the function of laryngeal muscles. Sustained oscillations were established and maintained with the flow of heated and humidified air through the trachea. The instantaneous air velocity above the glottis which is the summation of a periodic velocity and the turbulent component, was measured with a constant temperature hot-wire probe at various locations. The phase averaged velocity was used to construct the patterns of jet flow at selected time frames of the oscillation cycle. Results suggest that supraglottal air velocity is highly spatially and temporally dependent. Cycles of local air velocity with double peaks were not uncommon and a case is provided. For one phase-averaged phonatory cycle, a 9 x 13 velocity measurement grid demonstrated strongly non-uniform velocity surfaces for eight phases of the cycle, with greater velocities located anteriorly.

Introduction

The glottis is a three dimensional space that deforms dynamically during phonation. As a dynamic valve, it creates pulsatile flow through the glottis. The flow field consists of instantaneous air pressures and particle velocities spatially distributed within the glottis. From an acoustic point of view, the instantaneous velocities through the glottis, summed or integrated together over the glottal orifice, creates the glottal volume velocity, characteristics of which correspond strongly to theoretically and empirically determined spectral characteristics (Fant, 1970, 1986; Gauffin and Sundberg, 1989). From a biomechanical point of view, the calculation of driving forces on the vocal folds requires the knowledge of both the pressure distribution and the velocity pattern within the glottis (Alipour and Titze, 1988).

Experimental investigations of steady flow in physical models of the larynx (Ishizaka and Matsudaira, 1972; Scherer, 1981; Scherer, Titze and Curtis, 1983; Scherer and Titze, 1983; Binh and Gauffin, 1983; Scherer and Guo, 1990b; Scherer and Guo, 1991) have provided empirical pressure-flow relationships for various glottal shapes, angles, and gaps, applicable to computer simulations of speech production. These studies were limited by the one-dimensional flow assumption and did not measure particle velocities.

Theoretical studies using methods of computational fluid dynamics (CFD) for steady or unsteady two-dimensional flow through the static glottis (Gauffin and Liljencrants, 1988; Liljencrants 1991; Iijima, Miki and Nagai, 1992; Alipour and Patel, 1991, 1992; Scherer and Guo, 1990a; Guo and Scherer, 1993) have suggested that glottal flow separates within (or near the exit of) the narrow glottal constriction (also ref. Scherer, 1993; Pelorson, Hirschberg, and Auregan, 1993) and forms a jet, and flow patterns and pressure distributions within the glottis and in the supraglottal region can be estimated.

Study of unsteady laryngeal airflow with models incorporating moving boundaries is at the beginning stages. Experimental work with flow visualizations (Shadle, Barney, and Thomas, 1991) have suggested the formation of supraglottal vortices and an unsteady jet. Also, flow with moving boundaries has been examined using the Phoenics code, a CFD program based on the finite volume method (Ni and Alipour, 1993; Alipour and Fan, 1993). This method resulted in frequency dependent velocity and pressure distributions at various phases of the oscillation cycle.

There has been much experimental work on laryngeal fluid mechanics, but few studies have measured particle

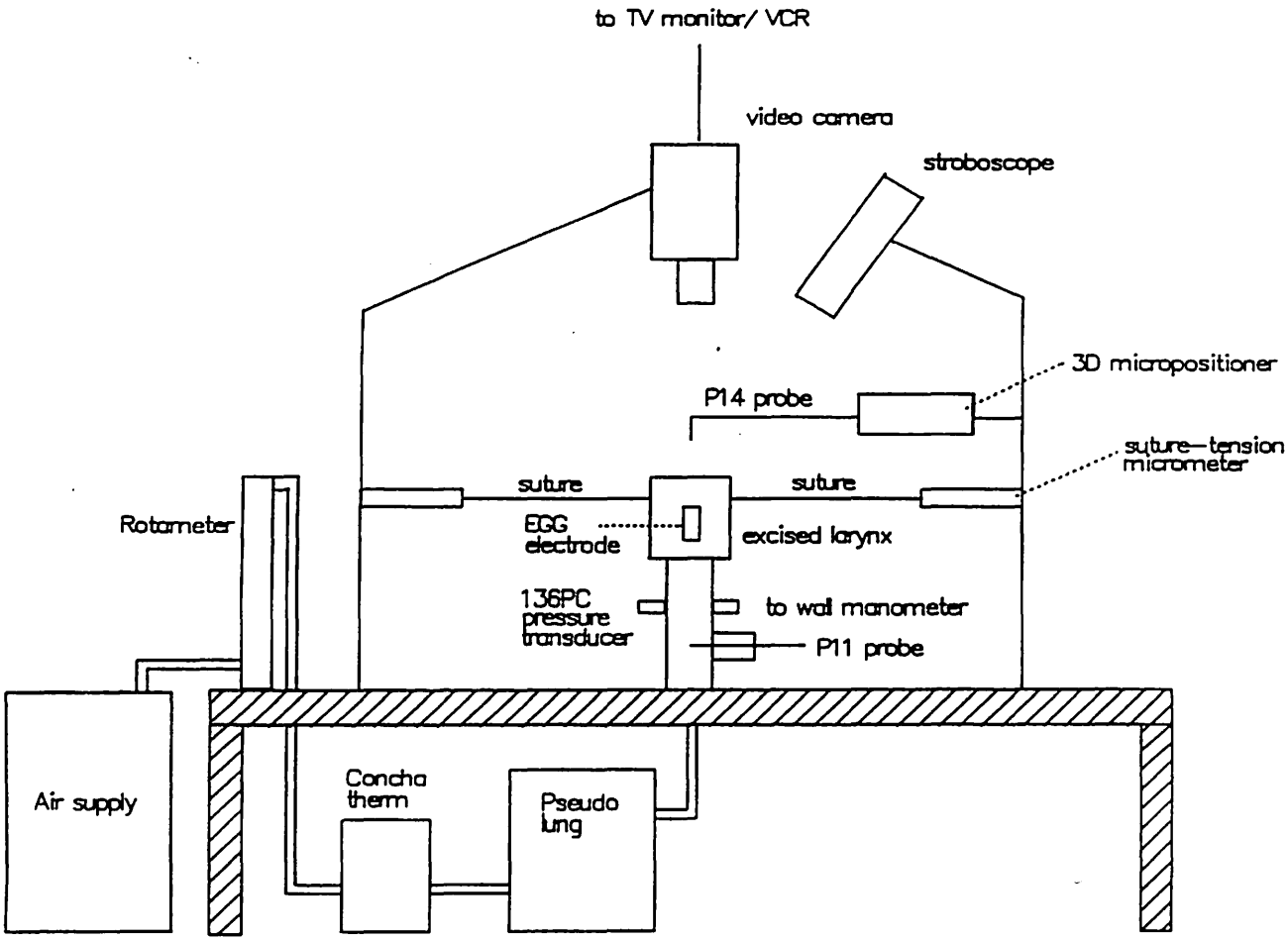


Figure 1 Schematic of the excised larynx experimental setup (not to scale).

velocities. The first attempt to quantify the glottal velocity profile appears to have been made by Berke, Moore, Monkewitz, Hanson, and Gerratt (1989). Using constant-temperature hot-wire anemometry, they measured particle velocity waveforms in an *in vivo* canine model at five locations along the midline of the glottis. Their results questioned the uniformity of the velocity profile across the glottis and hypothesized the existence of a jet flow out of the glottis. Although the data were limited in scope, they indicated the need for the study of laryngeal air velocities, and demonstrated that the canine larynx is a valuable and feasible model for the study of pulsatile flow during phonation (see also Slavit, Lipton and McCaffrey, 1990).

The purpose of this study was to quantify the jet flow at specific locations 1 cm above the glottis, and determine flow patterns at various phases within the phonatory cycle.

Methodology

Excised larynx[1]

Excised canine larynges were acquired from other laboratories at the University of Iowa Hospitals and Clinics. They were kept in saline solution prior to experimentation. The extraneous tissues outside of the larynx were trimmed and the false vocal folds were removed to expose the true vocal folds. After achieving humidified air of 100 % humidity and 37 C using a Concha Therm III Servo Control Heater unit (RCI Laboratories) and air from the building's compressed and filtered air supply, the larynx (with tracheal tissue) was mounted on a pseudo-tracheal rigid tube (i.d. of 17.5 mm) so that the glottis was easily viewed by a camera and accessible by the equipment (Figure 1). Adduction and tension controls were established by connecting cartilages to micrometers with sutures. Sutures attached to the thyroid cartilage anteriorly and the cricoid cartilages posteriorly were used to control vocal fold length. A suture was sewn to each arytenoid cartilage near the vocal processes and passed

[1] These experiments used excised canine larynges. Exact replication of excised larynx experiments is impossible, even using the same larynx at different times. This is due to the wide variability of tissue size and morphology across larynges, and variety of bench setups. Although fine details of phonation may vary, gross dynamic and mechanical aspects should give similar trends across larynges as parameters are changed. In this experiment, the intent is not to determine all expectations of particle velocities relative to phonation, but to offer new information concerning particle velocities relevant to a greater understanding of phonation.

laterally through the other to control adduction (Figure 2). This also stabilized the larynx position. Adduction, as indicated by the EGG signal, was within normal limits expected in human phonation (see Figures 5, 9, and 10). The adduction procedure was a variation of the method used by Durham, Scherer, Druker, and Titze (1987), where they used a pair of three-prong needles to control adduction and vocal fold elongation. In this experiment, supraglottal tubing (vocal tract) was not used. Although tubing may create realistic inertance conditions to influence the flow, the question of vocal tract inertance effects was beyond the scope of this project and not feasible with the set up used.

Larynx #1 was from a female dog weighing 26 kg with an *in situ* vocal fold length of 20.5 mm. Larynx #2 was from a 25 kg male dog with an *in situ* vocal fold length of 15.0 mm. Larynx #3 was from a 22 kg male dog with an *in situ* vocal fold length of 14.0 mm.

A pair of electrodes from research electroglottograph (EGG) made by Synchrovoice Inc. was attached to the thyroid cartilage with push pins to pick up the EGG signal during oscillations. The third (reference) electrode was placed on the posterior surface of the larynx. The EGG signal has been correlated to contact area and can be used to identify the approximate opening and closing phases of the oscillations (Childers, 1985; Scherer, Druker, and Titze, 1988). This signal was also used in the data analysis as a trigger signal to process the instantaneous velocity signal to obtain the phase-averaged velocity waveform.

Pressure and mean flow

The mean pressure in the subglottal region (measured at 10 cm below the vocal folds) was monitored with a wall mounted well-type water manometer (Dwyer No. 1230-8; accuracy of reading was ±1 mm) and the mean flow rate was monitored with an in-line rotameter (Gilmont J197; accuracy of measurement was ±3% relative to visual monitoring in the flow range used in this experiment). The time varying subglottal pressure was measured with a piezoresistive pressure sensor (Microswitch 136PC01G1; accuracy of measurement was ±2.3% within the range used in this experiment). This pressure was measured at the same location as the manometer tap (Figure 1) for ease of calibration. The subglottal pressure signal was amplified with a bridge amplifier (Bioamp model 205) and calibrated against the manometer at regular intervals (Figure 3). The pressure was corrected for its 1 ms delay (relative to the EGG signal) by a pressure transient method in association with simultaneous microphone recordings. Bandwidth for the pressure transducer was estimated to be 0-800 Hz[2]. Because there was no simulated supraglottal vocal tract, the supraglottal pressure was assumed to be nearly atmospheric, and the subglottal pressure was then equivalent to the translaryngeal pressure.

Particle velocity

The instantaneous velocity which is the summation of a periodic velocity and the turbulent component, was measured with a constant-temperature hot-wire anemometer system (Dantec 56C01). A straight miniature probe (Dantec 55P11) was positioned centrally in the subglottal air tube 12

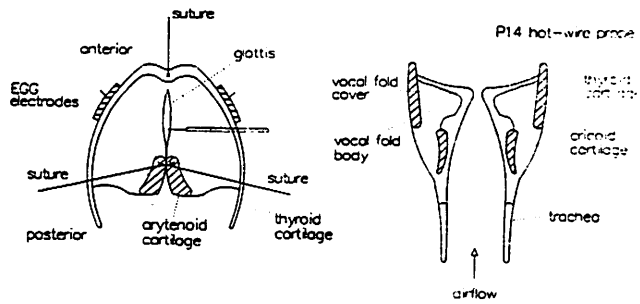


Figure 2 Schematic of top view (left) and coronal cut (right) of excised larynx. The top view shows the glottal slit at maximum.

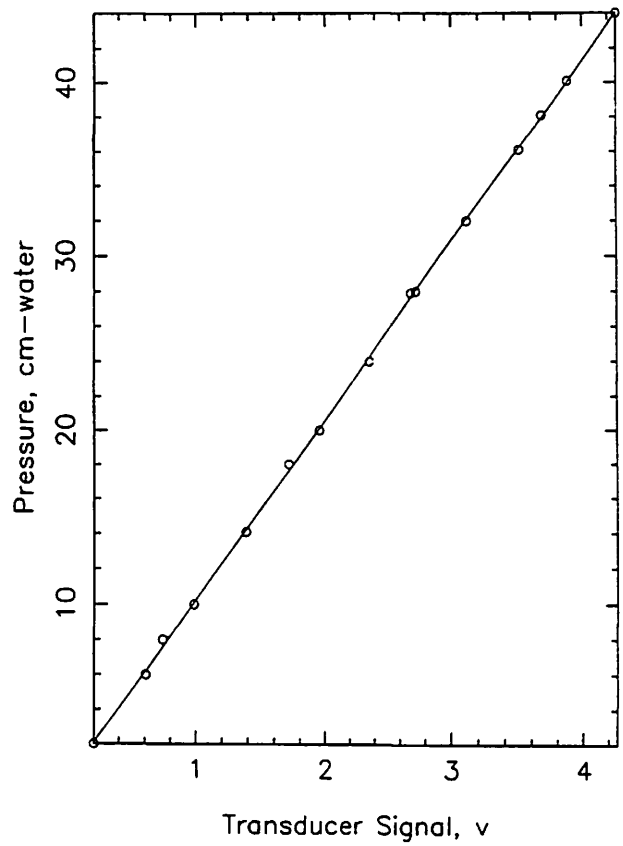


Figure 3 A typical calibration curve of the pressure transducer.

[2] Subglottal resonances (obtained by insertion of a microphone into the subglottal space with the use of dry air and a latex larynx) were measured at 650, 1100, & 1650 Hz. The variations in the subglottal pressure signals of this study near their peaks refer at least in part to subglottal resonance effects, as seen in Figures 7 and 8.

cm below the vocal folds (Figure 1) to measure time varying velocity in the subglottal region. The velocity in the pulsatile jet above the glottis was measured at various locations with a miniature hot-wire probe (Dantec 55P14) with the sensor wire parallel to the glottal slit. The probe was mounted on a three-dimensional micropositioner (Deadal model 393M) to gain access to any location in the glottal jet (Figure 2). Measurements were made to quantify aspects of the jet and its behavior during phonation.

The hot-wire signal was calibrated to velocity against a Pitot tube sequentially replaced over the center of an air jet exiting a 0.6 cm diameter uniform tube[3]. Humidified air at 35-37 degree Celsius was used during both the calibration and the experiments. Since the sensor wire was perpendicular to the flow direction, the velocity calculated from the pitot tube was directly used in the calibration of the hot-wire signal. Figure 4 shows a typical calibration curve of the P14 probe. The best fit for velocity calibration was obtained with the least-square polynomial model. The low velocity portion of the calibration (below approximately 1 m/s) was not as reliable as the rest of the calibration range. The operating range for laryngeal velocities was typically above 1 m/s. Velocity error estimates were calculated to be within ±1.0% of any measurement.

Recording and digitizing

During the experiment, analog data from the hot-wire probes, EGG, and the pressure transducer were monitored on a digital oscilloscope (DATA 6000 from Data Precision of Analogic Corp.) and simultaneously recorded on DC channels of 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. The multiple channels of data were digitized at 20 kHz for one second per channel. During the experiment, the top view of the larynx and vocal folds were recorded on video tape for later analysis of the images during oscillations. For stroboscopic images, a Pioneer DS-303ST stroboscope was used and was manually controlled to match for fundamental frequency.

Phase averaging

To separate the periodic component of the velocity signal from the turbulent noise, a numerical method called 'phase averaging', similar to low-pass filtering, was employed. In this method, a trigger signal (the EGG waveform signal) was used as reference. Every sample of the EGG signal was a reference point. Every velocity point corresponding to (occurring at the same moment in time as) a reference point on the trigger EGG cycle was averaged with its same point on other cycles across many cycles for at least one second. This process generated another signal without the turbulence. Figure 5 shows the phase-averaged velocity, supraglottal instantaneous velocity, and EGG, from top to bottom, respectively. The phase averaging process is limited if jitter in the trigger signal is too great. Waveform examination suggested that this did not appear to be a problem in this study.

Supraglottal hot wire probe positioning

Figure 6 compares the mean and maximum velocity of the glottal jet (in larynx # 1) as the hot-wire probe was positioned over the midline and center of the glottis and moved closer to the vocal folds from above. For this case, the larynx had an Fo of 159 Hz, subglottal pressure of 18.0 cm H_2O, and a bulk flow of 305 cc/s. As the supraglottal position decreased from the glottis from 10 mm to 5.5 mm, the measured particle velocity approximately 30 % increased. A distance of 10 mm (±0.2) between the probe and the vocal folds gave an intermediate velocity measure and was used for the rest of the velocity measurements. It is noted that it would not have been feasible to traverse the probe across the glottal

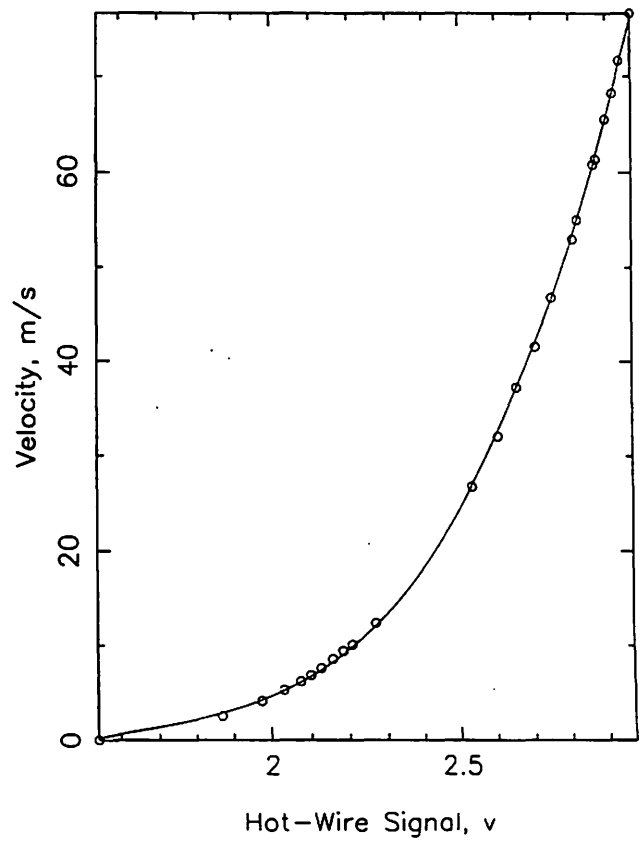


Figure 4 A typical calibration curve of the Dantec hot-wire type P14.

[3] The Pitot tube technique of determining particle velocity is a standard procedure in steady flow, and allows the calculation of the velocity component on line with the axis of the pitot tubing bore (White, 1979).

opening for probe positions closer than 7 mm due to the interference of the thyroid cartilage with the instruments. In addition, the closer the hot wire probe tip came to the glottal exit location, the greater was the likelihood of probe wire damage due to fluid droplets or other particles in the flow. Berke et al. (1989) also used the distance of 10 mm above the glottis, allowing closer data comparison with their study.

Experiment 3 described below involves measuring velocities at numerous locations above the vocal folds at the nominal glottal distance of 10 mm. The estimated distance was 10 mm ± 0.2 mm, resulting in a possible error of ± 20% for velocity measurements in Experiment 3.

Experimental cases

Three experiments will be reported in this paper. Experiment 1 used excised larynx #2 and involved placing the supraglottal hot wire probe 10 mm above the glottis on the midline and at a position of 2 mm (±0.1 mm) from the anterior commissure. The operational glottal length was 14.2 mm, with a fundamental frequency of 131 Hz, subglottal pressure of 10.0 cm H_2O, and a bulk flow of 470 cc/s. The object of that experiment was to determine the supraglottal velocity signal at that location and discuss it in comparison with the other phonatory signals simultaneously obtained.

Experiment 2, using the same larynx under the same phonatory conditions, placed the probe 7 mm (±0.1 mm) from the anterior commissure (near the center of the glottis) in order to compare the changes in the supraglottal velocity pattern from Experiment 1.

Experiment 3 produced supraglottal velocity surfaces in order to quantify the shape of the glottal flow at various phases of the oscillation cycle. By locking the strobe light with the EGG signal and changing the delay, the maximum opening was observed on the video in a stationary (freeze) mode. The area of the opening was divided into a 9x13 point grid with 0.25 mm spacing laterally (the X direction) and 1 mm spacing longitudinally (the Y direction). The hot-wire probe was located at every grid point for at least 5 seconds and moved to the next one using the X- Y micropositioner. It is estimated that the accuracy of positioning the probe over a prescribed location, taking possible minor movement of the glottis into consideration, was ± 0.01 mm in either direction. The data were digitized later and processed on a VAX station computer. For this experiment, the glottis produced a maximum width at midglottis of 1.0 mm, with phonation at 156 Hz, 14.0 cm H_2O subglottal pressure, and bulk flow of 230 cc/s.

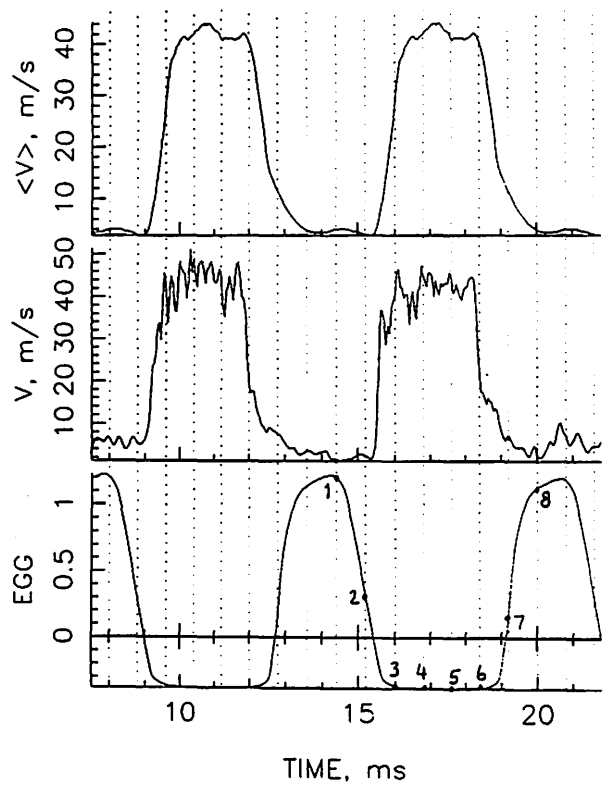


Figure 5 Velocity waveform measured 10 mm above the vocal folds and over the midline 7 mm from anterior commissure. The top trace is the phase average of the middle trace. The markings on the EGG trace refer to the 8 phases discussed in Experiment 3 (see text).

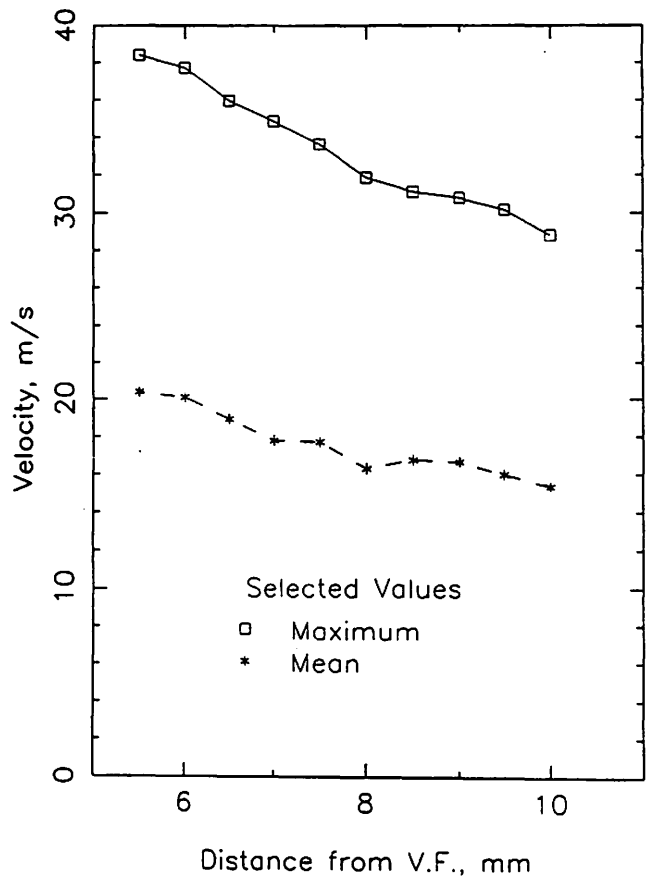


Figure 6 Effect of distance from the glottis on the maximum and mean velocities of the glottal jet measured over the midline at maximum glottal opening location.

The pressures and flows for these experiments were within the ranges of human values. Although the subglottal pressure in Experiment 3 was 14.0 cm H_2O, relatively high for human speech, it resulted in clear canine phonation.

Results and Discussion

Experiment 1

Figure 7 shows the phase-averaged supraglottal jet velocity Vj, velocity in the center of the trachea Vs, subglottal pressure Ps, and EGG, for Experiment 1. The jet velocity signals were phase averaged to reveal their underlying waveforms. The supraglottal jet velocity (top trace of Figure 7) was measured with the hot-wire probe positioned over the midline and 2 mm from the anterior commissure. The width of the velocity pulse at this location was shorter (2 ms at the 15% height of the Vj signal) than the overall glottal open portion (4.1 ms at the 20% height of the EGG signal), and therefore reflects restriction of flow near the anterior commissure due to anterior closure prior to medial closure of the glottis. The mean flow rate for this case was 470 ml/s and the mean subglottal pressure was 10.0 cm water. The subglottal pressure increased during the glottal closing portion of the cycle (interpreted from the EGG signal) and continued to build up until the vocal folds began to open and then decreased to a minimum. The subglottal velocity fell to a non-negligible minimum, indicating that the vocal folds did not close completely, or there was subglottal tracheal expansion at the time of glottal closure and maximum subglottal pressure.

Experiment 2

Figure 8 shows similar data for the different hot-wire position of Experiment 2. The supraglottal jet velocity was measured in the same larynx over the midline at about 7 mm from the anterior commissure (near the center of the glottal length) with a flow rate of 750 ml/s. This figure indicates two peaks of velocity within a vibratory cycle for the supraglottal velocity measure Vj. The first peak is shown within the location of glottal opening as suggested by the EGG signal. The two velocity peaks occur during the time when the subglottal velocity is increasing or near maximum, and the drop in the velocity between the two peaks may be related to the phase difference of the inferior and superior sections of the vocal folds during the oscillation cycle, and the decrease in the translaryngeal pressure.

The double velocity peaking phenomenon is not necessarily present over the center of the glottis, however. Figure 5, taken from data from Experiment 3 (below), shows single velocity pulses without double peaks. Also, the velocity pulses are relatively broad, apparently due to the

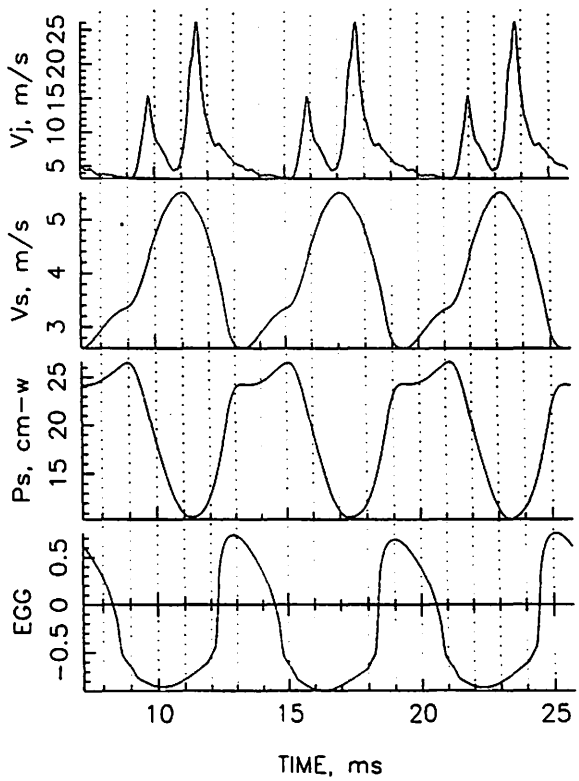


Figure 7 The waveforms of a typical result. The signals are: electroglottograph EGG, subglottal pressure Ps, velocity in trachea Vs, and supraglottal jet velocity Vj, measured 2 mm from the anterior commissure.

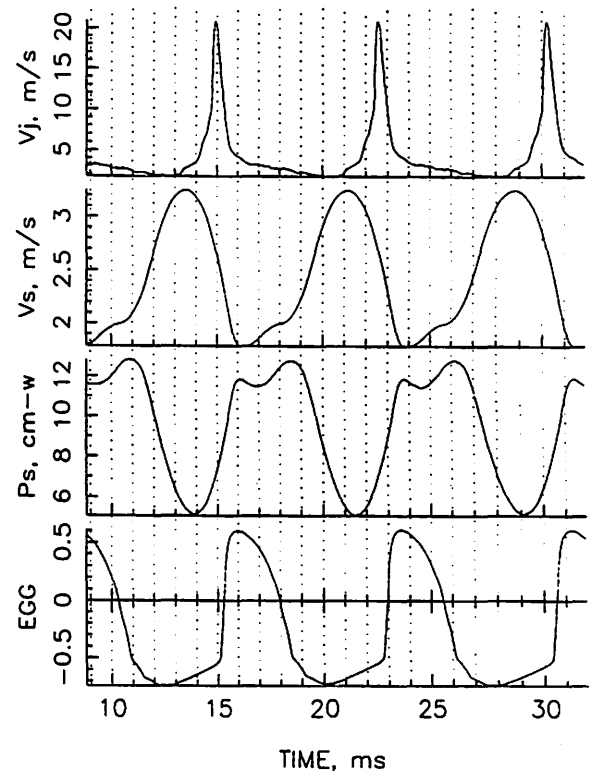


Figure 8 Results similar to Figure 7 with jet velocity measured in middle of the glottal opening.

time the glottis is open at the center of the glottis, compared with the shortened open interval at the anterior section of the glottis as shown in Figure 7.

These interpretations, although apparently reasonable, need to be viewed critically, however. EGG is an integral scaler variable related to glottal kinematics, and the velocities and pressures are local field variables, suggesting complicated interrelationships of these recordings. Additionally, the 10 mm distance between the glottis and the downstream velocity probe creates difficulties in making assumptions about time delays and particle trajectories (concerns which may be heightened by the inclusion of inertance effects in the extended case of a vocal tract above the glottis). It is reasonable, however, to hypothesize that particle velocities are in general related to spatially dependent flow impedances within the glottis, and using velocity probes directly at the exit of the glottis would resolve some of these concerns. For Experiment 3, the glottis did not appear to close all the way, and there was constant opening between the vocal processes. The amplitude of vibration of the glottis at the anterior-posterior center was 1.0 mm.

Although the study by Berke et al. (1989) used *in vivo* canines with phonatory conditions of glottal closed times and subglottal pressures that were about twice as large as in the present study, their data are similar in a number of ways. For the same glottis-probe distance of 10 mm, they also reported double peaked velocity waveforms above the glottal midline. The onsets, offsets, peaks and valleys of the velocity waveforms are essentially in the same relationship to the EGG and subglottal pressure signals as in the current study. "The velocity wave form had a notched peak in most samples. The velocity peaked initially and then dropped acutely. As the cords began to close ... the velocity again rose to a second peak [the second peak in both studies appear just as the EGG begins to move toward glottal closure as shown in Figure 8]. ... [The] notch generally coincided with the ... maximal glottal area. [This should correspond, as it does in Figure 8, to the near minimum of the EGG signal.] Following the second peak, the velocity tapered off rapidly but less abruptly than it rose on opening. The ... moment of lower margin closure, coincided with the shoulder in the diminishing velocity wave form ... The rise of velocity is closely aligned with the fall in [subglottal pressure] at vocal cord opening. At maximal opening ... [the subglottal pressure] was close to its nadir, which corresponds to the notch in the velocity wave form." (Berke et al., 1989, pp 310-311). These comments also hold for the results of the present study.

Experiment 3

Figure 9 shows the maximum velocity surface in the cross section of the glottal jet for Experiment 3. The maximum velocities were determined from the phase averaged velocity. The figure demonstrates the non-uniformity of this glottal jet. In this case, the maximum glottal opening was 1.25 mm, and the measurement grid included a 0.25 mm extension on each side. There were variations in the velocity both in the X and Y directions with the highest velocity above a location near the anterior commissure. Maximum velocity occurred for the location of 3 mm from the anterior commissure, whereas maximum glottal opening (as measured from the video tape recording) was 5 mm from the anterior commissure. These findings of maximum particle velocity measurements above the anterior glottis, anterior to maximum vocal fold displacement, were also reported by Berke et al. (1989). The greatest particle velocity was approximately 40 m/s. Average bulk flow was 230 ml/s and subglottal pressure was 14 cm water. Figure 10 shows the mean velocity surface calculated from the phase averaged velocities (time averaged) which appear smoother. It has the same trend as the maximum velocity but with smaller values.

Figures 9 and 10 suggest that, for the larynx used, the maximum velocity and the largest average velocity occur relative to the anterior portion of the larynx. This may be due to spatial variations in glottal impedance, with the least impedance anteriorly. The shape of the velocity surface

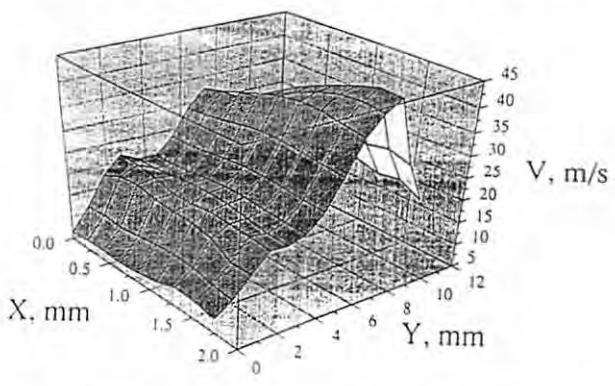


Figure 9 The maximum velocity surface in the glottal cross-section at a mean flow rate of 230 ml/s and subglottal pressure of 14 cm H_2O, larynx #2. The x axis refers to the lateral direction, and the y axis refers to the anterior posterior direction.

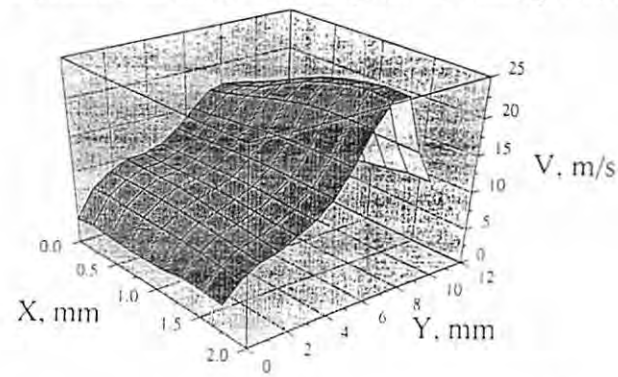


Figure 10 The time-averaged (mean) velocity in the glottal cross-section the same condition as Figure 9.

shown in Figure 10 suggests that the mean velocities at some y positions (anterior-posterior) appear relatively uniform in the transverse (lateral) direction. The measured velocities include flow through the glottis as well as any displacement flow due to vertical tissue oscillations. The latter would create non-zero flow superior and lateral to the glottal opening. It is pointed out that the single sensor hot-wire probe is not sensitive to flow direction and this creates uncertainty in regions of possible flow reversal and displacement flow.

To better understand the pattern of the glottal flow within a glottal cycle, one cycle of the phase-averaged velocity for every grid point was divided into eight segments and eight sets of velocity data were extracted from these data sets. Figure 5 shows the eight locations relative to the EGG signal. The starting point (#1) for this process was chosen at the peak of the EGG waveform which corresponded to maximally closed vocal folds. Figure 11 shows the eight simulated snap shots of particle velocity at every 1/8th of the period. Figure 11 indicates an increase of flow over the anterior commissure with higher values to the left for the first half of the cycle, switching toward the right during the second half of the cycle. The maximum flow appears to be attained at 5/8ths of the period (90 degree before glottal closure), corresponding to a skewing of the volume velocity as seen in normal human phonation.

Summary and Conclusions

This study examined pulsatile flows from the glottis during phonation. Air particle velocities from excised canine larynges were measured using hot-wire anemometry, and were accompanied by measures of subglottal pressure, subglottal air particle velocity, and the electroglottographic signal. Phase averaging of the air velocity signals were performed to remove turbulence and show primary flow patterns. Recordings of velocity were made over the glottal midline as well as over a 9x13 measurement grid 10 mm

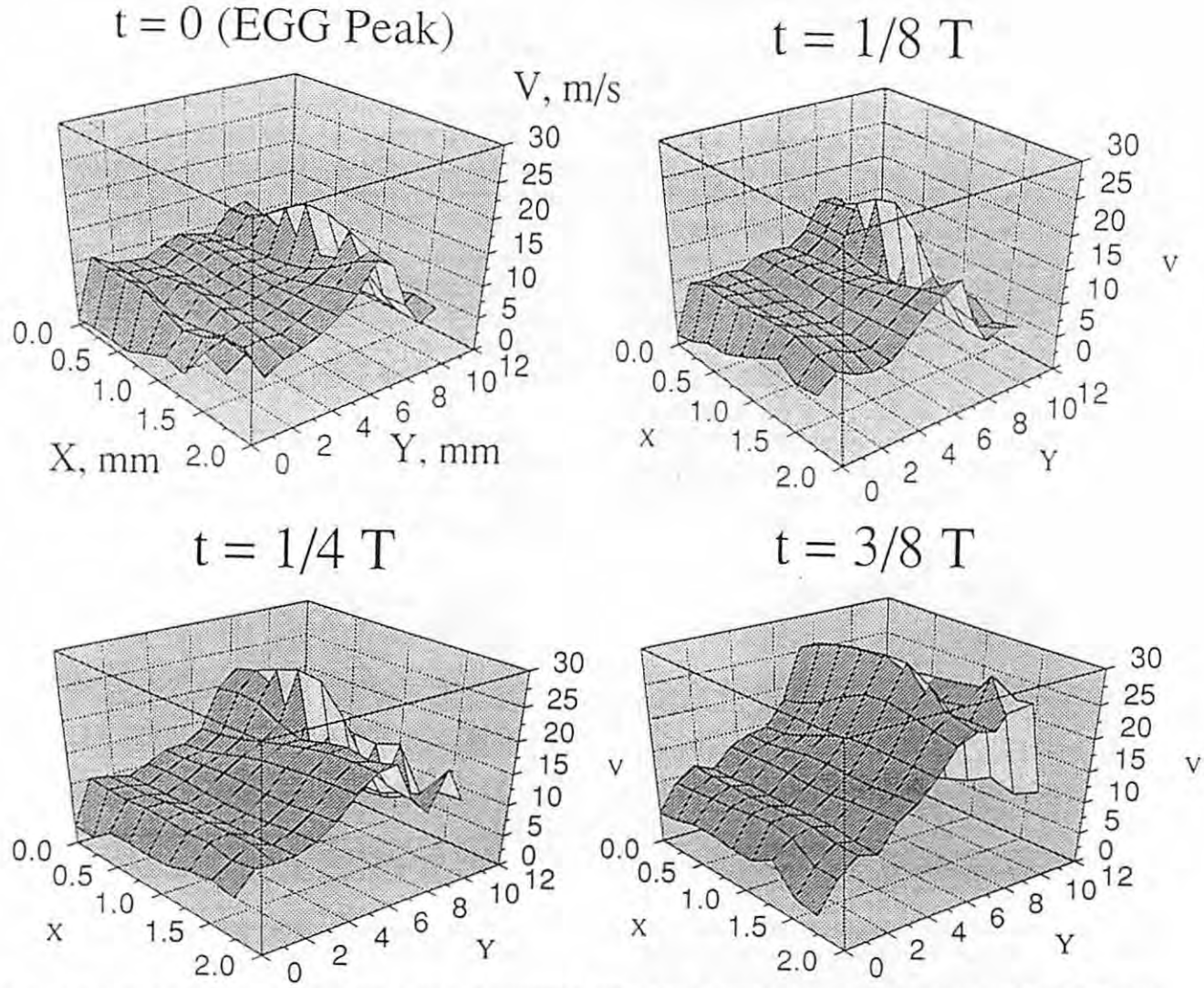


Figure 11 (above and opposite). Simulated snap shots of flow patterns at every 1/8th of a period with the same oscillation conditions as Figure 9.

above the glottis. Results of this study suggest the following:

1. The phase-averaged air particle velocity signal above the glottis may have single or double peaks within their cyclic patterns.

2. Phase-averaged particle velocity patterns above the glottal midline appear to conform to explanations strongly related to glottal impedance changes, inferred from the electroglottographic and subglottal pressure signals (and supported by work by Berke et al., 1989).

3. Phase-averaged maximum and average velocity surfaces above the glottis are highly non-uniform. The velocity surface patterns can significantly alter shape throughout a phonatory cycle.

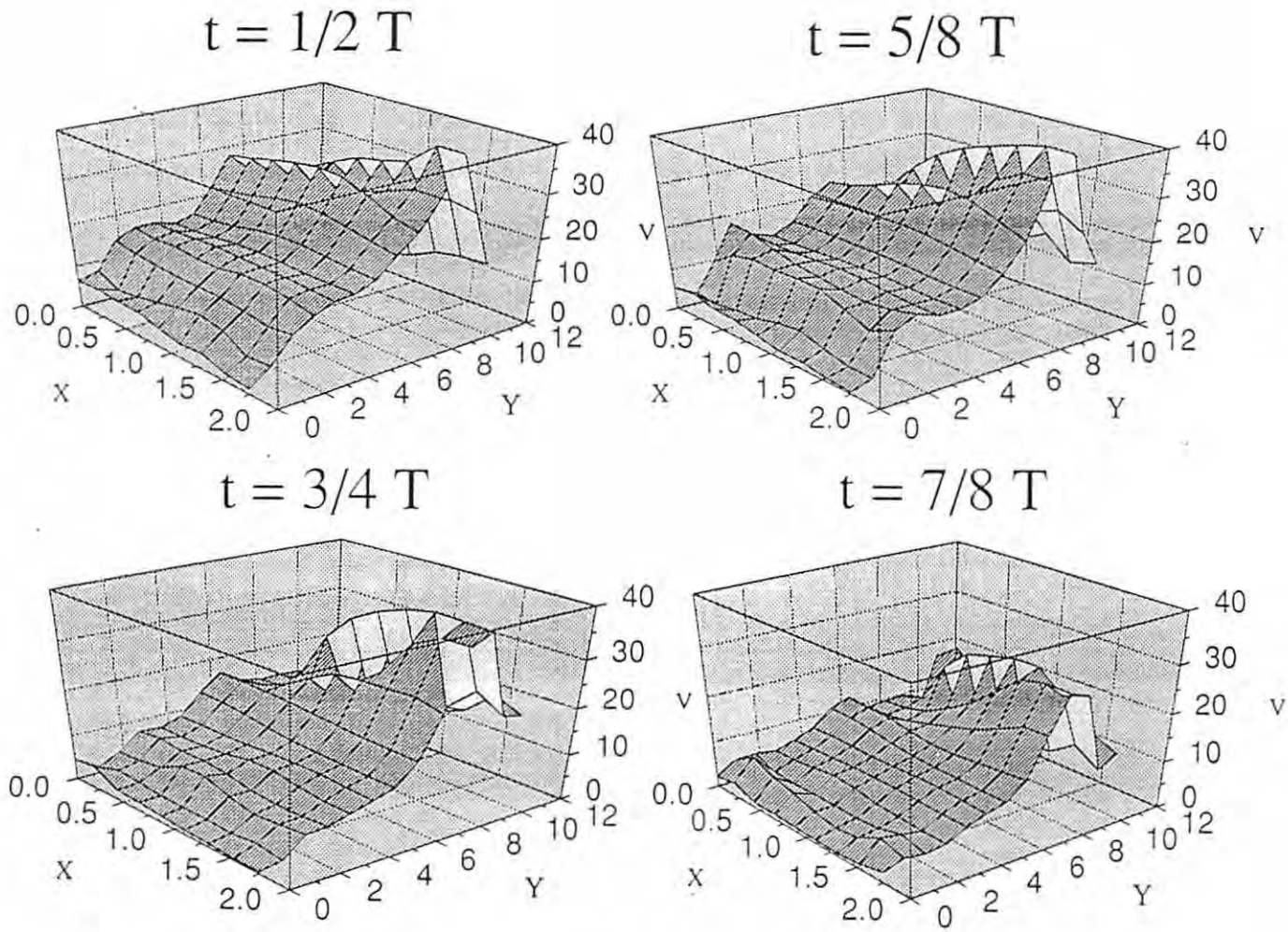
4. Greatest velocity values occurred above the anterior glottal region, apparently anterior to the location of cyclic maximum glottal excursion.

5. Maximum flow occurred at a phase location of 90 degrees before glottal closure, corresponding to skewing of the flow pulse.

These results suggest that particle velocities above the glottis are far from instantaneously uniform. Spatial distributions of these velocities need to be considered in any theory of phonatory aeroacoustics.

Acknowledgements

The authors would like to thank Jim Knowles, Daryl Lorell and Katie Sackett for their assistance in larynx preparation and data collection. They would also like to thank Elaine Stathopoulos and two anonymous reviewers for very helpful editorial reviews. This work was supported by the National Institute on Deafness and other Communication Disorders, Grant No. DC00831-02.



References

Alipour, F. and Titze, I.R. (1988). "A Finite element simulation of vocal fold vibration," in the Proc. Fourteenth Annual Northeast Bioeng. Conf. Durham, N.H., IEEE publication #88-CH2666-6, pp. 186-189.

Alipour, F. and Patel, V.C. (1991). "Numerical simulation of laryngeal flow," in the Proc. ASME Winter Annual Meeting, Edited by R. Vanderby, BED-Vol. 20, pp. 111-114.

Alipour, F. and Patel, V.C. (1992). "Steady flow through modeled glottal constriction," in the Proc. Int. Conf. on Engineering Application of Mechanics, Edited by Ali Meghdari, Vol. 1, pp. 93-100.

Alipour, F. and Fan, C. (1993). "Pulsatile Flow in a Three- Dimensional Model of Larynx," in the Proceedings of 1993 ASB Meeting, Edited by A.A. Biewener and V.K. Goel, pp. 221-222.

Berg, Jw. van den, Zantema, J.T., and Doornenbal, P. Jr. (1957). "On the air and the Bernoulli effect of the human larynx," J. Acoust. Soc. Am., 29(5), pp. 626-631.

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

Berke, G.S., Moore, D.M., Monkewitz, P.A., Hanson, D.G., and Gerratt, B.R. (1989). "A preliminary study of particle velocity during phonation in an in vivo canine model," Journal of Voice, Vol. 3, No. 4, pp. 306-313.

Childers, D.G. (1985). "A critical review of electroglottography" CRC Critical Reviews in Biomedical engineering, Vol. 12, issue 2, pp. 131-161.

Durham, P.L., Scherer, R.C., Druker, D.G., and Titze, I.R. (1987). Development of excised larynx procedures for studying the mechanisms of phonation. Technical Report VABL-1-1987, Voice Acoustics & Biomechanics Laboratory, Department of Speech Pathology/ Audiology, University of Iowa, Iowa City.

Fant, G. (1970). "Voice source dynamics," Speech Transmission Laboratory QPSR, No. 2-3, pp. 1737.

Fant, G. (1986). "Glottal flow: models and interaction," J. Phonetics Vol. 14, pp. 393-399.

Gauffin, J. and Liljencrants, J. (1988). "The role of convective acceleration in glottal aerodynamics," in O. Fujimura (Ed.) Vocal Physiology: Voice Production, Mechanisms, and Functions, (pp. 219-226), Raven Press, Ltd., New York.

Gauffin, J. and Sundberg, J. (1989). "Spectral correlates of glottal voice source waveform characteristics," J. Speech Hear Res. Vol. 32, pp. 556-565.

Guo, C.G. and Scherer, R.C. (1993). "Finite element simulation of glottal flow and pressure," J. acoust. Soc. Am. Vol. 94, No. 2, Pt. 1, pp. 688-700.

Iijima, H., Miki, N., and Nagai, N. (1992). "Glottal impedance based on a finite element analysis of two-dimensional unsteady viscous flow in a static glottis," IEEE Trans. on Signal Processing, Vol. 40, No. 9, pp. 2125-2135.

Ishizaka, K. and Flanagan, J. L. (1972). Synthesis of voiced sounds from a two-mass model of the vocal cords. Bell Syst. Tech. J., 51(6), pp. 1233-1268.

Ishizaka, K. and Matsudaira, M. (1972). Fluid mechanical considerations of vocal cord vibration. SCRL-Monograph8, Speech Communication Research laboratory, Santa Barbara.

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

Ni, J. and Alipour, F. (1993). "Animation and Numerical Simulation of Laryngeal Airflow," in Z. Tang (Ed.) New Advances in Computer Aided Design & Computer Graphics, (pp. 849-854), International Academic Publishers, Beijing, China.

Pelorson, X., Hirschberg, A., and Auregan, Y. (1993). "Fluid dynamics aspect of speech: 1) A modified two-mass model," Stockholm Music Acoustics Conference.

Scherer, R.C. (1981). Laryngeal Fluid Mechanics: Steady Flow Considerations Using Static Models. Ph.D. thesis, University of Iowa, Iowa City.

Scherer, R.C. (1983). "Pressure-flow relationships in a laryngeal airway model having a diverging glottal duct," J. Acoust. Soc. Am., 73(S1), S46(A).

Scherer, R.C. and Titze, I.R. (1983). "Pressure-flow relationships in a model of the laryngeal airway with a diverging glottis," in D.M. Bless and J.H. Abbs (Eds.) Vocal Fold Physiology: Contemporary Research and Clinical Issues, (pp. 179-193) College-Hill Press, San Diego.

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

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 Functions, (pp. 279-291), Raven Press, Ltd., New York.

Scherer, R.C. and Guo, C.G. (1990a). "Effect of vocal fold radii on pressure distributions in the glottis," J. Acoust. Soc. Am. Suppl. 1, 88, S150.

Scherer, R.C. and Guo, C.G. (1990b). "Laryngeal modeling: translaryngeal pressure for a model with many glottal shapes", ICSLP Proceedings, 1990 International Conference on Spoken Language Processing, Volume 1, The Acoustical Society of Japan, Japan, 3.1.1 - 3.1.4.

Scherer, R.C. and Guo, C.G. (1991). "Generalized translaryngeal pressure coefficient for a wide range of laryngeal configurations", in J. Gauffin and B. Hammarberg (eds.), Vocal Fold Physiology: Acoustic, Perceptual, and Physiological Aspects of Voice Mechanisms, (pp. 83-90), Singular Publishing Group, San Diego, California.

Shadle, C.H., Barney, A.M., and Thomas, D.W. (1991). "An investigation into the acoustic and aerodynamics of the larynx," in J. Gauffin and B. Hammarberg (Eds.) Vocal Fold Physiology, (pp. 73-80), Singular Publishing Group, San Diego, California.

Slavit, D.H., Lipton, R.J., and McCaffrey, T.V. (1990). "Glottographic analysis of phonation in the excised canine larynx," Annals of Otology, Rhinology & Laryngology Vol. 99, No. 5, pp. 396-402.

White, F.M. (1979). Fluid Mechanics, McGraw-Hill, Inc., New York.

NCVS Status and Progress Report - 7
December 1994, 131-137

Stress-Strain Response of the Human Vocal Ligament

Young Min, M.D.
Department of Otolaryngology-Head and Neck Surgery, The University of Iowa
Ingo Titze, Ph.D.
Fariborz Alipour, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa

Abstract

The longitudinal elastic properties of the human vocal ligament were quantified by stress-strain measurements and by modeling the response mathematically. Human ligaments were obtained from surgery and autopsy cases. They were dissected, mounted, and stretched with a dual-servo ergometer to measure force versus elongation and to convert the results into stress and strain. To calculate a longitudinal Young's modulus, the stress-strain curves were fitted with polynomial and exponential functions and differentiated. Young's modulus was separately defined in the low and high strain regions. The mean Young's modulus for the low strain region was 33.1 kPa± 10.4 kPa. In the high strain region, A and B parameters for an exponential fit in the high strain portion were 1.4 kPa ± 1.0 and 9.6 ± 1.2, respectively. The stress-strain and Young's modulus curves showed the typical hysteresis and nonlinearity seen previously in other vocal fold tissues (muscle and mucosa), but the nonlinearity was most profound for the vocal ligament.

Introduction

The development of a biomechanical simulation model of phonation is a critical step toward understanding normal and abnormal vocal fold (VF) physiology. Quantification of the mechanical properties of the tissue layers that comprise the VF is needed to construct an accurate model. One such property is the longitudinal elasticity of the vocal ligament, which is assumed to be the stress-carrying member for high pitches (Titze, 1994)[1].

Several investigations have been directed toward measuring the elasticity of the VF. Methodologies have varied among studies in terms of direction of application of force, type of VF tissue sample, species of donor (mainly dog and human), and data analysis. Direction of force has been either longitudinally[2,3,4,5] or transversely[6,7]. When Hirano[6] noted that the VF consists of different tissue layers and suggested that elasticity varies from one layer to another, some investigators dissected the layers of interest - the mucosa, the muscle, and the ligament. The mucosa consists of the epithelium and the superficial layer of the lamina propria, which makes up the entire vocal fold *cover* in the dog larynx. The ligament, which is prominent in the human vocal fold but not the dog, consists of intermediate and deep layers of lamina propria. This ligament is composed of dense collagenous and elastic fibers. The densest fibers are in the deep layer of the lamina propria, which is considered part of the vocal fold *body* (the rest being thyroarytenoid muscle). Elasticity has then been quantified in terms of stiffness ratios, force-elongation curves, stress-strain curves, and Young's modulus.

The results of these studies were as follows: Van den Berg and Tan[2] measured the elongation of VFs as a function of longitudinal force. These authors stated "the vocal ligaments were carefully dissected and the extensibilities ...were measured," although the method of vocal ligament dissection was not discussed. The findings of the study were as follows: 1) there was a difference in elasticity of the right and left VF; and 2) the elasticity of the entire VF and the ligament were similar, particularly during increased elongation. Based on this second finding, it was suggested that the collagen fibers of the vocal ligament absorbed the longitudinal force.

Bauer and König[3] investigated the elasticity of the human VF in response to longitudinal force and reported the mean longitudinal elastic moduli of the male and female VFs as 61.9 kPa and 132 kPa, respectively. Perlman et al.[8,9] measured the longitudinal elastic moduli of the canine VF. Young's moduli ranged from 80-120 kPa for the

thyroarytenoid muscle (low to high strain) and 200-400 kPa for a 1 mm thick mucosa (low to high strain). The mucosa included the epithelium, which is rather stiff. Major gender differences were not observed in the dog.

In the Perlman et al. study, the tissue was elongated in a stepwise manner, resulting in stress relaxations during measurement. This methodology was subsequently improved upon by Alipour-Haghighi and Titze[5] who obtained continuous stress-strain data. This technique provided substantially more data points, such that it was possible to observe the nonlinear elastic property and hysteresis of the cover and the body for the canine VF. Their results will be systematically compared with our results in a later section.

Hirano[6] measured the degree of transverse elongation of the mucosa, ligament, and body of one human VF. This study found the VF mucosa to be most compliant layer, then the ligament, and finally the muscle (which was the stiffest layer). Tran et al.[7] also examined the transverse elastic property of the human VF. For unstimulated VF, the mean transverse Young's modulus was 12.7 kPa. We believe that the transverse and longitudinal elastic properties are related, but the precise relationship has yet to be established.

The purpose of our investigation was to model the longitudinal elastic property of the human vocal ligament. The tangent Young's modulus has been derived because it quantifies the small-amplitude elastic property of tissue in vibration.[10]

Method

Tissue Profile

Four hemilarynges from 3 surgical patients (2 males, 1 female, mean age 65 years) and 4 hemilarynges from 2 autopsy cases (1 male, 1 female, mean age 38 years) were obtained. The surgical larynges were those resected during total laryngectomy for supraglottic cancer or intractable dysphagia with at least one vocal fold being free of pathology. The absence of gross pathology on all vocal folds included in the study was confirmed under a dissecting microscope.

The donors' age, gender, vocal fold identification (right or left), and history of vocal fold pathology or surgery were noted (Table 1). For the surgical specimen, the time interval from laryngeal resection to force-elongation measurement were noted; for the autopsy cases, the time interval from death to force-elongation measurement was also noted. For one of the surgical donors and both autopsy donors, right and left vocal folds were available for comparison. The time interval from surgical resection of larynx to the measurement of force-elongation ranged from 4-22 hrs. The time interval from death of the autopsy donor to the measurement of data ranged from 13-94 hrs.

To maintain fluid balance and temperature similar to the *in vivo* state, the surgical hemilarynges were submerged in Krebs-Ringer solution at 37± 1°C with a pH of 7.4. One exception was a hemilarynx that was refrigerated at 4°C in the Krebs-Ringer solution overnight.

Two of the autopsy hemilarynges were frozen and thawed prior to obtaining the force-elongation data to determine how the freezing process would affect the data.[1] These hemilarynges were wrapped in Krebs-Ringer-soaked gauze, enclosed in a plastic bag, and quick-froze in liquid nitrogen. They were stored in a freezer (-4°C) for 10-80 hours and thawed in a refrigerator (+4°C) overnight prior to the dissection of the ligaments.

Table 1.
Surgery and Autopsy Specimen Profile

Specimen #	Age	Gender	Vocal Fold	Time of resection to measurement, hr	Time ofdeath to measurement, hr
S 1	64	M	R	22	-
S 2	61	M	L	4	-
S 3	67	F	R	5	-
S 4	67	F	L	5	-
A 1	30	M	L	-	20
A 2*	30	M	R	-	94
A 3	46	F	L	-	13
A 4*	46	F	R	-	83

S 1-4: surgical specimens, A 1-4: autopsy specimens
* Specimen was frozen and thawed
R-right, L-left

Vocal Ligament Dissection

The larynges were bisected mid-sagittally through the thyroid cartilage at the anterior commissure. The reference length of the vocal ligament was defined as the *in situ* length since a previous study found this reference length to be the most consistent reference measure for stress-strain.[8] Using a caliper (± 0.1 mm accuracy), the reference length of the ligament was measured from the insertion of vocal fold at the thyroid cartilage to the tip of vocal process.

Under a dissecting microscope, the vocal ligament was isolated by resecting the epithelium, superficial lamina propria, and muscular layers while the tissue remained moist with Krebs-Ringer solution. The lateral border of the ligament was dissected up to the lateral edge of the vocal process of the arytenoid cartilage, where the whitish fibers of the lamina propria were visualized. The medio-inferior aspect of the vocal ligament was dissected down to the inferior border of the tip of the vocal process of the arytenoid

[1] Since the number of normal larynges from surgical cases were small, the autopsy cases were also obtained for study to consider frozen larynges for additional specimen.

cartilage to ensure that it was above the elastic conus. Using ophthalmic scissors, the epithelial and superficial layers of the lamina propria (i.e. the mucosa) were carefully resected by dissecting at the plane of the superficial layer, which was loosely attached to the underlying intermediate layer.[6] However, as noted by Hirano[6], the border between the deep layer of lamina propria and muscular layer was not clear, since the collagenous fibers of the vocal ligament insert into the muscular layer. Therefore, the muscle fibers were dissected out as much as possible under the microscope. To verify that the intermediate and deep layers of lamina propria were isolated, the vocal ligaments were sectioned coronally and stained histologically with Weigert stain after the force-elongation measures were obtained (see center plate - photo 15). It was evident that the epithelium and superficial layers were successfully resected out. Although all ligaments had a few muscle fibers attached to the edges of the tissue sample, the stray muscle fibers constituted only about 1-5% of the total cross-sectional area of the specimen.

Following the protocol described by Perlman and Titze[8], the resected vocal ligament remained attached to a small portion of the thyroid cartilage at the anterior commissure and a portion of the arytenoid cartilage of the vocal process. A 2-0 Tevdek polyester suture was attached to each cartilages and the vocal ligament was vertically mounted to an ergometer in a water-jacketed chamber containing the Krebs-Ringer solution maintained at 37°C (Figure 1). The ligament was mounted close (within ± 1 mm) to the previously measured reference length without any slack in the suture.

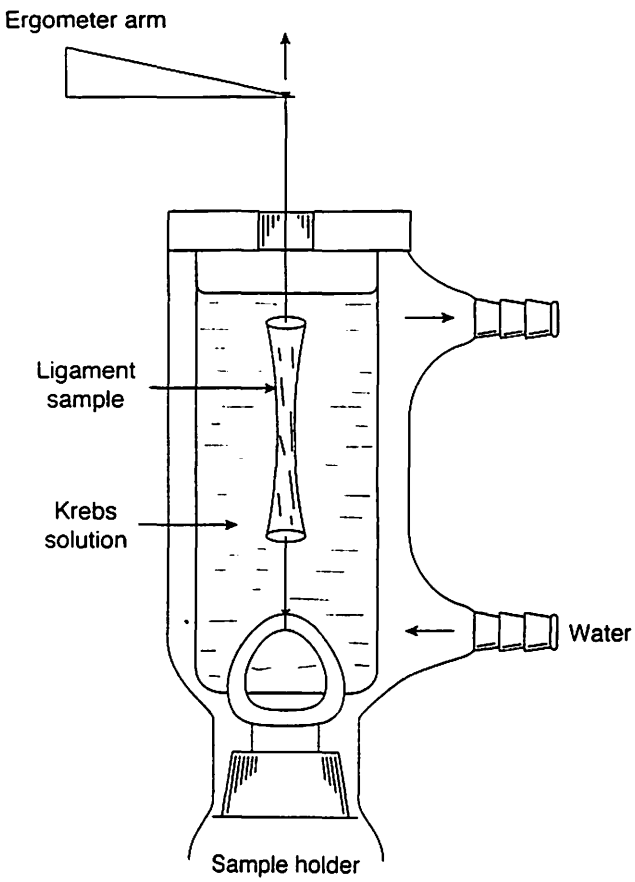


Figure 1. Diagram of tissue mounted to an ergometer. Heated water is circulated around the jacket to maintain temperature at 37°C.

Data Acquisition

Following the method described by Alipour-Haghighi and Titze[5], the force-elongation data was obtained by 1 Hz sinusoidal stretch-and-release of the ligament, using a dual-servo ergometer (Cambridge Technology Model 305). The accuracy of the force was ± 50 mg and the displacement accuracy was ±.02 mm. A sinusoidal displacement was applied to the ergometer arm so that the ligament was stretched and released at a frequency of 1 Hz for the duration of 5 seconds to obtain 5 cycles of force-elongation curve. The force and elongation signals were then transmitted via a 16-bit A/D converter at 1000 Hz per channel to a VA station computer workstation.

Since the cross-sectional area measurement of a soft, thin tissue sample is difficult, an average cross-sectional area was calculated from the following equation: Area = (mass/density) ÷ reference length.[11,12] After the force-elongation data were obtained, the cartilages attached to the ligament were dissected. Mass and density were measured using a Mettler balance (model AE 100, ± 0.1 mg accuracy) and a density meter for the Mettler balance (model ME-40290).

Data Analysis

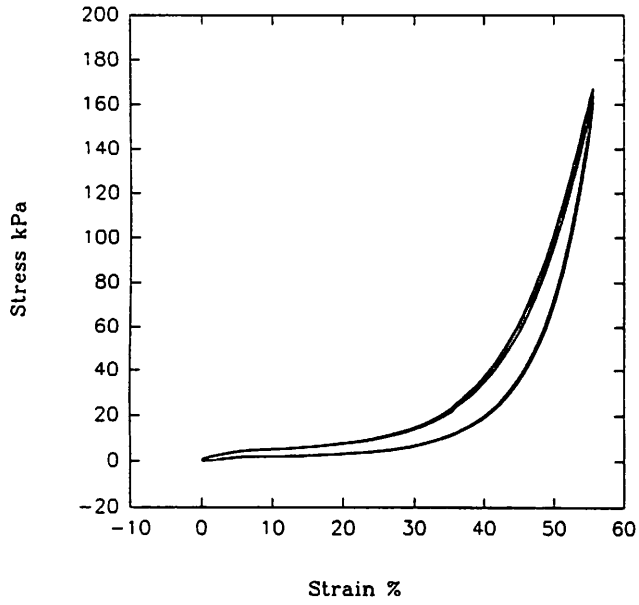
Stress was calculated from the equation, Stress = Force/Area, and strain in % was calculated from the equation, Strain = 100 x (final length - reference length)/reference length.

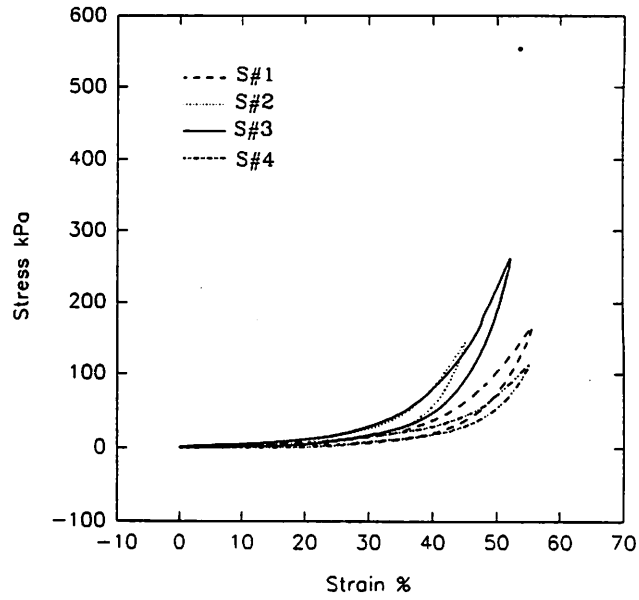
Results

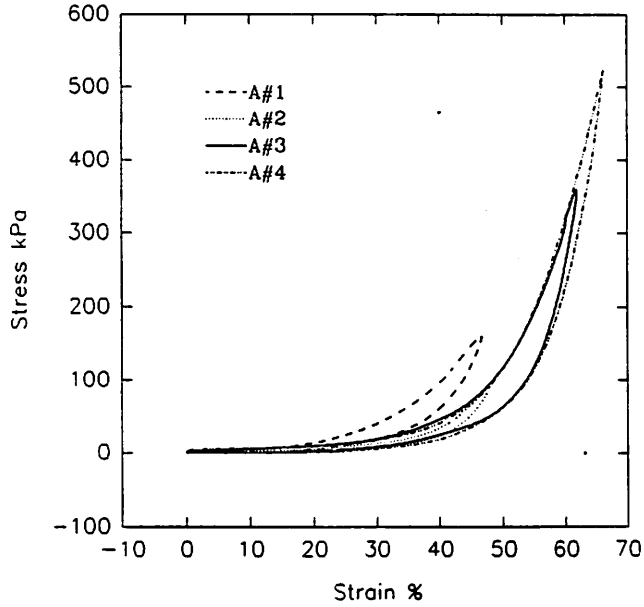
Stress-strain Data

The results reported are based on the first complete cycle of the 5 repetitions of the force-elongation curve. Since all 5 cycles were consistently similar, the first cycle was chosen as a representative result for that tissue (Fig. 2).

The stress-strain curves of all tissue samples showed the typical nonlinearity of collagen (Figure 3 and 4). Although the general shape of the stress-strain curve appeared exponential, the curve could be divided into two portions, low strain (<15%) and high strain (<15%). Since the low strain portion of the stress-strain curve was nearly linear, it could be approximated by a single Young's modulus (stress/strain ratio). The mean and standard deviation for the low strain Young's modulus of the surgical ligaments was 36.1 kPa ± 10.6 kPa (N=4) and of the autopsy ligaments was 30.1 kPa ± 10.7 kPa (N=4) (Table 2). The mean of combined Young's moduli was 33.1 kPa ± 10.4 kPa (N=8).







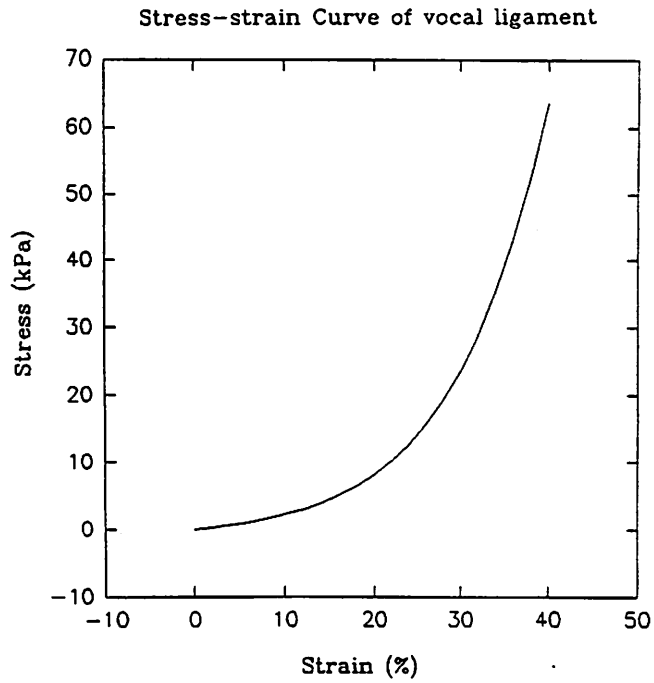


Figure 2 (top left). Successive 5 cycles of stress-strain curves obtained at 1 Hz for the surgical ligament #1. Figure 3 (top right). Stress-strain curves of 4 surgically obtained vocal ligaments. Figure 4 (bottom left). Stress-strain curves of 4 vocal ligaments from autopsy cases. Ligaments A#2 and A#4 were measured after freezing and thawing process. Figure 5 (bottom right). Stretch portion of the averaged human ligament stress-strain curves plotted from the exponential equation: Stress=A($e^{B\,strain}$ - 1). A and B parameters of each ligament were calculated and the means of A and B used in the plot.

At high strain, the stress-strain curve could best be approximated by an exponential function. We applied the equation, stress = A($e^{B\,strain}$ - 1),[13] as previously described by Alipour & Titze[5] to the entire stretch portion of the data. The A and B parameters were determined for individual specimen (Table 2).[2] The means and the standard deviations for parameters A and B for the combined surgical and autopsy ligaments were 1.4 ± 1.0 and 9.6 ± 1.2, respectively. A plot of the modeled strain-strain curve is given in Fig. 5.

[2] Although the overall curve seem best described by an exponential function, this model was not sufficient. While the high strain portion of the stress-strain data was best characterized by the exponential equation, this model was not as accurate for the low strain portion. Instead, the linear model was more appropriate. This observation was consistent throughout other ligaments.

Table 2.
Modeling High-Strain Portion of the Stress-Strain Curve

Specimen #	E, kPa	A parameter, kPa	B parameter
S 1	40.1	0.99	9.3
S 2	35.4	0.93	11.2
S 3	47.1	1.7	9.6
S 4	21.9	0.97	8.7
Mean ± SD	36.1 ± 10.6	1.1 ± .37	9.7 ± 1.1
A 1	2.2	3.8	8.2
A 2*	21.0	0.35	11.6
A 3	36.1	1.1	9.4
A 4*	21.2	1.4	9.0
Mean ± SD	30.1 ± 10.7	1.7 ± 1.5	9.6 ± 1.5
Combined Mean ± SD	33.1 ± 10.4	1.4 ± 1.0	9.6 ± 1.2

E- Young's modulus for strain < 15%
A & B are parameters for the exponential equation: stress= $A(e^{B\,strain} - 1)$
S 1-4: surgical specimens, A 1-4: autopsy specimens
* Specimen was frozen and thawed

In addition, the tangent Young's modulus of the human ligament was calculated as the derivative of the stress-strain curve, Young's modulus = $ABe^{B\,strain}$, and plotted in Figure 6. For an exponential curve, the derivative is another exponential, but the ordinate values are changed and the curve does not go through the origin.

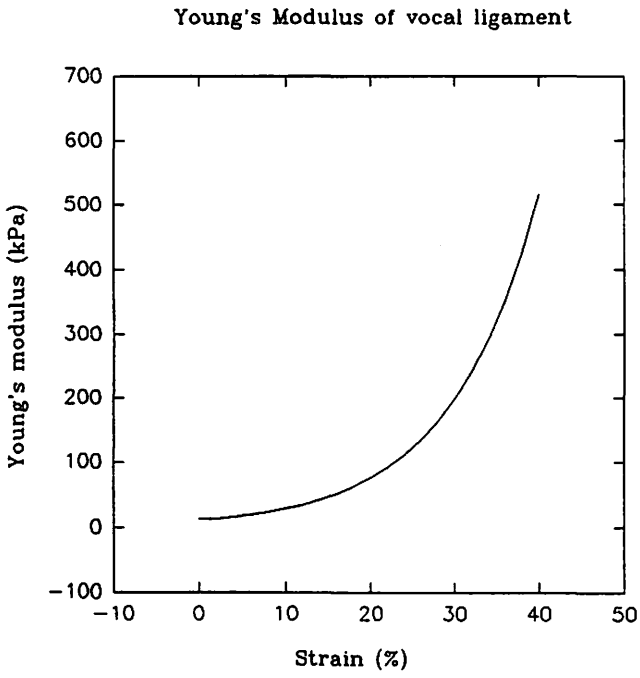


Figure 6. Young's modulus of human ligament plotted using the equation: Young's modulus=$ABe^{B\,strain}$. The means of A and B parameters of ligaments were applied in the equation.

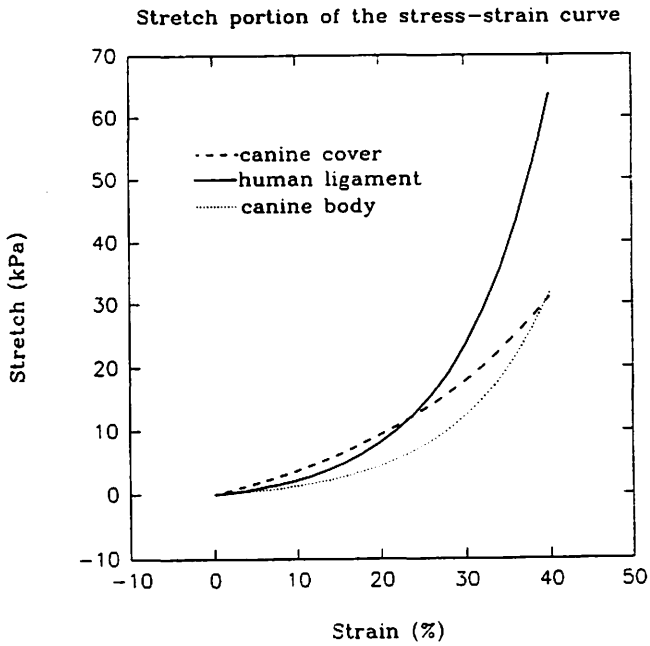


Figure 7. Stress-strain curves of the canine cover, the human ligament, and the canine body. Curves plotted by applying the A and B parameters to the exponential equations.

Discussion

The primary purpose of this investigation was to determine the elasticity of the human vocal fold (VF) ligament so that the stress-strain curve could be used for the development of a biomechanical model of human phonation. The strain-strain curve of the human ligament was nearly linear at low strain and the Young's modulus was relatively small (13-52 kPa). However, Young's modulus reached 135 kPa at approximately 25% strain and 600 kPa at 40% strain. The value at 25% strain agrees well with earlier measurements on human tissue by Bauer and König[3].

We were also interested in comparing the elasticity of the ligament with the elasticity of the canine cover and the body reported by Alipour-Haghighi and Titze[5]. At low strain, the Young's modulus of the canine cover was 41.9 kPa, somewhat greater than that of the ligament measured here (33.1 kPa). However, the canine thyroarytenoid muscle (body) had a low strain Young's modulus of only 20.7 kPa. Therefore, assuming canine and human tissues to be similar, the mean longitudinal Young's modulus for noncontractile vocal fold tissues decreases laterally from mucosa to the ligament to the thyroarytenoid muscle, at low strains. To illustrate the relationship between these three tissue layers, the stress-strain curves were overlapped in Figure 7. Note the order of decreasing stress, at 10% strain, from canine cover (mucosal tissue), to ligament, to canine body (muscle tissue). In addition, the mean Young's modulus for the same tissues is plotted in Figure 8 (next page). This makes the low strain differences even clearer.

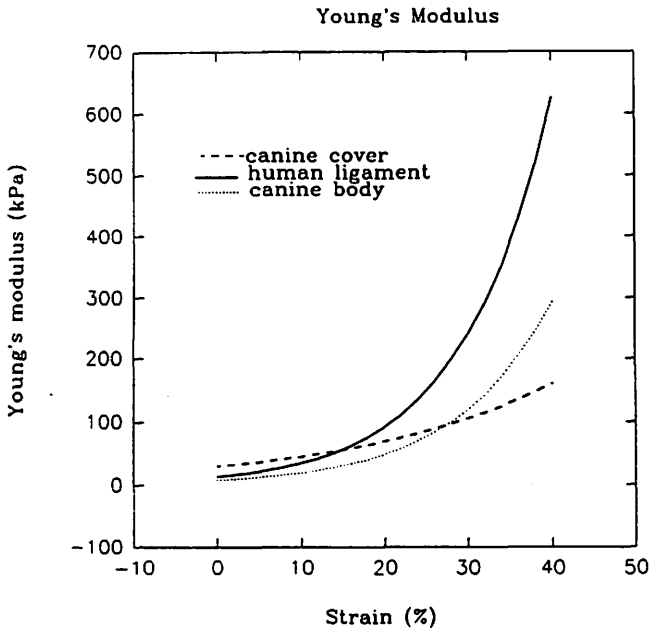


Figure 8. Young's modulus curves of the canine cover, the human ligament, and the canine body. Curves plotted by applying the A and B parameters to the Young's modulus equation.

The situation is rather different, however, for high strains. At 40% strain, for example, the ligament stress dominates over the mucosal stress and the passive muscle stress. This result supports the much hypothesized concept that, although most of the VF vibration is observed in the cover, the vocal ligament plays a major role in F_0 control at high strain. High ligament stress can occur simultaneously with a loose cover layer for the propagation of a mucosal wave[2,5].

This study illustrates that the elasticity of the ligament is most nonlinear, while mucosal tissue is least nonlinear, with passive muscle tissue being between the two. These distinct elastic properties support Hirano's[12] notions that the elasticity of VF layers is due to their histologic differences; in particular, the amount of collagen and elastin fibers in each layer. Hirano[12] observed that mucosal tissue has few elastin or collagen fibers, whereas the ligament has many elastic and collagenous fibers. The linear elasticity at low strain and nonlinear elastic properties at high strain are compatible with the unique collagen structure. As noted in other types of ligament, the collagen fibers are helical in structure, with an accordion-like pattern.[12,14] This type of fiber pattern could function as a buffer so that initial elongations can occur without much resistance.[14] At high strain, however, the unique collagen structure absorbs high stress and functions like a "shock absorber"[14] along the length of the fibers.

Histologically, some investigators[15,16] have observed age-related changes in the vocal ligament. After the age of 40, there was atrophy of elastic fibers and increase in collagen fibers, with the loss of parallel orientation at the edge of the VF. It was suggested that these changes may result in a stiffer VF with increased age. The mean age of our surgical donors was approximately twice that of the autopsy donors, but our Young's moduli and the exponential parameters for the stress-strain curves of surgical and autopsy larynx donors were similar. This lack of age effects is not conclusive, however, because the number of samples in our study was small.

In the literature, there are conflicting reports of gender difference in the elastic property of VF. Van den Berg and Tan[2] found both gender groups gave comparable results for both the entire VF and the ligament, but Bauer and König[3] reported that the VF elastic modulus for females was approximately twice that of the males. Our results on vocal ligaments support the van den Berg and Tan's[2] finding that the elastic property between the two gender groups were similar.

Tissue viability is a major issue in elasticity measurement of VF muscle because of the changes associated with rigor mortis.[4] For the ligament tissue, however, viability is not critical. The Young's modulus and the exponential parameters of surgical ligament sample #1, which were measured 17 hrs later than the rest of the surgical ligaments, showed no significant difference. Furthermore, there were no dissimilarity between the elastic property of the surgical versus the autopsy ligaments. This finding was consistent with the results of elasticity investigations in the frozen versus the fresh knee ligament.[17] This result is compatible with the notion that, while the muscle tissue is composed of living cells, the ligament is composed mostly of extracellular fiber, protein and water.[18] Therefore, ligament tissue should be less sensitive to oxygenation and vascularity, other than during development and following injury when fibroblast cells need to replace the destroyed ligament fibers.

Acknowledgements

This work was supported by Grant #P60 DC00976 from the National Institute on Deafness and Other Communication Disorders.

References

1. Titze, I.R. Principles of Voice Production. Englewood Cliffs, NJ: Prentice Hall, 1994.

2. Van den Berg J, Tan TS. Results of experiments with human larynges. Practica Oto-Rhino-Laryngolo 1959;21:425-450.

3. Bauer H, König W. Physikalische und histologische untersuchungen an normalen und einem weiblichen virilisierten kehlkopf. Archiv fur klinische und experimentelle Ohren-Nasen-und Kehlokopfheilkunde. 1967;188:358-363.

4. Perlman AL, Titze IR, Cooper DS. Elasticity of canine vocal fold tissue. J Speech Hear Res 1984;27:212-9.

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

6. Hirano M. Structure and vibratory behavior of the vocal fold. In: Sawashima M, Cooper FS, eds. Dynamic aspects of speech production. Tokyo, Japan: University of Tokyo, 1977:13-30.

7. Tran QT, Berke GS, Gerratt BR, Keiman J. Measurement of young's modulus in the in vivo human vocal folds. Ann Otol Rhinol Laryngol 1993;102:584-591.

8. Perlman AL, Titze IR. Measurement of viscoelastic properties in live tissue. In: Titze IR, Scherer RC, eds. Vocal fold physiology: biomechanics, acoustics and phonatory control. Denver, CO: Denver Center for the Performing Arts, 1985:273-81.

9. Perlman AL, Durham PL. In vitro studies of vocal fold mucosa during isometric condition. In: Baer T, Sasaki C, Harris K, eds. Laryngeal Function in Phonation and Respiration. Boston, Mass: College-Hill, 1987:291-303.

10. Vennard JK. Elementary fluid mechanics. 2nd ed. New York, NY: John Wiley & Sons, 1947.

11. Vennard JK. Elementary fluid mechanics. 2nd ed. New York, NY: John Wiley & Sons, 1947.

12. Hirano M, Kikita Y, Ohmaru K, Kurita S. Structure and mechanical properties of vocal fold. Speech Lang 1982;7:271-297.

13. Hamming RW. Numerical Methods for Scientists and Engineers. New York, NY:McGraw-Hill, 1973.

14. Viidik A. Simultaneous mechanical and light microscopic studies of collagen fibers. Z Anat Entwicklungsgesch 1972;136:204.

15. Hirano M, Kurita S, Nakashima T. Growth, development, and aging of human vocal folds. In: Bless D, Abbs J, eds. Vocal fold physiology: contemporary research and clinical issues. San Diego, CA: College-Hill Press, 1983:22-43.

16. Kahane, JC. A survey of age-related changes in the connective tissue of the human adult larynx. In: Bless DM, Abbs JH, eds. Vocal fold physiology. San Diego, CA: College Hill Press, 1983:44-49.

17. Woo SL-Y, Orlando CA, Camp JF, Akeson WH. Effects of postmortem storage by freezing on ligament tensile behavior. J Biomech 1986;19:399-404.

18. Banga I. Structure and function of elastin and collagen. Budapest, Hungary: Akademiai Kiado, 1966.

NCVS Status and Progress Report - 7
December 1994, 139-147

Fundamental Frequency and Tracheal Pressure during Three Types of Vocalizations Elicited from Anesthetized Dogs

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

Abstract

Electrical stimulation of the midbrain was used to elicit a variety of vocalizations from 6 anesthetized dogs. This study was conducted to investigate the ranges of and relationships between fundamental frequency of the vocalizations (F_0) and tracheal pressure (P_t) produced during the vocalizations. The vocalizations were described according to type (growl, howl, and whine), and F_0 and P_t, as well as patterns of laryngeal muscle activity, were examined for each vocalization type. Natural-sounding growl and howl vocalizations were elicited from 5 dogs; 3 dogs also produced whines. With few exceptions, F_0 was categorically different for the three vocalization types (low for growls, average for howls, very high for whines). P_t values overlapped for the three vocalization types, although, on average, howls were produced with greater P_t than growls. Patterns and degrees of laryngeal muscle activity varied across and within vocalization types, but general findings were consistent with the presumed function of most of the muscles. Laryngeal muscle activity may help explain some of the variability in the acoustic and aerodynamic data.

An understanding of the neural control of phonation and respiration is essential to our ability to assess and manage disorders of laryngeal and respiratory control. In human subjects, perceptual, acoustic, and aerodynamic characteristics of voice can be assessed noninvasively, as can some aspects of laryngeal and respiratory biomechanics, muscle activity (through electromyography), and central nervous system control (through cortical evoked potentials and imaging techniques such as PET and MRI). However, detailed study of neurophysiologic events under controlled experimental conditions are not possible in the intact human.

Fortunately, animal models are available for the *in vivo* study of vocalization. Such models involve techniques that result in activation of respiratory and/or laryngeal muscles. Activation may result from peripheral or central stimulation. The method described by Berke and colleagues (1) involves peripheral stimulation. While warm humidified air is channelled through the larynx, the recurrent laryngeal nerves and external branches of the superior laryngeal nerves are stimulated, resulting in vocalization. Central stimulation involves either chemical or electrical stimuli. Chemical stimulation involves the injection of excitatory amino acids into the midbrain periaqueductal gray (PAG) and results in repetitive vocalizations for periods of 1 to 2 minutes. This and similar techniques have been used in the cat (2-4) and squirrel monkey (5-7). Electrical stimulation involves delivering a current through a bipolar concentric or monopolar electrode to the midbrain. This technique has been used in a variety of species, including bats (8), rats (9), monkeys (10-14), cats (13, 15-20), and dogs (21, 22).

In our laboratory, we study evoked vocalization by electrical stimulation of the midbrain in fully anesthetized dogs. Dogs are the preferred species for these experiments because the size and morphology of the larynx is comparable to human, and many biomechanical details of vocal fold tissue and vibration have been discovered using canine larynges (e.g., 23-26). During initial studies of evoked vocalization in the dog, we observed that different "types" of vocalization resulted from stimulating slightly different sites in the effective region of the midbrain or by varying stimulus parameters. The vocalizations sounded like a howl, growl, whine, or, less often, bark. We have examined patterns of respiratory and laryngeal muscle activity as a result of midbrain stimulation (19), have developed a functional "map" of behavioral responses to stimulation of a large area

of the midbrain (22), and have investigated the influence of experimentally varying degrees of subglottal (tracheal) pressure (P_t) on the fundamental frequency (F_0) of the elicited vocalization (21).

Although the dog's vocalizations observed in this study sounded natural and of different qualitative types, we recognize the need to provide objective evidence on this aspect of the experiments. Without such analysis, it might be that categorization of the "type" of vocalization is not reliable between listeners, or that such categorization does not correspond in a systematic way with objective acoustic features. If these types of vocalization are qualitatively different from one another, one would expect, for example, that the distribution of the F_0 of all vocalizations observed would exhibit modes corresponding to the perceptually identified type of vocalization. On the other hand, a normal distribution of F_0 would refute the notion of categories of vocalization types. Instead, perhaps human listeners force categories on vocalizations, depending upon whether the F_0 is relatively low (growl), average (howl), or high (whine). Therefore, one goal of this research was to elicit a wide range of vocalization types in anesthetized dogs by stimulating various regions of the midbrain.

The physiologic and aerodynamic contributions to vocalization are expected to vary in some systematic ways with the different types of canine vocalizations, or with some acoustic aspect of the vocalizations. In this report, we will describe typical patterns of muscle activity during each type of vocalization. In addition, the pressure generated by the respiratory system (tracheal pressure, P_t) for each vocalization will be examined for differences across vocalization types. The acoustic measure fundamental frequency (F_0) will be analyzed for each vocalization, and also examined for differences across vocalization types. Finally, correlations between F_0 and P_t during elicited vocalizations will be explored.

Method

Six adult mongrel dogs were sedated with an intramuscular injection of a ketamine (7.5 ml/kg) and xylazine (2.5 mg/kg) mixture.[1] A catheter was then placed in the femoral vein for hydration and periodic injection of pentobarbital (25 mg/kg) to induce and maintain a surgical level of anesthesia (absence of corneal reflexes and deep pain responses). The head was positioned in a modified Kopf 1430 stereotaxic instrument (Tajunga, CA), which was attached to a heavy rigid table that supported the animal's body. With the animal in a prone position, the surface of the cortex was exposed by drilling a circular opening in the skull and carefully retracting the dura mater. A single monopolar microelectrode or an array of 2 or 4 etched metal microelectrodes was advanced into the midbrain. Electrical stimuli were 2-second trains of 0.2 ms pulses at 200 Hz from a constant current stimulator (Grass 88 with SIU6). Current levels were varied as necessary (from 0.2 to 1.5 ma) to evoke vocalizations with a minimum of body movements.

After the electrode was advanced into the brain, the animal was rotated to a supine position and the larynx and trachea were exposed by a midline neck incision. In addition, a brace was used to open the oral airway maximally, and the larynx was viewed intraorally. Bipolar hooked-wire electrodes, made from two insulated stainless steel wires (0.08 mm diameter), which were bonded together along their length ("bifilar"), were placed carefully in bilateral pairs of four intrinsic laryngeal muscles: thyroarytenoid (TA), cricothyroid (CT), lateral cricoarytenoid (LCA), and posterior cricoarytenoid (PCA). Electrode placements were guided by visual inspection and confirmed by stimulating the muscles through the EMG electrodes (100 ms trains of 2 ms pulses, 50 Hz, 0.3-3.0 ma) and observing the appropriate laryngeal movement. After confirming the position of the EMG electrode in the intended muscle, the electrode wires were connected to a differential amplifier (bandpass 30 Hz to 5 kHz). A cannula was placed in the trachea and coupled with a pressure transducer (Microswitch 143PC03G) to sense tracheal pressure. An electret lapel microphone (Realistic) was attached to a maxillary canine tooth to record the animal's vocalizations. The larynx was visualized throughout the experiment with a 10 mm rigid boroscope and charge-coupled-device (CCD) camera (Fibertron ICV 700TMC, Carrollton, TX) mounted on the table. Examples of laryngeal views and acoustic signals during various vocalizations were videotape recorded.

A signal corresponding to the stimulus applied to the midbrain, EMG signals, tracheal pressure, and acoustic signals were recorded simultaneously on 2 FM instrumentation reel-to-reel tape recorders (HP 3968A, DC - 2.5 kHz). Data also were monitored on-line using a computer data acquisition program (Codas-Windaq from Dataq Instruments, Akron, OH) and desktop computer.

Stimuli were delivered to the midbrain at a variety of locations until vocalization was elicited. Stimulation sites generally were in a region dorsomedial to the medial lemniscus just anterior to the most anterior border of the inferior colliculus. Several trials of vocalization at each midbrain site were recorded at various current levels. Electrode sites and current levels were varied until a satisfactory number of vocalizations were elicited. After data collection, the animal was euthanized with a lethal dose of pentobarbital.

Data Analysis

The acoustic signal containing vocalizations and the corresponding tracheal pressure signal were low-pass filtered at 5 kHz (8-pole Butterworth characteristic) and then digitized at a rate of 10 kHz. The portions of the responses that were selected for analysis corresponded with relatively steady-state vocalizations. The fundamental frequency of

[1] Dogs #1 and #2, not included in the total number of experimental animals, were involved in pilot studies to work out methodological details of this study.

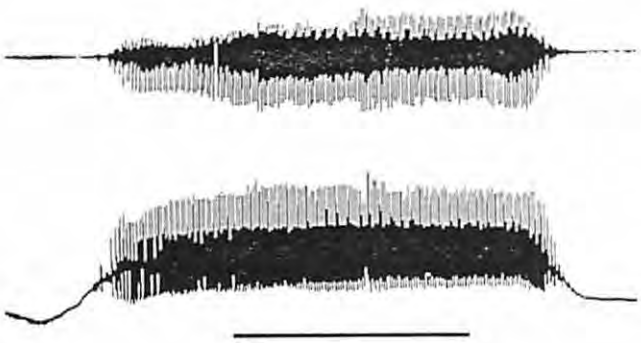


Figure 1. Example of a vocalization categorized as a growl from Dog 6. The top trace is the acoustic signal and the bottom trace is tracheal pressure (P_t). The fundamental frequency (F_0) of the vocalization was determined by spectral analysis to be 112 Hz. The portion of the P_t signal to be measured in this example is indicated by the bar (0.8 s) beneath the data. Average P_t in this example is 2.9 kPa.

the vocalization (F_0) was determined by conducting a fast Fourier transform (FFT) using a Hanning window on the acoustic data. The first peak in the spectrum corresponded with the voice F_0. Pressure was determined by calculating the mean value of the data points within the portion of interest. A typical record indicating the portion of the P_t signal selected for analysis is provided in Figure 1.

Vocalization type was determined independently by 2 listeners. The acoustic channel of the data tape in was played in free field. Listeners used their own experiences and knowledge of dog's vocalizations to categorize productions. Generally, the perceptual characteristics of the vocalization types can be described as follows: (1) growl -- rough-sounding vocalization that was usually low in pitch and loudness; (2) howl -- smooth, rich-sounding sustained vocalization; (3) whine -- an extremely high pitched "thin" sustained vocalization, usually low in loudness. The 2 listeners were in agreement for 96% of the data included in the study. In the few cases of disagreement, 1 listener reviewed the vocalization and made the final decision. Approximately 40 additional vocalizations were elicited that were excluded from the analysis because they were too brief or weak to be categorized and measured with confidence.

In a separate analysis, the stimulus, vocalization, pressure signals, and EMG data were low-pass filtered at 500 Hz and digitized at a sampling rate of 1 kHz. These data were inspected visually for patterns of EMG activity.

Results

Natural-sounding vocalizations were elicited from 5 of the 6 dogs during electrical stimulation of the midbrain. Vocalizations elicited from one dog (#4) were not natural sounding. Visual inspection of this animal's larynx revealed vocal nodules; voice analyses were not conducted for this animal, but the EMG data were useful. Pressure data were missing for one dog (#3) because of equipment problems; F_0 data were used from this animal.

All 5 animals with natural-sounding vocalizations produced growls and howls as a result of electrical stimulation of the midbrain. Three of these animals also produced whines. No dog produced a recognizable bark. Some vocalizations were perceived as a combination of types. That is, a whine and a howl ($\underline{n} = 28$), or less often, a growl and a howl ($\underline{n} = 8$), were produced concurrently. Spectra of the 3 vocalization types (growl, howl, whine) and 1 combination vocalization (howl + whine) are illustrated in Fig. 2, panels a-d.

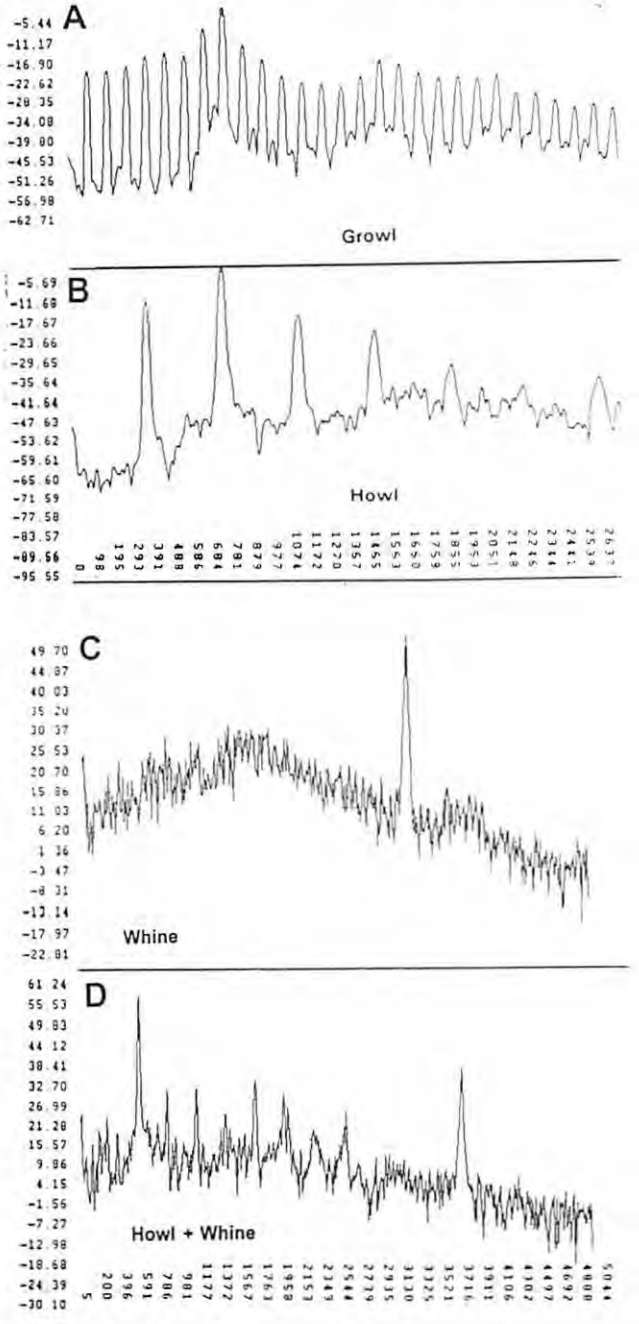


Figure 2. Spectra of 4 vocalizations. Amplitude is in Volts; frequency is in Hz. Panel A: growl produced by Dog 6; F_0 (first spectral peak) = 117 Hz. Panel B: howl by Dog 6; F_0 = 303 Hz. Panel C: whine by Dog 7; F_0 = 3198 Hz. Panel D: howl + whine combination; F_0 =561 Hz, secondary peak = 3706 Hz.

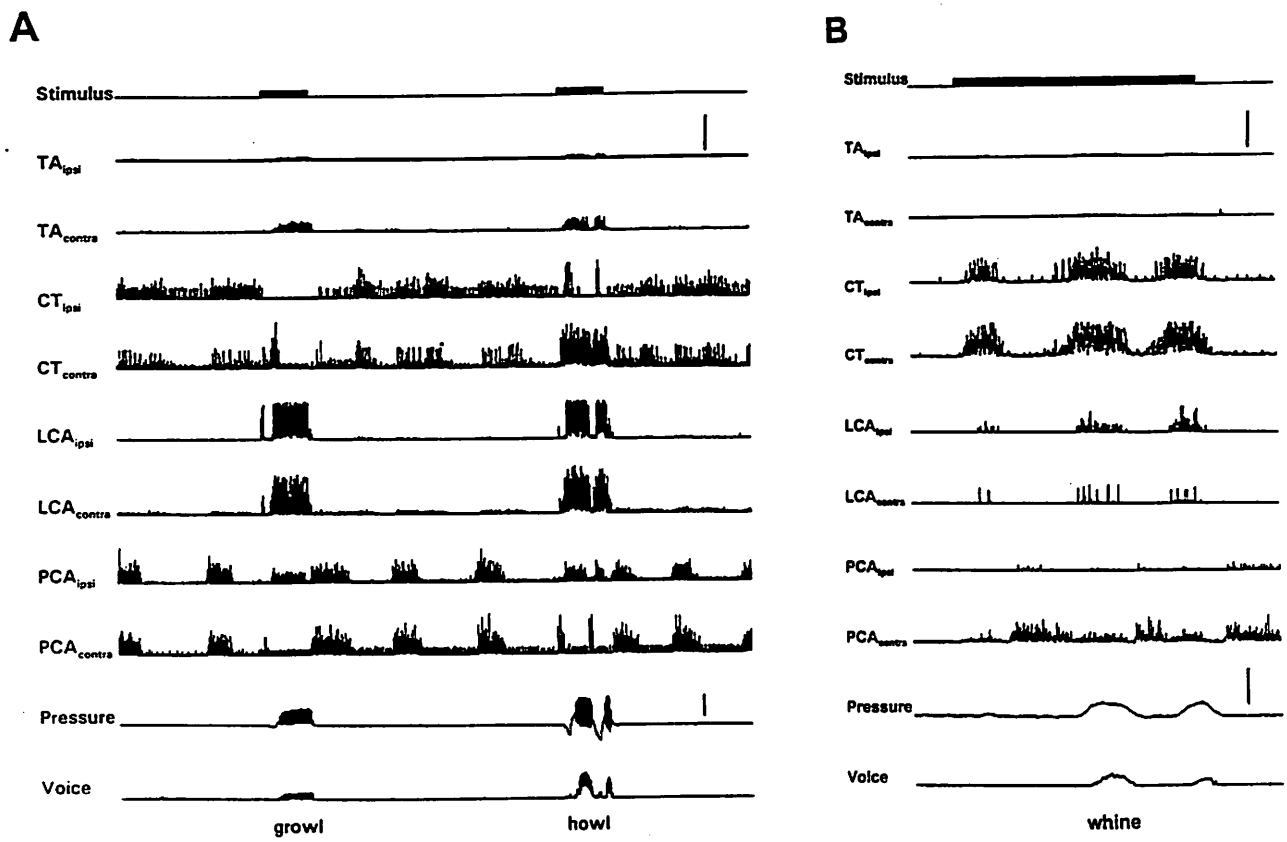


Figure 3. Examples of typical patterns of muscle activity for growl (Panel A, left), howl (Panel A, right), and whine (Panel B) vocalizations. Ipsilateral (ipsi) and contralateral (contra) muscles are referenced to the side of the midbrain that was stimulated. TA = thyroarytenoid, CT = cricothyroid, LCA = lateral cricoarytenoid, PCA = posterior cricoarytenoid, Pressure = tracheal pressure (P), Voice = acoustic signal. Each stimulus signal, represented by the solid bars in the first channel of Panels A and B, is 2.0 s in duration. In Panel A, the calibration bar on the right side of the TA_{ipsi} channel represents 0.5 mV for TA_{ipsi} and PCA_{ipsi}, and 1.0 mV for the other muscles. The calibration bar on the right of the Pressure channel is 5 kPa. In Panel B, the calibration bar on the right side of the TA_{ipsi} channel represents 1 mV for the TA, LCA, and PCA muscles, and 2.5 mV for both CT muscles. The pressure calibration is 2 kPa.

Table 1.
The maximum and minimum (separated by a hyphen) F_0 (in Hz) for each vocalization (and the number of these, n) produced by each dog is listed by type of vocalization. In a few cases, F_0 or a group of F_0's was distinctly different from others for the same type of vocalization; these are listed separately.

Dog	GROWL	n	HOWL	n	WHINE	n
3	141 - 210	12	258 - 588	36	478, 1010	2
					3178 - 3725	8
5	78 - 119	9	134, 156	2		
	183	1	295 - 678	60		
6	58 - 124	30	239 - 449	26	690 - 786	4
					3554 - 3745	6
7	78 - 117	11	142 - 918	89	732, 771	2
					2959 - 3330	21
8	112-156	9	127	1		
			171-317	26		

Combination vocalizations were produced by Dogs 3, 6, and 7 (n = 19, 6, and 11, respectively); these were not included in the summary analyses and figures.

Analysis of the EMG data revealed some characteristic patterns of muscle activity accompanying various types of vocalization. However, there was a substantial amount of variability in the details of the EMG patterns over repeated trials and animals. The most common patterns of muscle activity for a howl, growl, and whine, each elicited with a 2-second stimulus, were selected for illustration in Figure 3, panels a and b. Note that during a growl (first vocalization in Figure 3a), LCA activity but little or no CT activity occurred. During howl productions, either one or both CT muscles often were active (for example, see the second vocalization in Figure 3a). CT muscles were highly active during a whine, but the LCA muscles often were less active than during the other vocalization types (Figure 3b). Surprisingly, TA activity was low or absent during all types of vocalization elicited in this study. However, high levels of TA activity were elicited from each dog during a gagging or

Table 2.
The mean, standard deviation (in parentheses), maximum and minimum (separated by a hyphen) values for tracheal pressure (P_t, in kPa; 1 kPa = 10.2 cm H_2O) for each vocalization (and the number of these, n) produced by each dog are listed by type of vocalization. P_t data for Dog 3 are missing because of equipment problems.

Dog	GROWL	n	HOWL	n	WHINE	n
5	.656 (.202) .471 - 1.010	11	1.809 (.785) .672 - 3.33	62		
6	1.732 (.669) .711 - 3.197	30	2.735 (.623) 1.147 - 3.767	26	2.454 (1.175) 1.547 - 2.397	10
7	.952 (.218) .593 - 1.323	11	1.487 (.580) .531 - 2.928	89	.789 (.248) .369 - 1.189	23
8	.906 (.160) .666 - 1.123	9	2.339 (.748) .857 - 4.363	27		

retching response to electrical stimulation to a more posterior and ventral region of the midbrain.

Selected summary statistics for the F_0 values for each dog are listed in Table 1, and for the P_t are listed in Table 2. For F_0, the minimum and maximum value for each vocalization by each dog is listed by type of vocalization (Table 1). Instances of F_0 that were distinctly different from others for the same type of vocalization are listed separately. It is interesting to note that F_0 was measurable for all but one vocalization: a growl production by Dog 5. In this case, the FFT analysis did not detect a periodic component from the acoustic record.

All vocalizations for which both P_t and F_0 data were available are plotted in Figure 4. Thus, data from 4 dogs comprise this plot of 347 vocalizations. Examination of Table 1 and Figure 4 reveals that there is little overlap in F_0 between the voice types. Voice F_0 is lowest for the growl and highest for the whine productions. The overlap between F_0 for the growl and howl evident in Figure 4 results mostly from differences between dogs. Of the eight whine productions with F_0 less than 2.9 kHz, two had an additional major peak in their spectra that was typical of the other whines.

The ranges for P_t are not nearly as exclusive as those for F_0 (see minimum and maximum values in Table 2 and Figure 4). That is, the values for P_t overlap for the 3 types of vocalization. To describe the results for this variable in regard to voice type, means and standard deviations in addition to minima and maxima for each dog were calculated and are included in Table 2. Average P_t was greater for howls as compared to growls for each dog. The standard deviation was greater for howls than growls for 3 of the 4 dogs. It appears from Figure 4 that whines were produced with relatively low P_t, however, it is clear from Table 2 that this was true only for 1 of the 2 animals who produced whines.

To assess the frequency of occurrence of values of P_t (grouped into 0.1 kPa intervals), the P_t data for all vocalization types are presented as a frequency histogram in Figure 5. From this figure, it is clear that most of the vocalizations were produced with P_t values from 0.8 to 1.2 kPa (about 8 to 12 cmH_2O), although values ranged from 0.4 to 4.4 kPa.

A relationship between F_0 and P_t is unclear by viewing the data in Figure 4. However, their appears to be a linear relationship between these variables for the growl. Inspection of data from one animal who produced growls with a wide range of P_t values (Dog 6, illustrated in Figure 6 - see next page) supports this observation. The P_t range for the other dogs' growls were too limited to discern the nature of the relationship between these variables. No clear relationship between the two dependent variables was evident for howls and whines (e.g., see howls in Figure 6).

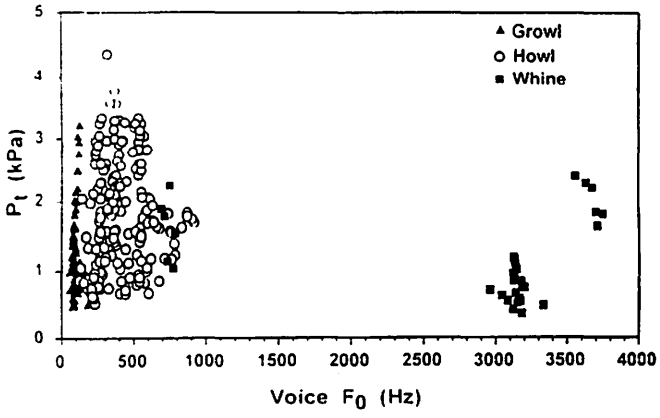


Figure 4. Tracheal pressure (P_t in kPa) and voice fundamental frequency (F_0 in Hz) are plotted for all vocalizations included for analysis produced by 4 dogs (N=347). Vocalizations categorized as growls are indicated by closed triangles, howls are open circles, and whines are closed squares.

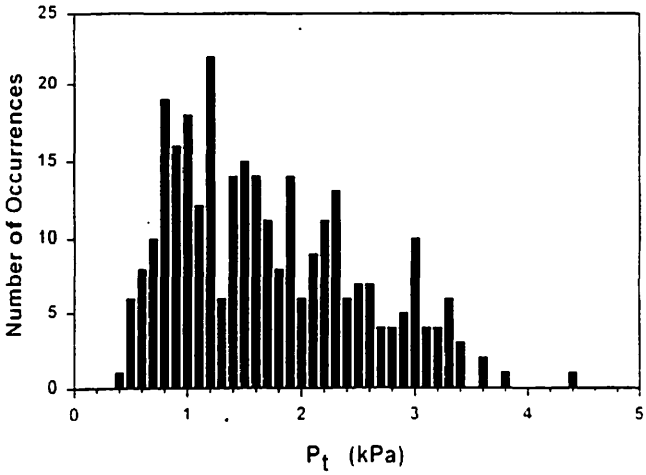


Figure 5. Histogram illustrating the frequency of occurrence of values of P_t (in intervals of 0.1 kPa) for all vocalizations.

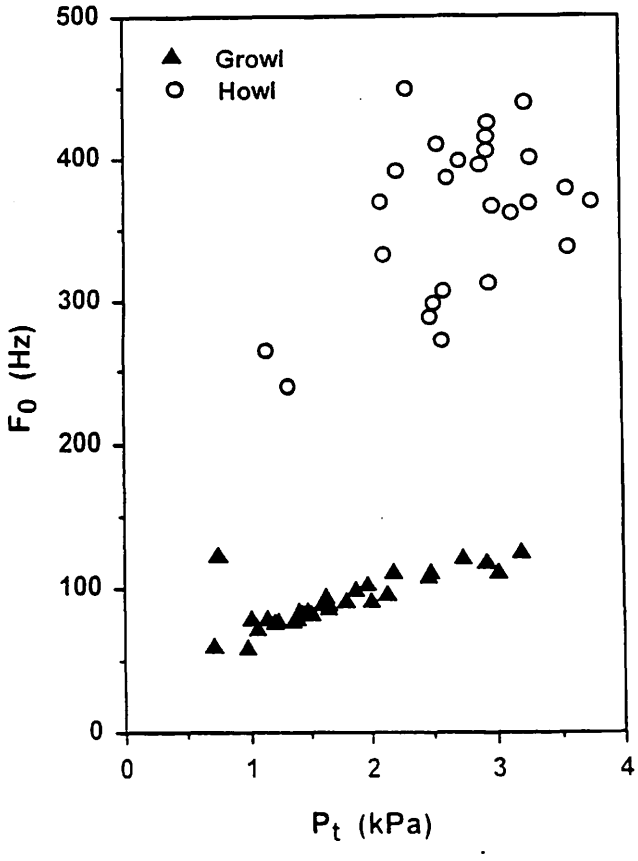


Figure 6. Fundamental frequency (F_0) and tracheal pressure (P_t) plotted for each vocalization by Dog 6. Note linear relationship between F_0 and P_t for growls (closed triangles), but not for howls (open circles).

Discussion

Electrical stimulation of multiple sites of the midbrain of anesthetized dogs resulted in the productions of natural-sounding growls, howls, and whines in 3 dogs, and only howls and growls in 2 dogs. The analyses conducted provided objective evidence to support the "categorization" hypothesis of vocalization types with regard to F_0 and patterns of laryngeal muscle activity, but not P_t. Of course, there are other variables, such as vocal fold vibratory pattern and spectral and temporal characteristics of vocalization, that could affect our categorical perception of vocalization types in dogs.

Other researchers have identified a variety of vocalization types elicited by stimulation of the midbrain or brainstem in animals. De Lanerolle & Lang (27) reviewed the literature (and contributed to it; 4) in this area for cats: In studies conducted from 1909 to 1968, meows (mews), hisses, growls, piercing cries, and purrs have been elicited by electrical stimulation. Different types of vocalizations elicited from midbrain electrical stimulation have also been described in the monkey (e.g., barks, chucks, peeps, cackling, coos; 6, 11, 14).

By understanding which regions of the brain are responsible for various types of vocalizations, we can begin to develop a model for the motor coordination of vocalization. In addition, information relating to this issue for a variety of animals is important for determining if such a model is valid across species or if certain details of the neural control of vocalization is species specific. Kirzinger & Jürgens (6) identified specific regions of the ventrolateral midbrain that, when lesioned, affected acoustic characteristics of "peep" and "cackling" vocalizations by the squirrel monkey; these areas included the substantia nigra and overlying reticular formation. This region is typical of the area we have explored to produce a variety of vocalization types in the dog (22). Kanai & Wang (17) found that they could elicited different types of vocalizations in decorticated or decerebrated cats by electrical stimulation of different parts of the midbrain and brainstem. Signs of rage including loud screams or growls resulted from stimulation to the dorsomedial, ventromedial, and supraoptic nuclei of the hypothalamus, and low-pitched growls resulted from stimulation of the reticular formation throughout the lower brain stem. De Lanerolle (4) used electrical stimulation in the ventrolateral pons in cats to elicit growls, hisses, meows, and meow-growls. In preliminary studies, de Lanerolle and Lang (27) reported that stimulation to the preoptic area resulted in meow and hiss-growl vocalizations, and stimulation to the PAG evoked meows, growls, and hisses. The present study cannot address this issue because vocalizations were elicited from only a few tracks representing a small region of the midbrain, typically 2 mm in the anterior-posterior dimension, and 3 mm in the medial-lateral dimension. Nonetheless, we examined our data for a possible relationship between vocalization types and relative site of midbrain stimulation, and found none. Although the extant literature does not unequivocally relate specific areas of the brain to specific types of vocalizations, it is clear that the midbrain PAG and associated pathways play an important role in mammalian vocalization.

Qualitative analysis of patterns of laryngeal muscle activity during the three types of vocalizations are consistent with the presumed function of most of the muscles. That is, vocalizations with lower F_0 involve little or no CT activity, whereas vocalizations with higher F_0 involve high levels of CT activity. The CT muscle is known to lengthen and stretch the vocal folds, thus facilitating faster vibration. LCA muscle activity was apparent during growl and howl vocalizations, but much less so during whine productions. The LCA muscles assist with vocal fold adduction, especially at the area of the vocal processes of the arytenoid cartilages. Tight vocal fold approximation would interfere with the whine production. The minor contribution of the TA muscles to any type of vocalization in any of the dogs was surprising, and is unlike what one would expect in the human voice. The TA and CT muscles interact to affect F_0 of the voice (28) in

at least some humans, but the dog may rely primarily on CT muscle activity for this function. The PCA muscles act as laryngeal abductors, and generally are active before and after vocalization. The appearance of ipsilateral PCA activity during the growl and howl productions pictured in Figure 3a is puzzling. We considered that this EMG activity may derive from a different muscle, but careful inspection of the timing of the activity and shapes of the waveforms suggests that the PCA may in fact contract along with laryngeal adductors during vocalization in the dog. Further scrutiny of laryngeal EMG data in these and other experimental animals is needed to confirm this supposition.

Each of the three types of vocalizations identified in this investigation had a characteristic range for voice fundamental frequency, but were produced with overlapping values for tracheal pressure. Previous research has identified a linear relationship between F_0 and P_t on theoretical (29) and experimental (21; howl-like vocalizations in dogs) grounds. The experimental study (21) employed the midbrain stimulation method described here, but provided an artificial source of air pressure to the larynx through an opening in the trachea. The electrical stimulus to the midbrain did not vary, and thus elicited consistent and robust responses by laryngeal and respiratory muscles. Air pressure was varied experimentally, and a linear relationship between F_0 and P_t was demonstrated clearly in seven dogs. In the present experiment, laryngeal muscle activity was free to vary. Therefore, it is not surprising that there was no consistent linear relationship between F_0 and P_t for these data. Those vocalizations whose data produce the scatter in the data presumably were produced with varying degrees and patterns of laryngeal muscle activity, especially of the TA and CT muscles (28). Currently, we are analyzing the EMG data quantitatively to reveal specific characteristics of muscle activity associated with different types of vocalizations. This analysis will allow us to examine the extent to which muscle activity explains the variability in the present results.

In addition to growls, howls, and whines involving essentially exclusive fundamental frequency values, each had a unique and characteristic vocal quality. We believe that these vocalization types may correspond with vocal registers in the human voice. In fact, the dog's growl sounded similar to a vocal fry, or pulse register vocalization. Howls could be considered a modal type of vocalization, and, if compared to human voice, might be characteristic of chest or head voice. Some howls may have been produced in a manner similar to loft, or falsetto, register voice in humans, but this potential subgroup of howls was not differentiated in our perceptual analyses. In contrast to growls and howls, the whine does not seem typical of a human vocal register. It is unlikely that the whine vocalization by the dog is produced by the same mechanism as the so-called "whistle" (also called flageolet, flute, or bell) register in human singers. Whereas vocal fold oscillation accompanies whistle register production (30), it seems highly unlikely that the vocal folds oscillate at frequencies approximating 3000 Hz for the whine. Indeed, the F_0 produced by women using whistle register is quite a bit lower (988-1,568 Hz; 31) than the F_0 produced by the dogs during the vast majority of the whine productions. In addition, it is clear from Figure 2c that the spectrum is quite different from whistle register productions by sopranos. The spectrum of sung whistle register has a harmonic structure (31); the spectrum for the whine may indicate a pure tone (characteristic of a true whistle; 32, 33). Unfortunately, we cannot confirm the absence of harmonics because of the frequency limitations of our acoustic recording equipment. Based on the evidence presented, we presume that the mechanism for the whine, as for a true whistle, involves "vortex shedding" (34). That is to say, as air from the lungs flows past the vocal folds (acting as obstacles to flow), vortices (circular patterns of airflow) are formed at a frequency corresponding to the F_0. Thus, one could conclude that dogs have a true whistle register, but that this term, as used for singers, is a misnomer for human voices.

Barking is the most common type of vocalization produced by domestic dogs, yet we did not elicit a single identifiable bark in this series of experiments. Although we have elicited barks from dogs in other experiments using the same stimulation technique (e.g., 22), it is rare. The reason for this discrepancy between spontaneous vocalizations by awake, freely moving dogs and evoked vocalization in anesthetized, mechanically stabilized dogs is unclear. It is unlikely that these dogs simply never barked, because a few of them were heard to bark in the animal care center before they were anesthetized. Anesthesia should not effect whether vocalization types are elicited with electrical stimulation; de Lanerolle (4) demonstrated that the same types (and frequency distributions of these types) of vocalizations were evoked from both anesthetized and unanesthetized cats. Perhaps our inability to evoke barking relates to the observation that purring is rarely evoked via electrical stimulation in the cat (4, 17). De Lanerolle (4) suggested that purring may be mediated by an separate pathway than that involved in other types of vocalization. Alternatively, he suggested that different stimulus parameters may be needed to evoke purring than to evoke other types of vocalization. Both of these possibilities should be considered for the absence of barking in the present study as well. Another explanation for this unexpected finding could be that biomechanical constraints prohibited the production of recognizable barks. Recall that the jaw was displaced maximally so that the mouth opening was wide, providing a clear view of the larynx. Freedom of the upper airway articulators may be necessary to "shape" the acoustic signal into a bark. In addition, barks may be too brief to provide an adequate acoustic signal. Vocalizations that were brief were excluded from analysis in this study because F_0 extraction was not valid and a determination of vocalization type was difficult. Another explanation in-

volves the neural control of vocalization. The vocalizations studied in this and similar studies were elicited by stimulating areas of the central nervous system no higher than the midbrain. Perhaps barking involves the coordinated activation of several sites in the brain, including cortical centers. In fact, barking elicited in previous experiments has been a late response (i.e., occurs after the termination of stimulation) and may be related to seizure activity of neurons at the electrode site. Such activity may eventually activate neural systems are from the site of stimulation. The pathways and systems involved in bark production may or may not include the midbrain PAG. The PAG may be involved only in vocalizations associated with emotive responses, whereas the bark may be used by dogs as a greeting or warning. We cannot determine, using the current model of evoked vocalization, whether the PAG is involved in the production of barking.

This study has allowed us to examine selected acoustic, aerodynamic, and electromyographic characteristics of different types of vocalizations elicited from anesthetized dogs. In summary, we measured and evaluated the voice fundamental frequencies and tracheal pressures that occurred during the vocalization of growls, howls, and whines. It does not appear that different types of vocalizations are produced simply by varying P_t. Rather, each vocalization type occurs within a relatively restricted F_0 range, and is produced with a particular quality of voice that is easily identified. Qualitative study of laryngeal muscle activity indicated that certain patterns were observed for each voice type. However, quantitative analysis is needed to see if the variability in the acoustic and aerodynamic data can be explained, and to reveal the variety of EMG patterns that occur with such vocalizations.

This research has helped expand our knowledge base regarding the mechanisms of voice production, especially for vocalizations that fit in different qualitative categories and that have widely different F_0 values. By using an animal model, we were able to use invasive procedures and ask more probing questions than are possible with human models. Moreover, such experiments are necessary to explore the differences between vocalizations between human and dog (e.g., apparent differences in TA activity) and to evaluate the utility (advantages and limitations) of using dog larynges to develop models of human vocalization.

Acknowledgement

This research was supported by Grant No. P60 DC00976 from the National Institutes on Deafness and Other Communication Disorders. A version of this research was presented at The Voice Foundation's 23rd Annual Symposium: Care of the Professional Voice, June 1994, Philadelphia. We extend our gratitude to Brad Story for informative discussions about the physics of whistling, and to an anonymous reviewer for helpful suggestions on a previous version of this manuscript.

References

1. Berke GS, Moore DM, Hantke DR, Hanson DG, Gerratt BG, Burstein F. Laryngeal modeling: Theoretical, in vitro, in vivo. Laryngoscope 1987;97:871-881.

2. Bandler R, Carrive P. Integrated defense reaction elicited by excitatory amino acid injection in the midbrain periaqueductal grey region of the unrestrained cat. Brain Res 1988;439:95-106.

3. Davis PJ, Zhang SP, 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 1993;607:61-80.

4. de Lanerolle NC. A pontine call site in the domestic cat: Behavior and neural pathways. Neuroscience 1990;37:201-214.

5. Jürgens U, Richter K. Glutamate-induced vocalization in the squirrel monkey. Brain Res 1986;373:349-358.

6. Kirzinger A, Jürgens U. The effects of brainstem lesions of vocalization in the squirrel monkey. Brain Research 1985;358:150-162.

7. Lu CL, Jürgens U. Effects of chemical stimulation in the periaqueductal gray on vocalization in the squirrel monkey. Brain Res Bull 1993;32:143-151.

8. Schuller G, Radtke-Schuller S. Neural control of vocalization in bats: mapping of brainstem areas with electrical microstimulation eliciting species-specific echolocation calls in the rufous horseshoe bat. Exp Brain Res 1990;79:192-206.

9. Yajima Y, Hayashi Y, Yoshii N. Ambiguus motoneurons discharging closely associated with ultrasonic vocalization in rats. Brain Res 1982;238:445-450.

10. Dressnandt J, Jürgens U. Brain stimulation-induced changes of phonation in the squirrel monkey. Exp Brain Res 1992;89:549-59.

11. Larson CR. The midbrain periaqueductal gray: A brainstem structure involved in vocalization. J Speech Hear Res 1985;28:241-249.

12. Larson CR, Wilson KE, Luschei, ES. Preliminary observations on cortical and brainstem mechanisms of laryngeal control. In: DM Bless, JH Abbs, eds. Vocal Fold Physiology San Diego: College-Hill Press, 1983:82-95.

13. Magoun HW, Atlas D, Ingersoll EH, Ranson SW. Associated facial, vocal and respiratory components of emotional expression: An experimental study. J Neurol and Psychopathology 1937;17:241-255.

14. Müller-Preuss P, Ploog D. Inhibition of auditory cortical neurons during phonation. Brain Res 1981;215:61-76.

15. Garrett JD. The response of laryngeal mechanoreceptors innervated by the superior laryngeal nerve during evoked phonation in the cat. Unpublished doctoral dissertation, University of Iowa, 1986.

16. Ingram WR, Ranson SW, Hannett FI, Zeiss FR, Terwilliger EH. Results of stimulation of the tegmentum with the Horsley-Clarke stereotaxic apparatus. Arch Neurol and Psychiatry 1932;28:513-541.

17. Kanai T, Wang SC. Localization of the central vocalization mechanism in the brain stem of the cat. Exp Neurol 1962;6:426-434.

18. McGlone RE, Richmond WH, Bosma JF. A physiological model for investigation of the fundamental frequency of phonation. Folia phoniat 1966;18:109-116.

19. Solomon NP, Grayhack JP, Liu K, Luschei ES. Patterns of laryngeal and respiratory activation during elicited phonation [Abstract]. Asha 1992;34:175.

20. Testerman RL. Modulation of laryngeal activity by pulmonary changes during vocalization in cats. Exp Neurol 1970;29:281-297.

21. Hsiao T.-Y, Solomon NP, Luschei ES, Titze IR, Liu K, Fu T-C, Hsu M-M. The effect of subglottic pressure on fundamental frequency of the canine larynx with active muscle tensions Annals of ORL 1994;103:817.

22. Liu K, Solomon NP, Luschei ES. Phonation evoked from canine midbrain stimulation: A mapping study [Abstract]. Asha 1993;35:139.

23. Alipour-Haghighi F, Titze IR. Elastic models of vocal fold tissues. J Acoust Soc Am 1991;90:1326-1331.

24. Berke GS, Gerratt BR. Laryngeal biomechanics: An overview of mucosal wave mechanics. Journal of Voice 1993;7:123-128.

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

26. Slavit DH, Lipton RJ, McCaffrey TV. Glottographic analysis of phonation in the excised canine larynx. Ann Otol Rhinol Laryngol 1990;99:396-402.

27. de Lanerolle NC, Lang FF. Functional neural pathways for vocalization in the domestic cat. In: JD Newman, ed. The physiological control of mammalian vocalization New York: Plenum Press, 1988:21-41.

28. Titze IR, Luschei, ES, Hirano M. Role of the thyroarytenoid muscle in regulation of fundamental frequency. J Voice 1989;3:213-224.

29. Titze IR. On the relation between subglottal pressure and fundamental frequency in phonation. J Acoust Soc Am 1989;85:901-906.

30. Walker JS. An Investigation of the whistle register in the female voice. J Voice, 1988;2:140-150.

31. Miller DG, Schutte HK. Physical definition of the "Flageolet Register." J Voice 1993;7:206-212.

32. Shadle CH. The acoustics of fricative consonants. Unpublished doctoral dissertation, Massachusetts Institute of Technology, 1985.

33. Olson HF. Music, Physics and Engineering. 2nd ed. New York: Dover, 1967.

34. Roberson JA, Crowe CT. Engineering Fluid Mechanics. 3rd ed. Boston: Houghton Mifflin, 1985.

NCVS Status and Progress Report - 7
December 1994, 149-154

Nonlinear Dynamics of the Voice: Signal Analysis and Biomechanical Modeling

Hanspeter Herzel, Ph.D.
Institute of Theoretical Physics, Technical University, Germany
David Berry, Ph.D.
Ingo Titze, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa
Ina Steinecke, M.S.
Institute of Theoretical Physics, Humboldt University, Germany

Abstract

Irregularities in voiced speech are often observed as a consequence of vocal fold lesions, paralyses, and other pathological conditions. Many of these instabilities are related to the intrinsic nonlinearities in the vibrations of the vocal folds.

In this paper, bifurcations and chaos in voice signals are analyzed using narrow-band spectrograms. We study sustained phonation of patients with laryngeal paralysis and data from an excised larynx experiment. These spectrograms are compared with computer simulations of an asymmetric 2-mass model of the vocal folds.

Introduction

The vocal folds, together with glottal airflow, constitute a highly nonlinear self-oscillating system. According to the accepted myoelastic theory of voice production, the vocal folds are set into vibration by the combined effect of subglottal pressure, the visco-elastic properties of the folds, and the Bernoulli effect [1,2]. The effective length, mass and tension of the vocal folds are determined by muscle action, and in this way the fundamental frequency ("pitch") and the waveform of the glottal pulses can be controlled. The vocal tract acts as a filter which transforms the primary signals into meaningful voiced speech [3].

Normal voiced sound appears to be nearly periodic, although small perturbations (in the order of a percent) are important for the naturalness of speech. Under certain circumstances, however, much larger irregularities are observed in vocalizations which are often associated with the term "roughness". In the phonetic literature phenomena such as "octave jumps", "biphonation" (two independent pitches), and "noise concentrations" have been reported for decades [4,5]. However, only within the past few years has it been suggested that these observations might be interpreted as period-doubling bifurcations, tori, and chaos, respectively [6-10]. It has been shown using Poincare' sections and estimations of attractor dimension and Lyapunov exponents that newborn cries [6] and vocal disorders [7-10] are intimately related to bifurcations and low-dimensional attractors.

However, even sustained vowels are not really stationary and, hence, estimations of attractor dimensions appear to be problematic in many cases. Therefore, in this paper we focus on spectrograms, allowing the detection of bifurcations in systems with slowly varying parameters [4, 5, 11, 6, 12]. Spectrograms are based on many subsequent short-time power spectra of overlapping segments. The spectral amplitude is encoded as a grey scale and, in this way, the (dis)appearance of spectral peaks due to bifurcations of the underlying dynamical system can be monitored. For example, period-doubling is reflected in additional peaks in the middle of the harmonics of the original pitch.

In the following sections we present spectrograms from patients with laryngeal paralysis, from an excised larynx experiment, and from computer simulations. Special

attention is devoted to bifurcations due to tension imbalance of the left and right vocal fold which is relevant for unilateral laryngeal nerve paralyses.

Voice Disorders

On the one hand, subharmonics and irregularities appear occasionally in normal vocalizations (newborn cries [6], "vocal fry" phonation [9], and speech [13]. On the other hand, vocal instabilities due to bifurcations are often symptomatic in voice pathology [7-10]. Laryngeal stroboscopy reveals that various voice disorders lead to irregular vibratory patterns of the vocal folds resulting in a rough voice quality [14]. The corresponding acoustic signals often show sudden jumps to subharmonic regimes (period-doubling or -tripling) and low-frequency modulations (tori) [7, 8, 10]. It is the aim of current research to understand the underlying physiological mechanisms of these bifurcations with the aid of high-speed films, excised larynx experiments [15], and biomechanical modeling [16, 17].

In the following we focus on a specific disorder: unilateral laryngeal nerve paralysis. In this case the rest position, the stiffness, and the effective mass of the affected fold may deviate drastically from normality. Often paralysis leads to turbulent fricative noise due to incomplete closure of the folds corresponding to a "breathy voice" [14]. Frequently, subharmonics and low-frequency modulations have also been observed [18, 19, 10, 20]. Here, the left-right asymmetry is presumably the origin of the instabilities [18, 21, 17].

In the remainder of the paper we study left-right asymmetry from three perspectives: we analyze sustained vowels from patients with unilateral paralyses, discuss excised larynx experiments with artificially induced asymmetry, and present simulations of an asymmetric 2-mass model of the vocal folds.

Fig. 1 shows a spectrogram of a sustained vowel from a male patient with unilateral recurrent nerve paralysis. The fundamental frequency F_0 of about 150 Hz and its harmonics at 300 Hz and 450 Hz appear as dark horizontal lines. The audible rough voice quality results from the strong subharmonic components. Particularly, period-tripling, i.e. subharmonics at 1/3 F_0, 2/3 F_0,...occur. Moreover, around 1700 ms another subharmonic regime is found: there are subharmonics at one fourth of the pitch F_0. Comparable transitions between subharmonic regimes will be discussed later in connection with our 2-mass model.

In Fig. 2 a sustained vowel of another patient with recurrent nerve paralysis is shown. His voice is characterized by low-frequency modulations typically between 20 and 30 Hz. The modulations appear in the spectrogram as side-bands of the fundamental frequency (≈ 20 Hz) and its harmonics. These toroidal oscillations result in a very rough

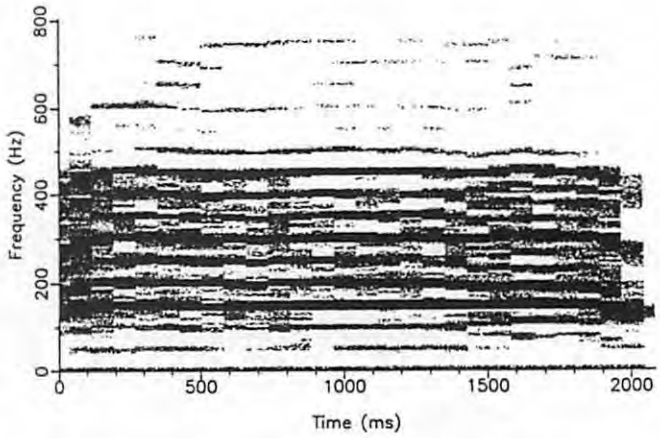


Figure 1. Computer spectrogram of the vowel [i:] from a patient with paralysis showing subharmonics. In order to get sufficient spectral resolution to resolve subharmonics and side-bands, segments of 4096 points (about 200 ms) have been used in all spectrograms.

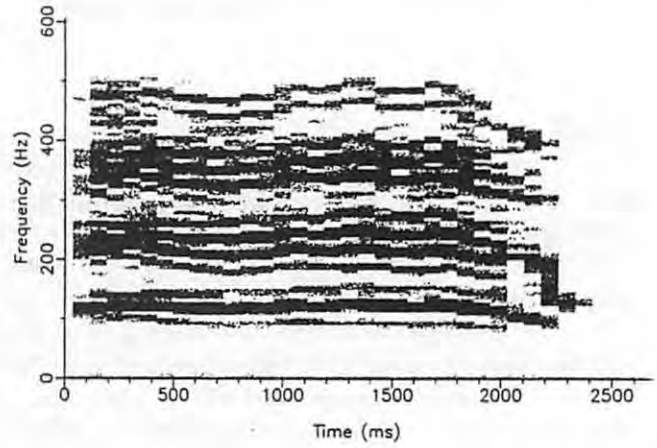


Figure 2. Spectrogram (vowel [e:], recurrent nerve paralysis) displaying low-frequency modulations visible as side-bands.

voice quality. Moreover, the voice sounds breathy due to turbulent air flow.In this case,two sources of hoarseness--complicated vibratory patterns of the folds and turbulence—superimpose on each other.

Excised Larynx Experiments

Experiments with human or animal larynges serve as a link between the human voice source in vivo and computer models [1, 22, 18]. They allow controlled and systematic parameter variations and easy observation of vibratory patterns.

We have examined three larynges from large (about 25 kg) mongrel dogs coming from coronary research units at The University of Iowa. The dissected larynges were

mounted on an apparatus described in detail elsewhere [22, 15]. Heated and humidified air was supplied from below as the driving force of the oscillations. The device was attached to several micrometers to control the adduction and the elongation of the vocal folds. To facilitate observation of vocal fold movement, a strobe light with adjustable frequency was placed above the glottis. The data were recorded on a color video system and afterwards digitized with 16 bit resolution and a sampling rate of 20 kHz.

In our experiments instabilities have been studied for varying subglottal pressure and for asymmetric adduction and elongation of the vocal folds. 2-parameter bifurcation diagrams can be found elsewhere [15]. Here we only summarize briefly the various dynamic regimes which have been observed for overcritical asymmetry and pressure:

* symmetric periodic phonation in head-like and chest-like registers
* periodic vibrations of the lax fold only
* subharmonics, modulations, and irregular vibrations with both folds involved
* whistle-like sound
* aphonia, i.e. vibrations ceased for very strong asymmetric tension

Typically, the parameter ranges of these regimes overlap, i.e. hysteresis is observed. Sometimes spontaneous transitions between different dynamical regimes appeared without external parameter changes. The spectrogram in Figure 3 shows such transitions from normal phonation to intermittent regimes with a low-frequency modulation of about 20 Hz, as evidenced by the spacing of the sidebands about the harmonics. These modulations have been found due to lengthening of one of the folds. Note the qualitative similarity to the sustained vowel shown in Fig. 2.

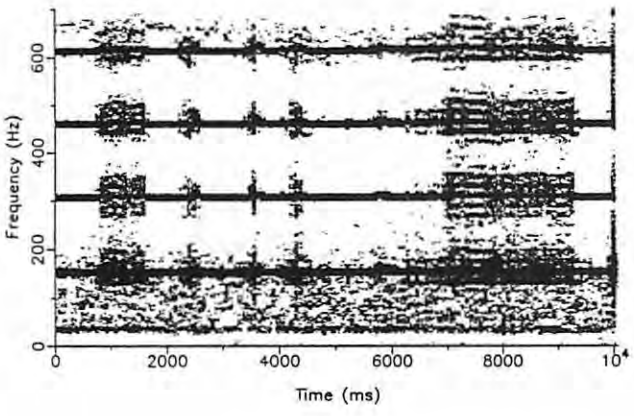


Figure 3. Spectrogram from an excised larynx experiment with asymmetric elongation of the folds. Around 1000 and 8000 ms episodes with low-frequency modulations of about 20 Hz can be seen.

Two-Mass Model

Computer models of speech production are valuable to understand the basic mechanisms of normal and pathological phonation. There are conceptionally simple 2-mass models [23, 18, 21, 9, 17] and sophisticated continuum models simulating the visco-elastic equations [24, 8, 16, 2]. In this section we discuss a simplified version of the intensively studied Ishizaka-Flanagan model [23]. For such relatively simple models extensive bifurcation analysis can be carried out [25, 9, 17].

In our model each fold is represented by two oscillators (defined by mass m, stiffness k, damping constant r which are coupled by a spring with stiffness k_c (see Fig. 4). This realization enables the "vocal folds" to oscillate with a phase difference between the lower and the upper part. Another restoring force with stiffness c acts during collision of the left and the right vocal fold, i.e. if areas a_i become negative. We can describe the elongation of each of the four masses by Newton's law:

$$m_{i\alpha}\ddot{x}_{i\alpha} + r_{i\alpha}\dot{x}_{i\alpha} + k_{i\alpha}x_{i\alpha} + \Theta(-a_i)c_{i\alpha}\left(\frac{a_i}{2l}\right) + k_{c\alpha}(x_{i\alpha} - x_{j\alpha}) = F_i(x_{1l}, x_{1r}, x_{2l}, x_{2r}) \quad (1)$$

$$a_{i\alpha} = a_{i0\alpha} + l x_{i\alpha}$$
$$a_i = a_{il} + a_{ir} \quad (2)$$

$a_{i0\alpha}$ – glottal rest area
l – length of the glottis

$$i, j = \begin{cases} 1 & \text{– lower mass} \\ 2 & \text{– upper mass} \end{cases}$$

$$\alpha = \begin{cases} l & \text{– left side} \\ r & \text{– right side} \end{cases}$$

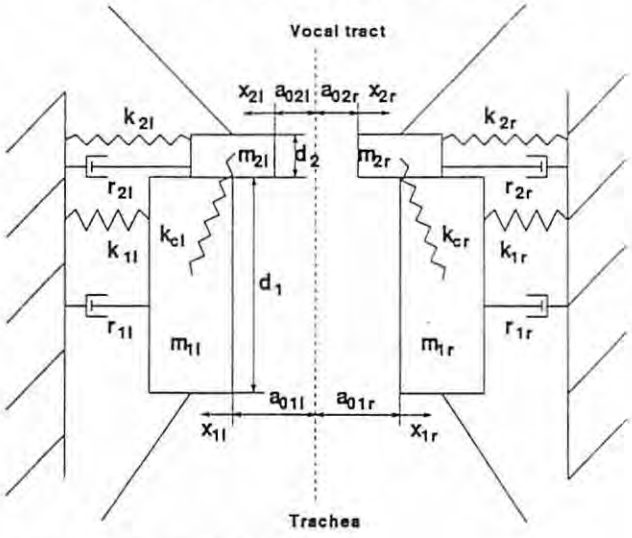


Figure 4. 2-mass model of vocal folds.

where $F_i(x_{1l}, x_{1r}, x_{2l}, x_{2r})$ is the force exerted by the pressure P_i on the corresponding part of the glottis.

$$F_i = l d_i P_i \tag{3}$$

d_i - thickness of part i

The air flow is considered to be laminar and is described by the Bernoulli equation. After passing the smallest gap in the vertical direction, a jet is built up and the pressure equals that of the supraglottal system which is set to zero [17, 2].

$$P_s = P_1 + \frac{\rho}{2}\left(\frac{U}{a_1}\right)^2 = \frac{\rho}{2}\left(\frac{U}{a_{min}}\right)^2$$
$$a_{min} = \min(a_{1l}, a_{2l}) + \min(a_{1r}, a_{2r}) \tag{4}$$

P_s - subglottal pressure
U - volume flow velocity
ρ - density of air

According to these assumptions we get the following pressure equations:

$$P_1 = P_s\left(1 - \Theta(a_{min})\left(\frac{a_{min}}{a_1}\right)^2\right)\Theta(a_1) \tag{5}$$

$$P_2 = 0 \tag{6}$$

$$U = \sqrt{\frac{2P_s}{\rho}}\, a_{min} \tag{7}$$

It is emphasized that in this manner the upper masses are only driven by the lower ones.

Following Ishizaka and Flanagan we choose the following parameters to model a normal voice.

$m_{1l}=m_{1r}=0.125$
$m_{2l}=m_{2r}=0.025$

$k_{1l}=k_{1r}=0.08$
$k_{2l}=k_{2r}=0.008$
$k_{cl}=k_{cr}=0.025$

$c_{1l}=c_{1r}=3_{k1}$
$c_{2l}=c_{2r}=3_{k2}$

$r_{1l}=r_{1r}=0.02$
$r_{2l}=r_{2r}=0.02$

$d_1=0.25$ $\quad a_{01}=a_{01l}+a_{01r}=0.05$
$d_2=0.05$ $\quad a_{02}=a_{02l}+a_{02r}=0.05$

$P_s=0.008$ (=8cm H_2O)
$\rho=0.00113$

All units are given in cm, g, ms and their corresponding combinations.

In the above model several aspects have been neglected: vertical motion, incompressibility and nonlinear tissue properties of the folds. Moreover, the vocal fold vibrations were separated from the vocal tract by assuming constant pressures below and above the glottis. Nevertheless, several features of the human voice source such as the waveforms of glottal pulses and the phonation onset can be captured quite realistically [9, 17]. However, we are aware that details such as polyps cannot be treated by 2-mass models.

Modeling Asymmetry

In this section the effect of asymmetric mechanical properties of the folds is studied as a first approach to unilateral paralysis. Following earlier studies [18, 21, 17] we introduce an asymmetry parameter Q:

$$k_{ir} = Q k_{il}$$
$$m_{ir} = m_{il}/Q \tag{8}$$
$$0 < Q < 1$$

In this way the eigenfrequency of the affected fold is reduced by the factor Q. Instabilities have been found for $Q < 0.6$ and subglottal pressure above $P_s = 0.013$ [17]. Typically, at the borderline of normal phonation abrupt jumps to subharmonic regimes are observed [17]. Fig. 5 shows a "spectral bifurcation diagram" for slowly decreasing Q. It can be seen that the pitch decreases with the ratio of the eigenfrequencies Q. Around 1000 ms subharmonics suddenly appear at one-half of the pitch. Around 2000 ms

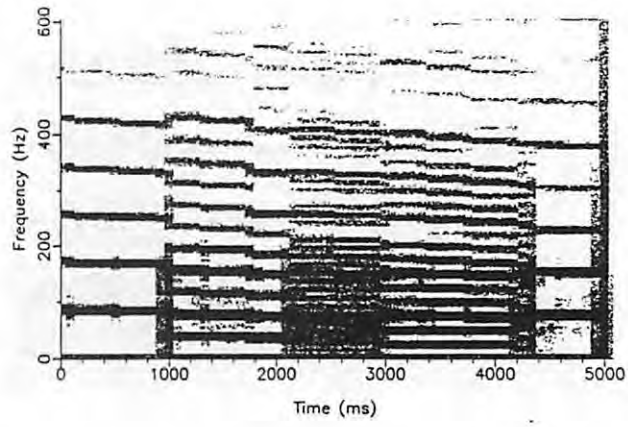


Figure 5. Spectrogram from the asymmetric 2-mass model (calculated from the flow velocity U). The asymmetry parameter Q was reduced every 400 ms by 0.005 from 0.56 to 0.50. In this way jumps between different subharmonic regimes including period-doubling and -tripling are induced.

another complicated subharmonic regime is reached. Inspection of the peaks of x_{lr} and x_{ll} reveals that during one cycle (about 5 times the original pitch period) five maxima of x_{lr} and eight maxima of x_{ll} occur, i. e. we can interpret this regime as a 5:8 resonance. At 3000 ms subharmonics of one third of the original pitch dominate corresponding to a 3:5 resonance in the above sense. The sequence of decreasing ratios from 2:3 to 5:8 to 3:5 is reminiscent of "Arnold tongues" in bifurcation diagrams of coupled oscillators.

Our simulations show that for sufficiently large pressures, which couple the left and right folds, the detuning of the eigenfrequencies induces transitions to subharmonic regimes comparable to observations as in Fig. 1. Comprehensive 2-parameter bifurcation diagrams also reveal the appearance of toroidal and chaotic oscillations [17].

Despite the qualitative resemblance of our simulations to observations, we have to keep in mind that a 2-mass model is only a crude approximation of the real vocal folds. However, in simulations of a three-dimensional model based on partial differential equations, similar bifurcations to subharmonic regimes and chaos have been found [8, 16]. Moreover, the calculation of empirical orthogonal functions from the continuum model reveal that the dynamics is often governed by only a few dominant modes. This can be regarded as a justification to study specific aspects of vocal fold vibrations such as left-right asymmetry with appropriate low-order models.

Summary and Discussion

Analysis of acoustic signals from newborn cries [6], voice disorders, and excised larynx studies reveals clear evidence of bifurcations and low-dimensional attractors [6-10]. Although the limited stationarity of such data allows no definite proof of chaos, nonlinear phenomena are plausible since the human voice source exhibits several essential nonlinearities:

* nonlinear stress-strain characteristics of vocal fold tissue [8, 2]
* highly nonlinear relation between pressure and glottal area (cf. Eq. (5))
* collision of the vocal folds
* vortices and jet instabilities

Simulations of the 2-mass model have shown that various bifurcations appear due to the desynchronization of the right and left fold for overcritical asymmetry. The corresponding instabilities are similar to observations of vocalizations of patients with unilateral paralysis. However, a more quantitative comparison would require the use of more sophisticated models and more detailed measurements of the biomechanical properties of vocal fold tissues, especially in the case of pathologies.

In forthcoming studies we plan to analyze the physiological mechanisms of vocal instabilities in detail using the continuum model [8, 16, 2], excised larynx experiments, and high-speed films of vocal fold vibrations.

Acknowledgements

Partial funding for this research was provided by the Deutsche Forschungsgemeinschaft and Grant No. P60 DC00976 from the National Institutes on Deafness and Other Communication Disorders. We thank J. Wendler and M. Cebulla for providing data from patients with paralyses.

References

[1] J. van den Berg, "Myoelastic-aerodynamic theory of voice production," J. Speech Hearing Res. 1, 227-244 (1958).

[2] Titze, I.R., and F. Alipour-Haghighi, Myoelastic Aerodynamic Theory of Phonation, (forthcoming book).

[3] G. Fant, Acoustic theory of speech production, (Mouton, The Hague, 1960).

[4] J. Lind (Ed.), Newborn infant cry (Almquist and Wiksells Boktrycken, Uppsala, 1965).

[5] A.W. Kelman, "Vibratory pattern of the vocal folds," Folia phoniat. 33, 73-99 (1981).

[6] W. Mende, H. Herzel, and K. Wermke. "Bifurcations and chaos in newborn cries," Phys. Lett. A 145, 418-424 (1990).

[7] H. Herzel and J. Wendler, "Evidence of chaos in phonatory samples," in EUROSPEECH, (ESCA, Genova, 1991), pp.263-266.

[8] I.R. Titze, R. Baken, and H. Herzel, "Evidence of Chaos in Vocal Fold Vibration," in Vocal Fold Physiology: Frontiers in Basic Science edited by I.R. Titze (Singular Publishing Group, San Diego 1993), pp.143-188.

[9] H. Herzel, "Bifurcations and chaos in voice signals," Appl. Mech. Rev. 46, 399- 413 (1993).

[10] Herzel, H., D.A. Berry, I.R. Titze, and M. Saleh, "Analysis of Vocal Disorders with Methods from Nonlinear Dynamics," J. Speech Hearing Res., in press.

[11] W. Lauterborn and E. Cramer, "Subharmonic routes to chaos observed in acoustics," Phys. Rev. Lett. 47, 1445-1448 (1981).

[12] M.A. Liauw, K. Koblitz, N.I. Jaeger, and P. Plath, "Periodic perturbation of a drifting heterogeneous catalytic system," J. Phys. Chem. 97, 11724-11730 (1993).

[13] L. Dolansky and P. Tjernlund, "On certain irregularities of voiced-speech waveforms," IEEE Trans. AU-16, 51-56 (1968).

[14] M. Hirano, "Objective evaluation of the human voice: clinical aspects," Folia phoniat. 41, 89-144 (1989).

[15] D. Berry, H. Herzel, and I.R. Titze, "Bifurcations in excised larynx experiments due to asymmetric vocal folds," in preparation.

[16] Berry, D.A., H. Herzel, I.R. Titze, and K. Krischer, "Interpretation of Biomechanical Simulations of Normal and Chaotic Vocal Fold Oscillations with Empirical Eigenfunctions," J. Acoust. Soc. Am., in press.

[17] I. Steinecke and H. Herzel, "Bifurcations in an asymmetric vocal fold model," J. Acoust. Soc. Am., submitted.

[18] K. Ishizaka and N. Isshiki, "Computer simulation of pathological vocal-cord vibrations," J. Acoust. Soc. Am. 60, 1194-1198 (1976).

[19] B. Hammarberg, B. Fritzell, J. Gauffin, and J. Sundberg, " Acoustic and perceptual analysis of vocal dysfunction," J. of Phonetics 14, 533-547 (1986).

[20] M. Ptok, G. Sesterhenn, and R. Arold, "Bewertung der laryngealen Klanggeneration mit der FFT-Analyse der glottischen Impedanz bei Patienten mit Rekurrensparese," Folia phoniat. 45, 182-197 (1993).

[21] M.E. Smith, G.S. Berke, B.R. Gerrat, and J. Kreimann, "Laryngeal paralyses: theoretical considerations and effects on laryngeal vibration," J.Speech Hear. Res. 35, 545-554 (1992).

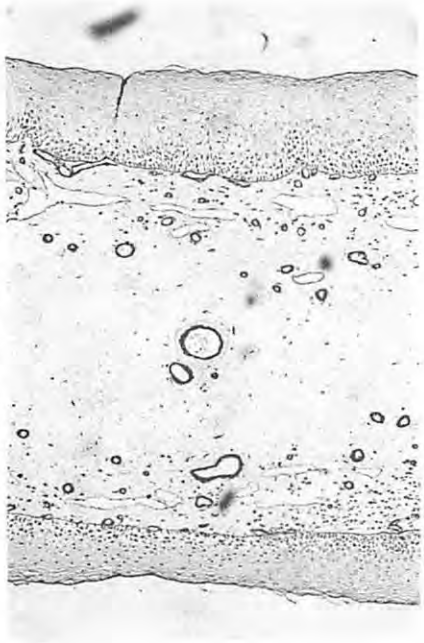
[22] T. Baer, "Investigation of the phonatory mechanism, " ASHA Reports 11, 38-47 (1981).

[23] K. Ishizaka and J.L. Flanagan, "Synthesis of voiced sounds from a two-mass model of the vocal cords," Bell Syst. Techn. J. 51, 1233-1268 (1972).

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

[25] H. Herzel, I. Steinecke, W. Mende, and K. Wermke, "Chaos and bifurcations during voiced speech," in Complexity, Chaos and Biological Evolution, edited by E. Mosekilde and L. Mosekilde (Plenum Press, New York, 1991), pp. 41-50.

Photo 1. Normal human vocal fold (25X) stained for collagen type 4 (arrows). Note the thin brown staining of the basement membrane around the arteriole and capillary structures which serves as a control to verify the accuracy of the staining technique. The basement membrane zone of the epidermis is represented as a very thin brown line. The lamina propria appears amorphous and non-descript.



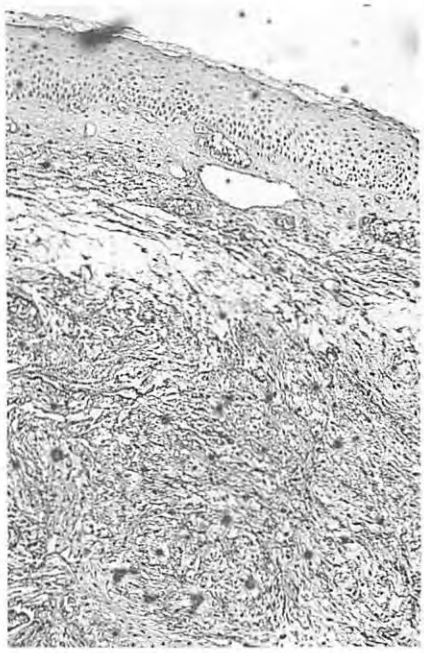
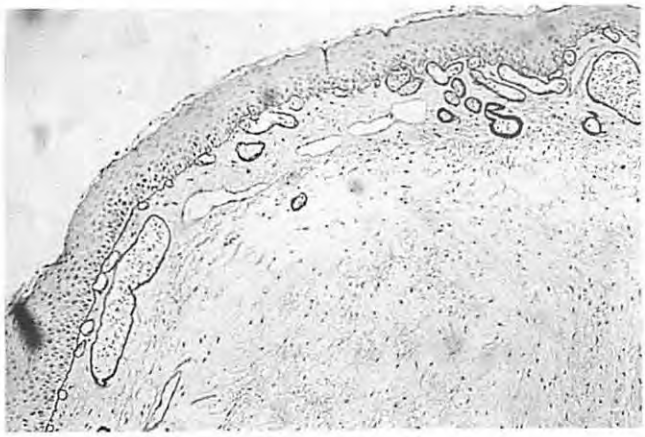
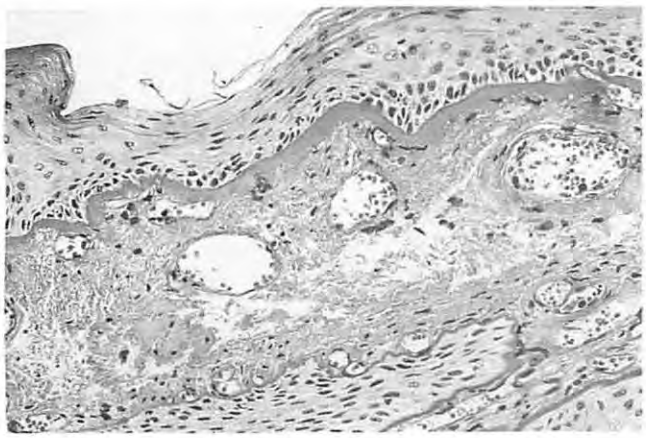


Photo 2. Human vocal fold lesion described as nodule showing fibronectin deposition. A - above) H&E stain at 25X. Note lack of lamina propria detail. B - far left) Fibronectin stain at 10X. Notice the intense brown stain throughout the lamina propria. (Arrows show region of intense fibronectin, compare with Photo 3, 4.) Note no fibronectin basement membrane involvement in this lesion. C - left) Collagen type 4 (arrows) stain at 10X. Normal basement membrane zone confirmed.





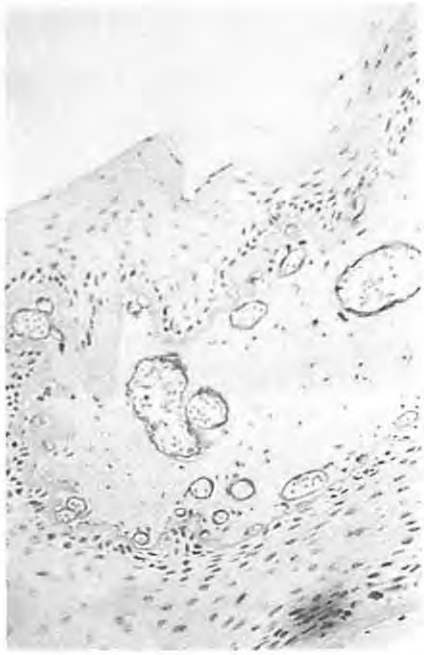
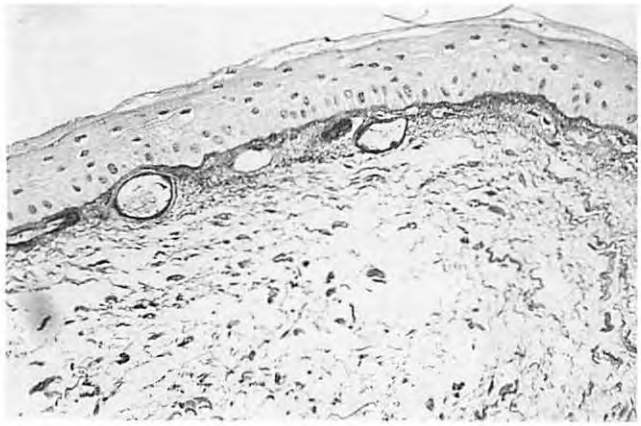


Photo 3. Another vocal fold lesion described as nodule showing both fibronectin and collagen type 4 deposition. A - above) Fibronectin stain at 25X. The basement membrane zone (arrows) is abundant with fibronectin. The lamina propria, although partially surgically crushed, displays much fibronectin (circles) throughout. Compare with the polyp in Photo 4. B - right) Collagen type 4 stain at 25X. This stain is interesting in two regards. The basement membrane zone (small arrows) stains broadly, but weakly. The weak staining probably indicates incomplete assembly of the protein. The broad band indicates extensive basement membrane zone disorganization and injury.[13] Large arrow indicates normal basement membrane zone around arteriole.



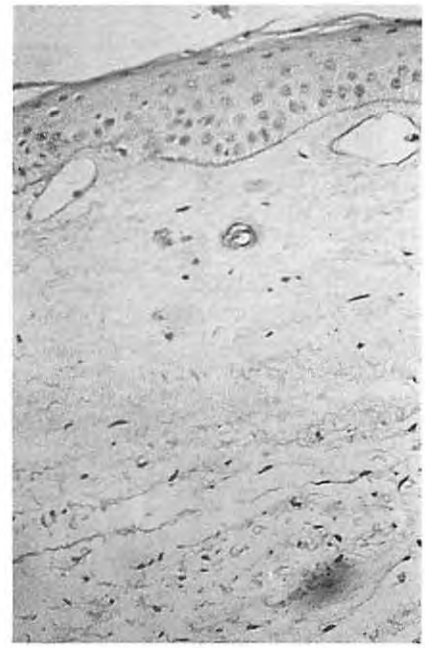
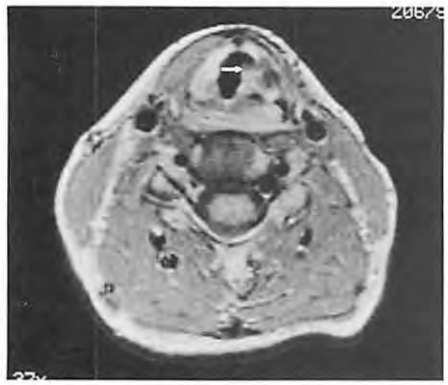


Photo 4. Human vocal fold lesion described as polyp. A - above) Fibronectin stain at 25X shows very little fibronectin in the lamina propria (arrows surround region of sparse fibronectin). Compare to Photos 2-3. Fibronectin is deposited normally in the basement membrane zone however. B - right) Collagen type 4 stain at 25X appears normal, with no increased thickness (arrows indicate collagen type 4 band).

Photo 5. Axial section shows left soft-tissue fullness on the supraglottic surface of the vocal fold (A - right: white arrow) just one cut superior to the implant. The coronal section (B - far right) demonstrates that the implant is in the proper vertical plane (black arrow-head) just lateral to the left vocal fold.



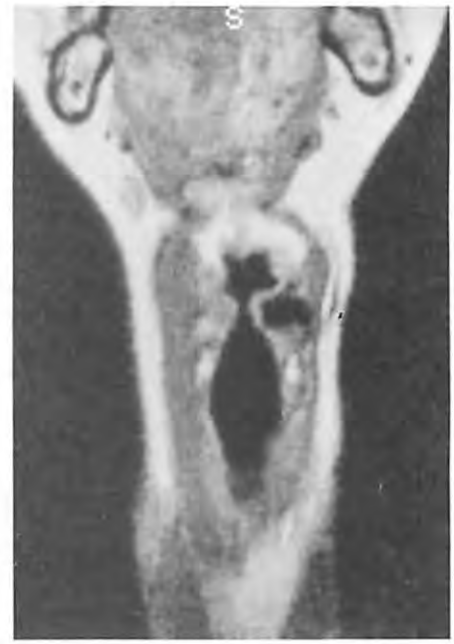
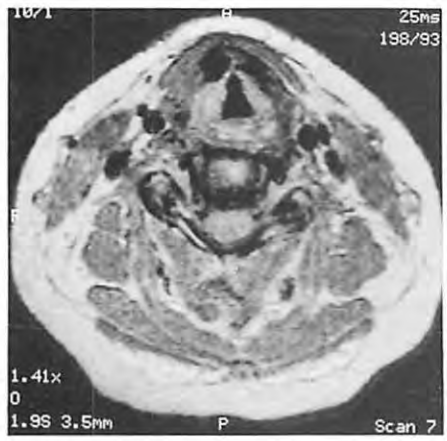


Photo 6 (right). (Axial section) Anterior placement of silastic shims was deliberate to avoid impinging the vocal process in this patient with mobile vocal folds, but implants were placed too far anteriorly to produce effective medialization of the membranous vocal fold.



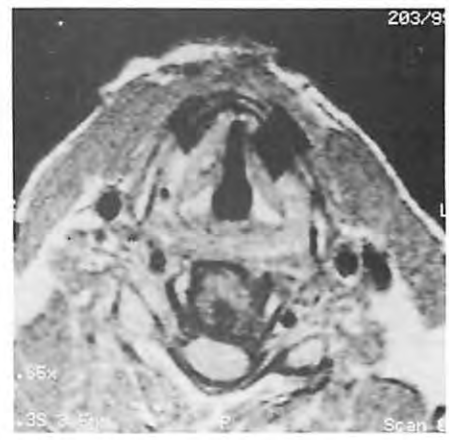
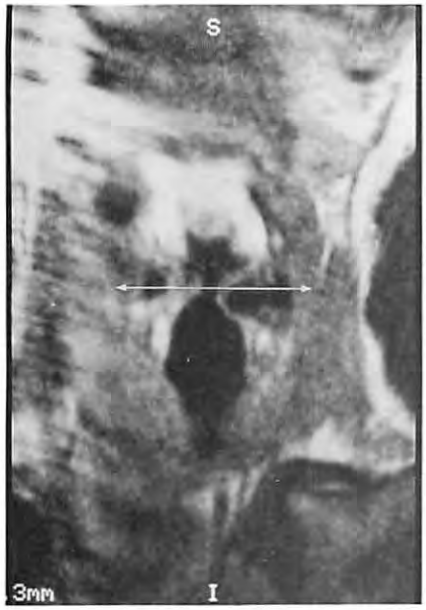
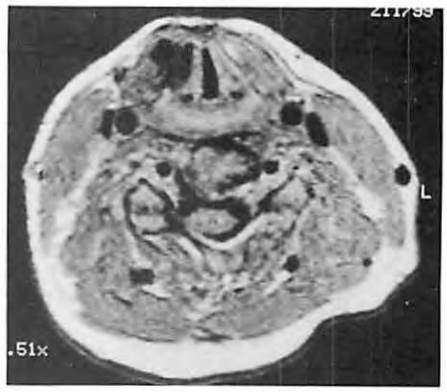
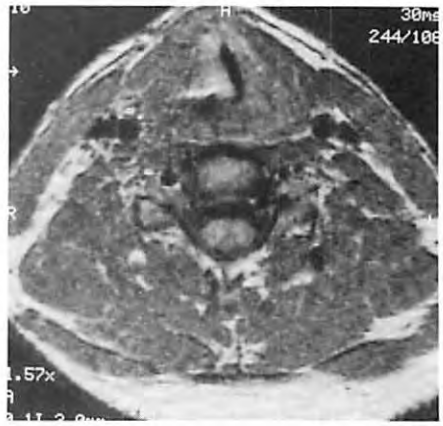
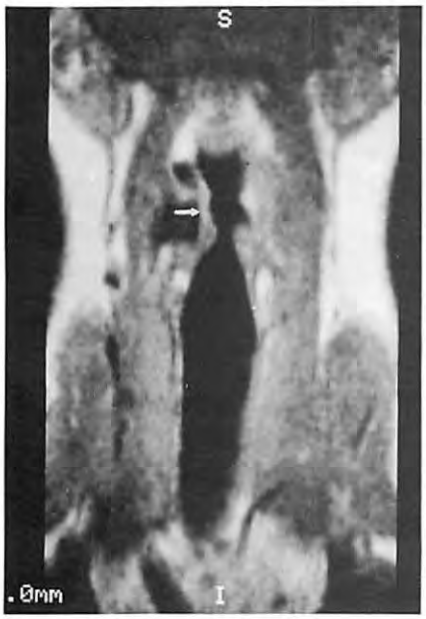


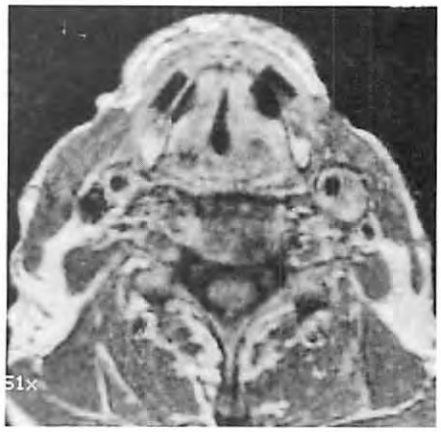
Photo 7 (right). Axial section shows bilateral implants displace the vocal folds but the left implant is too prominent anteriorly (white arrow) and the medial edge does not conform to the defect contour of a bowed vocal fold (A - right). Photo 7B (following page; upper left). The coronal view shows that displacement is on plane (white arrows depict glottic plane) but the left implant has too great a vertical dimension producing subglottic fullness.











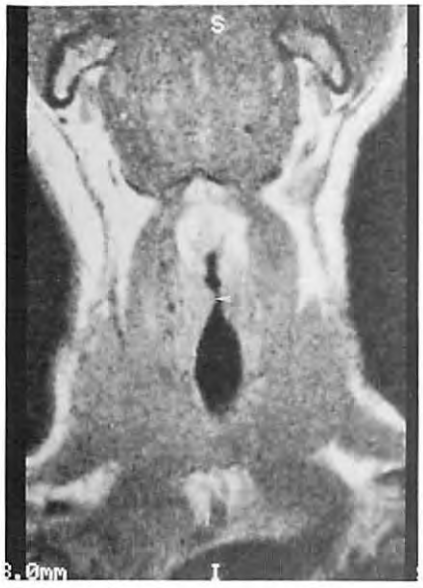
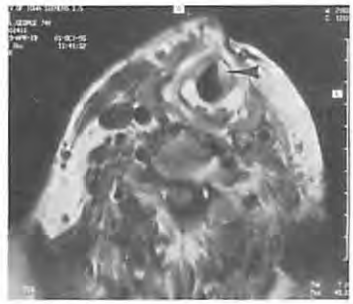


Photo 8. The axial section (A - top right) shows a sharply tapering medial edge of the implant that compensates for the anterior overlap of the implant (right hemilarynx). A more posteriorly placed window would be preferable to accommodate an anterior overlapping design. The coronal section (B - middle left) shows that even a slight superior overlap causes undesirable supraglottic displacement with obliteration of the ventricle (white arrow). Photo 9. In the axial plane (A - middle right), corrective over-closure of the posterior glottis is evident (right side). In the coronal plane (B - left), closure appears to be in the plane of the contralateral vocal fold (white arrowhead points to medialized true vocal fold). Photo 10 (above). The axial view shows the left implant designed to compensate for the laterally-placed vocal fold. Both implants are convex medially to compensate for vocal fold bowing.



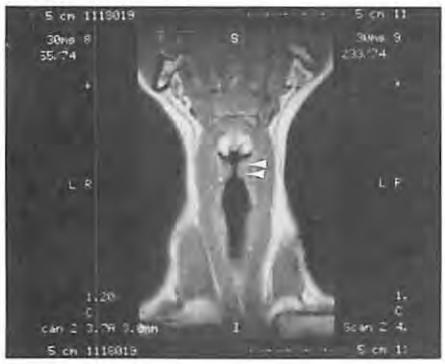
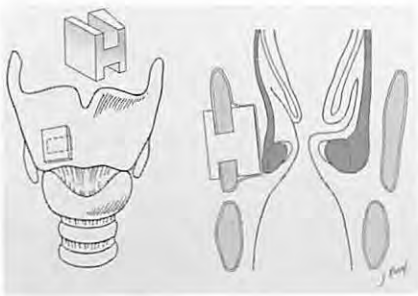
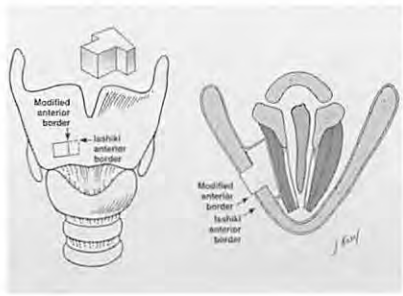


Photo 11A-C. (Axial sections).(A - upper left) The convex displacement of the vocal fold edge reflects the spherical shape assumed by the injected autogenous fat (black arrowhead). Medialization is minimal at the posterior glottis. (Photo 11B - above) This recently Teflon-injected larynx reveals medialization of the left true vocal fold but the implant (white arrow) is difficult to distinguish from the surrounding tissues. (Photo 11C - left) GAX-injected vocal fold shows foci of injectate filling vocal fold superior and inferior (black arrowheads) to the fibrous sulcus vocalis (white arrow). This demonstrates the difficulty in medializing scarred vocal fold segments by injection, without adjunctive surgical measures.





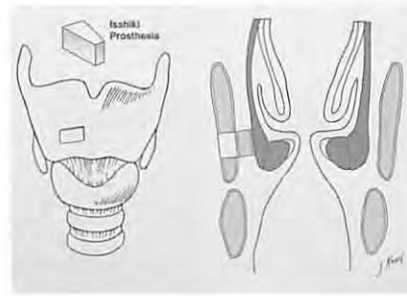
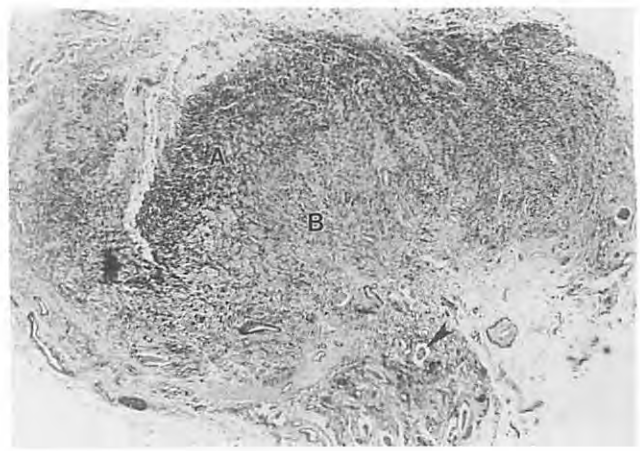


Photo 12 (left). Diagrammatic representation of the need for lower window placement when the prosthesis is designed to interlock the thyroid cartilage by overlapping the inner cartilage window surface in the vertical dimension. Photo 13 (middle). Diagrammatic representation of the need for more posterior window placement when the prosthesis is designed to interlock the thyroid cartilage by overlapping the inner cartilage window surface anteriorly. Photo 14 (right). Diagrammatic representation of the window placement as described by Isshiki. Note the wedge-shaped silastic shim is designed to fit flush without overlap and the cartilage filling the window is preserved.

Photo 15 (right). Coronal section of a dissected human vocal ligament stained with Weigert stain to visualize the elastin fibers in black, the collagen fibers in red, and the muscle fibers in light brown. A- Intermediate layer of the ligament with many elastin fibers. B- deep layer with many collagen fibers. Arrow head- muscle fibers present in border of the deep layer. No epithelium or superficial layers present. (This ligament was not one of the sample measured in the study since this practice specimen had no tissue damage due to stress-strain process to better appreciate the different tissue layers.)



Part II

Tutorial reports

NCVS Status and Progress Report - 7
December 1994, 155-170

Voice

Marshall Smith, M.D.
Department of Otolaryngology-Head and Neck Surgery, The University of Colorado Health Sciences Center
Wilbur James Gould Voice Research Center, The Denver Center For The Performing Arts
Steven Gray, M.D.
Division of Otolaryngology-Head and Neck Surgery, The University of Utah Medical Center

The voice is one of the first signs of life. Then, throughout life the voice is a primary means of expression and communication. It is an indicator of health, sickness, emotion, and age. The voice may provide means to earn a living. It may convey great artistic expression through skillful use. The voice has lifelong importance to normal oral communication and social well being. "Voice" may be defined in a broad (synonymous with speech) or narrow sense (Titze, 1994). This chapter focuses on "voice" in a narrow way, particularly pertaining to voice production. Other terms used in this context include "vocalization" - the sound produced by vocal fold vibration, and "phonation" - the physical and physiological processes of vocal fold vibration. Phonation is a key component of speech production, in conjunction with the functions of respiration, articulation, and resonance. Developmental anatomy and physiology of the larynx as it pertains to phonation is covered. The chapter then reviews the topic of "voice" clinically; disorders of phonation (or "dysphonias") seen in the pediatric age group, and their evaluation. Treatment of pediatric voice disorders, particularly medical and surgical approaches are discussed. These topics necessarily overlap to a great degree with other chapters on laryngeal diseases throughout this book for which any voice problem may be symptom, as well as chapters on "Disorders of Phonology, Phonation, and Fluency", "Early Identification of Speech and Language Disorders", and "Velopharyngeal Insufficiency". All these have relevance to the voice. The reader is referred to this material when appropriate.

Developmental Laryngeal/Phonatory Anatomy

1. Endolaryngeal/Histologic studies

The elegant studies of Hirano et al (1983), provide much insight into the development of the phonatory larynx. They studied 88 normal larynges in patients whose ages ranged from a few hours after birth to 69 years old. Several gross anatomic and histologic variables were studied. The length of the entire vocal fold was measured, and the length of the membranous portion (anterior fold) and cartilaginous portion (posterior fold including vocal process and arytenoid). It was found that, up to the age of 10 years, the length of the vocal fold did not vary much between males and females. At ten years of age the length of the membranous portion of the vocal fold gradually increases for males, up to 20 years of age. At age 10, the membranous fold is 6 to 8 mm long in males and females. In females the vocal fold will increase in length to 8.5 to 12 mm by age 20. In males the length will more than double, to 14.5 to 18 mm by age 20. The length increase was not dramatic during adolescence, but gradual (Figure 1 - next page). The study by Kahane (1978) on morphology of the prepubertal and pubertal larynx also documented the changes in vocal fold length with puberty. Measuring the entire vocal fold length (membranous and cartilaginous portions) before and after puberty, the average increase in female vocal folds was 4.2 mm and males was 10.9 mm, over twice as much.

The cartilaginous portion of the vocal fold also grows with age (Hirano et al, 1983). It increases from about 1.25 mm in newborns to 3 mm in males and 2.5 mm in females. If a ratio is made of the length of the membranous portion to that of the cartilaginous portion of the vocal fold, the ratio is about 1.5 in newborns (see Figure 2 - next page). It increases to about 4.0 in adult females and 5.5 in adult males. This ratio has value in understanding the functions of the larynx in children. In children, a larger portion of the glottis comprises the posterior glottis. This has been termed by Hirano et al (1986) the "respiratory glottis". Indeed, respiratory and protective functions of the larynx play a

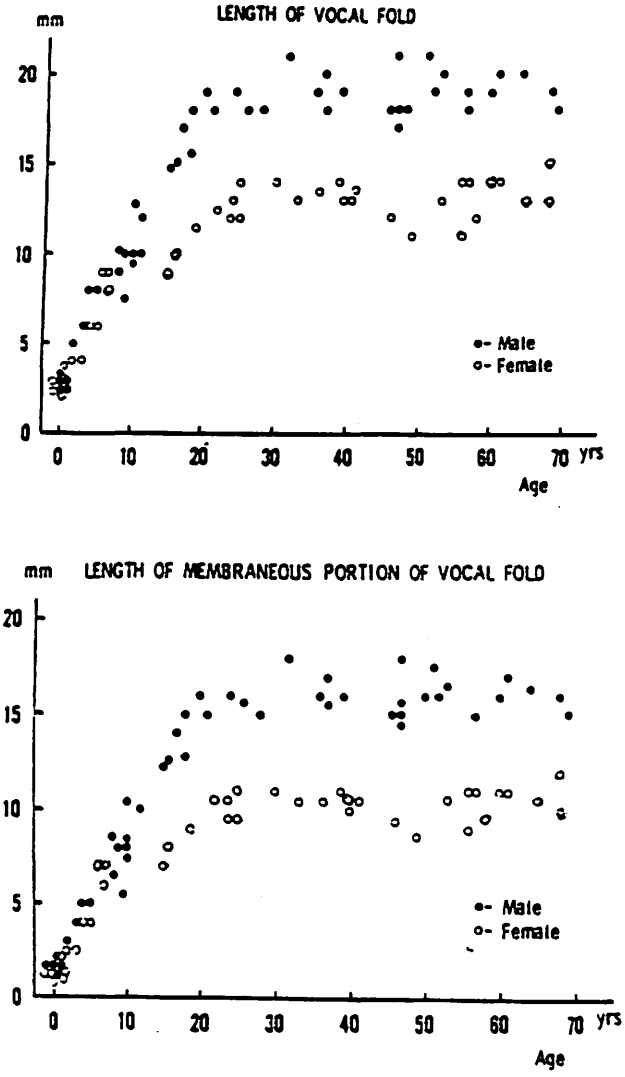


Figure 1. (A-top) Length of the entire vocal fold measured in millimeters for 48 males and 40 females ranging from a few hours to 60 years of age. (B-bottom) Length of the membranous portion of the vocal fold measured in millimeters for 48 males and 40 females ranging from a few hours to 60 years of age (From Hirano M, Kurita S, Nakashima T: Growth, development, and aging of the human vocal folds, in Bless DM, Abbs JH (eds.): Vocal Fold Physiology: Contemporary Research and Clinical Issues, San Diego, College-Hill Press, 1983, pg 26. Used by permission.)

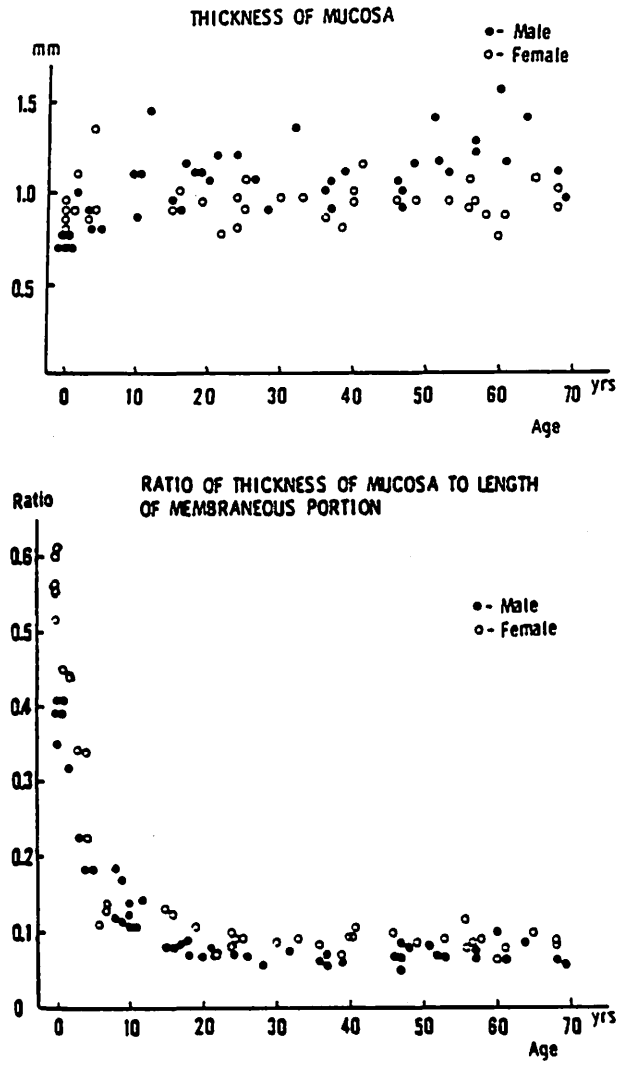


Figure 2. (A-top) Thickness of the mucosa at the midpoint of the membranous portion for 48 males and 40 females ranging in age from a few hours to 69 years. (B-bottom) ratio of the thickness of mucosa to the length of the membranous portion for 48 males and 40 females ranging in age from a few hours to 69 years. (From Hirano M, Kurita S, Nakashima T: Growth, development, and aging of the human vocal folds, in Bless DM, Abbs JH (eds.):Vocal Fold Physiology: Contemporary Research and Clinical Issues,San Diego, College-HillPress,1983, pg 30.Used by permission.)

larger role than phonation in infants and children. The membranous portion of the vocal folds is more susceptible to edema than adults, yet because the membranous folds (the anterior or "phonatory glottis") comprises a smaller percentage of the entire glottal area these obstructive effects are minimized, serving as a relative protection.

Hirano also studied the histology of the vocal folds in the developing larynx. He has reported extensively on the layered structure of the vocal fold (Hirano & Kakita, 1985). It is described as having five distinct layers. The first two layers, the vocal fold epithelium and superficial layer of the lamina propria (also known as Reinke's space) comprise the "cover". Underneath this is found the intermediate and deep layers of the lamina propria (or vocal ligament) and the thyroarytenoid or vocalis muscle. The deeper layers are called the "body". The complex stiffness interaction between the cover and body facilitates phonation through its range of pitches, loudness, and registers. This layered structure goes through extensive maturational changes (Hirano et al, 1983). Up to four years of age the intermediate and deep layers of the lamina propria are not differentiated. After four years an immature vocal ligament is observed. There is in childhood a much more extensive density of fibroblasts throughout the lamina propria than in the adult

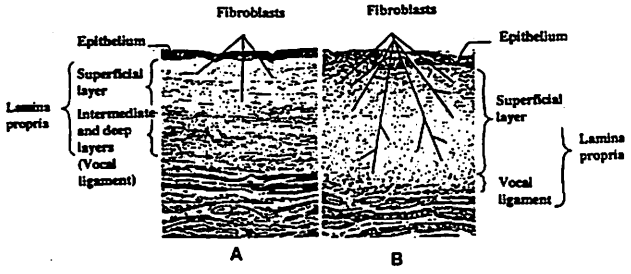


Figure 3. Histologic picture demonstrating the differences in density of fibroblasts in the lamina propria of the vocal fold mucosa of an adult (A) and a 4-year-old child (B). Hematoxylin-eosin stain. (From Hirano M, Kurita S, Nakashima T: Growth, development, and aging of the human vocal folds, in Bless DM, Abbs JH (eds.): Vocal Fold Physiology: Contemporary Research and Clinical Issues, San Diego, College-Hill Press, 1983, pg 36. Used by permission.)

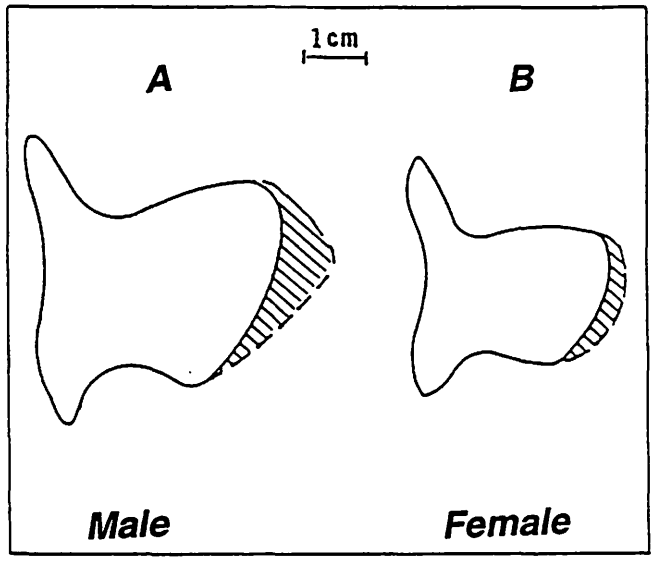


Figure 4. Tyroid cartilage configurations illustrating typical male (A) and female (B) morphology at puberty (solid lines) and adulthood (broken lines). The hatched areas represent the extent of postpubertal growth in the anteroposterior dimension and anterior thyroid height. (From Kahane JC: Growth of the human prepubertal and pubertal larynx. J Speech Hear Res 1982:25:446-455. Used by permission.)

(Figure 3). With growth, elastic fibers of the intermediate layer develop and fibroblasts decrease while collagen fibers of the vocal ligament form. By 16 years of age the layered structure of the adult is observed. The high density of fibroblasts in the submucosa of pediatric vocal folds implies that they may be prone to scar formation from surgical trauma, intubation, etc.

2. External Laryngeal Developmental Anatomy

In addition to the development of endolaryngeal structures, an appreciation of the growth and development of the larynx and vocal tract structures is helpful to understand the effects of growth on the voice. In the newborn, the larynx is positioned high in the neck with the cricoid at C3 to C4 (Dejonckere, 1984). This arrangement facilitates simultaneous respiration and swallowing during feeding (Bloom & Rood, 1981). The larynx gradually descends to the level C6 to C7 by fifteen years. The effect of these changes on the voice are not related to phonation but the vocal tract resonances (Bloom & Rood, 1981; Morrison et al, 1993). The frequency of vocal tract formants drops as the vocal tract enlarges. Changes with the vocal tract during puberty are different for male and female; males increase the size of the pharynx relative to the oral cavity more than females (Morrison et al, 1993).

The structure of the larynx itself changes markedly with age. In two reports, Klock (1968, 1986) published reports of extensive measurements on the anatomic dimensions of the larynx in infancy and childhood. It was found that, in general, the growth of the overall dimensions of the larynx is linearly related (directly proportional) to crown-heel length (somatic height). Laryngeal growth is thus related to age only as overall body growth relates; that is, a sigmoidal curve with acceleration between birth and three years, then deceleration, then rapid growth phase at puberty, especially in males. The thyroid ala in infancy are positioned in a curving semicircle of about 130 degrees. This narrows to 120 degrees in the prepubertal female and 110 degrees in the male (Kahane, 1978). Kahane (1978, 1982) also documented the changes in external laryngeal anatomy resulting from puberty. Significant regional growth localized to the anterior aspect of the thyroid cartilage was measured in laryngeal specimens of pubertal males, ie, formation of "Adam's apple" (see Figure 4). This results both in the increase in length of the anterior vocal folds, and change in the angle of the thyroid ala to 90 degrees. Other external laryngeal measurements showed less dramatic differences between pubertal males and females.

Developmental Physiology of the Voice

1. Phonation Mechanisms

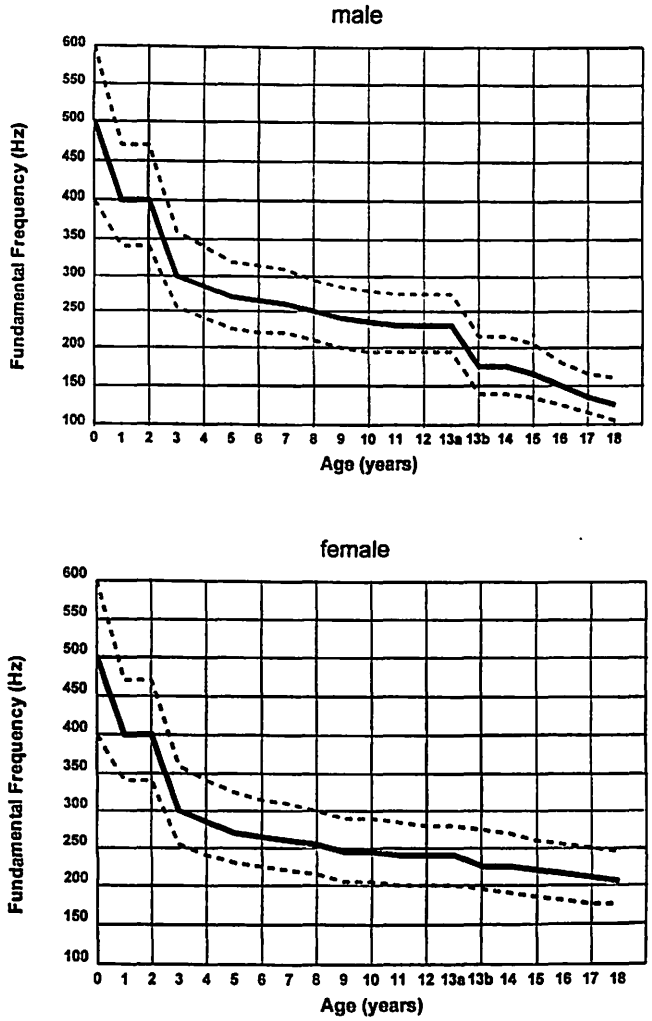
The acoustic output of the phonatory larynx is produced by the vibration of the anterior membranous edges of the vocal folds which periodically interrupt the airstream from the lungs. This fundamental aspect of phonation is creation and maintenance of mucosal traveling waves and their entrainment with the airflow (Berke & Gerratt, 1993). Vibration, or oscillation, is self-sustained by 1) the elastic recoil of vocal fold mucosa and, 2) by alternating pressures within the glottis that separate and bring together the folds (Titze, 1979; Cooper, 1987; Titze, 1994). This traveling mucosal wave, observed with high speed cinematography or stroboscopy, is created by the interaction of the airflow with

the vocal fold mucosal cover. The inferior-to-superior movement of the traveling wave is influenced by the pressure of the airstream (lung pressure), the thickness of the vocal folds, the approximation of the vocal folds, and the elasticity of vocal fold tissue.

While vocal fold oscillation is described as "periodic", that is, repeating itself at regular intervals, it in increasingly recognized that usually this does not precisely occur, in the voices of children and adults. Cycle-to-cycle variations in frequency and amplitude periodicity (termed "jitter" and "shimmer" respectively) have been measured in normal children's voices (Glaze et al, 1988; Glaze et al, 1990) as well as adults (Baken, 1988). Other types of sudden qualitative changes in vibration, known as "bifurcations" create unusual repeating patterns or aperiodicity (Titze, 1994). Recent studies have observed these patterns in voices of normal adults (Klatt & Klatt, 1990), infants (Mende et al, 1990) and disordered voice populations (Herzel et al, 1994). These features describe a system that is "chaotic" (both deterministic and unpredictable) and vocal fold vibration is now being investigated in these terms (Titze et al, 1993).

2. Pitch (Frequency)

The most notable feature of the pediatric voice is pitch change. The pitch drops throughout infancy, childhood, and adolescence, for both males and females (Figure 5a,b,c). The frequency of the infant's cry is about 500 Hz, and this drops by about one-half by age eight to ten. Several investigators have measured differences in the frequency of voices in prepubescent males and females (Glaze et al, 1988; Hasek et al, 1980), beginning at 5 to 8 years. Wilson's (1987) composite data suggest that the frequency of male voice is 5 to 10 Hz lower than female beginning about 8 years of age. Given Hirano et al's observation (1983) that membranous fold length is similar at these ages, the explanation for these findings is unclear, but may be related to height/weight differences, vocal tract changes, or cultural factors (Glaze et al, 1988). Since adrenal androgens begin secretion in males between 6 and 10 years of age (Tanner, 1971), subtle hormonal influences may also be present. Further decrease occurs through adolescence for males and females, with the male voice change at puberty most noticeable. Titze (1994) has shown how the changes in pitch are due mostly to changes in membranous length of the vocal folds (Figure 6). There are interesting irregularities in the



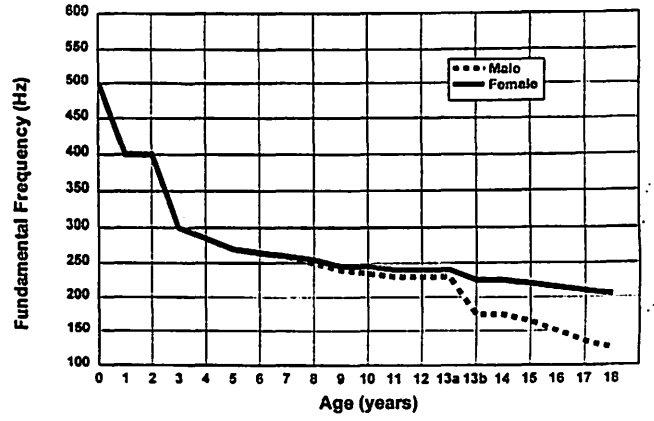


Figure 5. (A - upper left) Male vocal fundamental frequency average and range of normal limits, 0 through 18 years of age, with transition (a,b) at 13 years. (B - lower left) Female vocal fundamental frequency average and range of normal limits; 0 through 18 years of age. (C - above) Male and female average vocal fundamental frequency changes compared from 0 through 18 years, with transition (a,b) at 13 years. (From data compiled by Wilson DK: Voice Problems in Children, 3rd edition. Baltimore, MD, Williams and Wilkins, 1987.)

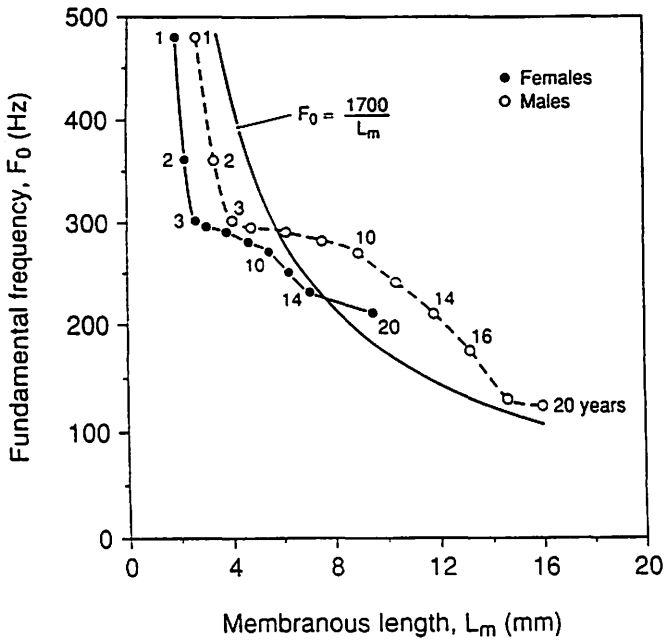


Figure 6. Fundamental frequency as a function of the membranous length of the vocal fold, with age as a parameter. (From Titze IR: Principles of Voice Production, Englewood Cliffs, NJ, Prentice-Hall, 1994, pg 180. Used by permission.)

curve at 3 and 10 years of age that may be related to changes in tissue stiffness as the vocal ligament develops.

The male adolescent voice goes through a transition, usually between 13 and 14 years of age, where the pitch drops. This is due to the anterior growth of the thyroid cartilage in response to testosterone, that causes an increase in vocal fold length. An additional change in laryngeal structure is an increase in bulk of the thyroarytenoid muscle. This causes increase in vertical thickness of the vocal fold as well as bulging of its medial contour (Titze, 1994). With this change glottal closure occurs over a larger portion of the glottal cycle, and amplitude of vocal fold vibration increases, resulting in a richer quality to the voice.

3. Loudness (Intensity), Aerodynamics, and Breathing

The cry of an infant or scream of a distressed child attest to the fact that children can create loud voices. Further consideration reveals that levels of acoustic power (averaging 70 decibels for conversational speech) comparable to adults are generated with a much smaller phonatory and respiratory mechanism. Several physiological principles underlie this; including the dependence of vocal intensity on frequency and lung pressure, and differences in the pediatric respiratory system. Titze (1994) explained that vocal intensity increases about 8 to 9 decibels per octave increase in fundamental frequency. A 3 year old child, speaking at 300 Hz is doubling the 150 Hz frequency of the adult male-female. However, other issues play into the ability to drive shorter vocal folds at a faster rate, namely lung pressure. In recent work, Stathopoulos and Sapienza (1993) studied vocal intensity variations during phonation in 4 year old and 8 year old children and adults. They found that for comparable soft, comfortable, and loud phonation tasks the children generated lung pressures 50 to 100 percent greater than the adults. In association with this, rib cage excursion for 4 year olds was equal to and 8 year olds was nearly twice that of adults. Because lung volume excursion in children is about half that of adults, they explained that children are required to move their rib cages more to achieve the same lung volume displacement, and have greater lung volume excursion relative to vital capacity during phonation. Similar findings, also reported in speech breathing research by Hoit et al (1990), lead to the conclusion that children work harder to use their voice. They have a shorter maximum phonation time than adults (Harden & Looney, 1984; Wilson, 1987) and take more frequent breaths during speech (Hoit et al, 1990). This increased respiratory effort expended decreases to adultlike patterns by 10 years of age (Hoit et al, 1990).

Causes of Voice Problems

When is the voice a problem? This question is not always straightforward, since dysphonia is often not perceived by the child. Though acute voice changes prompt concern, parent and family members may not regard as a problem a child's voice when it has not changed suddenly. Voice problems become noticeable to others when the child enters a larger social circle, such as starting school.

Voice problems are usually categorized as abnormalities in vocal quality, pitch, or loudness (Wilson, 1987). By far, the most common are abnormalities in quality; that is, a voice that is hoarse, rough, harsh, or breathy. Pitch abnormalities relate to voice that is felt to be too high or low, not corresponding to the child's age or gender. Loudness abnormalities relate to voice that is too soft or too loud: It is also fairly common that a perceived "voice problem" is actually a problem with speech resonance. In the series reported by Campbell and Stool (1993), of 203 children referred to a specialty clinic for a voice problem, 38 (19%) were found to have velopharyngeal insufficiency.

1. Incidence

Studies on the incidence of voice disorders in children have been nearly all done in school-aged children (Wilson, 1987). They rate voice disorders to be fairly common, ranging from 6 to 23% (Senturia & Wilson, 1968; Silverman & Zimmer, 1975; Bayes, 1966) with an average of 6-9% (Wilson, 1987). These data were recently corroborated by McAlister et al (1994) who recorded voices of 58 ten year old children twice, at two month intervals. Perceptual ratings identified eight with hoarseness (14%). All were boys. This preponderance of male over female school chil-

dren with voice problems, usually in ratios of 2 or 3 to 1 has been observed by others (Bayes, 1966; Senturia & Wilson, 1968; Leeper et al, 1980). Voice disorder of sufficient severity to warrant clinical treatment occurs in 2 to 3% of cases (Wilson, 1987). This corresponds to the average percentage of voice patients in the case load seen by school-based speech pathologists (Wilson, 1987).

As in other areas of pediatric otolaryngology, the causes of pediatric voice disorders include a few very common entities (eg. vocal nodules) and a large list of uncommon diagnoses. These have been separated by Cohen et al (1983) into neurological, congenital and genetic anomalies, tumors, infection and granulomatous diseases, trauma, endocrine and metabolic, physiological, psychogenic, and iatrogenic categories. The reader is referred to this reference for further study. It is similarly helpful to separate causes of voice disorders into those likely to be encountered in various age groups; including newborn to 6 months, 6 months to 5 years, 5 to 13 years, and 13 to 18 years (see Table 1). Though not comprehensive, this gives the clinician direction in sorting through a large differential diagnosis.

Table 1.
Differential Diagnosis of Common Pediatric Voice Problems Categorized by Age

0 to 6 months	**5 years to 13 years**
traumatic - intubation	behavioral - nodules
iatrogenic - surgical	infectious - respiratory
neurogenic - central or peripheral neuropathy	inflmmatory - allergy, GER
neoplastic - hemangioma, cysts	**13 years to 18 years**
congenital - webs, clefts	behavioral
infectious - herpes	male - mutational/transitional voice
6 months to 5 years	female - nodules
traumatic - foreign bodies, intubation	psychogenic - aphonias
infectious - respiratory	infectious - respiratory
neoplastic - papillomas	inflammatory - allergy, GER
behavioral - nodules	

In the newborn with an abnormal cry, with or without stridor, a variety of causes of dysphonia may be found. Congenital laryngeal abnormalities include anterior glottic web (Cohen,1985), laryngeal clefts-posterior (Cohen,1975) or anterior (Cohen & Thompson,1990), laryngeal anomalies associated with genetic defects such as cri-du-chat syndrome (Ward et al,1968) or lipoid proteinosis (Savage et al,1988). There are a variety of neurogenic causes of vocal fold paralysis. Generally, bilateral vocal fold paralysis has central neurologic origin, such as Chiari malformation, meningomyelocele, or hydrocephalus. Airway concerns predominate in these cases. Unilateral vocal fold paralysis causing dysphonia is often of peripheral origin. It may be related to cardiac disease (ie. "cardiovocal syndrome" reported by Condon et al (1985)), birth trauma (eg. neck traction during forceps delivery (Cohen et al,1982)), or iatrogenic trauma. Any surgical procedure in the chest or neck that encounters the course of the vagus or laryngeal nerves, such as tracheoesophageal fistula repair, esophageal atresia, or surgery of the great vessels may cause vocal fold paralysis. Ductus arteriosus ligation has been reported to cause left vocal fold paralysis in 4 to 5.2 percent of cases (Davis et al,1988;Fan et al,1989). Right vocal fold paralysis has been seen after the Blalock-Taussig shunt procedure (M Smith, personal observation). Traumatic causes of dysphonia in infants usually relate to the sequelae of intubation injuries with granuloma, laryngeal stenosis, or arytenoid dislocation. Neoplastic causes include hemangioma, papilloma, cystic hygroma, and saccular or other laryngeal cysts. Neonatal infectious laryngitis related to herpes virus infection has been reported (Opitz et al,1989).

In the infant and young child (6 months to 5 years) other causes of dysphonia may become manifest. A variety of viral and bacterial infectious illnesses often result in temporary hoarseness. Any other acute voice change in these patients, especially if associated with persistent cough or stridor, must always require consideration of foreign body aspiration (Hanokoglu,1986; Phillips et al,1989). Laryngeal papillomas may become manifest as hoarseness during this age group. Behavioral causes of dysphonia can begin at early ages with vocal nodule formation (see Figure 7).

In children 5 to 13 years of age vocal nodules predominate as the major cause of dysphonia (Dobres et al,1990). Other infectious and inflammatory causes of dysphonia are common. Chronic hoarseness from gastroesophageal reflux laryngitis, now increasingly recognized as a cause of voice problems in adults (Ward & Berci,1982; Olson,1991) has also been reported in this age group of children (Putnam & Orenstein,1992). Another of this group of patients with a high incidence of vocal nodules are those with velopharyngeal valving problems (McWilliams et al,1973), from laryngeal hyperadduction in efforts to close the velopharyngeal sphincter.

In the adolescent years, behavioral and psychogenic causes of chronic dysphonia predominate. The incidence of vocal nodules in males drops abruptly, however both males and females are still at fairly equal risk (Dobres et al,1990). In males, problems with voice transition at puberty may yield hoarseness and persistent high-pitch, termed puberphonia or mutational (adolescent male transitional) voice disorder. Psychogenic and functional voice disorders also may be seen (Froese & Sims,1987; Harris & Richards,1992; Smith et al,1993).

Several special categories of children who may present with voice problems should be recognized. In immunosuppressed children being treated for leukemias,

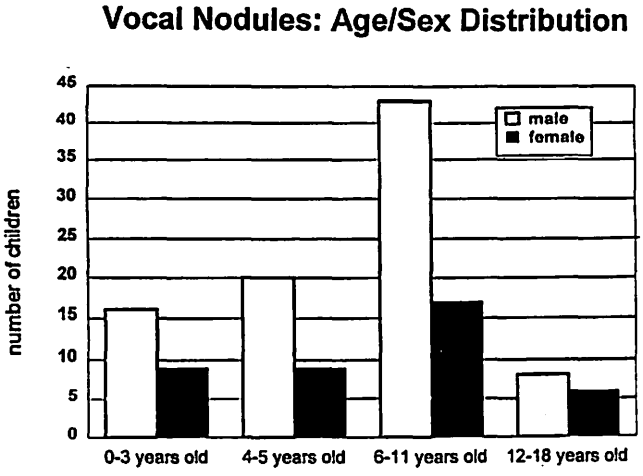


Figure 7. Vocal nodules in children grouped by age and sex, from the series of Dobres et al (1990).

voice problems related to fungal infection (Hass et al,1987) or chemotherapy induced laryngeal neuropathy (Annino et al,1992) may be seen. Traumatic brain injury patients may have voice problems in connection with speech deficits, usually characterized by decreased loudness (McHenry et al,1994). Likewise, the voices of children with cerebral palsy may have a variety of problems related to voicing interruptions, abnormal speech breathing, impaired pitch control, breathiness or hyperfunction (Wilson,1987). Children with mental retardation also have voice problems. The most examined group in this regard is Down's syndrome. Their voices may be rough, harsh, or have excessive pitch variability (Wilson,1987). Children with hearing loss are well known to demonstrate abnormalities in voice use. These are greatly dependent on the severity and age of onset of the hearing loss, and also embedded in associated language and articulation difficulties (Wilson,1987). Their voices are described as high-pitched, monotone, harsh, breathy, and tense (Wirz,1992).

Evaluation of Voice Problems

The approach to evaluation of a child with a voice problem differs from the adult in several respects (Morrison, et al, 1993). Voice disorders in children often co-occur with speech, language, and developmental delays. Children have limited ability to cooperate, so that any procedures or assessment techniques must be performed in the least-invasive and nonthreatening manner. Parents, family members, and other care-givers are also included in the acquisition of historical information and involved with therapy.

The assessment of voice problems requires a comprehensive medical and behavioral evaluation. Of necessity, this brings together specialists from a variety of disciplines; otolaryngology-head and neck surgery, speech-language pathology, pediatrics, psychology, and social work. Other medical specialties that may be involved include neurology, gastroenterology, genetics, and endocrinology. This team approach is described in the chapter by Campbell et al.

The history of a voice problem often involves individuals other than the child; parents, school teachers, etc. In detailing the history of a voice problem we find useful the questionnaire reported by Maddern et al (1991). Complete examination is performed by the otolaryngologist. A hearing test is obtained. In cases of adolescent male dysphonias, physical maturity is assessed (Andrews,1991).

Laryngeal examination is a crucial aspect of the voice evaluation. The flexible fiberoptic laryngoscope has become the tool of choice for the evaluation of voice problems, in adults and children (Koufman,1991). Even very young children can tolerate this examination well with proper preparation of the child and parent, use of topical nasal anesthesia, and an engaging, nonthreatening atmosphere. These methods have been well described (D'Antonio et al, 1986; Chait & Lotz,1991; Lotz et al, 1993). The use of video equipment has advantages of: engaging the child's attention, providing visual feedback for parent, and facilitating communication between physician, speech pathologist, and others. Several points may be emphasized in the course of this examination. After the fiberscope is passed beyond the turbinates, the velopharynx is examined for velopharyngeal insufficiency. The adenoid pad and tonsils are seen as the fiberscope passes beyond the nose into the pharynx. In the hypopharynx, the laryngeal and pharyngeal mucosa are inspected for erythema and mucosal thickening, especially the arytenoid mucosa and posterior glottis. Diffuse edema of the laryngeal mucosa, "cobblestoning" of the posterior pharynx, and edematous nasal mucosa may be a sign of allergy. Laryngeal mobility is inspected, including coughing and sniffing maneuvers. The larynx is observed during connected speech, phrases, such as counting to ten. These tasks are helpful for observing supraglottal hyperfunction, seen both in functional dysphonias and as compensatory behavior in organic lesions (Koufman,1991). The vocal folds themselves are examined for irregularities, swelling, or lesions. Such abnormalities of the vocal fold when seen unilaterally should be suspected for congenital cysts (Monday et al,1984;Bouchayer & Cornut,1992). During sustained vowel /i/ phonation the glottal closure is seen. Incomplete posterior glottal closure (posterior "chink") is common in children as it is in many adult females (Glaze et al,1990;Hirano & Bless,1993). During fiberscopic visualization, a determination is made regarding whether stroboscopic lighting is helpful to elucidate the problem, eg. suspicion of abnormal mucosal wave vibration (due to scar, sulcus) in the absence of nodules, unilateral cord lesion, etc. Laryngostroboscopy has limitations in young children, who may not be able to phonate an adequate length of time to

obtain stroboscopic pictures. The child's maximum phonation time should be at least 5 seconds. Older children generally can tolerate rigid oral telescopic examination of the larynx very well.

The voice assessment conducted by the speech-language pathologist is ideally conducted during the same visit as the otolaryngologist. This is covered in the chapter by Campbell et al. The child's voice is rated perceptually during a variety of speech tasks and sustained vowels. The voice is recorded for acoustic measures. Aerodynamic and glottographic measurements may also contribute information to the problem. These are easily conducted with current instrumentation. They have certain advantages: documentation of the problem, corroboration of findings with the laryngeal imaging examination, documentation of treatment efficacy, use as biofeedback during therapy.

There are occasions when direct visualization of the larynx under anesthesia is required for diagnosis of pediatric dysphonia. Indications for this include: 1) inability to examine the larynx fiberscopically, 2) when detailed visualization is required to determine the presence of glottic web, laryngeal cleft, arytenoid fixation, or vocal fold lesion such as papilloma, sulcus vocalis, cyst, etc. 3) perform laryngeal electromyography for paralysis assessment (Koch et al,1989).

Management of Voice Problems

The management of voice problems requires an understanding of the pathophysiology and mechanism of dysphonia, as well as the natural history of the problem. The severity of the voice problem is also taken into account; if the voice calls attention to itself, interferes with communication, or causes the speaker to be unhappy. The multifactorial nature of causes of dysphonia, in a bio-psycho-social perspective, may require consideration (Morrison et al,1993). This information guides answers to the questions of if, when, and how to intervene. Management options include behavioral, medical, and surgical therapies. Treatment of the common causes of pediatric dysphonia; vocal nodules, laryngeal paralysis, laryngeal webs, laryngeal papillomas, and laryngeal stenosis will be considered in this context.

Vocal Nodules and Functional Voice Disorders

Given the understanding that vocal nodules are an organic manifestation of laryngeal hyperfunction (Koufman & Blalock,1991), the mainstay of treatment of vocal nodules is behavioral voice therapy. A recent study by Murry and Woodson (1994) confirmed this. They studied 59 adult patients treated with 1) a combined voice therapy approach (speech pathologist and otolaryngologist), 2) voice therapy (speech pathologist only), or 3) vocal nodule surgery followed by voice therapy. By posttreatment perceptual analysis of voice recordings overall improvement was highest in the first group, followed by the second group. No patients in the first two groups had surgery. The surgical group had the poorest outcome. Behavioral voice therapy is clearly indicated in adolescents and adults, but the decision to treat is less clear in children.

Clinical experience of the natural history of vocal nodules in children suggests spontaneous remission by puberty (Toohill,1975). There is less data to document the efficacy of behavioral treatment of vocal nodules in children. Because of these factors, "therapeutic nihilists" advocate no intervention (Toohill, 1975; Kay, 1982; Koufman & Isaacson, 1991). Some behavioral treatment programs have reported "success" (Johnson, 1985; Nilson & Schneiderman, 1983) by various means of assessment. The topic is one of active debate (Kahane & Mayo,1989; Sander,1989). An interesting case was recently reported (Allen et al,1991) of a 9 year old with vocal nodules and dysphonia that persisted despite conventional voice therapy. He was treated with EMG biofeedback to reduce laryngeal hyperfunction and demonstrated resolution of the nodules after 6 months. Common behavioral therapy strategies for vocal nodules are outlined in approaches by Maddern et al (1991), Wilson (1987), and Andrews (1991).

Surgery is rarely recommended due to risks of vocal fold scarring, but there are some who advocate surgical treatment of vocal nodules using microdissection techniques when nodules are identified, after a trial of voice therapy (Benjamin & Croxson,1988; Bouchayer & Cornut,1992). The optimal age for voice surgery in these children is considered to be 9 - 11 years (Bouchayer & Cornut, 1992), at which age compliance with therapy to limit voice abuse is considered improved. Bouchayer & Cornut (1992) also reported on a high incidence of congenital cysts, polyps, and sulci seen only on microlaryngoscopy in children previously felt to have vocal nodules. No other studies have corroborated this finding; at this time it seems reasonable to recommend that such be considered when a unilateral lesion is seen on fiberoptic laryngoscopy, since typical vocal nodules come in pairs. In contrast to the European approach of voice therapy followed by surgery if warranted, in the U.S. a variety of conflicting philosophies have existed. A survey of 535 otolaryngologists on management of pediatric vocal nodules found that surgery followed by voice therapy was recommended by 9.4% and surgery alone recommended by .8% of respondents (Moran & Pentz,1987). 59% recommended voice therapy alone and 31% recommended no intervention or other treatment. At this time, we recommend voice therapy for those children old enough to comply with behavioral management, with surgical treatment generally viewed as an uncommon exception.

While vocal nodules may be considered a "functional voice disorder" with organic manifestations (Koufman

& Isaacson,1991), functional voice disorders without lesions may also be seen in children, presenting as dysphonia or aphonia (Bridger & Epstein,1983; Froese & Sims,1987; Koufman & Isaacson,1991; Harris & Richards,1992; Smith et al,1993b). Another functional voice disorder in this age group is the transitional male adolescent dysphonia (puberphonia). Voice therapy is indicated, but may require a prolonged course in aphonic children, especially adolescents (Harris & Richards,1992; Smith et al,1993b). Psychosocial considerations should be addressed in the treatment of these patients.

Laryngeal Paralysis

Unilateral vocal fold paralysis (UVFP) and associated glottal insufficiency may be a cause of dysphonia in children. Associated airway and feeding concerns are managed initially. Several large reviews have reported that the natural history of laryngeal paralysis is toward spontaneous "recovery". Cohen et al (1982) reported on 38 children with UVFP; though 14 were lost to follow up, 13 recovered completely. Emery and Fearon (1984) reported on 30 cases of UVFP in children. Peripheral nerve injuries usually recovered, in 60% of cases. In laryngeal paralysis it is important to distinguish between neurogenic (return of vocal fold movement) and phonatory (improvement in voice) recovery, since they do not always equate (Hirano et al,1987). This is demonstrated in the series of Emery and Fearon (1984); in eight cases followed for an average of six years, vocal fold mobility did not return yet in six of the eight patients the voice recovered. Only one patient had a surgical procedure (Teflon® injection). The very infrequent reports of surgical management of dysphonia in children from laryngeal paralysis attest to the trend toward natural improvement in voice through compensatory means, with or without recovery of vocal fold movement.

Surgical management of glottal insufficiency and laryngeal paralysis have received much attention. Though Teflon® injection of the vocal fold has been commonly performed in adults, it has several disadvantages in children, including irreversibility and unknown long-term effects (Tucker,1985). A variety of phonosurgical techniques as alternatives to vocal fold Teflon® injection have been popularized (Isshiki,1992). These laryngeal framework procedures have had some application in children (Smith & Gray,1994) though experience is limited. The arytenoid adduction procedure (Isshiki et al,1978), while appropriate in theory for closure of posterior glottic defects, has technical problems that preclude its use in young children (Smith & Gray,1994). This procedure is recommended in adolescents with UVFP and large glottal gap, and if concomitant aspiration symptoms are present.

Reinnervation of the paralyzed larynx is another surgical option for dysphonia from UVFP. Two variations of this approach have been investigated; ansa-hypoglossi to recurrent laryngeal nerve anastomosis (Crumley,1990), and ansa-strap nerve-muscle pedicle implantation into laryngeal muscle (lateral cricoarytenoid or thyroarytenoid) reviewed by Tucker (1979) and Goding (1991). Both of these techniques aim to reinnervate laryngeal muscle to prevent atrophy, increase tone, improve glottal adduction and voice. In the series reported by Tucker (1986) of eight children with UVFP, three underwent nerve-muscle pedicle procedure, in addition to voice therapy, with "good" voice results.

Laryngeal Papillomas

Recurrent respiratory papillomatosis (RRP) is the most common benign laryngeal neoplasm in children. These papillomas can cause hoarseness, stridor, and respiratory distress, which may necessitate surgical intervention to maintain an adequate airway. Impairment of the voice may occur as a result of RRP or the treatment. While the disease is present the voice may be affected because of interference of the papilloma with vocal fold closure. However, after the disease is in remission, the voice may be affected as a complication of the surgical treatment previously used. Crockett et al (1987) described glottic complications as a result of surgical intervention for RRP. Glottic scarring, anterior and posterior glottic webbing occurred in patients who had frequent and multiple procedures.

One of the difficulties in treatment of RRP is that aggressive removal of the papillomas to provide the best airway may result in eventual injury to the vocal folds. Since severe papilloma disease obscures normal anatomic laryngeal landmarks and structures, it can be difficult to limit the surgical excision to the epithelial layers. This can be particularly true in the anterior glottis, the posterior glottis, and the membranous vocal folds. Wetmore and colleagues (1985) reported that the incidence of soft-tissue complications increased in patients who required six or more laser operations. They concluded that deep vaporizations of papilloma resulted in higher incidence of glottic complications.

Papilloma removal in the 1980's focused on the concept of total or near-total eradication of papilloma with each surgery (Crockett et al, 1987; Wetmore et al,1985). In the late 1980's we conducted a study of voice characteristics in eight patients who had experienced extensive, repeated surgery for removal of papilloma and were then in remission for at least two years (Gray et al, 1991). Results demonstrated reduced frequency and intensity range, and laryngeal stroboscopic findings consistent with stiff vocal folds, due to scar. Subjectively, patients did not feel their voice performed normally. Such findings raised concerns about whether the voice difficulty experienced by these patients arose from the disease or from the treatment.

An awareness of these voice disorders has led some surgeons to a more conservative approach for papilloma

removal (Benjamin & Parsons, 1988). This approach is directed at two points of surgical technique. The first is use of more precise instrumentation for papilloma to minimize damage to the underlying subepithelial tissues, eg. lamina propria. Advances have been made in reducing the laser spot size, the use of a micromanipulator, and adjusting laser parameters to minimize tissue destruction (Ossoff et al, 1991). The second consideration involves sparing a small area of papilloma unilaterally at the anterior and posterior glottis if that area is involved. Small cup forceps may be used to debulk papilloma from these areas, avoiding laser injury (Benjamin & Parsons, 1988). Using this approach, Ossoff et al (1991) presented 22 patients (14 children and 8 adults) where the delayed soft-tissue complications were significantly reduced. Only three children had problems related to posterior glottic web or vocal fold scar. From this study, they concluded that these problems in patients undergoing papilloma surgery are related more to the surgical technique used and not to the number of procedures performed. This report underscores the need to consider the effect of these endoscopic laryngeal procedures on the voice, though the focus of the treatment involves airway and papilloma management.

Congenital Glottic Web

Congenital glottic webs are rare in children, and are nearly always located in the anterior glottis. In the largest series reported by Cohen (1985), 51 patients with congenital anterior glottic webs were seen over a 32 year period. Voice problem at birth was the most commonly reported symptom. Cohen categorized glottic webs based on observation and estimation of web extent in the glottic lumen. Those involving less than 35% often required no treatment. Airway symptoms increased with increasing web size. Only four of thirty-two patients with 50% or less of the glottis involved had a tracheotomy, but all sixteen patients with greater than 50% involvement had a tracheotomy. Treatment of the web involved a combination of endoscopic treatment with dilation, division or laser. Twelve patients underwent laryngofissure and placement of NcNaught or Silastic keel. Dedo (1979) reported success in managing a newborn with anterior glottic web using a Teflon® keel placed endoscopically. In this patient, however, problems with development of posterior glottic stenosis were encountered from keel irritation in the posterior commissure and arytenoids. Cohen (1985) recommended waiting until 3 years of age when the larynx is of sufficient size to avoid this complication of keel placement. He also emphasized that normal voice quality was rarely attained in his series. Most patients had hoarse or husky voice, weak voice, or "double voice". The comments of Zalzal et al (1991) on voice problems following pediatric laryngotracheal reconstruction also apply to these patients; the voice problems from the underlying abnormality may be compounded by surgery to correct the disease.

Laryngotracheal Reconstruction

Surgical techniques for repair of glottic and subglottic stenosis have made many advances (Cotton, 1987). These procedures, generally known as laryngotracheal reconstruction (LTR), have particular applicability to the pediatric population where laryngotracheal stenosis has been a prevalent clinical problem. The goal of these procedures is to restore the airway. However it can be readily appreciated that laryngeal surgery designed to address one aspect of laryngeal function may necessarily affect other functions of the larynx, such as phonation (Smith et al, 1992).

The association of voice problems in some children that have undergone cartilage graft LTR procedures on pediatric voice has been recognized (Sell & MacCurtain, 1988; Zalzal, 1988a; Zalzal, 1988b; Zalzal et al, 1991; Smith et al, 1993a; MacArthur, 1994). For example, Smith et al (1993a) reported on eight patients with voice problems following LTR. The voices were frequently rough, low pitched, or breathy. Two patients exhibited reverse or inhalatory phonation. Fiberoptic laryngoscopy and laryngostroboscopy demonstrated supraglottal phonation in three, glottal incompetence in two, arytenoid fixation in two, anterior commissure blunting or widening in three, vertical asymmetry of the vocal folds in two, and vocal fold scarring (ie. absent mucosal wave) in three. Most patients exhibit more than one abnormal finding. Though the study group was not controlled and was likely representative of LTR patients with more severe voice problems, it is comparable to other reports (Kearns et al, 1988; Zalzal et al, 1991; Zalzal et al, 1993; MacArthur, 1994) and is indicative of the consequences for voice that may result from LTR. This can be a significant long term disability for a child.

LTR may adversely affect phonation in various ways. Stenosis at the level of the free margin of the vocal folds and scar of the superficial lamina propria will inhibit vocal fold vibration and is difficult to reconstruct. It is also apparent that surgery designed to enlarge the laryngeal airway may adversely affect phonation, which requires glottal closure. Trends from recent studies indicate some additional factors that appear increase risk for a poor postoperative voice result in children who undergo LTR. These include the use of posterior cricoid cartilage grafts, combined use of anterior and posterior grafts, the long term placement of endolaryngeal stents, and multiple LTR procedures (Cotton, 1991; Zalzal, 1991; Smith et al, 1992; Zalzal, 1993).

Posterior glottic and subglottic stenosis can be successfully treated with posterior cricoid cartilage grafts (Zalzal, 1988; Cotton, 1991). This technique can also be used for treatment of impaired vocal fold mobility, such as

bilateral vocal fold paralysis (Gray et al,1994). In the series reported by Zalzal (1988), twelve patients were treated for posterior laryngeal stenosis with posterior cricoid cartilage grafts. The patient's voice quality was assessed by a household member who spent the most time with the patient before and after surgery (subjective perceptual assessment). Of the eight patients with preoperative normal voice quality, only two had a postoperative normal voice quality with six patients reported as having hoarse or husky voice quality. In another series reported by Zalzal et al (1991), sixteen patients had voice quality formally assessed. Only four of the nine patients who received posterior grafts had breathiness, yet these four were the only patients with a breathy postoperative voice. Smith et al (1992) reported on fifteen patients that underwent "single stage" LTR (no tracheotomy tube employed or removal of the tracheotomy tube at initial surgery). Of the twelve patients who were successfully extubated, at informal voice quality assessment three to six months postoperatively, seven had normal voice, four had moderate dysphonia, and one had severe dysphonia. All five of these patients had both anterior and posterior cricoid cartilage grafts placed. For three of these five, the surgery was a revision procedure.

The use of endolaryngeal stents to secure cartilage grafts in place has been well described. However, these stents appear to injure the voice, especially when used long term. In Cotton's (1991) large series of 61 patients that underwent posterior cricoid graft LTR, the duration of stenting was found to be correlated with postoperative voice assessment in that better voice results occurred when the duration of stenting was 12 weeks or less. Maddalozzo and Holinger (1987) reported in a series of 20 children that underwent LTR that hoarseness was not an infrequent problem in those that required stenting. In the report of Zalzal et al (1991) all sixteen patients had a stent placed, fifteen had aberrant voice quality. The authors failed to find a correlation, however, between stenting duration and postoperative voice quality.

Several animal studies have examined the effects of stent/intubation on the larynx, with implications for voice problems. In a goat animal study of the effect of long-term endolaryngeal stents on the larynx, disruption of laryngeal mucosa and underlying tunica elastica, particularly in the posterior glottis, was observed in preparations that underwent endolaryngeal stent placement for 3 months (Smith et al,1991). Squamous metaplasia of the posterior glottic mucosa was seen, as well as erosion of the vocal process of the arytenoid. Epithelial hyperplasia and fibrous proliferation in the submucosa anterior to the vocal process were observed. Leonard et al (1992) studied the effect of 7-day intubation in small dogs. Larynges harvested 5 weeks after extubation exhibited epithelial disruption, hypertrophy of the epithelial layer, and proliferation of subepithelial connective tissue. These changes, although mainly in the posterior glottis, also were observed in the membranous fold anterior to the vocal process. It would be expected that in the infant and pediatric larynx, endotracheal tubes and stents would tend to contact more of the membranous folds and anterior glottis. This has potential implications for injury to the membranous folds. Because of the abundant and diffuse distribution of fibroblasts throughout the superficial and deep lamina propria, the membranous folds ("phonatory glottis") of the pediatric larynx may be more susceptible to voice injury from surgical trauma, stents or intubation.

Techniques of LTR may alter the position and anatomy of the vocal folds and endolaryngeal structures through external surgical manipulation of their support, the thyroid cartilage and arytenoids. In this way these procedures may be also viewed as laryngeal framework surgery. Examples of this may be seen in the commonly employed techniques of LTR. Cartilage grafts, usually from autogenous rib, are popularly used in LTR (Cotton,1987; Cotton,1991). These may be placed in the anterior and/or posterior cricoid region. Anterior cartilage grafts alter the laryngeal framework by immobilizing the action of the cricothyroid muscles, which lengthen and tense the vocal folds. Highly placed anterior cartilage grafts may disrupt the anterior commissure and splay the vocal folds apart. Posterior cricoid cartilage grafts widen the posterior commissure. This separates the arytenoids and affects the ability of the vocal folds to approximate at the vocal processes. Posterior cricoid grafts may also impair action of the interarytenoid muscle. These effects on glottal closure could be more pronounced in the pediatric larynx since it has a relatively larger posterior glottis.

A summary of suggestions for minimizing or preventing phonation problems in laryngeal framework cartilage graft surgery for pediatric laryngotracheal stenosis is given in table 2.

Table 2.
Technical Factors in Laryngotracheal Reconstruction that Affect Voice from Smith et al (1992).

Laryngotracheal Reconstruction Procedures	Potential Adverse Effects on Phonation	Techniques
Anterior laryngotracheal split and/or graft	Anterior commissure disruption; vocal fold vertical asymmetry; cricothyroid muscle dysfunction; supraglottic collapse	Avoid complete laryngofissure, if possible; avoid graft placement in anterior commissure; perform exact alignment of anterior commissure
Posterior laryngotracheal split and/or graft	Increased glottic gap, impaired arytenoid adduction; arytenoid subluxation	Avoid excessive graft width; use gentle retraction of hemicricoid
Stents	Scarring of vocal fold mucosa; impaired arytenoid mobility	Minimize stenting time; use single-stage laryngotracheal reconstruction, if possible; stent below vocal folds, if possible

Conclusion

The pediatric otolaryngologist is called on to manage the voice in a variety of settings; from newborn to adolescent, from the common to the obscure, from the severely developmentally impaired child to the aspiring vocal professional. This chapter has emphasized several points. First, diagnosis and treatment of voice is approached from 1) a firm understanding of the anatomy and physiology of phonation, 2) knowledge of the myriad of disease processes that affect phonation, and 3) realization of the complex behavioral and neuromotor development of the laryngeal valve as a sound source for purposes of communication. The bio-psycho-social model of disease is pertinent in the approach to many voice problems (Morrison et al,1993). In treatment, the pediatric otolaryngologist realizes that 1) most voice problems are not managed surgically, but behaviorally and medically, and 2) any surgical procedure on the larynx may affect the voice, sometimes adversely. In the context of communication, the voice has a pivotal role. The goal of maximizing and preserving lifelong, healthy vocal use is an important one, as it's function, as part of our means of oral communication, "is basic to social/emotional development, to learning, to career choice, to interpersonal relationships, in sum, to all the things we are and do as human beings" (Casper,1985).

Acknowledgements

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

Selected References

Aronson AE: Clinical Voice Disorders, 3rd edition. New York, NY, Thieme, Inc, 1990.

This comprehensive textbook from a renowned speech pathologist at the Mayo Clinic approaches voice disorders from a detailed perspective that emphasizes pathophysiology, neurology, and psychology. It is encyclopedically referenced.

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

The physics of phonation are explained by a physicist and speech researcher who has spent his career studying vocal fold physiology and acoustics. It is written as a college textbook from courses taught to graduate speech pathology students and otolaryngology residents at the University of Iowa. Mathematical and physical principles are explained in a clear, but not overwhelming way. The chapter on principles underlying voice changes with age is particularly pertinent.

Wilson DK: Voice Problems in Children, 3rd edition. Baltimore, MD, Williams and Wilkins, 1987.

This textbook comprehensively reviews both voice and resonance disorders. It has an extensive bibliography. It covers organic and functional causes of voice disorders, and has several chapters on their treatment, geared toward behavioral voice therapy for children, based on the author's lifetime of work.

References

Andrews ML: Voice Therapy for Children. San Diego, CA, Singular Publishing Group, 1991.

Andrews ML: Intervention with young voice users: a clinical perspective. J Voice 1993;7:160-164.

Annino DJ, MacArthur CJ, Friedman EM. Vincristine-induced recurrent laryngeal nerve paralysis. Laryngoscope 1992;102:1260-1262.

Baken RJ. Clinical Measurement of Speech and Voice. Boston, MA, College-Hill Press, 1987.

Bayes RA: An incidence study of chronic hoarseness among children. J Speech Hear Disord 1966;31:171-176.

Benjamin B, Croxson G: Vocal nodules in children. Ann Otol Rhinol Laryngol 1987:96;530-533.

Benjamin B, Parsons DS: Recurrent respiratory papillomatosis: a ten-year study. J Laryngol Otol 1988;102:1022-1028.

Berke GS, Gerratt BR: Laryngeal biomechanics: an overview of mucosal wave mechanics. J Voice 1993:7:123-128.

Bloom LA, Rood SR. Voice disorders in children: structure and evaluation. Pediatr Clin North Am 1981;28:957-963.

Bouchayer M, Cornut G: Microsurgical treatment of benign vocal fold lesions: indications, technique, results. Folia Phoniatr 1992;44:155-184.

Campbell TF, Stool S: Principles of assessment of the child with a voice disorder. Paper presented at the Conference on Recent Advances in the Evaluation of the Difficult Voice Patient. April, 1993, University of Pittsburgh School of Medicine, Pittsburgh, PA.

Casper J. Disorders of speech and voice. Pediatr Ann 1985;14:220 229.

Chait DH, Lotz WK: Successful pediatric examinations using nasoendoscopy. Laryngoscope 1991;101:1016-1018.

Choi SS, Cotton RT: Surgical management of voice disorders. Pediatr Clin North Am 1989;36:1535.

Cohen SR: Cleft larynx: a report of seven cases. Ann Otol Rhinol Laryngol 1975;84:747-756.

Cohen SR, Geller KA, Birns JW, et al: Laryngeal paralysis in children: a long-term retrospective study. Ann Otol Rhinol Laryngol 1982;91:417-424.

Cohen SR, Thompson JW, Geller KA, Birns JW. Voice change in the pediatric patient: a differential diagnosis. Ann Otol Rhinol Laryngol 1983;92:437-443.

Cohen SR. Congenital glottic webs in children: a retrospective review of 51 patients. Ann Otol Rhinol Laryngol 1985;94 (Supplement 121):1-16.

Cohen SR, Thompson JW: Ventral cleft of the larynx: a rare congenital laryngeal defect. Ann Otol Rhinol Laryngol 1990:99;281-285.

Condon LM, Katkov H, Singh A, et al: Cardiovocal syndrome in infancy. Petriatrics 1985;76:22-25.

Cooper DS: The laryngeal mucosa in voice production. Ear, Nose, and Throat J 1988;67:332-352.

Cotton RT: The management and prevention of subglottic stenosis in infants and children. In Myers EN, Bluestone CD, Brackmann DF, et al (eds): Advances in Otolaryngology-Head and Neck Surgery, vol 1. St. Louis, Mosby, 1987,pp 241-260.

Cotton RT: The problem of pediatric laryngotracheal stenosis: a clinical and experimental study on the efficacy of autogenous cartilaginous grafts placed between the vertically divided halves of the posterior lamina of the cricoid cartilage. Laryngoscope 1991;101(suppl 56):1-45.

Crockett DM, McCabe BF, Shive CJ: Complications of laser surgery for recurrent respiratory papillomatosis. Ann Otol Rhinol Laryngol 1987;96:639-644.

Crumley RL: Nerve transfer technique as it relates to phonatory surgery. In Otolaryngology-Head and Neck Surgery, Update II, pp 100-106. St. Louis, MO, CV Mosby Co, 1991.

D'Antonio LL, Chait DH, Lotz WK, et al:Pediatric videonasoendoscopy for speech and voice disorders. Otolaryngol Head Neck Surg 1986;94:578-583.

Davis JT, Baciewicz FA, Suriyapa S, et al: Vocal cord paralysis in premature infants undergoing ductal closure. Ann Thorac Surg 1988;46:214-215.

Dedo HH: Endoscopic teflon® keel for anterior glottic web. Ann Otol Rhinol Laryngol 1979;88:467-473.

Dejonckere PH: Pathogenesis of voice disorders in childhood. Acta Otorhinolaryngologica (Belgica), 1984; 38:307-14.

Dobres R, Lee L, Stemple JC, Kummer AW, Kretschmer LW: Description of laryngeal pathologies in children evaluated by otolaryngologists. J Speech Hear Disord 1990:55;526-532.

Emery PJ, Fearon B: Vocal cord palsy in pediatric practice: a review of 71 cases. Int J Pediatr Otorhinolaryngol 1984;8:147-154.

Fan LL, Campbell DN, Clarke DR, et al: Paralyzed left vocal cord associated with ligation of patent ductus arteriosus. J Thorac Cardiovasc Surg 1989;98:611-613.

Froese AP, Sims P: Functional dysphonia in adolescence: two case reports. Can J Psychiatry 1987;32:389-392.

Glaze LE, Bless DM, Milenkovic P, et al: Acoustic characteristics of children's voice. J Voice 1988;2: 312-319.

Glaze LE, Bless DM, Susser RD: Acoustic analysis of vowel and loudness differences in children's voice. J Voice 1990;4:37-44.

Goding GS: Nerve-muscle pedicle reinnervation of the paralyzed vocal cord. Otolaryngol Clin North Am 1991;24:1239-1252.

Gray SD, Barkmeier J, Shive C, et al: Vocal function in papilloma patients. Presented at the American Society of Pediatric Otolaryngology, Waikoloa, HI, May 5-6, 1991.

Gray SD, Kelly SM, Dove H: Arytenoid separation for impaired pediatric vocal fold mobility. Ann Otol Rhinol Laryngol 1994;103:510-513.

Hanukoglu A, Fried D, Segal S: Loss of voice as sole symptom of subglottic foreign-body aspiration. Am J Dis Child 1986;140:973.

Harden JR, Looney NA: Duration of sustained phonation in kindergarten children. Int J Pediatr Otorhinolaryngol 1984;7:11-19.

Harris C, Richards C: Functional aphonia in young people. J Laryngol Otol 1992;106:610-612.

Hasek C, Singh S, Murray T: Acoustic attributes of children's voices. J Acoust Soc Am 1980;68:1262-5.

Hass A, Hyatt AC, Kattan M, et al: Hoarseness in immunocompromised children: association with invasive fungal infection. J Pediatr 1987:111;731-733.

Herzel H, Berry D, Titze IR, et al. Analysis of vocal disorders with methods from nonlinear dynamics. J Speech Hear Res 1994;37:1008-1019.

Hirano M, Kurita S, Nakashima T: Growth, development, and aging of the human vocal folds, in Bless DM, Abbs JH (eds.): Vocal Fold Physiology: Contemporary Research and Clinical Issues, pp 22-43. San Diego, College-Hill Press, 1983.

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

Hirano M, Kurita S, Kiyokawa K, et al: Posterior glottis: morphological study in excised human larynges. Ann Otol Rhinol Laryngol 1986;95:576-581.

Hirano M, Nozoe I, Shin T, Maeyama T: Electromyography for laryngeal paralysis. In Hirano, M, Kirchner JA, Bless DM,(eds). Neurolaryngology: Recent Advances, pp. 232-248. College-Hill, Boston, 1987.

Hirano M, Bless DM: Videostroboscopic Examination of the Larynx. Singular Publishing Group, Inc, San Diego, CA, 1993.

Hoit JD, Hixon TJ, Watson PJ, et al: Speech breathing in children and adolescents. J Speech Hear Res 1990;33:51-69.

Isshiki N: Laryngeal framework surgery. In Myers EN, Bluestone CD, Brackmann DE, et al (eds): Advances in Otolaryngology-Head and Neck Surgery, vol 5, pp 37-56. St. Louis, Mosby, 1991.

Isshiki N, Tanabe M, Sawada M: Arytenoid adduction for unilateral vocal cord paralysis. Arch Otolaryngol 1978;104:555-558.

Kahane JC: A morphological study of the human prepubertal and pubertal larynx. Am J Anat 1978; 151:11-20.

Kahane JC: Growth of the human prepubertal and pubertal larynx. J Speech Hear Res 1982;25:446-455.

Kahane JC, Mayo R: The need for aggressive pursuit of healthy childhood voices. Lang Speech Hear Serv Schools 1989;20:102-107.

Kay NJ: Vocal nodules in children - aetiology and management. J Laryngol Otol 1982;96:731-736.

Kearns DB, Grace AR, Parsons DS, et al: Laryngeal characteristics after reconstruction for subglottic stenosis: a clinical correlation. J Laryngo Otol (Suppl) 1988;17:39.

Klatt DH, Klatt LC: Analysis, synthesis, and perception of voice quality vibrations among female and male talkers. J Acoust Soc Am 1990;87:820-841.

Koch BM, Milmoe G, Grundfast KM: Vocal cord paralysis in children studied by monopolar electromyography. Pediatr Neurol 1987;3:288.

Koufman JA: Approach to the patient with a voice disorder. Otolaryngol Clin North Am 1991;24:989-998.

Koufman JA, Blalock PD: Functional voice disorders. Otolaryngol Clin North Am 1991;24:1059-1074.

Klock LE: The growth of the human larynx from birth to adolescence. thesis, Univ Washington, Seattle, 1968.

Klock L, Beckwith JB: Appendix: dimensions of the human larynx during infancy and childhood, in Bosma JF (ed.): Anatomy of the Infant Head. Baltimore, MD, Johns Hopkins University Press, 1986.

Leonard R, Senders C, Charpied G: Effects of long-term intubation on vocal fold mucosa in dogs. J Voice 1992;6:86-93.

Leeper HA, Leonard JE, Iverson RL: Otorhinolaryngologic screening of children with vocal quality disturbances. Int J Pediatr Otorhinolaryngol 1980;2:123-131.

Lotz WK, D'Antonio LL, Chait DH, et al: Successful nasoendoscopic and aerodynamic examinations of children with speech/voice disorders. Int J Pediatr Otorhinolaryngol 1993;26:165-172.

MacArthur CJ, Kearns GH, Healy GB: Voice quality after laryngotracheal reconstruction. Arch Otolaryngol Head Neck Surg 1994;120:641-647.

Maddalozzo J, Holinger LD: Laryngotracheal reconstruction for subglottic stenosis in children. Ann Otol Rhinol Laryngol 1987;96:665-669.

Maddern BR, Campbell TF, Stool S. Pediatric voice disorders. Otolaryngol Clin North Am, 24, 1125-1140, 1991.

McAlister A, Sederholm E, Sundbert J, Gramming P: Relations between voice range profiles and physiological and perceptual voice characteristics in ten-year-old children. J Voice 1994, 8, 230-239.

McHenry MA, Wilson RL, Minton JT: Management of multiple physiologic system deficits following traumatic brain injury. J Med Speech-Lang Path 1994;2:59-74.

McWilliams BJ, Lavorato AS, Bluestone CD: Vocal cord abnormalities in children with velopharyngeal valving problems. Laryngoscope 1973;83:1745-1753.

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

Monday LA, Cornut G, Bouchayer M, et al: Epidermoid cysts of the vocal cords. Ann Otol Rhinol Laryngol 1983;92:124-127.

Moran MJ, Pentz AL: Otolaryngologists' opinions of voice therapy for vocal nodules in children. Lang Speech Hear Serv Schools 1987;18:172-178.

Morrison M, Rammage L, Nichol H, et al: Pediatric voice disorders: special considerations. In The Management of Voice Disorders. San Diego, CA, Singular Publishing Group, 1994, pg 120-140.

Murry T, Woodson GE: A comparison of three methods for the management of vocal fold nodules. J Voice 1992;6:271-276.

Olson NR: Laryngopharyngeal manifestations of gastroesophageal reflux disease. Otolaryngol Clin North Am 1991;24:1201-1213.

Optiz JC, Kettrick MA, Sullivan BJ: Recurrent neonatal herpes presenting initially with hoarseness. Am J Perinatol 1989;6:307-309.

Ossoff RH, Werkhaven JA, Dere H: Soft-tissue complications of laser surgery for recurrent respiratory papillomatosis. Laryngoscope 1991;101:1162-1166.

Phillips DE, Childs D, Walsh S: Hoarse cry with fatal outcome. Br Med J 1989;299:847.

Putnam PE, Orenstein SR: Hoarseness in a child with gastroesophageal reflux. Acta Paediatr 1992;81:635-6.

Sander EK: Arguments against the aggressive pursuit of voice therapy for children. Lang Speech Hear Serv Schools 1989;20:94-101.

Savage MM, Crockett DM, McCabe BF: Lipoid proteinosis of the larynx: a cause of voice change in the infant and young child. Int J Pediatr Otorhinolaryngol 1988;15:33-38.

Sell D, MacCurtain F: Speech and language development in children with acquired subglottic stenosis. J Laryngol Otol (Suppl) 1988;17:35-38.

Senturia BH, Wilson FB: Otorhinolaryngic findings in children with voice deviations. Preliminary report. Ann Otol Rhinol Laryngol 1968;77:1027-1042.

Silverman EM, Zimmer CH: Incidence of chronic hoarseness among school-age children. J Speech Hear Disord 1975;40:211-215.

Smith ME, Gray SD, O'Connor DM, et al: Endolaryngeal trauma associated with long-term endolaryngeal stents (abstract). Otolaryngol Head Neck Surg 1991;105:206.

Smith ME, Mortelliti AJ, Cotton RT, et al: Phonation and swallowing considerations in pediatric laryngotracheal reconstruction. Ann Otol Rhinol Laryngol 1992;101:731-738.

Smith ME, Marsh JH, Cotton RT, et al: Voice problems after pediatric laryngotracheal reconstruction: videolaryngostroboscopic , acoustic, and perceptual assessment. Int J Pediatr Otorhinolaryngol 1993a;25: 173-181.

Smith ME, Darby KP, Kirchner K, et al: Simultaneous functional laryngeal stridor and functional aphonia in an adolescent. Am J Otolaryngol 1993b;5:366-369.

Smith ME, Gray SD. Laryngeal framework surgery in children. In Myers EN, Bluestone CD, Brackmann DE, et al (eds): Advances in Otolaryngology-Head and Neck Surgery, vol 8, St. Louis, Mosby, 1994.

Stathopoulos ET, Sapienza C: Respiratory and laryngeal measures of children during vocal intensity variation. J Acoust Soc Am 1993;94:2531-2543.

Tanner JM: Sequence, tempo, and individual variation in the growth and development of boys and girls aged twelve to sixteen. Daedalus 1971;100:907-930.

Titze IR: Comments on the myoelastic-aerodynamic theory of phonation. J Speech Hear Res 1980;23: 495-510.

Titze IR, Baken R, Herzel H:Evidence of chaos in vocal fold vibration. In I Titze (Ed.) Vocal Fold Physiology: New Frontiers in Basic Science, Singular Publishing Group, Inc., San Diego, CA, pg 143-188, 1993.

Toohill RJ: The psychosomatic aspects of children with vocal nodules. Arch Otolaryngol 1975;101:591-595.

Tucker HM: Vocal cord paralyis in small children: principles of management. Ann Otol Rhinol Laryngol 1986;95:618-621.

Tucker HM: Reinnervation of the paralyzed larynx: a review. Head Neck Surg 1979;1:235-242.

Ward PH, et al: The larynx in the cri du chat (cat cry) syndrome. Trans Am Acad Ophthalmol Otolaryngol 1968;72:90-102.

Ward PH, Berci G: Observations on the pathogenesis of chronic nonspecific pharyngitis and laryngitis. Laryngoscope 1982;92:1377-1382.

Wetmore SJ, Key JM, Suen JY: Complications of laser surgery for laryngeal papillomatosis. Laryngoscope 1985;95:798-801.

Wirz S: The voice of the deaf. In Voice Disorders and their Management, 2nd Ed. (M Fawcus Ed.) Singular Publishing Group, San Diego, CA, 1992.

Zalzal GH: Rib cartilage grafts for the treatment of postrior glottic and subglottic stenosis in children. Ann Otol Rhinol Laryngol 1988a;97:506-511.

Zalzal GH: Use of stents in laryngotracheal reconstruction in children: indication, technical considerations, and complications. Laryngoscope 1988b;98:849-854.

Zalzal GH, Loomis SR, Derkay CS, et al: Vocal quality of decannulated children following laryngeal reconstruction. Laryngoscope 1991;101:425-429.

Zalzal GH, Loomis SR, Fischer M: Laryngeal reconstruction in children: assessment of voice quality. Arch Otolaryngol Head Neck Surg 1993;119:504-507.

NCVS Status and Progress Report - 7
December 1994, 171-186

Laryngeal Function During Phonation

Ronald Scherer, Ph.D.
Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts

Introduction

The larynx performs many functions to aid communication and allow life. As an open flow valve, it permits breathing, blowing, and sucking, as well as yawning, voiceless consonant production, and musical instrument playing. As a transient closed valve it produces coughing and throat clearing. As a constant closed valve it participates in swallowing and effortful behaviors such as lifting and defecation. As a voiceless repetitive articulator, it valves airflow to produce staccato whistling. As a voiced repetitive articulator it produces laughter, the singing ornament trillo, and repetitive glottalization of vowels such as the admonition with rising pitch and intensity "a-a-a-ah!". As a partially closed voiceless valve, it produces whisper. As a partially closed voicing valve, the larynx produces vowels and prolonged voiced consonants. And as a speech coarticulator, it participates in the production of consonant-vowel strings. For example, the word "seat" begins with an open glottis for the /s/, the glottis acting as an open flow valve allowing airflow under lung pressure to travel through the glottis and through the anterior oral constriction. This is followed by vocal fold approximation for the /i/ vowel, the larynx acting as a partially closed voicing valve. And finally, the glottis opens as an open flow valve permitting impoundment of air pressure behind the anterior oral /t/ occlusion with the subsequent release of air, creating the characteristic aspiration.

The purpose of this chapter is to describe the larynx as a partially closed voicing valve, that is, as the organ that produces phonation. Neuromuscular behavior and biomechanical characteristics determine phonation duration, pitch, loudness, quality, register and vocal fold motion, through control of or changes in vocal fold length, vocal fold contour, arytenoid and vocal fold adduction, subglottal pressure, vocal fold tissue normalcy, and vocal tract shape. This chapter emphasizes the reasonable hypothesis that effective intervention to help people with voice concerns (whether these involve rehabilitation, surgery, pharmacology, training, or prevention) can be improved by an understanding of basic vocal function (1,2).

Overview of Basic Characteristics of Phonation

The basic perceptual characteristics of phonation are the presence and change of duration, pitch, register, loudness and quality. Each of these has one or two primary biomechanical control variables. These will be discussed briefly in this section, followed by discussions of selected topics in greater detail.

Duration of phonation refers to the length of time the vocal folds oscillate during the creation of sound. In the normal larynx, adduction is one of two main control variables; the vocal folds must be sufficiently close to permit oscillation. Subglottal pressure is also necessary so that there is sufficient force to move the vocal folds laterally at the beginning of each vibratory cycle. After phonation commences, to then cease phonation, the arytenoid cartilages can be moved apart (abducted), or moved further together (more highly adducted), both ceasing phonation if the degree of abduction or adduction is great enough. Although the simplest statement that explains the cessation of phonation might be that the glottal flow resistance is too low during the abducted state and too high during the adducted state, a better explanation is that there is insufficient subglottal pressure and too much vocal fold deformation and displacement to allow the mechanical possibility of vocal fold oscillation for either case.

Pitch perception, corresponding to the physical measure of fundamental frequency, and vocal register (the very low pitches of vocal fry, the conversational pitches of chest or modal register, and the very high pitches of falsetto), are highly dependent on vocal fold length (3,4) and the associated tension of the vocal fold mucosal cover. The string model for frequency (5,6,7,8),

$$F_o = \frac{1}{2L}\sqrt{\frac{T}{\rho}} \qquad (1)$$

where T is the tension of the vocal fold mucosal cover, rho is the density of the tissue, and L is the length of the vibrating vocal folds, is an explanatory model suggesting that tension of the vocal fold cover governs the fundamental frequency (tension T increases faster than vocal fold length L as length increases, and rho is essentially constant regardless of phonatory condition; 9,8).

The loudness of sounds corresponds to their acoustic intensity, and for phonation is primarily dependent upon subglottal pressure. An increase in subglottal pressure changes the characteristics of the airflow that exits the glottis (the glottal volume velocity) during vocal fold vibration, creating an increase in acoustic intensity (see below).

At the glottal level, vocal quality variation (separate from tissue or neurologic abnormalities) is primarily related to the perception of normal, breathy, and pressed (constricted) qualities. These are governed primarily by the closeness of the vocal folds, i.e., by adduction. Breathiness occurs when the vocal folds are slightly abducted such that they do not close completely during each cycle, allowing some of the glottal volume flow to be unmodulated and turbulent (10). A breathy voice also can be created by full vocal fold closure each cycle but with the posterior glottis open, allowing turbulent air to flow through the posterior glottis. If there is hyperadduction of the vocal folds, but with a significant opening of the posterior glottis, a pressed-breathy (or pressed-leakage) quality can result (11, cited and discussed in 12). Pressed or constricted voice quality without breathiness results from full glottal adduction, and little air flows through the glottis (13).

Intervention strategies to improve voice production require an appreciation of the basic mechanics of phonation. The remainder of this chapter offers a more thorough discussion of basic control characteristics of voice production.

Duration of Phonation

Figure 1 illustrates the adductory range of phonation and the corresponding distances between the vocal processes of the arytenoid cartilages, the ventricular folds, and the supero-medio eminences of the arytenoid cartilage apexes. Phonation takes place over only a small range of adduction (the phonatory adductory range) permitted by movement of the arytenoid cartilages (approximately 14% of the adductory range as shown in the example of Figure 1). From an adductory standpoint, therefore, duration of phonation depends on the potential for the vocal folds to be placed within the phonatory adductory range, and the length of time the vocal folds are actually placed within that range. To discontinue phonation, the arytenoid cartilages can be configured to produce sufficient over-compression of the vocal folds or, alternatively, sufficient abduction.

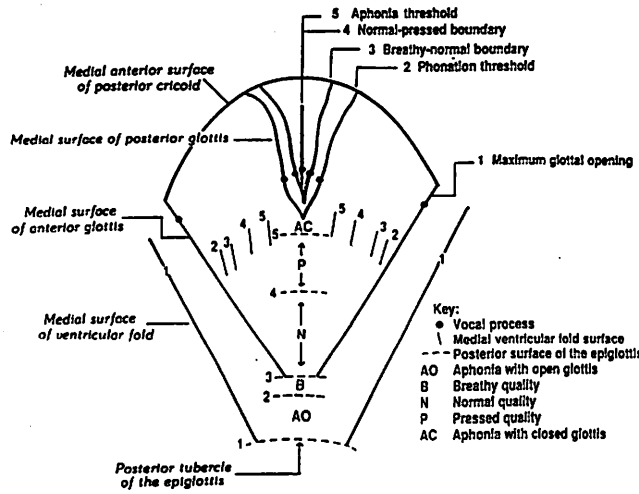


Figure 1. Adductory Range of the Larynx. This figure is a composite from photographs of glottal adductory positions obtained from a single adult male subject using rigid videolaryngoscopy. Positions of the medial surface of the anterior (membranous) glottis, posterior glottis, ventricular folds, and the posterior tubercle of the epiglottis are shown. Stages of glottal positioning include maximum glottal opening (position 1) to full adduction when phonation is not possible (position 5). The other positions (2, 3, and 4) are the position for adductory phonatory threshold and boundary locations between breathy, normal and pressed voice qualities in this subject, perceptually judged and checked against EGG recordings. The adductory phonation range is only about 14% of the entire adductory range (see text).

Phonation requires a certain minimal amount of subglottal air pressure to set the vocal folds into vibration (the phonatory threshold pressure, 14) and to maintain phonation (this maintenance phonation pressure is slightly less than the threshold pressure; 15,16,17). If the vocal folds are placed in the phonatory adductory range, the subglottal pressure must overcome the stiffness and damping of the vocal folds (14) to literally push them laterally and superiorly from below to start the first cycle. The threshold pressure typically varies with fundamental frequency, ranging from about 3 cm H_2O (0.3 kPa) for lower pitches to about 6 cm H_2O (0.6 kPa) for higher pitches (18) due to greater tension of the vocal fold cover with vocal fold lengthening.

There also may be an upper limit of subglottal pressure beyond which phonation is prevented or too chaotic to maintain normal voice quality (14). This upper limit should correspond to the upper limits of the phonetogram, a graph of the intensity range versus the fundamental frequency range of phonation for a particular person (19,20,136,21,22; the recommendation of the term Voice Range Profile has been accepted by the Voice Committee of the International Association of Logopedics and Phoniatrics, 123). Subglottal pressures have been measured as high as 30-50 cm H_2O in loud phonation (24; singing loudly at high

pitches may produce even greater subglottal pressures, 13), but typically is below 10 cm H_2O for conversational speech (24,25).

Too little subglottal pressure prevents the start of phonation and also can create the cessation of ongoing phonation. Intervention in voice surgery, therapy and training often attempts to regain normal thresholds or to lower existing thresholds, the latter possibly creating less effortful phonation (relative to the employed forces of the respiratory system). Establishment of lower phonation thresholds may correspond strongly to more physiologically efficient phonation.

The discussion directly above refers to phonation dependence upon subglottal pressure. More appropriately, phonation is dependent upon the translaryngeal pressure, defined as the subglottal air pressure minus the supraglottal air pressure for phonation on exhalation. The translaryngeal pressure determines the difference between the subglottal and supraglottal pressures acting on the inferior and superior surfaces of the vocal folds, respectively, and also drives the air through the glottis. If the translaryngeal air pressure were zero, the pressure would be equal on all surfaces of the vocal folds, air would not pass, and phonation would not exist. This can be approached, for example, with overly-prolonged voiced consonants such as /b/, /d/, and /g/ during their (total) occlusion within the vocal tract. Prolonging the voicing of these consonants allows build up of supralaryngeal air pressure until that pressure nearly equals the subglottal pressure, causing cessation of phonation as the translaryngeal pressure drops below the minimum sustaining pressure.

Thus, the creation and duration of phonation depend upon how close the vocal folds are to each other and the amount of translaryngeal air pressure. To cease phonation, the vocal folds can be over adducted or under adducted, or the translaryngeal pressure can be lowered by decreasing the subglottal pressure or by increasing supraglottal pressure through supraglottal occlusion. All four methods most likely are used in normal speech production. These mechanisms for phonation cessation potentially can be compromised by arytenoidal, respiratory, and articulatory dysfunction, or by abnormal adductory configuration due to vocal fold tissue change.

Fundamental Frequency of Phonation

Perceived pitch corresponds (nonlinearly) to the physical measure of fundamental frequency Fo (26), which in turn corresponds to the number of cycles per second of the glottal motion during phonation. For normal phonation, the motion of the vocal folds is similar from cycle to cycle, giving rise to nearly equal periods of time between glottal closures.

Each phonatory cycle releases a time-varying glottal flow (also called the glottal volume velocity) that generates sound. For normal phonation, there is consistent creation of sound typically at a primary location within the cycle's volume velocity waveform. Figure 2 shows two glottal volume velocity cycles of human phonation (top trace). The cycle period is T, which is 10 ms (and thus Fo = 1/T = 100 Hz). The glottal volume velocity (usually given in liters per second, L/s, or cubic centimeters per second, cm^3/s) begins to exit the glottis gradually, rises to a peak, and then "shuts off" relatively abruptly. Air exits the glottis from time A to time B, but the glottis is closed, or nearly so, from B to C.

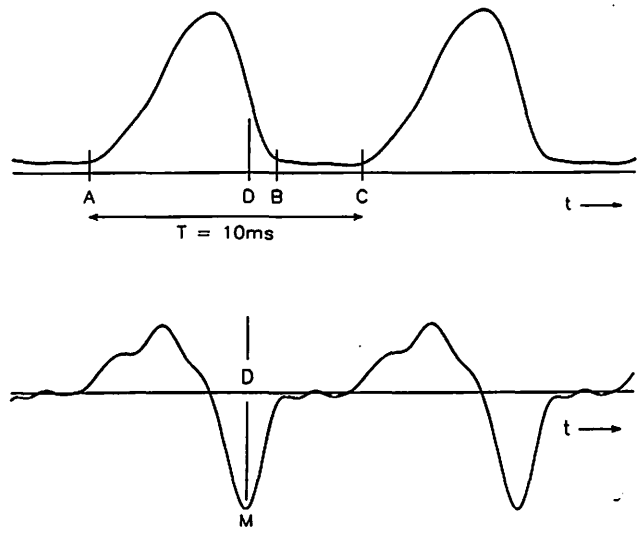
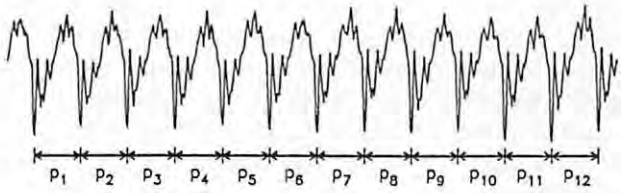


Figure 2. Glottal Volume Velocity Waveforms and Their Derivatives. The top trace is a glottal volume velocity signal, and the bottom trace is the derivative of the top trace, showing the instantaneous slopes of the top trace. Cycle period is T (from time A to time C). "Open" glottis time is from A to B, and "closed" glottis time is from B to C. The moment of the maximum flow shut off rate, or the maximum negative derivative of the glottal volume velocity signal, occurs at location D and corresponds to point M.

The lower trace of Figure 2 is the time derivative of the volume velocity signal of the upper trace. At any moment in time, the value on the lower trace equals the slope of the volume velocity signal at that moment. The fastest change of the volume velocity is near location D, and corresponds to the point M on the derivative waveform. The point M corresponds to the moment of time at which the greatest acoustic excitation is created (27,28).

Perceptual judgements of an unclear voice (rather than confusion of pitch per se) occur when the prominent moments of acoustic excitation during each cycle are not consistent from one cycle to the next (for a review of cyclic instabilities, ref. 26; also ref. 29,139). The time between primary acoustic excitations from one cycle to the next may vary (creating aperiodicities) if there are tissue abnormalities such as nodules, polyps, unilateral stiffness abnormali-



Microphone Signal

$P_1 = 6.60 \rightarrow Fo = 151.52$ Hz
$P_2 = 6.55 \rightarrow Fo = 152.67$ Hz
$\vdots$
$P_{12} = 6.50 \rightarrow Fo = 153.85$ Hz

N=12

Jitter :

$$J = \frac{100 \sum_{i=2}^{N} |Fo_i - Fo_{i-1}|}{(N-1)\overline{Fo}}$$

$$= 0.82$$

Figure 3. Aperiodicity of a Microphone Signal for a Prolonged Vowel. The periods of the microphone signal are indicated, and the calculation of the corresponding jitter is given. This is an example of a calculation of voicing perturbation.

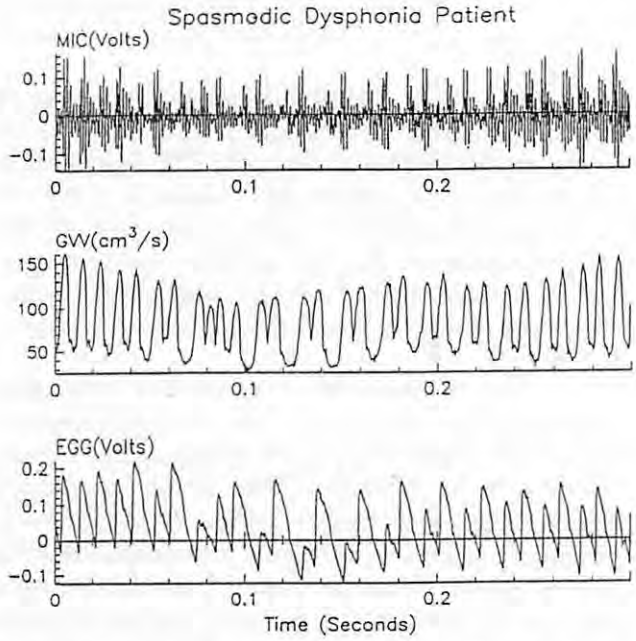


Figure 4. Voicing Period Alterations in a Spasmodic Dysphonia Patient. Microphone, glottal volume velocity, and EGG signals are given, top to bottom. Notice the change from essentially single cycles to cycle clustering of two (and one case of three) cycles. The clustering of two cycles can be seen in all three signals: apparent motion variation of vocal fold contact, double flow pulsing, and double acoustic excitation seen in the microphone signal. (Data courtesy of Dr. Kimberly Fisher).

ties, etc., apparently causing kinematic (vocal fold motion) inconsistencies from cycle to cycle. Consecutively varying periods of primary acoustic excitations also can be created by turbulent airflow through the glottis (as in breathy voice), creating added noise to the acoustic signal. Aperiodicities can be measured by jitter, one definition being the average cycle-to-cycle difference in period or equivalent frequency, with greater jitter corresponding to a greater sense of vocal roughness (ref. 26). Figure 3 illustrates quasi-periodic cycles of the acoustic output of a prolonged vowel with the calculation of the segment's jitter value.

Jitter may be caused by or related to neuromuscular innervation abnormalities (137,31,32,33) as well as the structural and turbulence causes given above. Because tension of the vocal folds is highly dependent on the passive lengthening of the vocal fold cover, as suggested by Equation 1, relatively fast abnormal innervation changes to the cricothyroid muscles may cause length and therefore tension changes, creating cycle to cycle fundamental frequency changes.

Pitch and vocal clarity are also affected by changes that occur over longer time lengths than a usual phonatory cycle. Diplophonia (the existence of two pitches simultaneously) and "subharmonics" (essentially integer subdivisions of the fundamental frequency) come from multi-cycle length modulations of the volume velocity signal, giving rise to primary acoustic excitations at varying time intervals as well as about twice (or more) the primary phonatory period (34,35,36,37,38,39,40). Figure 4 shows an example of the presence of a "subharmonic" in the glottal volume velocity signal of a spasmodic dysphonia patient. The cause of this dysfunction appears to be related to hyperadduction or asymmetry of laryngeal muscle function.

In vocal fry and "creaky" voice, pitch may be extremely low, dicrotic, or of varied roughness qualities depending on the complexity of the low frequency periodicities combined with higher frequency periodicities (41,42,43,44). Figure 5 illustrates three different examples of vocal fry (ref. 44; in that paper, 11 different types of vocal fry and creaky voice were illustrated). For each case, the microphone signal and the electroglottograph signal are displayed. These samples were created intensionally by a single normal adult male subject. The electroglottographic (EGG) signal corresponds significantly to the contact area between the vocal folds when they touch each other during each cycle (45,26). Figure 5a refers to vocal fry of low and specific pitch. Each cycle shows a single primary acoustic excitation. This excitation corresponds to the fast upward movement of the EGG signal which corresponds to glottal closure and therefore to the glottal volume velocity "shut off". Figure 5b illustrates the typical bimodal electroglottograph signal for vocal fry, with corresponding double excitations during each (low pitch) cycle. Figure 5c illustrates an extreme case of multiple motions of the vocal folds each cycle, with corresponding acoustic excitations, with two prominent frequencies, 7.5 Hz and 93 Hz.

Pitch can be altered by fluid engorgement (edema) conditions. The usual explanation for pitch drop in edema cases is that greater mass creates lower natural frequencies. For example, computer modelling of phonation using biomechanical characteristics of tissue mass, stiffness, and

damping relate fundamental frequency to the inverse of the vocal fold mass (e.g., 46). Figure 6 shows the results of increasing vocal fold mass (only) in the two-mass model of vocal fold function adapted from Ishizaka and Flanagan (47) by Smith (48; figure courtesy of Dr. Marshall Smith). Figure 6 shows a decrease in frequency by 5.6 semitones for a doubling of the mass of the vocal folds. The figure also suggests that if mass is increased, maintaining a desired fundamental frequency requires greater subglottal pressure if vocal fold length remains constant.

Subglottal pressure plays a significant role in the control of pitch, as just indicated by reference to Figure 6 (see also 30,49,50,51; data show that a change of 1 cm H_2O subglottal pressure results in an Fo change of 3-6 cm H_2O). The air pressure pushing against the undersurface of the vocal folds during the closed phase of the phonatory cycle deforms the vocal fold laterally and somewhat upwardly (52,53,16,54,55). The extent of excursion laterally is dependent upon the amount of the subglottal pressure and length of the vocal folds (6,56,57). Thus, for a constant anterior-posterior length of the glottis, a greater subglottal pressure will literally push the vocal folds to a greater lateral extent, creating a greater maximum stretch than for a lower subglottal pressure. Greater maximum stretch creates higher effective tension and thus a higher fundamental frequency (6). Intonational (i.e., pitch) changes during conversational speech appear to be due to a significant combination of the passive vocal fold stretch by both the cricothyroid muscles and the subglottal pressure (58,59; ref. especially to the general discussion at the end of the latter paper).

Tension of the vocal fold cover may be changed (and therefore fundamental frequency changed) by external adjustments affecting the length of the vocal folds. Anterior pull of the hyoid bone by suprahyoid muscles may help tilt the thyroid cartilage forward to position it closer to the anterior cricoid cartilage (similar to the function of the cricothyroid muscle), thus increasing vocal fold length and raising the fundamental frequency, as shown by Honda (60). Alternatively, the cricoid cartilage may be tilted down, shortening the vocal fold length, by an inferior tracheal pull from diaphragm lowering with higher lung volume levels (or by coactivation of the diaphragm during phonation), inducing a pitch drop unless compensated for by increased cricothyroid muscle activity (61).

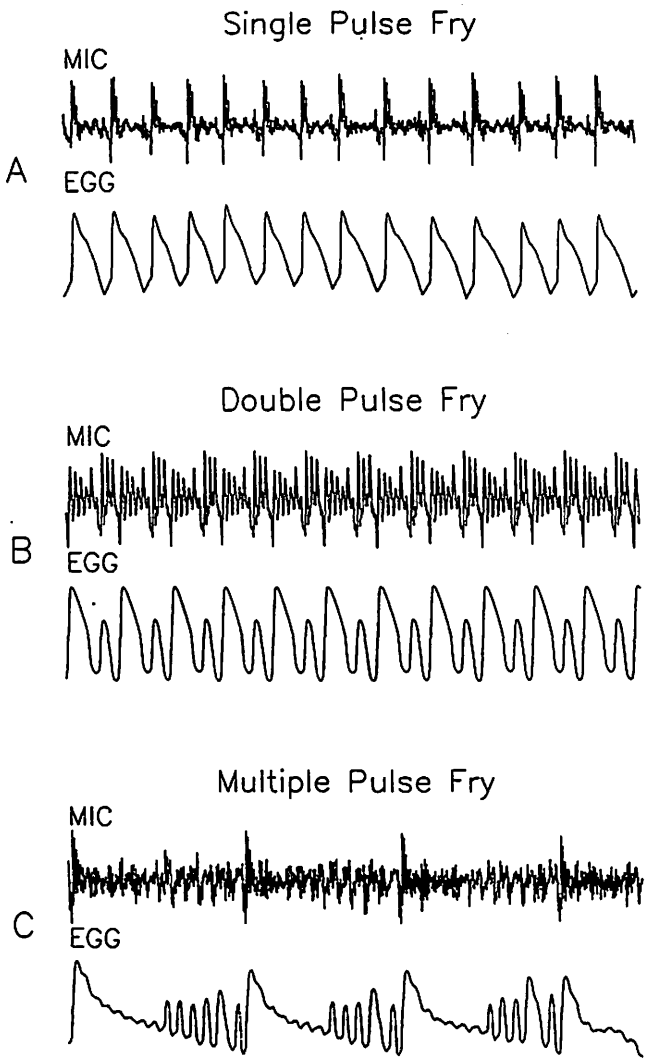


Figure 5. Vocal Fry. Examples of three different vocal fry: A, single glottal pulses at 30 Hz; B, double glottal pulses at 62 Hz for the period combining the two pulses; and C, a multiple pulse case with a low frequency of 7.5 Hz and a higher frequency of 93 Hz. In each grouping, the upper trace is the microphone signal and the lower trace is the electroglottograph signal. The examples were produced by a single adult male subject. See (44).

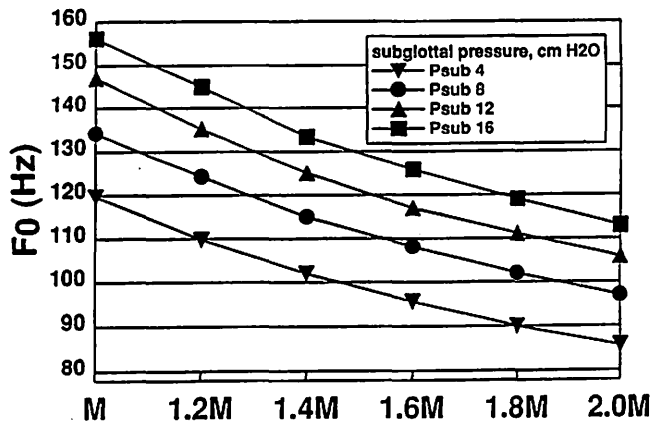


Figure 6. Vocal Fold Mass Effect on Fundamental Frequency. As the amount of mass of the vocal fold in vibration increases, the fundamental frequency decreases. The abscissa ranges from a value of M=0.16 gm to twice that. The figure also indicates that fundamental frequency is dependent upon subglottal pressure. Figure courtesy of Dr. Marshall Smith, who used an adaptation of the Ishizaka and Flanagan simulation (47,48).

The above discussion emphasizes the contribution to pitch control through tension change of the vocal fold cover, whether through passive stretch by contraction of the cricothyroid muscle, or dynamic stretch by increasing subglottal pressure. The fundamental frequency is also dependent upon the activity of the vocalis (thyroarytenoid) muscle (62,63). The vocalis acts antagonistically to the cricothyroid muscle relative to length change of the cover (64,65,62). Thus, if the cover only is vibrating, as in high soft pitch, increase in vocalis contraction should shorten and reduce the tension of the mucosal cover, and thereby lower the fundamental frequency. In addition, if the vocalis muscle participates in the motion of the vocal fold to a significant degree, as in loud low pitch phonation, increase in vocalis muscle contraction will increase the effective tension of the entire tissue in motion as the primary effect, and thereby raise the fundamental frequency. At intermediate pitches and loudness levels, where the cover and vocalis muscle both participate in the vibration of the vocal folds, pitch could be controlled or even remain constant for many combinations of cricothyroid muscle contraction (to passively stretch the cover), vocalis muscle contraction (to passively shorten the cover and actively increase the tension of the vibratory vocalis portion), and subglottal pressure (to vary the amount of vocal fold mass placed into vibratory motion). The effect on Fo change related to vocalis muscle involvement conceptually depends, then, on the relative amount of vocalis muscle participation in the vibratory mass, the tension within the vocalis muscle portion of the mass in motion, and the relative activity level of the vocalis muscle, as Titze et al. (62) analytically describe. The control of pitch production (as in intentional expression and performance) therefore relies primarily on cricothyroid and vocalis muscle contraction as well as subglottal pressure production (62,57). Intervention strategies dealing with pitch alteration need to take into consideration these physiologic and aerodynamic bases.

Various combinations of cricothyroid and vocalis muscle contraction, with adduction, may change vocal quality, especially relative to the medial shaping of the vocal fold. Greater vocalis muscle contraction tends to round (or medialize) the medial contour of the vocal fold (66,67,63), potentially decreasing the vibratory threshold pressure (14) and creating a relatively longer closed time of the vocal folds each cycle (68), which may change the voice spectra to reflect a brighter, louder sound.

Loudness and Quality of Phonation

Loudness and quality of phonation are perceptual correlates to the physical measures of intensity and acoustic spectra, respectively (69,26,13,70). Both perceptions depend upon the glottal volume velocity waveform characteristics.

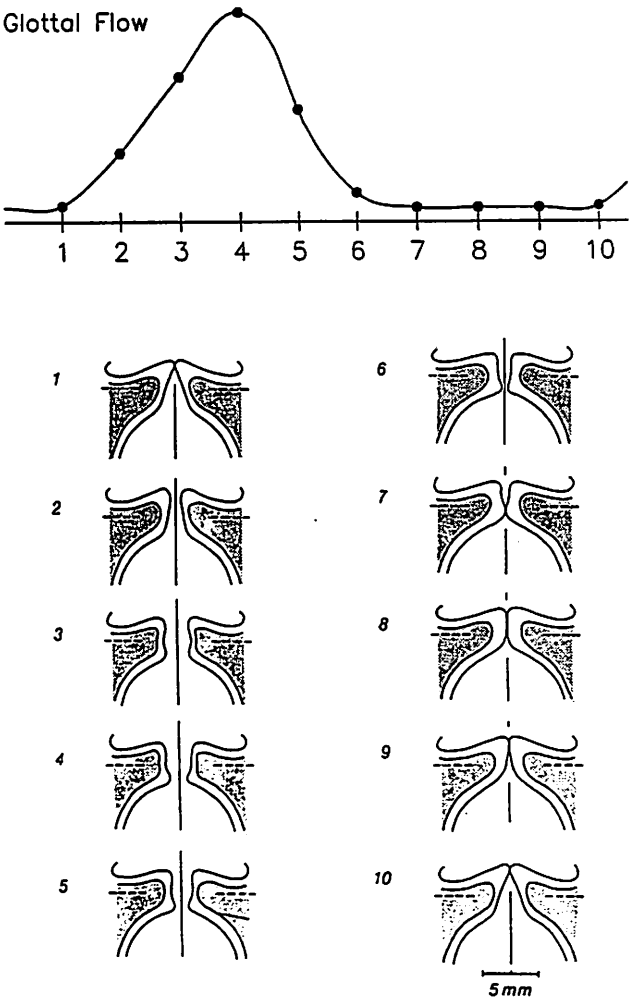


Figure 7. Glottal Volume Velocity Waveform and Corresponding Glottal Motion. The specific phases of the glottal cycle shown in the motion schematic are indicated on the glottal volume velocity waveform. (Adapted with permission from Hirano, 24)

Figure 7 illustrates a typical glottal volume velocity waveform with corresponding glottal motion. The relationship is schematic only. The general shape of the glottal volume velocity waveform shows that the flow typically begins more gradually than it is shut off, and the flow maximum is produced after the maximum of the glottal area occurs. This flow delay (or skewing to the right) characteristic (relative to the glottal area) is related primarily to the vocal tract inertive effect (71,72). The inertive effect refers to the fact that the air within the vocal tract has mass. When the glottis just opens, the air (being driven by the translaryngeal air pressure) moves through the glottis to meet the column of air of the vocal tract. The air coming through the glottis must literally move other air already within the vocal tract, and this requirement slows the motion of the air as it first comes out of the glottis (73,2). Corresponding to this event is the increase in air pressure just above the glottis as air moves through the glottis into the air

above, thus typically reducing the translaryngeal pressure drop (74,75,76,77,78,2). If the vocal tract were modelled as a uniform tube, greater glottal air flow skewing could be created from greater inertance by elongating the vocal tract (through larynx lowering or lip protrusion) or by narrowing the cross-sectional vocal tract area (79). Fant (75) analytically showed that skewing increases not only with an increase of the vocal tract inertance (which changes with certain vowels and is higher with a constriction at the false fold level), but also with an increase in the maximum glottal excursion during a phonatory cycle, a faster glottal closing time, a smaller subglottal pressure, and a smaller glottal kinetic flow factor k (the latter is derived in detail in 80 and 81), concepts fruitful for clinical and training considerations. After the moment of maximum flow, the airflow will reduce to zero or to its minimum value as the two vocal folds come together at the end of glottal closing.

There is a supplemental hypothesis that might contribute to the occurrence of the maximum flow taking place after the moment of maximum glottal opening. As Figure 7 shows, the glottis takes on a number of shapes during the vibratory cycle. A duct that expands in shape (a diffuser shape as seen in steps 4, 5 and 6 of Figure 7) can have less resistance to flow than one of the same but constant (uniform) diameter (82,83). The minimum flow resistance may occur when the glottis creates a diffuser shape of small angle (estimated to be 7 to 10 deg by Guo & Scherer, 84), which would most likely occur just past the time of the maximum glottal area. Therefore, when the glottis takes on the shape of a small-angled diffuser, the flow resistance may be less than at maximum glottal opening, and greater flow may then exit the glottis, supplementing the inertance effect of the vocal tract to skew the flow to the right relative to glottal area. The relative contribution of the diffuser glottal shaping with reduced flow resistance needs further investigation.

The glottal volume velocity waveform varies depending on the qualities of the voice. Figure 8 show an example of breathy, normal and pressed phonations in a normal adult male. The figure demonstrates typical variations of the glottal flow, the glottal flow spectrum, and the electroglottograph (EGG) waveforms for these qualities (the signals are not time aligned). A Glottal Enterprises wide-band pneumotach system was used to acquire the inverse filtered (glottal) flow, and a Synchrovoice Laryngograph was used to obtain the EGG signal. Breathy voice is characterized by a more sinusoidal flow waveform than for normal phonation, with a significant flow bias. The flow bias is seen as a shift of the waveform away from the zero flow baseline, indicating that there is always some flow exiting the glottis due to non-closure throughout the cycle (the bias is also called the "DC flow"). In pressed phonation, the peaks of the flow are significantly smaller than for normal phonation, and the amount of time within the cycle during which the air exits the glottis is relatively short compared to normal phonation, as indicated in the figure. Spectrally, the breathy quality example has energy primarily in the first two partials, whereas greater energy is distributed to higher frequencies in the normal and pressed quality examples. Up to 1000 Hz, the spectral slope is relatively steep at -17.3 dB/octave for breathy, less steep at -14.4 dB/octave for normal, and flattest at -10.8 dB/octave for the pressed condition. Notice the relative reduction in the level of the fundamental frequency compared to the first overtone from breathy to normal to pressed. Electroglottographically, the EGG waveforms suggest that there was a relatively large amount of contact (area) between the surfaces of the two vocal folds during the breathy voicing, as seen by the relatively large amplitude of the breathy waveforms (note that the higher the EGG waveform value, the more vocal fold contact area there is presumed to be, 45). However, the time during which the glottis was open, shown by the baseline length of the waveform, was relatively long compared to the other two qualities. The height of the pressed quality EGG waveform is shortest, suggesting relatively less dynamic contact of the vocal fold surfaces. This was explained by viewing the subject's larynx with stroboscopy; in pressed voice, the compression allowed only a restricted anterior glottal region to vibrate, thus resulting in the relatively short amplitudes because less of the total medial vocal fold surface participated in the vibration.

The intensity and spectrum of vocal sound depend on the glottal volume velocity waveform, as suggested in Figure 8 (next page). The overall intensity or sound pressure level (SPL) of the output sound will increase as the maximum rate of change of the glottal volume velocity "shut off" (literally the maximum negative derivative value of the glottal volume velocity waveform) increases (87,12,85,86,87). Greater maximum slope of the glottal airflow waveform has the spectral effect of raising the energy of the partials primarily within the region of the first formant (88), usually the most important spectral portion for overall SPL, and thus is like "turning [up] the volume control" (12, p. 562; see also 89). A doubling of the maximum negative flow derivative (doubling the slope of the flow "shut off") corresponds empirically to an approximate increase of 5-9 dB in overall SPL (12). Notice that an increase in the maximum flow derivative as a result of a vocal tract acoustic inertance effect, and not as a result of greater vocal effort of some kind or as a laryngeal configuration effect, is relevant to intervention strategies. The effect on overall SPL and source spectra constitute an interaction between the glottal flow source and the configuration of the vocal tract. A more efficient voice, therefore, may employ beneficial vocal tract shaping for intensity and quality control, although methods to do so need investigation.

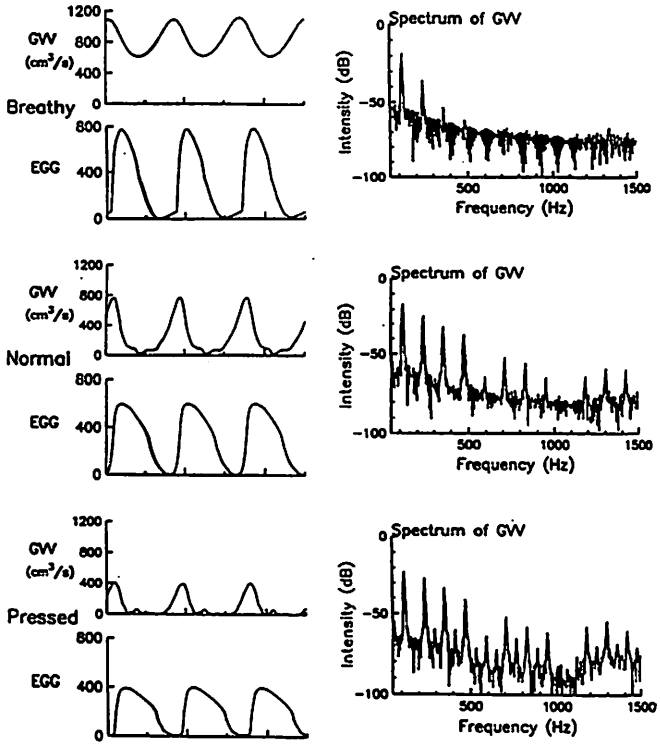


Figure 8. Breathy, Normal and Pressed Voice Qualities. Glottal volume velocity and EGG waveforms are shown, as well as the spectrum of the glottal volume velocity waveforms. The subject was a normal adult male.

It is of importance to note some other spectral effects of shaping differences of the glottal volume velocity waveform. If the amount of time the waveform shows air exiting the glottis (time A to time B in Figure 2) is reduced but with the same period of time (time A to time C in Figure 2) for the entire glottal cycle (i.e., a smaller open quotient, not dissimilar from changing adduction from a more breathy to a more normal quality voice), there will be a minor reduction (a few dB) of the intensity of the fundamental frequency and possibly a minor boost (a few dB) of the intensity of the first overtone (an octave above the fundamental frequency) (90,89). Also, the greater the amplitude of the volume velocity waveform (or the greater the area under the volume velocity waveform), the greater is the amplitude of the fundamental frequency (91,88,12). A doubling of the amplitude of the waveform corresponds to an increase of approximately 3-7 dB in the spectral level of the fundamental frequency. When the flow has nearly completely shut off, that is, when the flow has nearly reached the baseline just before glottal closure, there is a "shut off corner" that defines how sharp the final shut off is. The sharpness of this corner is relatively important to the energy generated in the overtones of the voice, according to the modeling by Fant and his colleagues (88,90,89). A very sharp corner versus a well rounded corner can cause the intensity of the overtones to differ by up to 10-20 dB, undoubtedly affecting the quality of the sound (90,89; changing glottal adduction from a normal voice quality to a breathy quality would round this flow shut off corner more). This concept needs exploration relative to how the vocal folds come together and to the perception of vocal quality, especially taken in the context of clinical intervention and vocal performance instruction.

Important variables controlling intensity and spectra of the glottal airflow are subglottal pressure and fundamental frequency. As subglottal pressure increases for a constant level of glottal adduction, the maximum airflow through the glottis increases. This follows from both aspects of greater maximum glottal width (see discussion above) and greater driving pressure across the same or larger size orifice opening (see e.g., 92). As the maximum value of the volume velocity waveform increases, the greater is the intensity level of the fundamental frequency, as discussed above (12). Also, if the peak flow increases, the maximum flow derivative should typically increase since the flow must reduce to zero (or near zero) from a greater value if the time during which the flow decreases remains the same (87). An increase in the maximum flow derivative will increase the overall SPL and augment the spectrum, as indicated above. In addition, the increase in subglottal pressure may cause the vocal folds to come back together faster (or perhaps alter their dynamic phasing) after their larger maximum excursion, creating a sharper flow shut off corner near baseline, raising the overall spectrum as discussed above. Therefore, greater subglottal pressure may contribute to increasing the flow peak, increasing the maximum flow derivative, and sharpening the baseline flow shut off corner. These effects change the flow spectrum shape by increasing the intensity level of the fundamental frequency and increasing the intensity of the voicing overtones, thus raising the overall intensity of the voice. Titze and Sundberg (86) show explicitly how a doubling of subglottal pressure (more precisely, a doubling of the difference between subglottal pressure and threshold pressure) raises source acoustic power by 6 dB. Fant (75) broke down the intensity increase relative to the contributions of increased air velocity, increased maximum glottal area during the cycle, reduced time the glottis is open, faster glottal closing, and an increased flow derivative, totalling approximately 6 dB in intensity gain for a doubling of subglottal pressure. That is, historically it has been shown that voice intensity has been strongly associated with subglottal pressure (93,94,95,96,97), and this discussion has attempted to give at least a partial explanation as to how the pressure increase affects the sound source volume velocity waveform and the resultant increase of intensity, based on a physical interpretation.

Intensity is affected strongly by the fundamental frequency of voice production. Titze and Sundberg (86)

explicitly show that the glottal power output will increase by 6 dB for an octave rise in fundamental frequency (all else the same), due to the dependence of the maximum flow derivative on the fundamental frequency. At this point one might ask why females, who speak at about 9-10 semitones above the fundamental frequency of men (26), typically do not sound louder than men, and indeed do not differ substantially from men in intensity (98). The higher frequency for the women (about a ratio of 1.7 to 1) is offset by a larger amplitude of glottal volume velocity (see above) for the males (a ratio of about 2 to 1) so that the SPL for females is only 1-2 dB lower than for males (86,98).

Because vocal sound is created by the glottal volume velocity, intervention strategies should attempt to improve the glottal airflow waveform. There may be combinations of voicing variables that produce optimally efficient voice production from acoustic and physiologic orientations. Titze (99) has demonstrated maximal intensity production in excised dog larynxes for vocal processes that are placed very near each other, with less intensity (ranging to a few dB) for greater and less adduction. Sundberg (13; see also 91) has proposed the term "flow mode", a type of phonation in which the glottal volume velocity amplitude is relatively large, suggesting relatively high efficiency of laryngeal function. This is further emphasized, perhaps even diagnostically, by examining the value of the flow amplitude divided by the subglottal pressure (this ratio is called the glottal permittance, ref. 87). This ratio allows a clear separation between a pressed (constricted) voice versus normal and "flow" phonations (for a limited number of subjects, 87) because of the greater flow amplitudes for the same subglottal pressure in the normal and "flow" types. It is of considerable importance (theoretically and clinically) that professional (classically trained) male singers produce SPL levels about 12 dB greater than male nonsingers for the same subglottal pressures, due primarily to glottal flow amplitudes that are 3-4 times greater (86). In addition, resonance training strategies have emphasized vocal tract adjustments to match lower voicing partials with the first formant (100) or to create an enhanced higher formant region by clustering formants number 3, 4 and 5 (101,102), strategies that boost the output energy of certain partials of the voice, creating a "bright" or "carrying" voice with desirable performance qualities. These tactics are not dissimilar to strategies employed in voice pathology (e.g., 103).

Vocal Fold Motion During Phonation

Although it is important to realize that the creation of the sound of voicing is within the airflow exiting the glottis, and not in the motion of the vocal folds per se, the motion of the vocal folds helps to determine the characteristics of the airflow.

Basic kinematic (motion) aspects of the vocal folds include the effects of increased subglottal pressure, vocal fold elongation, and glottal adduction. Increased subglottal pressure produces increased lateral and vertical pressures on the undersurface of the vocal folds, creating greater maximum excursion during the vibratory cycle. Elongation of the vocal folds by increased contraction of the cricothyroid muscles modifies the cross sectional shape of the vocal fold (with no change of total vocal fold mass) to form a thinner vertical vocal fold thickness (104,105,106,107), giving rise to potentially less vocal fold tissue contact during vocal fold closure. For the same subglottal pressure and glottal adduction, this would affect the glottal volume velocity waveform by increasing the duty cycle (the open quotient) and decreasing the maximum flow derivative. Increased adduction will affect the motion of the vocal folds by creating more contact between the vocal folds and a longer glottal closed time within each cycle (40, 108). Any abnormality of tissue morphology (swelling, stiffness or growth along the vocal fold), unilaterally or bilaterally, may alter the vibratory motion to create different glottal volume velocity waveforms from cycle to cycle, or from cycle group to cycle group (refer to the earlier discussion on perturbation). Severe laryngitis due to extreme edema creates the well known response of aphonia (inability or absence of phonation), which may be explained by the mass being too great to permit vibration, and the rounded glottal shaping causing less effective intraglottal pressures during phonation (109; see discussion below). Refer to Hirano and Bless (110) for photographs of vibratory patterns for a number of vocal pathologies including nodules, polyps, cysts, etc.

These kinds of structural changes to the vocal folds lead to the question: how is the vibration of the vocal folds maintained? That is, what allows them to remain in oscillation? A response to this question will now be explored.

The mechanical phonatory motion of the vocal folds depends upon the folds being driven by air pressures within the glottis (111,112,113,114,115,80,84), such air pressure forces working with the biomechanical characteristics of the vocal folds (mass, stiffness, damping) to overcome the damping losses within the tissue (47,116,117,118,119,120,55,121).

The intraglottal pressures are extremely important, then, in the maintenance of vocal fold oscillation (8,2,138). The translaryngeal air pressures drive the air through the glottis, whereas the intraglottal air pressures (perpendicular to the vocal fold surfaces) drive (push and pull on) the vocal folds. If the intraglottal pressures are positive, they act to push the vocal folds away from each other. If the intraglottal pressures are negative, they pull the vocal folds together. The polarity (positive or negative) of the intraglottal pressure depends upon both the dynamic air pressures directly above and below the glottis and the shape of the glottal airway, as will now be discussed(with reference to 78 and 2).

As the glottis nears closure during a vibratory cycle, the airflow through the glottis decreases relatively

quickly, as discussed earlier. The air that has already passed through the glottis continues to travel up the vocal tract, creating greater distance between itself (the air) and the glottis. This air momentum produces a negative rarefaction pressure directly above the glottis as the particles of air separate more and more from each other. Negative pressure at the glottal exit location requires that the pressure be negative within the glottis (at least within the glottal duct near the exit) if the glottis has not closed at the lower glottal edge. The intraglottal negative pressure, located at least near glottal exit, thus should facilitate the final closing of the glottis. The amount of the negative intraglottal pressure depends upon the value of the maximum negative supraglottal pressure.

As the glottis opens, the airflow through the glottis meets the mass of air directly above it, creating compression of the air and positive air pressure, as discussed above. The positive air pressure directly above the glottis creates positive air pressure within the glottis, at least near the glottal exit. The positive pressure due to the positive supraglottal pressure therefore facilitates glottal opening. Thus, during opening, supraglottal positive pressure aids glottal opening, and during closing, supraglottal negative pressure aids glottal closing, both pressures being considered as acoustic pressures.

The glottis takes on two primary shapes during (exhalatory) phonation, convergent and divergent (122,123,15,124). During glottal opening, the convergent glottal shape is produced with a wider opening at glottal entrance and a narrower opening at glottal exit (ref. Figure 7). Because of the convergent shape, a pressure drop (from a higher value to a lower value) is created from the positive tracheal pressure to the pressure existing at the glottal exit location (84), which may be close to positive (as discussed above) or near atmospheric. Thus the glottal convergent shape (itself) creates a positive pressure within the glottis, and this positive pressure pushes on the vocal folds, facilitating glottal opening.

During glottal closing, the divergent glottal shape may be prominent and is produced with a narrower opening at glottal entrance than at glottal exit. Divergent glottal duct shapes are similar to mechanical diffusers in that the expanding area allows pressures within the duct to increase between the upstream inlet and the downstream outlet (82,84). The pressure at glottal entry is therefore lower than at exit, and is therefore negative within the glottis if the glottal exit pressure is atmospheric (zero) or negative (per the discussion above on glottal closing). This negative pressure is due to the divergent shape of the glottis, with the minimal duct diameter formed at entry, creating the lowest pressure there (at entry) essentially in accordance with the Bernoulli energy equation. The negative intraglottal pressure would pull on the vocal fold surfaces, facilitating glottal closure.

The "Bernoulli effect" has referred to the existence of negative pressure in a narrowed location of a duct through which air (fluid) is flowing. This is a possible result of the tradeoff (described by the Bernoulli equation) between air pressure and air particle velocity for different size sections of a duct (lower pressure and higher velocity exist in a narrower section). The Bernoulli effect has been used historically to help "explain" phonation (Tonndorf was an important early proponent of this concept for the maintenance of vocal fold oscillation, see 125). That is, as the air particles speed up as they enter the glottis from the trachea, the air pressure supposedly becomes negative in the glottis because it constitutes a smaller opening than the trachea, and then proceeds to suck the folds together because they are separated. This explanation is incomplete, of course. During glottal opening, when the glottis is convergent and the minimal glottal diameter is at the glottal exit, the Bernoulli energy equation does help to explain the pressure reduction from the glottal entrance to the exit, but the intraglottal pressure is positive, not negative, despite the glottis being smaller than the trachea, since the pressure lowers from the positive tracheal value to the positive or near zero value at glottal exit. During glottal closure when the glottis forms a divergence (diffuser) shape, the pressure reduction from the trachea to glottal entry may follow the Bernoulli equation for the most part, but even the application of the Bernoulli equation within the glottis would essentially apply only in the special case of no flow separation from the glottal walls. The flowing air may separate from (come away from, not flow in the downstream direction near) the glottal walls because of pressure rise and air viscosity in the diverging glottal duct, creating a condition where the Bernoulli equation no longer applies (beyond the flow separation point; 84,112,113). The pressure within the glottis depends on the actual subglottal and supraglottal pressures. The greater the difference between the subglottal and the supraglottal pressures, the larger the pressure changes within the glottis as it changes shape during oscillation.

The actual interdependence among dynamic (time-varying) glottal shaping, translaryngeal pressures, and intraglottal pressures, and the application of equations of mechanics require precise empirical measurement of glottal shaping, flows, and pressures during phonation, research yet to be adequately performed. In addition, the relationship among these aspects, the particle velocities that make up the glottal flow, and the resulting acoustic signal (126,127,128,129,27,130,131,132,133,134,135) extend this matter to the necessary and interfacing aeroacoustic level of phonation.

Conclusion

This chapter has reviewed basic aspects of laryngeal function during phonation. The concepts underlying

the understanding of duration, frequency, intensity, spectra, and vocal fold motion during phonation are the bases with which to make effective intervention decisions for laryngeal and voice change. These underlying concepts are applicable to both clinical and training practices.

Acknowledgements

The author is grateful to and would like to thank CG Guo for his help with the figures, Kimberly Fisher for the use of her spasmodic dysphonia data, Marshall Smith for his figure dealing with vocal mass and fundamental frequency change, Tamara Field for help in preparing Figure 1, and David Kuehn for an anatomical discussion. This project was prepared with support from NIDCD grant P60 DC00976 RTC, and is dedicated to the memory and mentorship of Dr. Wilbur James Gould.

References

1. Scherer RC: Aerodynamic assessment in voice production. In Cooper JA: Assessment of Speech and Voice Production: Research and Clinical Applications. NIDCD Monograph Volume 1-1991. Bethesda, National Institute on Deafness and Other Communication Disorders, NIH, 1991, p 112-123.

2. Scherer RC: Physiology of phonation: A review of basic mechanics. In Ford CN, Bless DM: Phonosurgery: Assessment and Surgical Management of Voice Disorders. New York, Raven Press, 1991, p 77-93.

3. Hollien H, Moore GP: Measurement of the vocal folds during changes in pitch. J Speech Hear Res 3: 157-165, 1960.

4. Nishizawa N, Sawashime M, Yonemoto K: In Fujimura O: Vocal Physiology: Voice Production, Mechanisms, and Function. New York, Raven Press, 1988, p 75-82.

5. Backus J: The Acoustical Foundations of Music, ed 2. New York, W. W. Norton & Co., Inc., 1977.

6. Titze IR, Durham PL: Passive mechanisms influencing fundamental frequency control. In Baer T, Sasaki C, Harris KS: Laryngeal Function in Phonation and Respiration. San Diego, College-Hill Press, 1987, p 304-319.

7. Colton RH: Physiological mechanisms of vocal frequency control: The role of tension. J Voice 2/3: 208-220, 1988.

8. Titze IR: Principles of Voice Production. Prentice Hall, 1994.

9. Perlman AL, Titze IR: Development of an in vitro technique for measuring elastic properties of vocal fold tissue. J Speech Hear Res 31/2: 288-289, 1988.

10. Colton RH, Casper JK: Understanding Voice Problems. Baltimore, Williams & Wilkins, 1990.

11. Hammarberg B, Fritzell B, Gauffin J, Sundberg J: Acoustic glottogram, subglottic pressure, and voice quality during insufficient vocal fold closure phonation. Proceedings of the International Conference on Voice, Kurume, Japan, 28-35, 1986.

12. Gauffin J, Sundberg J: Spectral correlates of glottal voice source waveform characteristics. J Speech Hear Res 32: 556-565, 1989.

13. Sundberg J: The Science of the Singing Voice. Dekalb, Illinois, Northern Illinois University Press, 1987.

14. Titze IR: Phonation threshold pressure: A missing link in glottal aerodynamics. J Acoust Soc Am 91/5: 2926-2935, 1992.

15. Baer T: Investigation of Phonation Using Excised Larynges. Dissertation, Massachusetts Institute of Technology, Cambridge, MA, 1975.

16. Baer T: Observation of vocal fold vibration: Measurement of excised larynges. In Stevens KN, Hirano M: Vocal Fold Physiology. University of Tokyo Press, 1981, p 119-133.

17. Draper MH, Ladefoged P, Whitteridge D: Expiratory pressures and air flow during speech. British Med J 18: 1837-1843, 1960.

18. Verdolini-Marston K, Titze I, Druker D: Changes in phonation threshold pressure with induced conditions of hydration. J Voice 4/2: 142-151, 1990.

19. Coleman RF, Mabis JH, Hinson JK: Fundamental frequency- sound pressure level profiles of adult male and female voices. J Speech Hear Res. 20: 197-204, 1977.

20. Klingholz F, Martin F: Die quantitative Auswertung der Stimmfeldmessung. Sprache-Stimme-Gehor, 7: 106-110, 1983.

21. Titze IR: Acoustic interpretation of the voice range profile (phonetogram). J Speech Hear Res 35: 21-34, 1992.

22. Coleman RF: Sources of variation in phonetograms. J Voice 7/1: 1-14, 1993.

23. Bless DM, Baken RJ, Hacki T, Fritzell. B, Laver J, Schutte H, Hirano M, Loebell E, Titze I, Faure MA, Muller A, Wendler J, Fex S, Kotby MN, Brewer D, Sonninen A, Hurme P: International Association of Logopedics and Phoniatrics (IALP) Voice Committee discussion of assessment topics. J Voice 6/2: 194-210, 1992.

24. Hirano M: Clinical Examination of Voice. New York, Springer-Verlag, 1981.

25. Hirose H, Niimi S: The relationship between glottal opening and the transglottal pressure differences during consonant production. In Baer T, Sasaki C, Harris KS: Laryngeal Function in Phonation and Respiration. San Diego, College-Hill Press, 1987, p 381-390.

26. Baken RJ: Clinical Measurement of Speech and Voice. Boston, College-Hill Press, Little, Brown and Company, 1987.

27. Kakita Y: Simultaneous observation of the vibratory pattern, sound pressure, and airflow signals using a physical model of the vocal folds. In Fujimura O: Vocal Physiology: Voice Production, Mechanisms, and Function. New York, Raven Press Ltd., 1988, p 207-218.

28. Fant G: Some problems in voice source analysis. Speech Comm 13: 7-22, 1933.

29. Kiritani S, Hirose H, Imagawa H: High-speed digital image analysis of vocal cord vibration in diplophonia. Speech Comm 13: 23-32, 1993.

30. Baer T: Reflex activation of laryngeal muscles by sudden induced subglottal pressure changes. J Acoust Soc Am 65: 1271-1275, 1979.

31. Larson C, Kempster G, Kistler M: Changes in voice fundamental frequency following discharge of single motor units in cricothyroid and thyroarytenoid muscles. J Speech Hear Res 30: 552-558, 1987.

32. Kempster GB, Larson CR, Kistler MK: Effects of electrical stimulation of cricothyroid and thyroarytenoid muscles on voice fundamental frequency. J Voice 2/3: 221-229, 1988.

33. Titze IR: A model for neurologic sources of aperiodicity in vocal fold vibration. J Speech Hear Res 34/3: 460-472, 1991.

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

35. Moon FC: Chaotic Vibrations. New York, John Wiley & Sons, 1987.

36. Gerratt BR, Precoda K, Hanson D, and Berke GS: Source characteristics of diplophonia. J Acoust Soc Am 83: S66, 1988.

37. 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 89: 383-394, 1991.

38. Berke GS, Gerratt BR: Laryngeal biomechanics: An overview of mucosal wave mechanics. J Voice 7/2: 123-128, 1993.

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

40. Scherer R, Gould WJ, Titze I, Meyers A, Sataloff R: Preliminary evaluation of selected acoustic and glottographic measures for clinical phonatory function analysis, J of Voice 2: 230-244, 1988.

41. Hollien H, Moore P, Wendahl RW, Michel JF: On the nature of vocal fry. J Speech Hear Res 9: 245-247, 1966.

42. Keidar A: Vocal Register Change: An Investigation of Perceptual and Acoustic Isomorphism. Ph.D. Dissertation, The University of Iowa, 1986.

43. Titze IR: A framework for the study of vocal registers. J Voice 2/3: 183-194, 1988.

44. Scherer RC: Physiology of creaky voice and vocal fry. J Acoust Soc Am 86(S1): S25(A), 1989; this paper has an accompanying video tape.

45. Scherer RC, Druker DG, Titze IR: Electroglottography and direct measurement of vocal fold contact area. In Fujimura O: Vocal Physiology: Voice Production, Mechanisms, and Function. New York, Raven Press, Ltd., 1988, p 279-291.

46. Ishizaka K, Matsudaira M: Analysis of the vibration of the vocal cords. J Acoust Soc Jap 24: 311-312, 1968.

47. Ishizaka K, Flanagan JL: Synthesis of voiced sounds from a two-mass model of the vocal cords. Bell Sys Tech J 51/6: 1233-1268, 1972.

48. Smith ME, Berke GS, Gerratt BR, Kreiman J: Laryngeal paralysis: Theoretical considerations and effects on laryngeal vibration. J Speech Hear Res 35: 545-554, 1992.

49. Rothenberg M, Mahshie J: Induced transglottal pressure variations during voicing. Speech Comm 14/3-4: 365-371, 1986.

50. Cheng YM, Guerin B: Control parameters in male and female glottal sources. In Baer T, Sasaki C, Harris KS: Laryngeal Function in Phonation and Respiration. San Diego, College-Hill Press, 1987, p 219-238.

51. Baken, R.J. and Orlikoff, R.F. (1987). Phonatory response to step-function changes in supraglottal pressure. In Laryngeal Function in Phonation and Respiration, edited by T. Baer, C. Sasaki, and K.S. Harris, College-Hill Press, San Diego, 273-290.

52. Saito S, Fukuda H, Isogai Y, Ono H: X-Ray stroboscopy. In Stevens KN, Hirano M: Vocal Fold Physiology. University of Tokyo Press, 1981, p 95-106.

53. Saito S, Fukuda H, Kitahara S, et alii: Pellet tracking in the vocal fold while phonating - experimental study using canine larynges with muscle activity. In Titze IR, Scherer RC: Vocal Fold Physiology: Biomechanics, Acoustics and Phonatory Control. Denver, The Denver Center for the Performing Arts, 1985, p 169-182.

54. Fukuda H, Saito S, Kitahara S, Isogai Y, Makino K, Tsuzuki T, Kogawa N, Ono H: Vocal fold vibration in excised larynges viewed with an x-ray stroboscope and an ultra-high-speed camera. In Bless DM, Abbs JH: Vocal Fold Physiology, Contemporary Research and Clinical Issues. College-Hill Press, 1983, p 238-252.

55. Alipour-Haghighi F, Titze IR: Simulation of particle trajectories of vocal fold tissue during phonation. In Titze IR, Scherer RC: Vocal Fold Physiology: Biomechanics, Acoustics and Phonatory Control. Denver, The Denver Center for the Performing Arts, 1985, p 183-190.

56. Muta H, Fukuda H: Pressure-flow relationship in the experimental phonation of excised canine larynges. In Fujimura O: Vocal Physiology: Voice Production, Mechanisms, and Function. New York, Raven Press, Ltd., 1988, p 239-247.

57. Titze IR, Luschei ES, Hirano M: Role of the thyroarytenoid muscle in regulation of fundamental frequency. J Voice 3/3: 213-224, 1989.

58. Gelfer CE, Harris KS, Collier R, Baer T: Is declination actively controlled? In Titze IR, Scherer RC: Vocal Fold Physiology: Biomechanics, Acoustics and Phonatory Control. Denver, The Denver Center for the Performing Arts, 1985, p 113-126.

59. Gelfer CE, Harris KS, Baer T: Controlled variables in sentence intonation. In Baer T, Sasaki C, Harris KS: Laryngeal Function in Phonation and Respiration. San Diego, College-Hill Press, 1987, p 422-435.

60. Honda K: Relationship between pitch control and vowel articulation. In Bless DM, Abbs JH: Vocal Fold Physiology, Contemporary Research and Clinical Issues. College-Hill Press, 1983, p 286-297.

62. Titze IR, Jiang J, Drucker DG: Preliminaries to the body-cover theory of pitch control. J Voice 1/4: 314-319, 1988.

63. Choi HS, Berke GS, Ye M, Kreiman J: Function of the thyroarytenoid muscle in a canine laryngeal model. Ann Otol Rhinol Laryngol 102: 769-776, 1993.

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

65. Fujimura O: Body-cover theory of the vocal fold and its phonetic implications. In Stevens KN, Hirano M: Vocal Fold Physiology. University of Tokyo Press, 1981, p 271-288.

66. Hirano M: Phonosurgery: Basic and Clinical Investigations. Otologia (Fukuoka), 21: 239-442, 1975.

67. Hirano M: The laryngeal muscles in singing. In Hirano M, Kirchner JA, Bless DM: Neurolaryngology, Recent Advances. Boston, A College-Hill Publication, Little, Brown and Company, 1987, p 209-230.

68. Titze IR: A four-parameter model of the glottis and vocal fold contact area. Speech Comm 8: 191-201, 1989.

69. Plomp R: Aspects of Tone Sensation. New York, Academic Press, 1976.

70. Strong WJ, Plitnik GR: Music Speech Audio. Provo, Soundprint, 1992.

71. Rothenberg M: Acoustic interaction between the glottal source and the vocal tract. In Stevens KN, Hirano M: Vocal Fold Physiology. University of Tokyo Press, 1981, p 305-328.

72. Rothenberg M: An interactive model for the voice source. In Bless DM, Abbs JH: Vocal Fold Physiology, Contemporary Research and Clinical Issues. College-Hill Press, 1983, p 155-165.

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

74. Kitzing P, Lofqvist A: Subglottal and oral pressures during phonation - preliminary investigation using a miniature transducer system. Medical and Biological Eng,13/5: 644-648, 1975.

75. Fant G: Preliminaries to analysis of the human voice source. STL-QPRS 4/1982: 1-27, 1983.

76. Miller DG, Schutte HK: Characteristic patterns of sub- and supra-glottal pressure variations within the glottal cycle. In Lawrence VL: Transcripts of the XIIIth Symposium: Care of the Professional Voice. New York, The Voice Foundation, 1984, p 70-75.

77. Cranen B, Boves L: A set-up for testing the validity of the two mass model of the vocal folds. In Titze IR, Scherer RC: Vocal Fold Physiology: Biomechanics, Acoustics and Phonatory Control. Denver, The Denver Center for the Performing Arts, 1985, p 500-513.

78. Olson HF: Solutions of Engineering Problems by Dynamical Analogies, ed 2. New York, D. Van Nostrand Co., Inc., 1966.

78. Titze IR: Mean intraglottal pressure in vocal fold oscillation. J Phonetics 14: 359-364, 1986.

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

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

82. Kline SJ: On the nature of stall. ASME J Basic Eng 81: 305-320, 1959.

83. Miller DS: Internal Flow, a Guide to Losses in Pipe and Duct Systems. Cranfield, British Hydromechanics Research Association, 1971.

84. Guo CG, Scherer RC: Finite element simulation of glottal flow and pressure. J Acoust Soc Am 94/2/pt.1: 688-700, 1993.

85. Scherer R, Sundberg J, Titze I: Laryngeal adduction related to characteristics of the flow glottogram. J Acoust Soc Am 85 (S1): S129(A), 1989.

86. Titze IR, Sundberg J: Vocal intensity in speakers and singers. J Acoust Soc Am 91: 2936-2946, 1992.

87. Sundberg J, Titze I, Scherer R: Phonatory control in male singing: A study of the effects of subglottal pressure, fundamental frequency, and mode of phonation on the voice source. J Voice 7/1: 15-29, 1993.

88. Fant G, Liljencrants J, Lin Q: A four-parameter model of glottal flow. STL-QPSR 4/1985: 1-13, 1985.

89. Gobl C, Karlsson I: Male and female voice source dynamics. In Gauffin J, Hammarberg B: Vocal Fold Physiology: Acoustic, Perceptual, and Physiological Aspects of Voice Mechanisms. San Diego, Singular Publishing Group, Inc., 1991, p 121-128.

90. Fant G, and Lin Q: Comments on glottal flow modelling and analysis. In Gauffin J, Hammarberg B: Vocal Fold Physiology: Acoustic, Perceptual, and Physiological Aspects of Voice Mechanisms. San Diego, Singular Publishing Group, Inc., 1991, p 47-56.

91. Sundberg J, Gauffin J: Waveform and spectrum of the glottal voice source. In Lindblom B, Ohman S: Frontiers of Speech Communication Research. New York, Academic Press, 1979, p 301-322.

61. Sundberg J, Leanderson R, von Euler C: Activity relationship between diaphragm and cricothyroid muscles. J Voice 3/3: 225-232, 1989.

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

93. Ladefoged P, McKinney NP: Loudness, sound pressure, and subglottal pressure in speech. J Acoust Soc Am 35: 454-460, 1963.

94. Isshiki N: Regulatory mechanism of voice intensity variations. J Speech Hear Res 7: 17-29, 1964.

95. Isshiki N: Remarks on mechanism for vocal intensity variation. J Speech Hear Res 12: 665-672, 1969.

96. Rubin HJ, LeCover M, Vennard W: Vocal intensity, subglottic pressure, and airflow relationships in singers. Folia Phoniatr 19: 393-413, 1967.

97. Bouhuys A, Mead J, Proctor D, Stevens K: Pressure-flow events during singing. Ann NY Acad Sci 155: 165-76, 1968.

98. Holmberg EB, Hillman RE, Perkell J: Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. J Acoust Soc Am 84: 511-529, 1988.

99. Titze IR: Regulation of vocal power and efficiency by subglottal pressure and glottal width. In Fujimura O: Vocal Physiology: Voice Production, Mechanisms, and Function. New York, Raven Press, Ltd., 1988, p 227-238.

100. Raphael BN, Scherer RC: Voice modifications of stage actors: Acoustic analyses. J Voice 1: 83-87, 1987.

101. Bartholomew WT: A physical definition of "good voice quality" in the male voice. J Acoust Soc Am 6: 25-33, 1934.

102. Sundberg J: Articulatory interpretation of the "singing formant". J Acoust Soc Am 55: 838-844, 1974.

103. Perkins WH: Voice Disorders. New York, Thieme-Stratton Inc., 1983.

104. Hollien HF, Curtis JF: A laminagraphic study of vocal pitch. J Speech Hear Res 3: 361-371, 1960.

105. Hollien HF: Vocal fold thickness and fundamental frequency of phonation. J Speech Hear Res 5/3: 237-243, 1962.

106. Hollien HF, Colton RH: Four laminagraphic studies of vocal fold thickness. Folia Phoniat 21: 179-198, 1969.

107. Hollien HF, Coleman RF: Laryngeal correlates of frequency change: A STROL study. J Speech Hear Res 13/2: 271-278, 1970.

108. Scherer RC, Vail VJ, Rockwell B: Examination of the laryngeal adduction measure EGGW. In Bell-Berti F, Raphael LJ: Studies in Speech Production: A Festshrift for Katherine Safford Harris. American Institute of Physics, in press, 1994.

109. Scherer RC, Guo CG: Intraglottal pressures for pathological glottal configurations using a computational fluid dynamic approach. In preparation.

110. Hirano M, Bless DM: Videostroboscopic Examination of the Larynx. San Diego, Singular Publishing Group, Inc., 1993.

111. Berg Jw van den, Zantema JT, Doornenball P Jr: On the air resistance and the Bernoulli effect of the human larynx. J Acoust Soc Am 29: 626-631, 1957.

112. Scherer RC: Pressure-flow relationships in a laryngeal airway model having a diverging glottal duct. J Acoust Soc Am 73 (S1): S46(A), 1983.

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

114. Gauffin J, Binh N, Ananthapadmanabha TV, Fant G: Glottal geometry and volume velocity waveform. In Bless DM, Abbs JH: Vocal Fold Physiology: Contemporary Research and Clinical Issues. San Diego, College-Hill Press, 1983, p 194-201.

115. Ishizaka K: Air resistance and intraglottal pressure in a model of the larynx. In Titze IR, Scherer RC: Vocal Fold Physiology: Biomechanics, Acoustics and Phonatory Control. Denver, The Denver Center for the Performing Arts, 1985, p 414-424.

116. Ishizaka K, Matsudaira M: Fluid mechanical considerations of vocal cord vibration. SCRL Monograph No. 8, April, 1972.

117. Titze IR: The human vocal cords: A mathematical model. Part I. Phonetica 28: 129-170, 1973.

118. Titze IR: The human vocal cords: A mathematical model. Part II. Phonetica 29: 1-21, 1974.

119. Titze IR: On the mechanics of vocal fold vibration. J Acoust Soc Am 60/6: 136-1380, 1976.

120. Titze IR: Biomechanics and distributed-mass models of vocal fold vibration. In Stevens KN, Hirano M: Vocal Fold Physiology. University of Tokyo Press, 1981, p 245-270.

121. Alipour F, Titze IR: A finite element simulation of vocal fold vibration. In Proc. Fourteenth Annual Northeast Bioeng. Conf., Durham, N.H., IEEE publications #88-CH2666-6: 186-189, 1988.

122. Schonharl E: Die Stroboskopie in der praktischen Laryngologie, Thieme, Stuttgart, 1960.

123. Stevens KN, Klatt DH: Current models of sound sources for speech. In Wyke B: Ventilatory and Phonatory Control Systems, An International Symposium. Oxford University Press, 1974, p 279-292.

124. Hirano M: Structure and vibratory behavior of the vocal folds. In Sawashima M, Cooper FS: Dynamic Aspects of Speech Production. Tokyo, Japan, University of Tokyo Press, 1977, p 13-27.

125. Cooper DS: Voice: A historical perspective (Woldemar Tonndorf and the Bernoulli effect in voice production). J Voice 3/1: 1-6, 1989.

126. Kaiser J: Some observations on vocal tract operation from a fluid flow point of view. In Titze IR, Scherer RC: Vocal Fold Physiology: Biomechanics, Acoustics and Phonatory Control. Denver, The Denver Center for the Performing Arts, 1985, p 358-386.

127. Teager H, Teager S: Active fluid dynamic voice production models, or there is a unicorn in the garden. In Titze IR, Scherer RC: Vocal Fold Physiology: Biomechanics, Acoustics and Phonatory Control. Denver, The Denver Center for the Performing Arts, 1985, p 387-401.

128. McGowan RS: An aeroacoustic approach to phonation. J Acoust Soc Am 88: 696-704, 1988.

129. McGowan RS: Phonation from a continuum mechanics point of view. In Gauffin J, Hammarberg B: Vocal Fold Physiology: Acoustic, Perceptual, and Physiological Aspects of Voice Mechanisms. San Diego, Singular Publishing Group, Inc., 1991, p 65-72.

130. Berke GS, Moore DM, Monkewitz PA, Hanson DG, Gerratt BR: A preliminary study of particle velocity during phonation in an in vivo canine model, J Voice 3/4: 306-313, 1989.

131. Shadle CH, Barney AM, Thomas DW: An investigation into the acoustic and aerodynamics of the larynx. In Gauffin J, Hammarberg B: Vocal Fold Physiology: Acoustic, Perceptual, and Physiological Aspects of Voice Mechanisms. San Diego, Singular Publishing Group, Inc., 1991, p 73-80.

132. Alipour F, Fan C: Pulsatile flow in a three-dimensional model of larynx. In Biewener AA, Goel VK: Proceedings of 1993 ASB Meeting. 1993, p 221-222.

133. Davies POAL, McGowan RS, Schadle CH: Practical flow duct acoustics applied to the vocal tract. In Titze IR: Vocal Fold Physiology: Frontiers in Basic Science. San Diego, Singular Publishing Group, Inc., 1993, p 93-134.

134. Scherer R: Practical flow duct acoustics applied to the vocal tract: response. In Titze IR: Vocal Fold Physiology: Frontiers in Basic Science. San Diego, Singular Publication Group, 1993, p 134-142.

135. Alipour F, Scherer RC: Pulsatile airflow during phonation: an excised larynx model. (in review).

136. Gramming P: The Phonetogram: An Experimental and Clinical Study. Malmo, Sweden; Department of Otolaryngology, University of Lund, 1988.

137. Baer T: Vocal jitter: A neuromuscular explanation. In Lawrence V, Weinberg B: Transcripts of the Eighth Symposium on Care of the Professional Voice. New York, The Voice Foundation, 1979, p 19-22.

138. Jiang JJ: A Methodological Study of Hemilaryngeal Phonation and The Measurement of Vocal Cord Intraglottal Pressure and Impact Stress. Dissertation, The University of Iowa, Iowa City, 1991.

139. Pinto NB, Titze IR: Unification of perturbation measures in speech signals. J Acoust Soc Am 87(3): 1278-1289, 1990.

NCVS Status and Progress Report - 7
December 1994, 187-196

Common Voice Disorders: A Primary Care Approach

Henry Hoffman, M.D.
Michael Karnell, Ph.D.
Department of Otolaryngology, The University of Iowa

Introduction

Not all hoarse patients require visual inspection of the larynx or referral to an otolaryngologist. Many processes affecting the larynx are either self-limited over a several-day period or respond to a therapeutic trial rapidly enough that evaluation and treatment without visualizing the larynx may be sufficient.

A critical step, however, in the evaluation of patients with hoarseness lasting more than two weeks is a mirror examination of the larynx. Based on the patient's anatomy and level of cooperation as well as strength of gag reflex, successful inspection of the larynx may require supplementation of a mirror exam with flexible fiberoptic laryngoscopy which is usually performed transnasally. Although a few primary care physicians practice laryngeal imaging by mirror exam or fiberoptics, most general practitioners and internists deal with patients who have voice problems without looking at the larynx. In many ways, this approach may be considered similar to treating a female patient with pelvic complaints without a speculum exam of the vagina and uterine cervix.

Without the capacity to visualize the larynx, an added emphasis must be placed on history taking and physical exam of the surrounding structures to infer what may otherwise be directly seen by inspection of the vocal cords.

Background

Hoarseness as a medical complaint may appear trivial when considered in the context of many of the life-threatening diseases physicians commonly encounter. Although the vast majority of processes causing hoarseness are benign, the rough voice associated with laryngeal cancer, the breathy voice associated with laryngeal paralysis from a thyroid cancer, and the weak voice of a patient with Parkinson's disease are early tip-offs to diseases that warrant serious attention.

Even in the absence of associated systemic disease, hoarseness as an isolated problem interferes with the lives of many otherwise healthy patients. Harshness in a singer's voice resulting from vocal nodules may mark the end of a career. The negative impact of voice abnormalities on the careers of other professionals who depend on their voices such as clergy, teachers, salesmen, physicians and lawyers is acknowledged, but not yet fully quantified. The difficulties President Clinton experienced with his voice during the 1992 campaign clearly demonstrate how a relatively minor disease process may interfere with the work of a voice professional.

The largest study to date addressing the impact of voice disorders consists of an evaluation of 113 patients seen for chronic voice problems at weekly voice clinics held at the Universities of Iowa and Utah. Questionnaires were administered to identify lifestyle effects of voicing abnormalities.[1] Voice disorders in this group were primarily due to vocal nodules, laryngeal polyps, laryngitis, vocal cord paralysis, spasmodic dysphonia, and laryngeal trauma. Among these 113 patients, 52% related that their voice problem had negatively effected their past career options and 75% felt that their condition limited future career options. Over 70% felt that social interactions were negatively affected by their voice problem with 55% avoiding social situations when possible. Over half of the patients related that they experienced difficulty in being understood over the phone.

Whereas the social impact of chronic voice disorders has just recently become apparent, the overall importance of infections of the upper aerodigestive system (including laryngitis) has been known for many years. A yearly loss

of more than 250 million days of work with over 100 million physician office visits has been estimated to result for upper respiratory tract infections annually in the United States.[11,12,13,14]

Terminology

The terminology describing hoarseness and its treatment is as specialized as are the disciplines of the speech scientist and laryngologist. The **glottis** is the segment of the larynx which is composed of the vocal folds, extending from the ventricles to 1 cm below the free surface of the vocal cords. The glottis functions as a **valve** which opens and closes in an oscillating fashion to convert the DC current of air travelling up the trachea from the lungs into an AC current which generates the **voice.** The position, length, and tension of the vocal cords are dependent on contraction of the laryngeal muscles, which set the vocal folds into a position leading to passive opening and closing of the valve (vocal fold vibration) as the tracheal air column passes through. The configuration of the glottic valve and nature of the substance of the vocal cords determine the character of voice including pitch and loudness. Voice should be differentiated from **speech,** which is the end result of modulation of voice (the laryngeal output) by supraglottic, oropharyngeal and oral cavity structures including the palate, tongue, teeth and lips.

Hoarseness or "dysphonia" describes abnormal voice production which reflects impaired vocal cord or lung function. **"Aphonia"** represents complete loss of voice. An abnormal voice termed "hoarse" or "dysphonic" is more specifically characterized as breathy, weak, rough, or strained. A **breathy voice** may result from **"glottic incompetence"** with incomplete closure of the valve if one of the vocal cords is paralyzed. A **weak (asthenic) voice** may occur as a combination of inadequate pulmonary support and general muscle wasting such as is commonly seen in debilitated, elderly patients. A **rough voice** commonly results from mass lesions of the vocal cord(s) such as the thickenings and irregularities associated with chronic smoking and excessive use of the voice. A **strained voice** may reflect abnormal motor control of the larynx resulting in excessively tight closure of the glottis. Several causes of vocal strain exist, with the most common, termed **"muscle tension dysphonia,"** thought to be functional in nature.

The vocal cords are complex structures comprised of several interacting layers that permit wave-like motion of the mucosa and overlying mucosal blanket during voice production. The term **"vocal fold"** is used by some (in a pedantic fashion) to emphasize the complicated nature of what is more commonly referred to as **"vocal cord."**

Inflammation of the larynx is termed **laryngitis** and can be associated with pain, hoarseness, dysphagia, odynophagia, and airway obstruction. Laryngitis with associated hoarseness can be acute or chronic. Factors responsible for laryngitis are infectious (viral, bacterial, fungal), toxic (smoke, gastroesophageal reflux, caustic ingestion), traumatic and allergic.

The following discussion is organized according to presenting signs and symptoms as would be evaluated by a physician without the capacity to visualize the vocal cords. An emphasis is placed on those processes that require urgent referral as well as those that can be treated without imaging the larynx. Most cases of chronic hoarseness warrant referral to an otolaryngologist on a non-emergency basis. The discussion of these processes is therefore truncated.

Signs and Symptoms

Acute Hoarseness: Traumatic

Traumatic laryngitis may be the most common cause of hoarseness, and is caused by closing the glottis with excessive muscular force. It affects large segments of the population attending sporting events and other activities associated with loud or prolonged cheering or singing. For the great majority of these people, resuming normal voice use for several days after the event permits resolution of vocal cord edema without sequel. However, repetition of such injury to the larynx may result in changes that become permanent (see **Chronic Hoarseness - Traumatic**). Vocal cord hemorrhage may occur acutely with the development of a hematoma that does not resolve but matures to become a polyp. The larynx altered by inflammation from an active upper respiratory tract infection or esophageal reflux is more susceptible to injury from vocal abuse. Desiccation of the vocal cords from low humidity or dehydration, as well as irritation of the vocal cords from tobacco smoke, increases the likelihood of permanent injury developing following excessive voice use. Many medications may affect the larynx in a negative fashion primarily by systemic dehydration (as with diuretics) or diminishing surface moisture (antihistamines) (**Table 1**). Although continued use of some agents known to negatively affect the voice may be required to maintain health, many of these medications can either be eliminated or substituted with agents without such drying effects.

Medications including ingested emollients, inhaled steroids, and mists containing anesthetics or vasoconstrictors are popular among some voice professionals but are generally not recommended due to the potential for irritating the larynx.[9,10] Systemic mucolytics and wetting agents such as iodinated glycerol and guaifenesin may be of some benefit in thinning secretions. The placebo effect of these medications may be difficult to separate from true pharmacological benefit. Because of their few side effects, it is reasonable to offer these medications to patients with complaints of thick secretions or frequent throat clearing in association with hoarseness. Attention to possible sino-nasal disease or gastroesophageal reflux is warranted in patients complaining of secretions in their throat.

Table 1
Medications in Common Use Affecting Voice
(Attenuated List)

Adverse Effect	
Desiccation of Vocal Cords	Antihistamines; Diuretics; Tricyclic antidepressants; Other (many)
Alter Cord Substance	Birth control pills; Androgens
Inflammation of cord surface	Propellants in inhalants
Candida Laryngitis	Steroid inhalants
Induce Intracordal Hemorrhage	Aspirin; Anticoagulants
Positive Effect	
Mucolytics	Iodinated glycerol; guaifenesin
Antitussives	Dextromethorphan; codeine
Replacement therapy	Estrogen (menopause); Thyroxine (hypothyroid states)[9]

Treatment of hoarseness developing after an episode of vocal abuse consists of a period of voice rest as well as attention to adequate hydration and humidification, with consideration for referral should the hoarseness persist for more than one or two weeks. Complete voice rest entails communication by nonverbal methods, usually employing a writing pad and a bell or buzzer to gain the listener's attention. Radical voice rest to this extent is of questionable value and is a true hardship for the patient. Although complete voice rest is commonly used in the care of professional vocalists and postoperatively after surgical manipulation of the vocal cords, "relative voice rest" may be a more reasonable goal for most patients during an episode of acute hoarseness. The "arms length rule" is easily communicated by instructing the patient to speak in an easy, soft manner only to those within arms reach. Adequate hydration is ensured by avoidance of diuretics (including caffeine) and attention to drinking sufficient fluids "to make the urine the color of water." Humidification of the environment is best accomplished by use of a bedside vaporizer which should be cleaned between uses to prevent aerosolizing proliferating bacteria.

External trauma to the larynx is rare due to its protected location between the mandible and sternum which limits exposure unless the neck is extended. Even a minor degree of hoarseness associated with external neck injury may herald rapid progression to airway obstruction and therefore warrants immediate referral for inspection of the larynx with mirror exam or fiber optics. Palpation of the neck may reveal crepitus, indicating mucosal disruption which allows tracking of air into the soft tissue of the neck. Maintenance of the airway is the primary concern with further evaluation needed via CT imaging of the larynx and rigid endoscopy, done under general anesthesia. Surgical intervention by way of a tracheotomy may also be required to deal with laryngeal cartilage fracture or dislocation, hematoma, or soft tissue laceration.

Acute Hoarseness: Infectious

Infectious laryngitis is a common cause of hoarseness and is usually viral. Hoarseness arising in this setting is often associated with other symptoms that may overshadow those due to vocal cord inflammation. Several days of fever, pharyngitis, rhinitis, bronchitis and cervical adenitis are frequently associated with hoarseness. The diagnosis of laryngitis in this setting does not require visualization of the larynx unless additional symptoms such as stridor or significant dysphagia develop.

Treatment consists of adequate hydration, humidification, vocal conservation, and symptomatic treatment of associated symptoms such as an anti-tussive for cough or nasal decongestants for rhinitis. Suppression of a cough with codeine or dextromethorphan may protect the vocal cords from repeated trauma which could lead to edema and hemorrhage.

While the vocal cords are inflamed, trauma resulting from excessive use may lead to permanent laryngeal injury (e.g. polyps or nodules). Systemic steroids (e.g. Medrol Dosepak or burst and taper of prednisone) have been used by the voice professional to avoid canceling a speaking

or singing engagement, to diminish inflammation during an episode of laryngitis; this is controversial and should be dealt with on an individual basis. The patient with a single chance --a "once-in-a -lifetime" opportunity --to perform a significant engagement may justifiably receive this type of medical intervention. As a general rule, however, voice conservation is a better approach to take in the case of acute laryngitis regardless of whether this requires cancelling a performance. The services of speech pathologists with an interest in the vocal arts are invaluable in situations such as these. They may be helpful in suggesting technical modifications to diminish laryngeal injury with voice use as well as modifications of the speaking or singing environment (e.g. use of a microphone, avoidance of difficult songs, and lip syncing when singing with a group).

Viral laryngitis may result from infection of more than 200 viruses known to be responsible for upper respiratory tract infections. The supportive management of these clinically indistinguishable illnesses is the same, with the exception of influenza A infection which can be treated with amantadine or, in high risk individuals, prevented with vaccination.

Like viral laryngitis, **bacterial laryngitis** rarely presents in isolation and usually is associated with rhinosinusitis and/or laryngotracheal bronchitis.[15] Purulent rhinorrhea or cough is considered sufficient evidence of a bacterial infection to treat with antibiotics supplemented by the same supportive care given for viral laryngitis. Erythromycin combined with sulfonamide, amoxicillin with clavulonic acid, cefuronzime or clarithromycin are currently recommended antibiotic choices for bacterial laryngitis.[16]

Fungal laryngitis is still rare but becoming increasingly more common due to its association with AIDS as well as other forms of immunosuppression. It usually represents infection with Candida.[17,18] Although isolated laryngeal involvement in otherwise healthy newborns has been reported, involvement of the larynx generally represents a small part of a more generalized mucocutaneous candidiasis. Symptoms of hoarseness in an immunocompromised patient should raise concern about this diagnosis and warrants referral. Inspection of the larynx reveals a thick white covering to an erythematous and edematous larynx. Parenteral antifungal agents may be required for treatment, with either intubation or tracheotomy, in addition to endoscopic removal as well as artificial maintenance of the airway.

North American Blastomycosis is a rare cause of fungal laryngitis that is managed with antifungal agents with attention to maintenance of an adequate airway. Concern about this process exists in a disproportionate scale to its incidence because of its similarity in appearance both grossly and microscopically to squamous cell carcinoma.

Table 2
Causes of Laryngeal Paralysis

Adults	Children
Surgery	Neurologic
Idiopathic	Surgery
Malignancy	Idiopathic
Trauma	Trauma
Neurologic	(Birth)

Acute Hoarseness: Toxic Exposure

Inhalation of toxic gases such as anhydrous ammonia and **ingestion of toxic solids or liquids** such as hydrochloric acid may injure the vocal cords and the adjacent upper aerodigestive tract. Depending on the type of exposure, hoarseness may be the significant presenting symptom or may be a minor consideration in the face of a life threatening pulmonary or esophageal injury. Urgent referral should be made for visualization of the larynx with the workup extended according to signs and symptoms of pulmonary or esophageal injury. Pending imaging of the larynx, supportive treatment should be offered including humidification (by face mask acutely), voice rest, and attention to signs of progressive airway compromise.

Acute Hoarseness with Aspiration: Neurogenic

Unilateral vocal cord paralysis in the adult is frequently associated with hoarseness or vocal fatigue, is occasionally asymptomatic and is only rarely a cause of airway obstruction. Patients may complain of shortness of breath which results from excessive air escape during speech due to incomplete closure of the glottis. Extra effort is required to generate sufficient airflow to produce voice. Symptoms due to unilateral vocal cord paralysis may present as an acute process or more gradually. Urgent but not emergent referral is suggested if aspiration is not significant. More timely intervention either through immediate referral or maintenance of NPO status may be required in selected cases to diminish the risk of life threatening aspiration pneumonia.

Reflex closure of the glottis is an important part of the swallowing reflex which is impaired in the paralyzed larynx. Aspiration of liquids is a common symptom but is frequently offered as a minor complaint such as "when I drink water rapidly it goes down the wrong pipe and I cough". Modification of swallowing technique is generally adequate treatment. More severe forms of aspiration may occur if the

paralyzed vocal cord is positioned so as to leave the glottis widely open during the swallow. Additional neurologic deficits either generalized or more focal such as an ipsilateral superior laryngeal nerve paralysis can result in life threatening aspiration. Institution of nasogastric feedings generally permits a studied approach to the evaluation and treatment. More severe aspiration in those unable to tolerate their own oral secretions warrants emergent intervention to include vocal cord medialization to restore glottic competence. A tracheotomy may be required for pulmonary toilet with the understanding that the presence of a tracheotomy tube interferes with the mechanism of deglutition to further impair swallowing.

Key to the evaluation is identification of the etiology of the paralysis. Although some cases ultimately are labelled idiopathic and ascribed to a viral infection (**Table 2**), cancer must be excluded. Associated history, signs and symptoms may help to identify the cause of the paralysis and to direct evaluation. Hoarseness developing immediately after carotid endarterectomy in association with the finding of an immobile vocal cord may be assumed to be due to surgical injury of the recurrent laryngeal nerve without need for further workup. Laryngeal paralysis presenting along with an ipsilateral neck mass determined to be thyroid cancer with fine needle aspiration biopsy may be considered to reflect tumor involvement of the recurrent laryngeal nerve without need for further evaluation. In the absence of such localizing signs, we currently evaluate laryngeal paralysis with a CXR to rule out lung neoplasm followed by a single tailored MRI to follow tract the course of the vagus and recurrent laryngeal nerve from the midbrain, base of skull, neck and mediastinum. The absence of findings results in the diagnosis of idiopathic laryngeal paralysis with the understanding that long term followup may permit identification of an occult tumor in a small percentage of patients.

The treatment of unilateral vocal cord paralysis has advanced over the past decade to the point that a near normal voice may be restored in the great majority of cases following treatment with one of several surgical procedures that have improved on the standard technique of Teflon augmentation of the paralyzed vocal cord. When prognosis for return of function of the paralyzed vocal cord is uncertain, injection with gelfoam may improve swallowing, cough and voicing without inducing irreversible changes to the larynx. The Gelfoam resorbs within three months at which time reinjection or a more permanent treatment may be considered.

Posterior glottic scarring, traumatic arytenoid subluxation, and **cricoarytenoid arthritis** are processes that may mimic laryngeal paralysis and present with all the symptoms associated with an immobile vocal cord. Evaluation with laryngeal electromyography and/or direct laryngoscopy including palpation of laryngeal structures may reconcile this differential diagnosis. A trial of a burst and taper of systemic steroids may be used as a diagnostic and therapeutic trial to the benefit the patient of rheumatoid arthritis with an immobile vocal cords suspected to represent cricoarytenoid arthritis.

Acute Hoarseness with Airway Obstruction

Interference to the passage of air through the larynx commonly occurs at levels other than the glottis and often occurs without hoarseness. A supraglottic lesion such as a tumor or edema of the epiglottis permits unimpaired lung and laryngeal output. The voice is therefore normal or may be muffled as a "hot potato voice" if the mass above the vocal cords is sufficiently large. Similarly, subglottic narrowing will impair airflow but is likely to affect the voice only by decreasing the breath support without otherwise affecting vocal cord function. Any process that predominately affects an adjacent structure to cause airway obstruction may also extend to involve the glottis and cause hoarseness.

Acute Hoarseness with Airway Obstruction: Infectious

Inflammation of the larynx is common in the pediatric patient and, as in adults, is generally a self-limited viral infection without long term sequel once the hoarseness, cough, and symptoms related to inflammation in adjacent upper respiratory structures resolve. Because the larynx, subglottis and trachea is smaller, inflammation with attendant edema is much more likely to result in airway obstruction in children than in adults. As a result, clinical evaluation suggesting laryngeal infection in children is more likely to warrant immediate intervention including hospitalization.

Supraglottitis, croup and **bacterial tracheitis** are three processes commonly seen in the emergency department that share the potential to progress to complete airway obstruction but are managed differently.

Supraglottitis, formerly termed **epiglottitis,** is a rapidly progressive cellulitis of the supraglottis with marked edema which, if considered as a diagnosis in an adult, warrants admission for observation. A more aggressive approach is required for pediatric cases with emergent transfer to the operating room for either endotracheal intubation or tracheotomy. Although supraglottitis is most commonly seen between the ages 3 to 5, it can present at any age. It is generally caused by Hemophilus influenza type B in the pediatric age group but may be caused by Streptococcus viridans, Staphylococcus pyogenes, and pneumococcus in the adult. The process is potentially lethal as reported by Hawkins et al who, in 1973 reviewed previously published accounts totaling 62 adults with supraglottitis with a 32% mortality.[2] Subsequent series have noted a much lower mortality felt to result from improved physician awareness of the need for urgent intervention.[3]

The presentation in pediatric and adult patients is similar with the exception that airway obstruction occurs more rapidly in the child. The initial symptom is usually throat pain which may be accompanied by fever and malaise followed rapidly by the development of odynophagia and dysphagia. Odynophagia is present in all cases of supraglottitis and, as it worsens, causes drooling. At this stage the airway is likely to be compromised somewhat by the swollen epiglottis which the patient generally maneuvers in a way to maximally open the airway by sitting erect and leaning forward. A muffling of the voice due to increased mass above the vocal cords is generally present. Stridor is generally a late sign which predicts the onset of complete airway obstruction which, in the absence of acute intervention, is followed by death.

The key to successful treatment lies in making the diagnosis before airway collapse. Emergent referral to a specialist capable of difficult intubation, rigid bronchoscopy and tracheotomy in the supportive environment of the operating room is indicated. In pediatric patients maneuvers such as placing the patient supine, phlebotomy, inserting an IV, and manipulation of the oral cavity may precipitate complete airway obstruction. As a result, immediate transfer to the operating room without the need for any further intervention is indicated for the child with signs and symptoms strongly suggestive of rapidly progressing supraglottitis.

Although an adult presenting with airway obstruction may require similar treatment, the larger airway may permit a less urgent approach to the workup. Flexible fiber optic laryngoscopy may be safely performed in most adults in this scenario without inducing airway obstruction. Observation in hospital in a closely monitored environment without intubation or tracheotomy may be possible for adults with milder cases. Radiographic imaging with anteroposterior and lateral soft-tissue films of the neck may be useful in those cases not requiring urgent intervention and will demonstrate marked enlargement of the epiglottis and aryepiglottic folds with filling of the hypopharynx and obliteration of the valleculae. Blood cultures as well as a complete blood count may be useful in predicting recovery without the need for airway intervention. A white blood count greater than 15,000 with a left shift and positive blood cultures predict a more fulminant course.[3,4]

Observation of adult patients without intubation may be supplemented with general measures to limit progression of supraglottic edema with cool mist by face mask, inhalations of racemic epinephrine, and intravenous dexamethasone. Antibiotic therapy should be targetted initially toward Hemophilus and modified according to blood culture results. Popular drugs as initial treatment include ceftriaxone, cefuroxime, or ampicillin/sulbactam combination with chloramphenicol or aztreonam as alternatives in the penicillin allergic patients. Bacterial supraglottitis generally resolves rapidly following initiation of therapy, with dysphagia and odynophagia resolving within 24 hours and epiglottic swelling disappearing 2 to 3 days later.[5]

Croup is used to describe several different clinical syndromes with the common features of respiratory obstruction along with a barking cough. **Laryngotracheobronchitis (LTB)** and **spasmodic croup** are two separate syndromes worthy of discriminating.

Laryngotracheobronchitis most commonly presents in children less than 3 years of age with the dominating signs of biphasic stridor and a "croupy" cough. A mild upper respiratory tract infection generally precedes the onset of these symptoms which generally intensify at night and improve during the day. LTB generally occurs during the winter months and is due to infection of the larynx, subglottis, and lungs with parainfluenza or respiratory syncytial viruses. Other viruses may be implicated as well with bacterial superinfection with Streptococcus, Staphylococcus or Hemophilus developing with more extensive involvement.

The most concerning feature of LTB is biphasic stridor resulting from edema with narrowing of the subglottis. Soft tissue radiographs of the neck will demonstrate a characteristic narrowing in this region referred to as a "steeple sign" or "pencil point sign".

Supportive treatment with cool mist inhalations and racemic epinephrine treatments as necessary are generally sufficient to permit natural resolution of the disease process. The value of systemic steroids is controversial. Repeated mild episodes or progression of a moderate episode warrants referral to an Otolaryngologist to assess for possible predisposing factors such as subglottic stenosis.

Spasmodic croup generally presents as recurrent episodes of upper airway obstruction associated with a barking cough in children without other signs of toxicity. Airway intervention or even admission to the hospital for observation are rarely necessary. Symptoms generally resolve quickly when treated with exposure to high humidity (i.e. running the shower). The value of other treatment with medication such as systemic corticosteroids, antihistamines, and subcutaneous epinephrine is questionable.[19,20]

Acute Hoarseness with Airway Obstruction: Foreign Body Ingestion

Due to protective reflexes and anatomic relationships, foreign bodies rarely lodge in the larynx at the level of the glottis and therefore generally are unassociated with the onset of hoarseness. They more commonly come to rest above the glottis in the oropharynx, adjacent the larynx in the hypopharynx, or pass inferiorly into the lungs or esophagus. Depending on the size and location of the object, a laryngeal foreign body may present with dysphonia or aphonia in addition to stridor and pain. Treatment begins with emergent consultation for visualization of the larynx. In those patients without signs of airway compromise and in clinics where visualization of the larynx is not possible, a lateral neck

radiograph may help in demonstrating a radio-opaque structure. Endoscopic removal under general anesthesia is generally the treatment of choice. In cooperative adults, removal under local anesthesia as an office procedure employing indirect mirror examination and specially designed instruments may occasionally be attempted.

Acute Hoarseness with Airway Obstruction: Neurogenic

Bilateral vocal cord paralysis may occur after either neck injury or surgery (ie thyroidectomy) and therefore may present acutely. The onset of symptoms most often develop immediately after the injury or within several days. Rarely the immobile vocal cords may be positioned such that symptoms are not noted until long after the injury. More gradual onset of airway obstruction may develop in the face of bilateral vocal cord paralysis that develops in association with a malignancy or progressive neurologic disorder.

Airway obstruction with a normal voice will occur if the vocal cords are positioned in the midline with the glottis closed. Hoarseness as breathiness will occur without airway obstruction if the vocal cords are positioned laterally with the glottis open. Dysphagia with aspiration is most likely to occur with the vocal cords in this position. In an intermediate fashion, both hoarseness and airway obstruction develop if the position of the vocal cords permits the glottis to be partially open, resulting in both breathiness and inspiratory stridor.

The treatment of bilateral vocal cord paralysis is based on a compromise to either improve the airway or the voice. Surgery may be performed to enlarge glottis with removal or reposition part of the vocal cord or arytenoid cartilage to open the airway resulting in a breathy voice. Alternatively, the voice and swallowing may be improved at the expense of the airway with procedure to close the glottis with vocal cord medialization requiring a tracheotomy. Prognosis for return of function, associated symptoms (aspiration) and the desires of a fully informed patient are critical determinants to which among the many treatment options should be considered. The occupation of the patient is a critical consideration as to choice of treatment. The choices of a lawyer who requires a strong voice and can hide a tracheotomy tube under his collar may be significantly different from those of a professional scuba diver who cannot work with a tracheotomy tube. Experimental surgery performed to reinnervate the larynx to restore the potential to open the glottis for inspiration and to close the glottis for speech and swallowing has been under investigation for many years but is still not widely practiced.

Chronic Hoarseness

Chronic hoarseness generally warrants referral to an Otolaryngologist. The urgency with which the referral is made depends on the symptom complex. If upper airway obstruction is considered, evaluation by the otolaryngologist should be made within minutes. If cancer is a concern, the referral should be made within days. If vocal nodules are the most likely diagnosis, the referral is not urgent, unless the patient is a singer, in which case he may request an evaluation emergently in order to prepare for a performance.

Chronic Hoarseness: Neoplastic

Most patients with **laryngeal cancer** are smokers who, prior to the development of cancer, may have a rough voice associated with the chronic laryngeal irritation of tobacco smoke. Voicing abnormalities developing with a laryngeal cancer vary according to the location of the tumor in the larynx and may consist of voice changes identified as breathy, harsh or muffled. Additional presenting symptoms include hemoptysis, stridor, dysphagia, pain, airway obstruction and the appearance of a neck mass.

Glottic cancers developing on the vibrating surface of the vocal cord cause a dramatic change in voicing with a harshness that may have a breathy component if interference with glottic closure occurs. Because these symptoms usually present early in the development of the cancer, many glottic cancers are identified at an early stage that makes them readily curable. Glottic cancers rarely present with neck metastases both because the draining lymphatics in the vocal cords are sparse and because treatment is generally begun early in the course of the disease.

Supraglottic tumors develop in a "silent area" that generally does not produce symptoms until the tumor has enlarged to a significant size. A change in voice develops as a result of a tumor in this location generally only after it is sufficiently large to muffle the laryngeal output resulting in a "hot potato" voice similar to that seen with supraglottitis. Extension from the supraglottis to involve the vocal cord may occur at a later stage to cause a harsh or breathy voice. Supraglottic tumors may progress to a sufficiently large size without detection to present with acute airway obstruction. Hemoptysis and dysphagia are relatively late symptoms. Otalgia due to referred pain involving a branch of the vagus nerve often presents along with other symptoms. Because of the late development of symptoms associated with supraglottic tumors, many are identified only after regional lymphatic metastases to the neck develop.

Subglottic tumor are rare and account for 1% to 2% of all laryngeal cancers. These tumors often present with airway obstruction due to their location within the confines of the cricoid cartilage. In this region, narrowing presents with inspiratory and expiratory stridor due to the rigid, nondistensible nature of the cartilage cricoid ring.

Recurrent respiratory papillomatosis is a benign neoplastic process with a potential for malignant degeneration. It occurs commonly in a juvenile form which may disappear at puberty but may also persist throughout life. Wart-like growths occur in the upper aerodigestive tract with a predisposition for accumulation at the junction between

respiratory epithelium and either squamous epithelium or scar. As a result, the vocal cords are commonly involved with symptoms developing in a fashion similar to that of glottic cancer with hoarseness and occasionally airway obstruction.

The disease is felt to be viral in nature and affects not only grossly involved tissue, but adjacent normal appearing mucosa. As a result treatment with endoscopic removal is usually associated with recurrence. Symptoms may develop as soon as several weeks after treatment or as late as years after treatment depending on the rate of growth. The current treatment of choice is endoscopic removal employing carbon dioxide laser ablation with microscopic control. Biopsy to rule out cancer is important to perform in adult patients because of the potential for malignant transformation. Experimental work with photodynamic therapy as well as interferon has yet to show results that make these interventions first line therapy.

Chronic Hoarseness:Neurogenic

The voice is affected by a large number of disease processes involving both the central and peripheral nervous system in a fashion that may be predicted by the systemic symptoms. **Tremor**, which may compromise fine motor control by interference of digital coordination by oscillations in hand movement, may also cause rhythmic alterations in the pitch and amplitude of the voice. Successful treatment of the systemic process may result in associated voice improvement. Unfortunately, many neurologic diseases with associated voice problems cannot be cured. As a result, for many of these processes, the goal of the laryngologist and speech pathologist is not to restore normal function, but to maximize potential with a clear understanding of the limitations posed by the underlying disorder.

Spasmodic dysphonia is characterized by a stress-strained voice quality with intermittent breaks that often is severe to the point that the affected patient will retire from work and withdraw socially. Distinction between this process and muscle tension dysphonia may be difficult and, in our hands, is done through our voice clinic with a trial of voice therapy. Spasmodic dysphonia is generally resistant to improvement with voice therapy alone. Spasmodic dysphonia often presents in association with vocal tremor. The etiology of spasmodic dysphonia remains unclear although it is generally felt to represent a true dystonia rather than a psychiatric abnormality. Spasmodic dysphonia may be related to a disorder localized to the midbrain that is felt to represent a dystonia that may be isolated to the larynx or exist as part of a more generalized dystonia such a Meiges syndrome.

The diagnosis of spasmodic dysphonia is supported by fiber optic imaging of irregular and intermittent glottic closure with hyperadduction of the vocal cords more evident during connected speech than with sustained utterances. Laryngeal electromyography is of questionable value in supporting the diagnosis but some feel that characteristic patterns of activity may be elicited from the intrinsic laryngeal muscles.

Treatment is most effective with the use of Botulinum neurotoxin A (Botox®) placed within the substance of the laryngeal muscles to dampen the excessive action induced muscle contractions causing the characteristic voicing patterns. Systemic treatment with medications used for generalized dystonias such as artane is generally not as effective as Botox®.[6]

Myasthenia gravis is a disorder of muscle activity related to the presence of antibodies reacting with the receptor site of the neuromuscular junction that may initially present with weakness associated with swallowing or speech impairment. Although specific distribution of muscle weakness may be limited, there is usually weakness noted around the orbicularis oculi and upper facial muscles that may progress to more widespread involvement. In addition to an asthenic vocal quality, hypernasal speech may be apparent because of palatal weakness.

There is an urgency in diagnosing this disease in order that treatment be instituted. Myasthenia gravis is a potentially lethal disease which may result in death from respiratory failure. The diagnosis is made by identifying a response with improved muscle strength following the administration of neostigmine (Prostigmine) or edrophonium (Tensilon). To establish the diagnosis in a patient with laryngeal complaints (which requires more time to assess than eye weakness), evaluation of the patient 30 to 40 minutes after administration of the longer acting neostigmine is generally preferred to the use of edrophonium which acts for only 30 seconds to 2 minutes.[7]

Amyotrophic lateral sclerosis is a disease caused by loss of anterior horn neurons and is characterized by progressive weakness which commonly presents with involvement of the tongue, pharynx, and vocal cords. Atrophy of the sides of the tongue coupled with fasciculations as well as palatal weakness and a decreased gag reflex are tip-offs to this disease process. There is no treatment available to reverse the progression of this disease toward death, but supportive help with speech and swallowing therapy can be of significant benefit. Tracheotomy and enteral tube feedings are commonly required as the disease approaches its final stages.

The voicing pattern of patients with **Parkinson's disease** is distinctive as a monotonous breathy voice characterized as "hypophonic". The presentation of voice in an emotionless manner coupled with the masklike fascies common to these patients may mistakenly identify this process as a psychiatric disorder. The additional findings of cogwheel rigidity of the limbs, pillrolling tremor and disorders of posture and gait are distinctive of more advanced cases of Parkinson's disease.[8] Speech disability has been reported as the most significant aspect of this disease process interfering

with quality of life.[22] This interference in the ability to communicate has been identified in 75% of patients with the disease. In addition to the general medical management of Parkinson's disease, speech therapy may be of benefit.

Chronic Hoarseness: Psychogenic

Functional aphonia, or psychogenic aphonia most commonly occurs in females and is identified by the absence of voicing in a patient whose laryngeal examination is otherwise normal. The capacity to effect glottic closure may be identified through a number of mechanisms including eliciting a normal cough. Treatment is with psychiatric counselling and voice therapy. Occasionally resistant cases may be treated with procedures to interfere with the normal laryngeal mechanisms such as lidocaine injection to the a recurrent laryngeal nerve to induce a temporary laryngeal paralysis.[21]

Muscle tension dysphonia is a common voicing disorder most often seen in patients with stressful jobs and significant demands on use of their voice. A tight, strained voice is commonly seen which, in severe cases, may be difficult to differentiate from spasmodic dysphonia. Associated neck muscle tightness may be identified. Treatment is with speech therapy which may include innovative techniques such as circumlaryngeal massage and biofeedback.

Chronic Hoarseness: Traumatic/Environmental

Voice misuse (voice abuse) with or without other forms of laryngeal irritation may result in chronic changes to the vocal cords resulting in nodules, polyps, cysts, scarring and chronic edematous changes termed "polypoid chorditis". These disorders are most commonly diagnosed with laryngeal imaging with mirror examination. Disorders of the vocal cords that are not clearly apparent on mirror examination can often be identified with use the newer videostroboscopic units that not only permit more sophisticated assessment of the motion of the vocal cord structures during phonation, but also provide enhanced illumination and magnification. If cancer is suspected, a biopsy usually under general anesthesia is done. Those patients who are surgical risks or at relatively low risk for laryngeal cancer, may be followed with sequential videoendoscopies to determine if the nature of the vocal cord lesion is stable and less likely to require biopsy.

For the majority of laryngeal lesions felt to represent benign processes, a course of voice therapy is offered before surgical treatment in order that the vocal behavior that induced the abnormality initially be modified. Surgical treatment is reserved for those who have persistence of symptoms following voice therapy, who have modified their vocal behavior sufficiently to make recurrence of the lesion unlikely, and who accept the risks of surgical intervention which include permanent scarring of the vocal cords.

References

1. Shumrick KA and Shumrick DA: Inflammatory Diseases of the Larynx, ch 23 pp 249-278, in *The Larynx: A Multidisciplinary Approach,* ed Marvin P. Fried, Little, Brown and Co, Boston, 1988.

2. Hawkins DB, Miller AH, Sacks GG and Benz RT: Acute epiglottitis in adults. Laryngoscope 83:1211, 1973.

3. Ossoff RH, Wolff AP, and Ballenger JJ: Acute epiglottitis in adults: Experience with fifteen cases. Laryngscope 90:1155, 1980.

4. Jones H and Camps F: Acute epiglottitis: Supraglottitis. Practitioner 178:223, 1957.

5. Mustoe T and Strome M: Adult epiglottitis. Am J Otolaryngol 4:393, 1983.

6. Ford, Charles N: "Spasmodic Dysphonia" pp 448-450 in *Current Therapy in Otolaryngology — Head and Neck Surgery,* Fifth Edition, George Gates, ed, Mosby, St. Louis, 1993.

7. H. Richard Tyler: "Neurologic Disorders" ch 16, pp 173-178 in *The Larynx: A Multidisciplinary Approach,* Marvin P. Fried, ed, Little, Brown and Co, Boston, 1988.

8. Darley F, Aronson A and Brown J: Motor Speech Disorders. Philadelphia, W.B. Saunders, 1975.

9. Sataloff RT, Lawrence VL, Hawkshaw MJ and Rosen DC: "Medications and Their Effects on the Voice" Ch 15, pp 216-225 in *Vocal Arts Medicine — The Care and Prevention of Professional Voice Disorders,* eds Benninger MS, Jacobson BH, and Johnson AF, Thieme Medical Publishers, New York, 1994.

10. Van Lawrence L: Common medications with Laryngeal effects. ENT J, 66:318-322, 1987.

11. Rabinowitz HK: "Upper Respiratory Tract Infections" pp 793-809 in *Primary Care* — Vol 17, No.4, December 1990.

12. Koster F: Respiratory infections. In Barker LR, Burton JF, Zieve PD (eds): *Principles of Ambulatory Medicine.* Baltimore, William's and Wilkins, 1987, pp 347-355.

13. Lowenstein SR, Parrino TA: Management of the common cold. Adv Intern Med 32:207, 1987.

14. Merigan TC: Respiratory viral infections of adults. In Rubenstein E, Federman DD (eds) *Scientific American Medicine*. New York, Scientific American Inc, 1984, 7(XXV), pp 1-9.

15. Tucker H: "Chapter 11, Infections and Inflammatory Disorders" pp 231-244 in *The Larynx* Second Edition, Thieme Medical Publishers Inc. New York, 1993.

16. Fairbanks DNF: "Pocket Guide to Antimicrobial Therapy in *Otolaryngology — Head and neck Surgery* 7th Edition, American Academy of Otolaryngology — Head and Neck Surgery, Alexandria Virginia, 1993.

17. Jacobs RF, Yasuda K, Smith AL, Benjamin DR:"Laryngeal candidiasis presenting as inspiratory stridor" Pediatrics 1982;69:234-236.

18. Hass A, Hyatt AC, Kattan M, Weiner MA, Hodes DS. "Hoarseness in immunocompromised children: association with invasive fungal infection. J Pediatr. 1987;111:731-733.

19. Friedman E.M.:"Inflammatory Illnesses of the Pediatric Airway" Ch 12 pp 137-142 in *The Larynx: A Multidisciplinary Approach*. Little, Brown and Company Boston.

20. Inflammatory Laryngeal Disease in Children, Sylvan Stool pp 439-440, in *Current Therapy* by Gates.

21. Smith ME, Darbe KP, Kirchner K and Blager FB: Simultaneous Functional Laryngeal Stridor and Functional Aphonia in Adolescent, in NCVS Status and Progress Report - 4 June 1993, 169-176.

22. King JB, Ramig LO, Lemke JH and Horii Y: Parkinson's Disease: Longitudinal Changes in Acoustic Parameters of Phonation, in NCVS Status and Progress Report - 4 June 1993, 135-149.

Acknowledgement

The authors acknowledge grant P60 DC00976 from the National Institute on Deafness and Other Communication Disorders.

NCVS Status and Progress Report - 7
December 1994, 197-205

The Application of Laboratory Formulas to Clinical Voice Management

Katherine Verdolini, Ph.D.
Ingo Titze, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa

Abstract

In this paper, we discuss the application of mathematical formulas to guide the development of clinical interventions in voice disorders. Discussion of case examples includes fundamental frequency and intensity deviations, pitch and loudness abnormalities, laryngeal hyper- and hypoadduction, and phonatory effort. The paper illustrates the interactive nature of theoretical and applied work in vocology.

Persons without extensive background in mathematics and physics may sometimes be perplexed by computationally-based work in voice science. Long pages of formulas may seem formidable, and the work may seem only remotely applicable to the clinical situation.

A closer look at some of the computational work in voice science reveals greater applicability than might initially appear, and greater accessibility - even without significant prior training in mathematics. The purpose of this paper is to illustrate these principles by demonstrating how selected mathematical formulas can be used to confront some problems in clinical voice, and by showing how some of the computational work can be approached without extensive prior training in mathematics.

The initial formulas discussed are familiar to most readers with previous training in basic speech science. The later formulas may be new, because they have only recently been introduced. Together, the discussion demonstrates the interactive nature of clinical and mathematical science. In some cases, clinical problem-solving can be enhanced by appealing to mathematical formulations. In other cases, computational work is stimulated by clinical questions. We will highlight examples of each case.

The Formulas and Their Application

Case # 1: High fundamental frequency with vocal fold bowing.

In our clinical caseload, a common finding with vocal fold bowing is a high average fundamental frequency (F_0) in speech, particularly in adult males. We have found F_0s as high as 196 Hz in an adult male with bowing, as compared with a norm of about 110 Hz (Hollien & Shipp, 1972). Male patients may describe high pitch as the primary complaint prompting them to seek treatment, and may report being mistaken as female on the telephone.

The physiological basis for the F_0 abnormality with bowing was initially unclear to us. However, the following formula indicates that only a limited number of parameters need to be considered to explain this finding, as discussed next.

The formula: F_0. F_0 regulation in humans can be simplified by considering a vibrating string (see for example, Titze, 1994). The relevant formula is:

$$f_o = \frac{1}{2L}\sqrt{\frac{\sigma}{\rho}} \qquad (1)$$

In this equation, L represents the length of the membranous vocal folds in oscillation, σ represents longitudinal tissue stress (defined as force per unit area), and ρ indicates the density of vocal fold tissue. Increases in F_0 may be theoretically due to decreased vocal fold vibratory length (within a given subject), increased vocal fold stress (stiffness), or decreased vocal fold tissue density (actually not a factor in phonation with normal folds).

Clinical application. A re-evaluation of some of our patients with bowed larynges, with the parameters from Equation 1 in mind, suggested that a decrease in vocal fold vibratory length might explain an increased F_0 in some cases. Specifically, when we reviewed the videoscopic examinations of patients with bowing, in several cases a posterior vocal fold compression was apparent during phonation, with compression perhaps being an attempt to achieve glottal closure. Whatever the reason, the compression effectively shortened the length (L) of the membranous vocal folds available for oscillation, resulting in increased F_0. To decrease F_0 to more normal levels in such cases, a reasonable therapy approach would be to decrease compensatory vocal fold squeezing, thus increasing the vocal fold vibratory length.

Case #2: Need for intensified voice output: Option 1 - Supraglottal tuning.

Clinical case. Some patients, including school teachers, stage actors, and courtroom attorneys, may need a greater voice intensity than usually required for conversational purposes (about 65 dB at one meter; see for example, Fletcher, 1953). In such cases, supraglottic "tuning" of the vocal tract to glottal output could be used as a secondary mechanism to achieve greater than normal intensity levels, without increasing subglottal pressure or glottal resistance. The basis for the tuning phenomenon is discussed next.

The formula: Formant frequency. Equation 2 describes the parameters regulating formant frequencies for a uniform tube open on one end and closed on the other (see for example, Fant, 1970):

$$F_n = (2_n - 1)\left(\frac{c}{4} L\right) \qquad (2)$$

In this equation, F_n are the formant frequencies for any positive integer n, c is the sound of speed in air (about 350 m/s for warm, moist air; Titze, 1994), and L is the length of the vocal tract. All formant frequencies F_n decrease as the length L of the tube (vocal tract) increases, and formant frequencies increase as the length of the vocal tract decreases. Vocal tract tuning can be achieved by lining up a formant frequency with a source harmonic, usually the second or third harmonic in theatre speech (see Raphael & Scherer, 1987).

Clinical application. If fundamental frequency and formant frequencies are both manipulated so that F1 coincides with $2F_0$ or $3F_0$, voice output intensity can be increased by a maximum of about 6 dB (a quadrupling of output power) without changes in glottal adduction or in subglottal pressure (Titze & Sundberg, 1991). In fact, it is theoretically possible to increase output intensity to some degree with formant tuning even while decreasing respiratory and glottal effort, if tuning is optimized. According to one empirical report (Raphael & Scherer, 1987), $2F_0$ and F1 can be made to approximate for /o/ in the dipthong /ou/ by lengthening the vocal tract with an "inverted megaphone posturing" (Figure 1; from Lessac 1967, for the dipthong /ou/), which lowers the larynx and protrudes the lips. In that study, the effect of the inverted megaphone posture was to enhance F1 and presumably, boost the net output intensity.

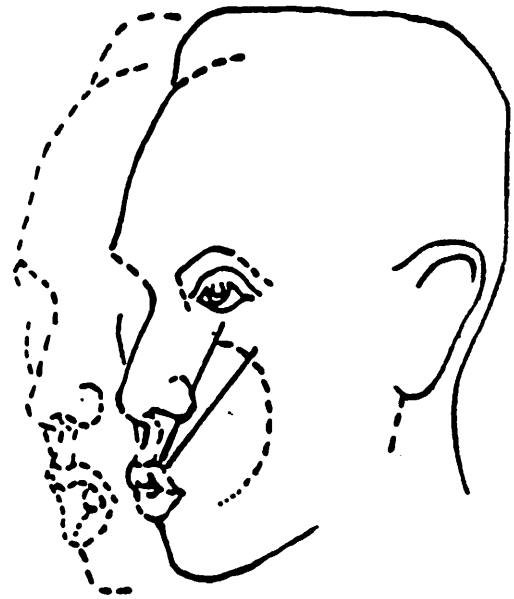


Figure 1. Inverted megaphone posture for /o/ in dipthong /ou/. Reprinted with permission from Lessac (1967).

Case #3: Need for intensified voice output: Option 2 - glottal power.

Clinical case. In some cases, vocal tract tuning may be inadequate for a functionally sufficient boost in vocal intensity. An example might be a weak voice secondary to abdominal paralysis. Vocal intensities might be expected to be particularly low in such cases, due to extremely low subglottal pressures. The next formula provides us with further information about aerodynamic parameters that can be manipulated to increase output intensity.

The formula: Glottal power. Equation 3 describes aerodynamic power available from the lungs. According to this formula (van den Berg, Zantema, & Doornenbal, 1957):

$$\mathcal{P} = P_s \times U \qquad (3)$$

$\mathcal{P}$ increases as subglottal pressure P_s increases and as mean glottal flow U increases. Flow is regulated by the degree of laryngeal adduction such that hyperadduction restricts flow.

Clinical application. From Equation 3, two complementary approaches could be taken to increase voice output in a patient with extremely weak voice, for example due to an abdominal paralysis. First, subglottal pressure P_s could be increased by teaching the patient to inspire to high lung volumes prior to speech, thereby exploiting the high passive lung pressures generated from recoil forces (see for example, Hixon, 1991). Second, transglottal flow could be increased by teaching the patient to limit laryngeal adduction during phonation, if hyperadduction is present as a compensatory strategy.

Case #4: Cases requiring manipulation of laryngeal adduction

Clinical case. A wide range of clinical voice cases are physiologically based in laryngeal adduction abnormalities. Many benign traumatic lesions, including nodules, polyps, granulomas, and generalized edema imply the possible presence of a chronic hyperadducted posturing during phonation, because of a relation between degree of adduction and intraglottal impact pressure (Jiang & Titze, 1994) and between impact pressure and lesion etiology for these conditions (Jiang & Titze, 1994; Hillman, Holmberg, Perkell, Walsh, & Vaughn, 1989; Verdolini-Marston, Hoffman, & McCoy, 1994). Another large set of clinical cases involves laryngeal hypoadduction, including adductory paralyses with the affected fold(s) away from midline, or vocal fold bowing. An equation for what we call "laryngeal economy," defined from a therapeutic perspective, may provide a unifying framework for many therapy techniques used to treat these conditions.

Conceptual background. In clinical cases involving laryngeal hyperadduction, a primary goal of therapy is to decrease adduction relative to the presenting physiology. In cases of laryngeal hypoadduction, the goal is often to increase glottal closure. Inherent in these observations is the notion that there is some optimal or idealized level towards which adduction migrates with therapy. That adduction level should produce a functionally adequate output intensity while at the same time minimizing the risk of laryngeal injury. We refer to this hypothetical configuration as "optimal laryngeal configuration" (O.L.C.), or a maximum of "laryngeal economy" (E_L), where a preliminary definition of economy is the ratio of glottal power divided by vocal fold impact pressure (force per unit area) as a function of laryngeal adduction. The maximum of the economy curve (corresponding to O.L.C.) would define the laryngeal adduction level producing the optimal trade-off between glottal power (the greatest possible) and vocal fold impact pressure (the least possible). We undertook an initial computational study to identify the idealized adduction level according to these parameters.

The formulas: A combination of formulas describing the vocal economy curve. As noted, the laryngeal economy curve would be the curve of acoustic power divided by the curve of vocal fold impact pressure, as a function of laryngeal adduction level. In our initial computations, acoustic power was represented as the power generated at the glottis with varying adduction. That is, sub- and supraglottic factors for voice output were not considered. However, ongoing work in our program by David Berry (unpublished data) indicates that the results are not qualitatively different when vocal tract influences are taken into account.

The first relevant equation was described by Titze and Sundberg (1992):

$$\bar{\mathcal{P}}_g = \rho_{6\pi c} \frac{Q_o}{Q_s} \dot{u}_m^2 \qquad (4)$$

In this equation, $\bar{\mathcal{P}}_g$ is glottal acoustic power, ρ and c are the density and speed of sound in air (1140 kg/m^3 and 350 m/s, respectively) Q_o is the open quotient, or the time that the glottis is open during oscillation divided by the total cycle time, Q_s is the skewing quotient, or the duration of increasing flow divided by the duration of decreasing flow (Figure 2), and $\dot{u}_m^2$ is maximum airflow deceleration rate, or the maximum rate at which glottal airflow is slowed during vocal fold closing (Figure 2).

Note that in Equation 4, there are several constants (ρ, 6, π, and c), and one parameter that we can treat a constant for this simplified discussion: Q_s (about 1.625, based on average data for males and females for quiet, normal, and loud phonation, at a spontaneous pitch; Hillman et al., 1989). If we wish to examine the general shape of the power curve, but are not concerned with the exact power values, we can consider all the constants together equal to an arbitrary value of "1," and thus simplify the equation:

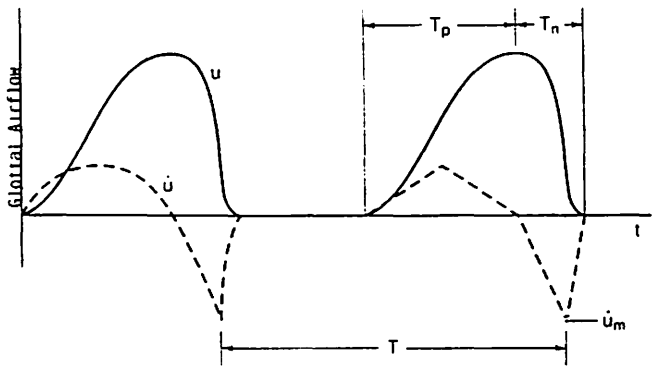


Figure 2. Glottal airflow u over time t. Airflow deceleration shown by $\dot{u}_m$, maximum airflow deceleration shown by $\dot{u}_m$. T_p is increasing airflow portion of the signal, grossly corresponding to vocal fold opening, and T_n is decreasing airflow portion, grossly corresponding to vocal fold closing. Skewing quotient Q_s is T_p/T_n. Reprinted with permission from Titze, 1994.

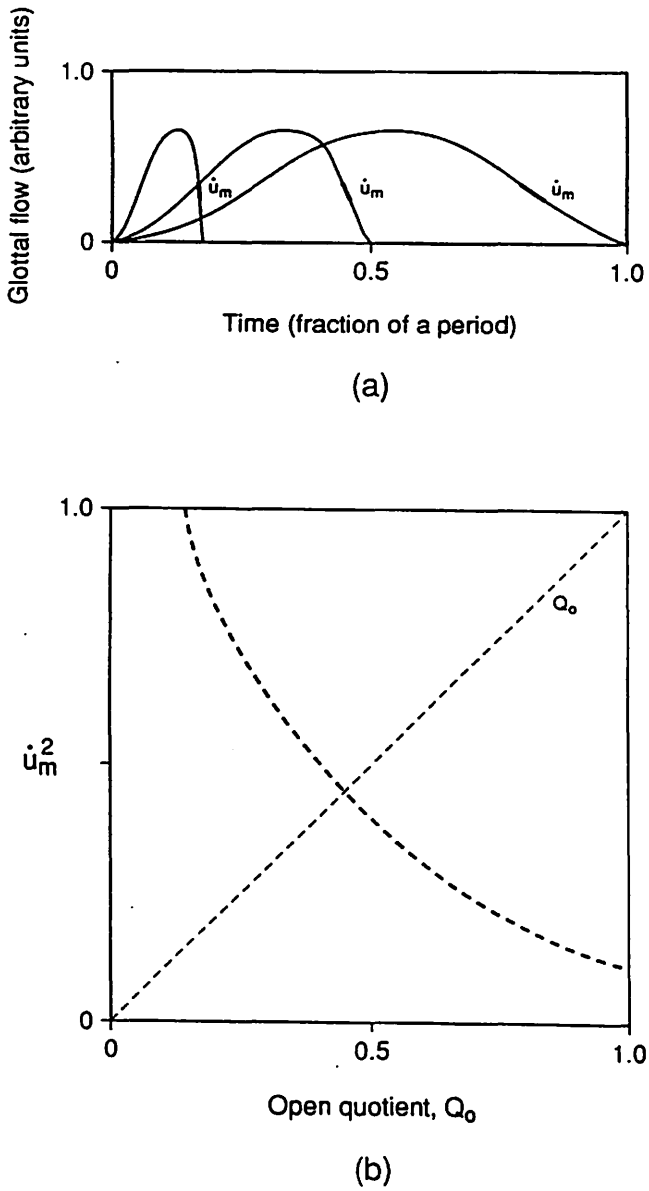


Figure 3 (a). Glottal flow over time, showing variations in maximum flow deceleration rate $\dot{u}_m$. (b). Variations in maximum flow deceleration rate squared ($\dot{u}_m^2$) and open quotient Q_o as a function of Q_o. Reprinted with permission from Titze, 1994.

$$\overline{P_g} \sim (Q_o)\,(\dot{u}_m^2) \qquad (5)$$

In further calculating the economy curve, we can take advantage of the fact that Q_o and $\dot{u}_m$ covary as indicated in Figure 3: longer vocal fold open times per cycle (greater Q_os) correspond to slower vocal fold closing rates (larger $\dot{u}_m$s; Titze, 1994). Multiplying the points on the $\dot{u}_m^2$ curve with the points on the Q_o curve, both as a function of Q_o

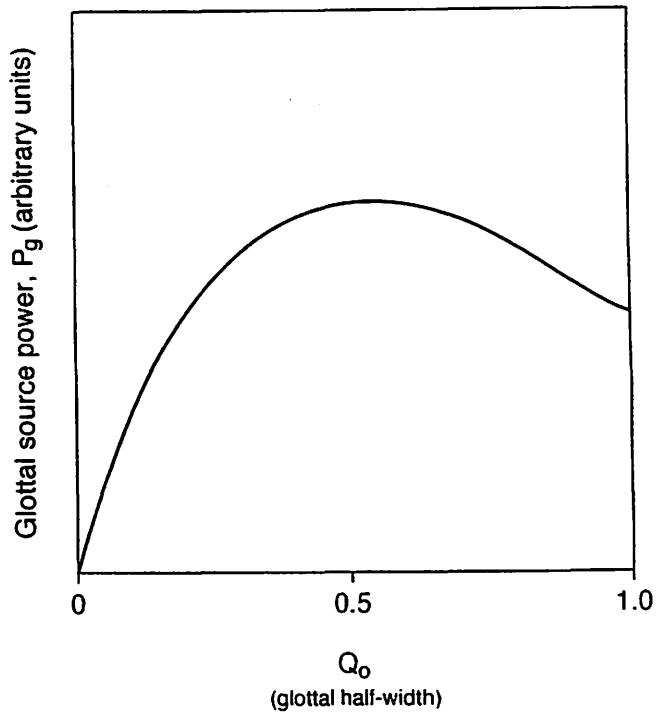


Figure 4. Glottal source power P_g as a function of open quotient Q_o and glottal half-width. Based on Titze, 1994.

(Figure 3), and as indicated in the simplified glottal power equation (Equation 5), we obtain the glottal power curve as a function of Q_o (Figure 4). We can further relate this glottal power curve to laryngeal adduction level, because of a direct relation between Q_o and glottal width, or abduction (Figure 4; Scherer & Titze, 1994). The resulting function shows a curvilinear relation of glottal power to adduction such that the greatest vocal fold output is obtained with an intermediate laryngeal configuration, with the vocal folds barely abducted (Titze, 1994). The curve in Figure 4 will provide the data points for the numerator in the laryngeal economy curve.

The denominator in the economy function is vocal fold impact pressure as a function of adduction. This function is based on empirical data from experiments with excised canine larynges (Jiang & Titze, 1994). In those experiments, intraglottal impact pressure was measured at three distinct glottal half-widths: -0.5 mm, 0.5 mm, and 1.5 mm (the negative glottal half-width was produced by pressing the arytenoids together; there was no tissue overlap, but vertical tissue "squishing"). The results for these three data points are shown in Figure 5. As seen, impact pressure decreases as glottal half-width increases.

Two types of curves could be generated to fit these discrete empirical data points. One would be a linear curve, described by the equation:

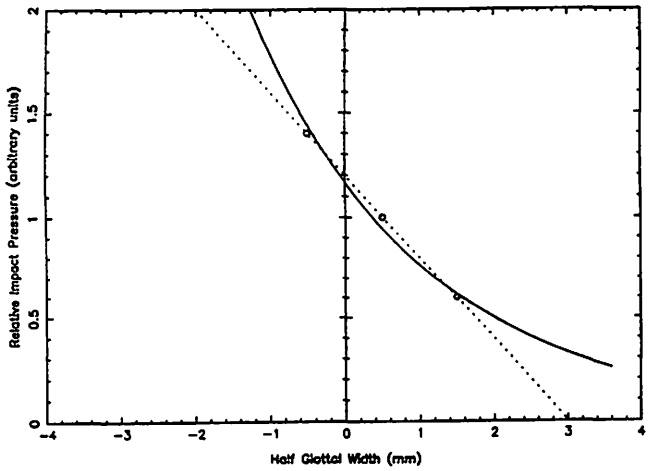


Figure 5. Relative intraglottal impact pressure as a function of glottal half-width, in mm, shown with open circle (data from Jiang & Titze, 1994). Linear and curvilinear solutions shown with dotted and solid lines, respectively (from David Berry, unpublished data).

$$y = mx + b \quad (6)$$

where y = impact pressure, m is the rate of change in impact pressure with each unit increase in glottal half-width x, and b is the amount of impact pressure when glottal half-width is zero. As shown in Figure 5, this curve has the undesirable property of impact pressure becoming negative for glottal half-widths of 3 mm or greater, an unphysiological outcome. Another solution would be a curvilinear fit, described by the equation:

$$y = e^{mx+b} \quad (7)$$

where e = 2.71 (Figure 5). This curve has the desirable property of avoiding negative values of impact pressure (curves from David Berry, unpublished data).

With the curves for the numerator and the denominator in the economy ratio, it is quite straightforward to show the ratio curve itself. Each point on the glottal power curve is divided by the corresponding point on the impact pressure curve (for a given level of adduction or glottal half-width), and the economy curve is generated. The result is shown in Figure 6.

This curve tracks the degree of trade-off between glottal power and vocal fold impact pressure with varying vocal fold adduction or glottal half-width. The maximum of this curve - representing the greatest glottal power for the least potential laryngeal injury - is obtained with slightly abducted vocal folds, slightly more abducted than the configuration required for maximum glottal power alone (see Figure 6). Pressed, or greatly adducted vocal folds, as well as hypoadducted vocal folds, result in relatively poorer trade-off (Figure 6).

Clinical significance. We assume that the maximum of the economy curve represents an idealized or "optimum laryngeal configuration" (O.L.C.) towards which adduction levels migrate with therapy across a broad range of voice disorders. Figure 7 (next page) illustrates these points graphically.

Based on empirical evidence from studies that we have conducted, O.L.C. roughly corresponds to the laryngeal configuration used to produce what we call "resonant voice," based on work by Lessac (Lessac, 1967; Peterson, Verdolini-Marston, Barkmeier, & Hoffman, 1994; Verdolini, Druker, Palmer, & Samawi, in review), and that one study indicates may be as effective in the treatment of nodules as the more traditional quiet, breathy voice (Verdolini-Marston, Burke, Lessac, Glaze, & Caldwell, in press). Other reports describe a similar laryngeal configuration as producing "flow" phonation (Gauffin & Sundberg, 1989). In sum, the O.L.C. should be pertinent as a reference point in therapy for both hyper- and hypofunctional cases.

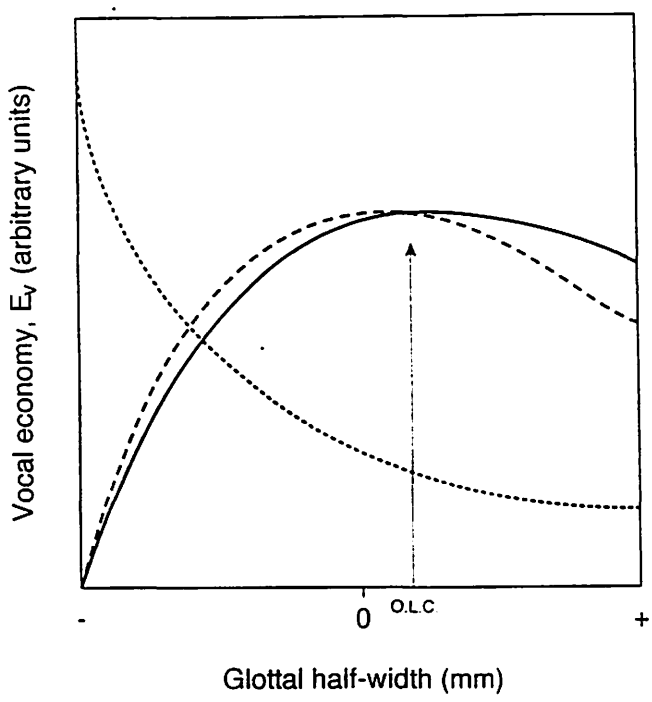
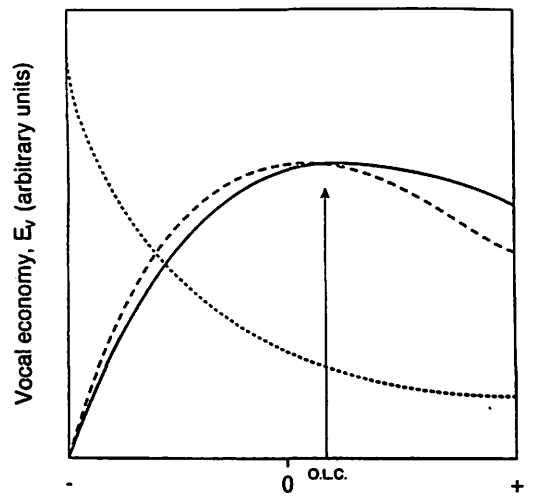


Figure 6. Vocal economy curve $_{EV}$ as a function of glottal half-width in mm indicated as solid line. Glottal power curve shown with dashed line, and intraglottal impact pressure with dotted line.



Clinical condition:	Hyperadducted		Hypoadducted
Voice therapy type:	Shout	Resonant Flow	Falsetto Yawn Breathy

Figure 7. Vocal economy curve $_{EV}$ described in Figure 6, indicating hyper- and hypoaddducted clinical conditions and voice therapy techniques [Lou Silverman Voice Treatment which uses "shout" voice, Ramig (in press); resonant voice, flow mode, falsetto, yawn-sigh, breathy

The laryngeal economy curve not only shows the optimum trade-off value for laryngeal adduction, but it also provides an organizing framework for many different techniques in voice therapy. Take for example the case of acute edema (incipient nodules) presumed to be secondary to extreme - but abbreviated - hyperadduction in time. In this case, the therapeutic prerogative would be to minimize vocal fold impact pressures, setting aside the issue of output intensity, to minimize further immediate intracordal trauma. Barring absolute voice rest, Figure 7 tells us that the best approach would be to phonate with widely abducted vocal folds, as with breathy voice (Casper, Colton, Woo, & Brewer, unpublished observations; Colton & Casper, 1990; Peterson et al., 1994), falsetto mode (Hollien, 1974), yawn-sigh (Boone & McFarlane, 1994), and preferably, quiet voice (to minimize vocal fold vibratory amplitudes). These approaches to voicing would not be maximally "economical," as described here, but it should promote maximal short-term laryngeal healing.

Consider on the other hand the case of voice in Parkinson's disease, with vocal fold bowing (Hansen, Gerratt, & Ward, 1984). To overcome the characteristic incomplete vocal fold closure, the clinical prerogative would be to increase vocal fold adduction through maximum effort phonation tasks as in Lou Silverman Voice Treatment (Ramig, in press), resulting in O.L.C. or in cases of "overshoot," an even more adducted laryngeal configuration.

Case #5: Phonatory effort

Clinical case. Excessive phonatory effort is a common complaint across a wide range of voice disorders. In fact, effortful phonation can be considered a general indicator of a voice problem, with rare exceptions (Verdolini, 1994). Assuming that perceived phonatory effort depends to a significant degree on respiratory effort, the formula below provides us with several parameters to consider in responding to this problem clinically.

The formula: Subglottal pressures required for phonation. Equation 8 tells us that the subglottal pressure P_s required for phonation is determined by the following relations (Titze, 1988; Verdolini-Marston, Titze, & Druker, 1990):

$$P_s = \frac{2k}{T} Bc \frac{w}{2} \qquad (8)$$

In this formula, k is an empirical constant (about 1.1), T is vocal fold thickness, B is vocal fold tissue damping (which reflects the fluid viscosity of the tissue), c is the speed of the mucosal wave along the medial edge of the vocal fold, and $w/2$ is the pre-phonatory glottal half-width. The equality in the equation represents the minimal subglottal pressure required to initiate and sustain vocal fold oscillation; it is called the phonation threshold pressure (PTP). The formula tells us that PTP decreases as vocal fold thickness increases, as vocal fold tissue damping B decreases (as the tissue becomes more watery, as opposed to dried or sticky), as the mucosal wave propagation speed c decreases (with decreasing stiffness, i.e. vocal fold shortening), and with decreased glottal half-width $w/2$ between the vocal folds.

Clinical applications. From this formula, a series of useful clinical applications follow. If phonatory effort covaries with PTP, this effort should decrease as the vocal folds are shortened; shortened vocal folds are thicker and more lax, increasing T and decreasing c (with a resultant decrease in pitch within a given register).[1] Effort should also decrease as the vocal folds become more watery, hypothetically achieved by hydration treatments discussed below, and as the vocal folds come closer together during phonation. Theoretically, the summary solution for patients complaining of vocal fold effort would be to lower the pitch (within a given register), to hydrate by increasing water intake, humidifying the envi-

[1] In fact, a higher pitch may often require decreased subglottal pressure if the pitch increase corresponds to a shift from chest to falsetto register (see for example, Kunze, 1962).

ronment, or taking mucolytic medications, and to approximate the vocal folds.

Thus far, there is partial empirical support for one of the relations in Equation 8, and that is the proposed relation between vocal fold tissue viscosity and subglottal pressures during phonation and specifically, PTPs. Two experiments have been conducted with laryngeally healthy subjects to assess the effect of hydration treatments - argued to affect vocal fold tissue viscosity by internal or external means - on PTPs (Verdolini-Marston et al., 1990; Verdolini, Titze, & Fennell, in press). In these experiments, hydration treatments consisted of 4-hour exposures to high humidity environments (80-100% humidity), maximum tolerated water consumption (16-128 oz), and mucolytic medications (4 tsp of Robitussin expectorant). Together, the treatments were assumed to decrease vocal fold tissue viscosity either superficially, by increasing the fluidity of respiratory secretions, or internally, by adding water to vocal fold tissue. The combined results across these two experiments are shown in Figure 8. These results showed a decrease in PTPs following hydration treatments, presumably because of a decrease in vocal fold tissue viscosity, consistent with theoretical predictions. The hydration effect was particularly marked at high pitches in both experiments. The caveat is that although the hydration effect was statistically reliable, reductions in PTPs with hydration treatments were small in magnitude. It was further not clear from the experiments which aspect of the hydration treatments were responsible for the PTP reductions: humid environment, water intake, or mucolytic medications.

Regarding subjects with voice disorders, another experiment was conducted assessing hydration effects in subjects with nodules and polyps (Verdolini-Marston, Sandage, & Titze, 1994). In that experiment, hydration and placebo treatments were each delivered over a 5-day period. Hydration treatments consisted of a 2-hr daily exposure to a 90-100% relative humidity environment, and instructions to take 3 tsp of Robitussin expectorant and to drink eight 16-oz glasses of water or more per day (all subjects came near to or exceeded the recommended amounts). Placebo treatments consisted of a 2-hr daily exposure to an "air-filtered" candlelit environment (with air filters actually deactivated), with instructions to take 3 tsp of a placebo medication and to perform 8 sets of 20 forefinger flexions or more per day (again, subjects approximated or exceeded the recommended amounts). PTP measures were made one day following the termination of treatments. The results indicated that PTPs were numerically lowest following the hydration treatment as compared with baseline and placebo conditions, but only at high pitches (Figure 9). However, the hydration effect was statistically unreliable even at the high pitch, in part due to a great variability in individual subjects' responses to treatment.

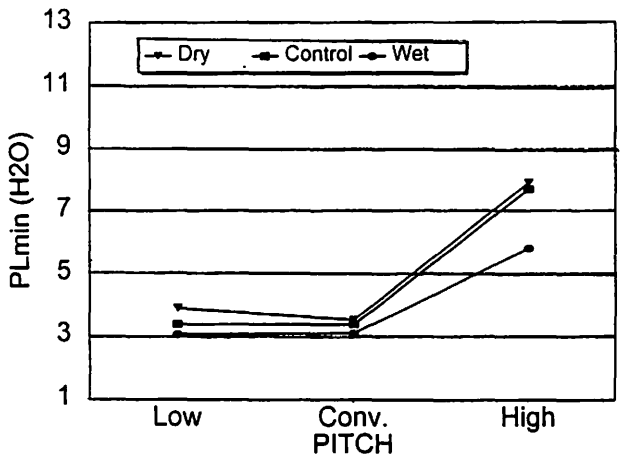


Figure 8. Phonation threshold pressure PTP in cm H_2O as a function of pitch, for dry, control, and wet treatment conditions. Average data based on Verdolini-Marston et al (1990) and Verdolini et al (in press).

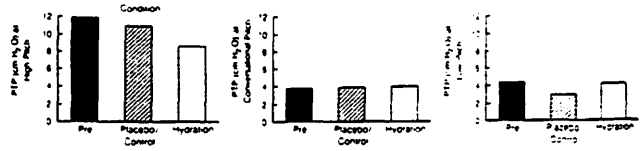


Figure 9. Phonation threshold pressure PTP in cm H_2O in pre-treatment, post placebo-control treatment, and post hydration treatment, at high, conversational, and high pitches. Reprinted with permission from Verdolini-Marston et al, 1994.

Why would statistically reliable hydration effects be obtained for laryngeally healthy subjects, but not for subjects with nodules or polyps? We think that the difference is not necessarily due to different physiological effects across subjects with normal larynges versus nodules or polyps, but rather due to a critical difference in the experimental protocols. In the study with subjects with nodules and polyps (Verdolini-Marston et al., 1994), PTPs were measured one day following treatment termination, as compared with immediate post-treatment measures in the studies with normal subjects (Verdolini-Marston et al., 1990; Verdolini et al., in press). Various aspects of the data indicated that the effect of hydration treatments wears off with time. Thus, it was likely that hydration effects were not clearly shown by subjects with nodules and polyps because much of the effect had worn off by the time the measures were made in that experiment, 24 hr after treatment termination. Of course, this interpretation requires further investigation. In the meantime, it should be kept in mind that if hydration treatments are effective, they may be effective only so long as they continue to be administered.

Summary

In conclusion, in this paper we have illustrated how some clinical problems can be addressed by appealing to analytical methods in voice science. We have also given an example of how these analytical methods are stimulated by clinical questions. The hope is that readers without extensive background in mathematical science can be encouraged to further consult basic science in confronting clinical issues.

Acknowledgements

This work was supported by Grant No. P60 DC00976 from the National Institute on Deafness and Other Communication Disorders. An earlier version of the paper was presented at the 1993 American Speech-Language-Hearing Association Convention, Anaheim, California, November 19-22. The authors thank Ms.s Mary Jo Yotty and Linnie Southard for secretarial support and Mr. Mark Peters for assistance with graphics.

References

Boone, D.R., & McFarlane, S.C. (1994). The voice and voice therapy (5th ed.). Englewood Cliffs, N.J.: Prentice Hall.

Casper, J.K., Colton, R.H., Woo, P., & Brewer, D. (1989, 1990). Investigation of selected voice therapy techniques. An amalgamation of two papers presented at the 18th and 19th Symposia: Care of the Professional Voice, Philadelphia, PA.

Colton, R.H. and Casper, J.K. (1990). Understanding voice problems: A physiological perspective for diagnosis and treatment. Baltimore: Williams and Wilkins.

Fant, G. (1970). Acoustic theory of speech production. The Hague: Mouton.

Fletcher, H. (1953). Speech and hearing in communication. N.Y.: D. Van Nostrand Company, Inc.

Gauffin, J. & Sundberg, J. (1989). Spectral correlates of glottal voice source waveform characteristics. Journal of Speech and Hearing Research, 32, 556-565.

Jiang, J.J., & Titze, I.R. (1994). Measurement of vocal fold intraglottal pressure and impact stress. Journal of Voice, 8, 132-144.

Hansen, D.G., Gerratt, B.R., and Ward, P.H. (1984). Cinegraphic observations of laryngeal function in Parkinson's disease. Laryngoscope, 94, 348-353.

Hillman, R.E., Holmberg, E.B., Perkell, J.S. Walsh, M., & Vaughn, C. (1989). Objective assessment of vocal hyperfunction: An experimental framework and initial results. Journal of Speech and Hearing Research, 32, 373-393.

Hixon, T.J. (1991). Respiratory Function in Speech and Song. San Diego: Singular Publishing Group, Inc.

Hollien, H. (1974). On vocal registers. Journal of Phonetics, 2, 125-143.

Hollien, H., & Shipp, T. (1972). Speaking fundamental frequency and chronologic age in males. Journal of Speech and Hearing Research, 15, 155-159.

Holmberg, E.B., Hillman, R.E., and Perkell, J.S. (1988). Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. Journal of Acoustical Society of America, 84 (2), 511-529.

Lessac, A. (1967). The use and training of the human voice: A practical approach to speech and voice dynamics. New York: Drama Book Publishers.

Peterson, K.L., Verdolini-Marston, K., Barkmeier, J.M., & Hoffman H.T. (in press). Comparison of aerodynamic and electroglottographic parameters in evaluating clinically relevant voicing patterns. Annals of Otology, Rhinology, and Otolaryngology.

Ramig, L. O. (in press). Speech therapy for patients with Parkinson's disease. In W. Koller & G. Paulson (Eds.), Therapy of Parkinson's disease. New York: Marcel Dekker, Inc.

Raphael, B.N., & Scherer, R.C. (1987). Voice modifications of stage actors: Acoustic analyses. Journal of Voice, 1, 83-87.

Scherer, R.C., & Titze, I.R. (1987). The abduction quotient related to vocal quality. Journal of Voice, 1, 246-251.

Titze, I.R. (1988). The physics of small-amplitude oscillation of the vocal folds. Journal of Acoustical Society of America, 83, 1536-1952.

Titze, I.R. (1994). Principles of voice production. Englewood Cliffs, N.J.: Prentice Hall.

Titze, I.R., & Sundberg, J. (1992). Vocal intensity in speakers and singers. Journal of Acoustical Society, 91, 2936-2946.

van den Berg, J.W., Zantema, J.T., & Doornenbal, P. (1957). On the air resistance and the Bernoulli effect of the human larynx. Journal of the Acoustical Society of America, 29, 626-631.

Verdolini, K. (1994). Voice disorders. In D. Spriestersbach, H. Morris, & J.B. Tomblin (Eds.), Diagnostic methods in speech-language pathology (pp. 247-306). San Diego: Singular Publishing Group.

Verdolini, K., Druker, M.S., Palmer, P.M., & Samawi, H. (in review). Physiological study of "resonant voice." Journal of Speech and Hearing Research.

Verdolini, K., Hoffman, H.T., & McCoy, S. (in press). Nonspecific laryngeal granulomas: A case study of a professional singer. Journal of Voice.

Verdolini, K., Titze, I.R., & Fennell, A. (in press). Dependence of phonatory effort on hydration level. Journal of Speech and Hearing Research.

Verdolini-Marston, K., Titze, I.R., and Druker, D.G. (1990). Changes in phonation threshold pressure with induced conditions of hydration. Journal of Voice, 4 (No.2), 142-151.

Verdolini-Marston, K., Burke, M.K., Lessac, A., Glaze, L., & Caldwell, E. (in press). A preliminary study on two methods otf treatment for laryngeal nodules. Journal of Voice.

Verdolini-Marston, K., Sandage, M. & I.R. Titze (1994). Effect of hydration treatments on laryngeal nodules and polyps and related voice measures. Journal of Voice, 8 (No.1) 30-47.

NCVS Status and Progress Report - 7
December 1994, 207-227

Neurological Disorders and the Voice

Marshall Smith, M.D.
Department of Otolaryngology-Head and Neck Surgery, The University of Colorado Health Sciences Center
Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts
Lorraine Olson Ramig, Ph.D.
Department of Communication Disorders and Speech Science, The University of Colorado-Boulder
Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts

Introduction

The laryngeal mechanism is subject to highly complex, extensive neural control, therefore, it is not surprising that disorders of the nervous system have effects on the voice. The phonatory function of the larynx, as well as its roles in respiration and swallowing, can be affected by a wide range of neurological disorders. These span the entire spectrum of central to peripheral etiologies and include trauma, cerebral vascular accidents, tumors and diseases of the nervous system. Except for the more common problem of laryngeal nerve paralysis and the unusual, but disabling, entity of spasmodic dysphonia, neurologically based voice disorders until recently have been a neglected topic for both basic and clinical research in the fields of otolaryngology, speech pathology and neurology. It is now more apparent that laryngeal dysfunction is a component of many neurological disorders and should be a critical consideration in patient assessment and management (1).

While there have been advances in knowledge of neurological control of vocalization (2-5), it is readily apparent that there is little information on the laryngeal pathophysiology of many neurological disorders. The integration between the basic science and clinical realm in neurolaryngology is far from complete. Advances in clinical and research laryngology have begun to bridge this gap. Laryngeal imaging has greatly assisted our ability to visually examine and record laryngeal function (6-8). In this way, the laryngeal examination has become more accessible as part of a neurological assessment. Other voice examination techniques include acoustic analysis, glottography, and aerodynamic measures (9,10). These also have emerging roles in documentation, diagnosis, and assessment of treatment efficacy for neurologic voice disorders (11,12).

The incidence of neurological disorders of the voice ranges from the common to the obscure, depending on the disorder. The "immobile vocal fold" is a frequently encountered clinical problem in otolaryngology. Though speech problems are often associated with the common neurological conditions of stroke and head injury, the association of voice disorders with these conditions has not been systematically described. Eighty-nine percent of the 1.5 million patients with Parkinson disease in the United States have voice disorders.(13) Essential tremor is the most common movement disorder, it affects the voice in four to twenty percent of cases.(14) Less often encountered are the voice disorders associated with neurological diseases such as Huntington's Disease and cerebellar ataxia. Koufman reviewed a series of one hundred consecutive patients seen in a practice with voice complaints.(15) Sixteen had neurologic disorders of the voice; 8 with vocal fold paralysis and 8 with a variety of other neurological diseases including Parkinson's disease, vocal tremor, and myasthenia gravis.

One framework for studying neurological diseases and the voice has come from studies of speech disorders. Research by Darley et al. (16,17) identified different clinically distinguishable types of dysarthrias that could be correlated with pathology in specific areas of the central and peripheral nervous system. Because many of these neurologic disorders also involve dysphonias, a similar "site-of-lesion" approach was adopted as a tool for classification of neurologic voice disorders. (7,18,19) A general outline of this classic approach is given in Table 1.

Aronson has categorized neurologic voice disorders according to the nature of the voice production, whether constant or variable.(20) Five types were described: 1) Relatively constant neurologic voice disorders, including

Table 1.
Traditional Classification of Neurological Voice Disorders Based on Etiology

Classification	Examples
I. Flaccid paresis/paralysis (lower motor neuron)	
A. muscle	myopathies, muscular dystrophies
B. neuromuscular junction	myasthenia gravis
C. peripheral nerve	
1. recurrent laryngeal nerve	trauma, tumor, idiopathic, iatrogenic - surgery, drugs, radiation
2. superior laryngeal nerve	collagen-vascular, Guillan-Baree'
3. upper vagus nerve	
D. brainstem nucleus	stroke (Wallenberg, syndrome), arnold-chiari malformation, sytingobulbia, tumor, trauma
II. Spastic paresis (Upper motor neuron)	pseudobulbar palsy
III. Dyskinetic movement disorders (extrapyramidal system)	Parkinson disease essential tumor, dyskinesias, myoclonus, choreas, laryngeal dystonias
IV. Ataxias (cerebellar)	degeneration, hemorrhage, infarction
V. Apraxias (cortical and subcortical)	trauma, stroke, tumor, cerebal palsy
VI. Mixed disorders	amyotrophic lateral sclerosis (ALS), multiple sclerosis, Shy-Drager syndrome

Table 2.
Classification of Neurologic Voice Disoders Based on Phonatory Dysfunction

Classification	Examples
I. Adduction/Abduction Problems	
A. Hypoadduction	all lower motor neuron laryngeal paresis/paralysis Parkinson disease Parkinsonism Shy-Drager syndrome progressive supranuclear palsy Traumatic brain injury
B. Hyperadduction	pseudobulbar palsy Huntington's disease adductor laryngeal dystonia (spasmodic dysphonia)
C. Malabduction	abductor laryngeal dystonia (spasmodic dysphonia)
II. Phonatory Instability	
A. Short-term (cycle-to-cycle perturbation, aperiodicity, subharmonics)	nearly all neurologically-based voice disorders
B. Long-term (tremor)	essential voice tremor Parkinson disease dystonic tremor ALS vocal "Flutter" palatopharyngolaryngeal myoclonus other respiratory/vocal tract sources
III. Mixed Disorders (adduction, abduction, instability)	amyotrophic lateral sclerosis (ALS) multiple sclerosis ataxic (cerebellar) dysphonia mixed (abductor/adductor/tremor) spasmodic dysphonia progressive supranuclear palsy Shy-Drager syndrome
IV. Miscellaneous Disorders	Apraxic dysphonia of cortical dysfunction Involuntary phonation of Tourette syndrome

flaccid, spastic (pseudobulbar), mixed flaccid-spastic, and hypokinetic (parkinsonian) dysphonias; 2) Arrhythmically fluctuating neurologic voice disorders, including ataxic, choreic, and dystonic dysphonias; 3) Rhythmically fluctuating dysphonias, including dysphonias of essential voice tremor and palatopharyngolaryngeal myoclonus; 4) Paroxysmal neurologic voice disorder, the sudden aberrant phonatory bursts of Gilles de la Tourette's syndrome; 5) Neurologic voice disorders associated with loss of volitional phonation, such as apraxia of speech and phonation associated with cerebrovascular accidents. A strength of this classification scheme is its focus on the resulting voice production. But it does not give direction to the clinician regarding phonatory function to guide treatment interventions.

Recently, Ramig and Scherer(21) proposed a system for considering neurological voice disorders with specific application to treatment. Rather than relating the neurological disorder to the site of neural damage, this classification system focuses on the existing phonatory dysfunction and resulting voice characteristics. The following categories of neural-based phonatory dysfunction were proposed: adduction problems (hypoadduction and hyperadduction), stability problems (long-term, e.g., tremor and short-term, e.g., hoarseness) and coordination problems (phonatory incoordination, e.g., dysprosody). These categories were used to organize approaches to treatment of neurological voice disorders which focused on modification of laryngeal physical pathology with corresponding changes in perceptual characteristics of voice and improved functional speech production. This classification system, in a modified form (Table 2) will be used to organize the description of neurological voice disorders in this chapter on the dimensions of adduction/abduction and stability. It should be recognized that phonatory function in neurologic disorders can vary and show features of hypoadduction, hyperadduction, malabduction, instability, and/or incoordination at the same time; disorders such as these are included in this system as mixed disorders. While variability within patients as well as

during the course of a disease is common in neurological disorders, this classification system provides an outline for observing and interpreting the effects of neurologic disorders on voice that can then guide treatment interventions, be they behavioral, medical, and/or surgical. Treatment of neurologic voice disorders will be mentioned briefly in this chapter but covered extensively elsewhere in this book.

Disorders of Adduction

Disorders of vocal fold adduction accompanying neurological disorders can range from inadequate adduction (hypoadduction) to excessive adduction (hyperadduction) to inappropriately timed abduction (malabduction).

Hypoadduction

Hypoadduction is observed in many forms and degrees in many neurological disorders. Damage to any component of the motor unit (muscle, neuromuscular junction, nerve or nucleus) may result in vocal fold hypoadduction. The most common example of this is vocal fold paralysis. It is also recognized that hypoadduction has been found in neurologic diseases of central origin, including Parkinson disease (22,23) and closed head injury (24). A primary consequence of hypoadduction is reduced vocal loudness and breathiness.

Hypoadduction Associated With Disorders of the Motor Unit

Muscle: Inflammatory myopathies, including polymyositis and dermatomyositis, are unusual conditions that may affect laryngeal function, particularly swallowing.(25) Muscular dystrophies are inherited diseases that yield progressive weakness that is variable in age of onset, distribution and disability.(25,26) The main laryngeal symptomatology concerns swallowing, yielding oropharyngolaryngeal weakness. Ramig (27) reported increased acoustic aperiodicity and weak, hoarse and nasal voices in patients with myotonic muscular dystrophy.

Myoneural Junction: Interruption of transmission of nerve impulses at the neuromuscular junction may lead to flaccid paresis or paralysis of the vocal folds. Myasthenia gravis (MG) is a neuromuscular disease manifested by weakness and fatigability of voluntary muscles. The incidence of MG has been reported between 2-10 per 100,000.(28,29) Overall, it is seen twice as frequently in women as men, with an even higher ratio in the third and fourth decades. Though neonatal and juvenile forms exist, voice and speech symptoms as initial or secondary symptoms occur primarily in the adult form. An association has been made with thymomas.

MG frequently presents with otolaryngological symptoms. The most common presenting symptoms are ocular (ptosis and diplopia). It is important to recognize that other head and neck symptoms may be present as an initial manifestation of MG. Carpenter et al reviewed 175 patients with MG.(30) 30% had primary symptoms of dysphagia, dysarthria, or dysphonia. Dhillon and Brookes reviewed 48 patients with MG.(31) Nine of their patients had only head and neck symptoms initially, six of these complained of dysphonia, usually in conjunction with dysphagia. Calcaterra et al had an even higher percentage of patients in their series present with voice complaints.(32) Of fifty patients with MG, 26 had initial speech complaint; 13 with dysarthria, 5 with hypernasality, 4 with stridor, and 4 with vocal weakness. These and other reports(33) make the neurolaryngologist and speech pathologist aware that voice and laryngeal involvement may occur as an initial or secondary symptom in MG.

Examination of the larynx in patients suspected of MG should involve observation of the velopharynx for inadequate closure, as well as glottal insufficiency - incomplete adduction or bowing of the vocal folds. The patient should be "stressed" with repetitive vocalizing tasks (reading or counting out loud for five minutes) to induce fatigability of involved structures. Fatigability may also be assessed by testing for stapedial reflex decay.(34) The diagnosis is made by a Tensilon® (edrophonium) test. Edrophonium, a short acting anticholinesterase drug, administered intravenously results in prompt reversal of symptoms. Improvement in voice and resonance characteristics on spectrograms following edrophonium injection has been documented.(35)

Because of the concern for progression to respiratory involvement that may occur in MG, prompt neurologic referral is warranted if MG is suspected.(36) Treatment of MG involves anticholinesterase drugs, eg. pyridostigmine. Steroids are used for severe cases. The effect of thymectomy in treatment of MG is controversial, but is generally recommended in younger patients.(37)

Peripheral Nerve (Recurrent laryngeal nerve): Vocal fold paralysis is a common problem resulting in vocal fold hypoadduction, probably the most common neurological voice disorder seen by the otolaryngologist.(15) This section will generally concern unilateral vocal fold paralysis (UVP), since this commonly affects the voice and is much more frequently seen than bilateral vocal fold paralysis. Many surveys have reported on the etiologies of vocal fold paralysis. The results of nine surveys involving 1019 patients were compiled by Terris et al.(38) 36% of cases were from neoplasm with 55% of these due to lung cancer. 25% were post-surgical, mainly from thyroidectomy. Other causes included idiopathic (14%), inflammatory or medical (13%), and central neurologic (6%). Medical etiologies should be sought in obscure cases; such as vincristine (a chemotherapeutic drug) induced laryngeal neuropathy(39), or post-polio syndrome associated laryngeal paralysis.(40,41) The

distinction between "idiopathic" and "inflammatory" causes of laryngeal paralysis may be blurred; in many cases "idiopathic" UVP is felt to be an isolated neuropathy caused by virus.(42,43)

The patient symptoms in UVP vary widely. Complaints of hoarse voice and breathy voice are frequent, although it is important to recognize that the voice may be fairly normal.(44) Problems with swallowing, choking, coughing, and aspiration can be found and may be of more immediate concern in considering intervention. The voice in a patient with UVP is usually perceived as reduced in loudness, hoarse and breathy. Diplophonia may be present. Acoustic findings that often accompany UVP include increased aperiodicity, reduced pitch range and variation, reduced vocal intensity, and decreased harmonics to noise ratio.(44) Variations in glottographic patterns may be seen.(45,46) Glottal aerodynamics often demonstrate high mean airflow rates, and an offset of airflow from baseline ("DC leak"), reflecting glottal insufficiency.(44,47).

There are several factors that influence glottal dynamics and resulting voice in UVP. These include 1) individual variations in laryngeal framework anatomy(48) and peripheral innervation,(49) 2) the degree and extent of reinnervation,(50) 3) the degree of laryngeal muscle atrophy,(51,52) 4) the time variation of the second and third factors, and, very importantly, 5) individual vocalizing behaviors and laryngeal biomechanics that attempt to compensate for the glottal insufficiency. The particular voice symptoms are created by the complex interaction of the above factors that influence laryngeal biomechanics; particularly in UVP relating to the geometry of the glottic (anterior and posterior) aperture, the three-dimensional position and orientation of the arytenoids, the vertical position of the folds relative to each other, muscle stiffness asymmetries and resulting body-cover relationships, and the interaction of these with glottal airflow and pressures.(7,53)

Laryngoscopy (flexible and rigid telescope) is of critical importance in the evaluation of UVP, as in other voice disorders. Laryngostroboscopy is also very helpful for viewing the effects of UVP on vocal fold position and movement, and glottal closure, mucosal wave asymmetries, and vocal process vertical position.(54) Tomography has also been used in UVP to demonstrate vocal folds on different vertical levels.(55)

The relation between the location of peripheral laryngeal nerve lesion and the position of the affected vocal fold in UVP seen on laryngoscopy has been a source of controversy. A common notion, promulgated as the "Wagner-Grossman theory", related the position of the vocal fold in vocal fold paralyses to the activity of the cricothyroid muscle.(56) Clinical relevance has been ascribed to this theory by relating glottal configuration to the topognostic localization of site of laryngeal nerve injury.(43,57) *In-vivo* studies have yielded conflicting results.(56,57) Several recent clinical studies have questioned this dogma, pointing to the wide variability in glottal position at various locations of injury.(58,59,60) A factor complicating this issue is that, in many cases, the exact site and degree of nerve injury are unknown.(7) Neuroanatomical studies have shed light on this controversy. Sanders et al described multiple and variable ipsilateral and crossed interconnections in the peripheral innervation of the human larynx between the superior and recurrent laryngeal nerves.(49) These variations may help explain the spectrum of glottal configurations seen in UVP.

The management of dysphonia secondary to UVP includes behavioral therapy and a variety of surgical techniques. These are covered extensively in other chapters of this book.

Peripheral nerve (Superior laryngeal nerve): Superior laryngeal nerve (SLN) paralysis is less common than recurrent laryngeal paralysis.(61) Causes include surgical trauma (eg. thyroidectomy, carotid endarterectomy), blunt neck trauma, and idiopathic cases. It's presentation may be subtle. Voice complaints may be only of a mild hoarseness, vocal fatigue, or difficulty with projection or singing despite normal speaking voice.(62) The laryngeal examination findings in unilateral SLN paralysis have been described as a rotation of the anterior commissure away from and posterior glottis toward the affected (paralyzed) side.(7,61) This maneuver may only be evident by having the patient attempt to phonate at high pitch.(63) Shortening of the vocal fold, unilateral bowing, or scissoring of the vocal folds may be observed.(64) The asymmetry of laryngeal vibration in SLN paralysis may be seen on laryngostroboscopy.(65) This asymmetry has been demonstrated in mathematical simulations(53,66) and animal phonation models of SLN paralysis.(67)

Nuclear: Injury to the lower motor neurons of the tenth nerve nucleus results in flaccid laryngeal paralysis and hypoadduction.(18,19) The injury can occur via vascular brainstem infarction of the posterior inferior cerebellar artery. This vascular accident when associated with symptoms of dysphagia and dysarthria, ipsilateral Horner's syndrome, ipsilateral face and contralateral body pain-temperature impairment, is known as Wallenberg's (lateral pontomedullary) syndrome. Other causes include arnold-chiari malformation, syringobulbia, tumor, and trauma.

Hypoadduction Associated With Parkinson Disease

Hypoadduction is most frequently associated with the extrapyramidal disorder of Parkinson disease (PD), a nigrostriatal disorder characterized by dopamine deficiency (68). Rigidity, tremor and reduced range of motion are the primary physical pathologies associated with PD. One in every one hundred individuals over 60 and one in every one

thousand individuals under age 60 have PD. While PD is frequently considered a disease of the elderly, patients have been diagnosed as early as age 35. Patients can live 10-20 years after diagnosis. At least 89% of the 1.5 million patients with PD in the United States have a voice disorder.(13)

The voice disorder in PD is generally characterized by reduced loudness, monotone, hoarseness and tremor. Reduced volume may be the first sign of PD.(20) These voice characteristics are frequently observed in the context of a generalized dysarthria which includes imprecise articulation and short rushes of speech. Acoustic studies have reported increased fundamental frequency(69) yet perceptual ratings have noted a decreased pitch.(70) This discrepancy has not been resolved,(44) yet may be sex-related. In a study of 34 PD patients Hirose and Joshita(71) found that 11 of 24 males had a higher speaking pitch, yet none of the 10 females had higher pitch while eight showed a lower pitch levels than controls. While the subjective impression is that PD patients have reduced intensity, Canter reported sound pressure level data within normal limits.(69) Recently, Ramig et al (72) reported sound pressure levels for 45 patients with idiopathic PD. The average level was 68 dB for maximum duration sustained /a/ phonation, 65.5 dB for reading, and 64.5 dB for conversation. Male and female data were comparable for these tasks and all data were measured at mouth-microphone distance of 50 cm. On laryngeal examination, bowed vocal folds, tremor and supraglottic hyperfunction have been reported.(22,23) Evidence of laryngeal hypoadduction has been observed in patients with early PD.(23,73) Electromyographic studies of laryngeal muscles in PD are limited. One study (71) reported thyroarytenoid recordings of a single subject with PD. Loss of reciprocal suppression of the TA during inspiration was found, indicative of parkinsonian rigidity. No neurogenic changes were seen to indicate a peripheral neuropathy. Though recordings of reciprocal muscles (posterior cricoarytenoid) were not made, it was felt that the hypokinetic laryngeal movement in PD was related to deterioration in the reciprocal adjustment of the antagonistic muscles.

While dopaminergic (levadopa) treatment has positive effects on the general function of PD, voice and speech symptoms are not consistently alleviated by these pharmacological treatments.(22,74) More recently, neurosurgical interventions (e.g., adrenal and fetal cell transplants) have been conducted.(75) Findings following fetal cell transplant support improved limb function and reduction of medication, but changes in laryngeal and speech mechanism function are not clear.(76)

Speech therapy has historically been ineffective for PD.(77,78) However, a recent approach by Ramig and associates (79-82), the Lee Silverman Voice Treatment, which focuses on increased phonatory effort, has proven effective. Significant functional changes in vocal loudness with co-occurring changes in sound pressure level, subglottal air pressure and vocal fold adduction have been reported following intensive voice treatment (sixteen sessions in one month). Details of this treatment are summarized elsewhere.(82)

Hypophonia Accompanying Closed Head Injury

Traumatic brain injury yields complex neurologic insults. They usually affect the cortical and subcortical structures as well as others. These injuries may result in voice and speech problems. Voice problems in these patients are described as having decreased loudness. A recent study found decreases in laryngeal airway resistance and increased glottal airflow in traumatic brain injury patients with voice and speech problems.(24) Laryngeal biomechanics have not been studied via endoscopy in these patients, but information suggests that laryngeal hypoadduction and glottal insufficiency are likely to be found.

Hyperadduction

Hyperadduction is observed in many forms and degrees in various neurological disorders. It may be constant (such as in the strain-strangle phonation of pseudobular palsy), or variable (such as the alternating spastic voice in Huntington's disease and spasmodic dysphonia). Laryngeal involvement may be found in all varieties of hyperkinetic movement disorders (chorea, essential tremor, myoclonus, tics, and tardive dyskinesia).(83) It may also be observed as a compensation for hypoadduction. Hyperadduction may involve the true vocal folds and, in many cases, the false folds and supraglottal region as well.

Hyperadduction Accompanying Upper Motor Neuron Lesions

Pseudobulbar palsy comprises a clinical picture of dysarthria with accompanying dysphonia, dysphagia, and signs of increased jaw jerk and snout reflexes.(84) The primary voice symptom is harshness and hoarseness and strain-strangle.(44) Another clinically distinguishing feature is emotional lability, a reduced threshold for laughing or crying.(20) The term pseudobulbar palsy is felt to be inaccurate, since the sites of pathology are supranuclear to cranial nuclei, involving bilateral interruption of corticobulbar pathways, most frequently in the internal capsules.(85) The most frequent cause is stroke; other possible etiologies include encephalitis, trauma, and neoplasm.(85)

Voice problems in pseudobulbar palsy occur in the context of dysarthria of speech. The major findings have been summarized by Darley et al as prosodic excess, prosodic insufficiency, articulatory-resonatory incompetence, and phonatory stenosis (harsh, strain/struggle, pitch breaks).(16,17) In a perceptual evaluation of 30 patients, the voices of patients with pseudobulbar palsy were rated gener-

ally as monopitch, harsh, and low pitched. Many exhibited a strained voice quality and some had pitch breaks or tremor. The resonance quality may be hypernasal. No studies have reported on laryngoscopy or other measures of phonatory function in these patients. The variability in voice findings in pseudobulbar palsy likely reflects involvement of various structures in the supranuclear neural pathways controlling phonation. No documented studies on behavioral treatment of the dysphonia and dysarthria of pseudobulbar palsy are available.(20)

Hyperadduction Accompanying Quick Hyperkinesia of Huntington's Disease

Choreic movements accompany Huntington's disease, an autosomal dominant disease, associated with loss of neurons in the caudate nucleus. The incidence of Huntington's disease is 4 to 8/100,000, with the average age of onset 38 years.(83,86) Abnormal, involuntary, abrupt, jerky movements may affect any part of the body. When choreiform movements effect the laryngeal mechanism, voice characteristics include irregular pitch fluctuations and voice arrests. Forced inspirations and expirations, harsh voice quality, excess loudness variations, strain-strangle phonation are classically observed.(16,17) Acoustic analysis reported octave drops (low frequency segments) in sustained phonation of patients diagnosed with Huntington's disease as well as individuals "at risk" for HD;(27,87) voice arrests (both adductory and abductory) and reduced vowel durations were also observed. Clinical endoscopic assessment in one patient revealed rapid adductory and supraglottal jerking movements at rest and during speech production (Ramig LO, personal observation).

The medical treatment of Huntington's disease involves pharmacologic attempts to control the choreic movements with anti-dopaminergic agents, phenothiazines, benzodiazepines, or anti-seizure medications.(83) The effects of these on voice in HD has not been documented. We have treated an HD patient who demonstrated hyperadductory voice arrests. Laryngeal thyroarytenoid botulinum toxin injections were effective in improving voice quality and ease of phonation.(Smith ME, personal observation)

Hyperadduction accompanying dystonia (including "adductor spasmodic dysphonia")

Spasmodic dysphonia (SD) is a rare but challenging voice problem, generally characterized by effortful, strain/strangle voice quality with frequent voice breaks. It is not a disease, but instead a constellation of voice symptoms and signs.(88) The actual incidence of SD is unknown.(89) SD is now considered more of a descriptive term for several types of voice disorders, including adductor SD and abductor SD. In this disorder, a variety of types of laryngeal dysfunction may be seen, including hyperadduction, malabduction, and phonatory instability.(90-94) These may vary between patients or be found in combination in an individual. The individual's voice symptoms may vary depending on the type and degree of these factors. This section will describe the neural bases involving SD and features seen in its most common form, adductor SD.

In recent years, increasing evidence has accumulated to support the concept that, in a large number (exact figures unknown) of individuals that have the perceptual voice symptoms and findings of SD, the voice disorder is considered to be a manifestation of dystonia. Dystonia is a movement disorder characterized by sustained muscle contractions that may cause twisting or repetitive movements or abnormal postures. These may be sustained or intermittent.(83) Dystonias are classified as focal (isolated to a small group of muscles in one body part), segmental (contiguous muscle groups), or generalized (widespread muscle involvement). They are also grouped by etiology, primary or secondary. Primary dystonias may be inherited (usually autosomal dominant) or sporadic. Secondary dystonias may be 1) associated with other hereditary neurologic disorders (eg. Wilson's disease, Huntington's disease), 2) environmental (eg. drug-induced tardive dystonic reaction from phenothiazines), 3) associated with parkinsonism, or 4) psychogenic. Research to date indicates that "spasmodic dysphonia" may be considered in most cases as a focal primary, function-specific, action-induced laryngeal dystonia.(20,83,95-97)

Evidence for the association of spasmodic dysphonia and dystonia comes from several sources, including clinical observations, neurophysiologic studies, and absence of response to voice and behavioral therapy. Important clinical observations have made the association of spasmodic dysphonia with other neurologic diseases, including movement disorders such as Meige's disease.(92,98) An extensive survey was done of the large population of dystonia patients at Columbia-Presbyterian Medical Center.(83) It was found that of 2,556 cases of dystonia laryngeal involvement was seen in 562 (22%). The male/female ratio (41% / 59%), ratio of primary dystonia to secondary dystonia (83% / 17%), positive family history of dystonia (20%), and other characteristics matched closely with the general dystonia population.

Neurophysiologic studies of spasmodic dysphonia have been conducted in attempts to establish a neural basis for SD and to localize the neurologic source of abnormalities.(94) Various findings include abnormal electroencephalograms(99), altered or delayed latencies of evoked auditory brainstem responses(100), blink reflex response abnormalities(101,102), altered cerebral regional blood flow(103), and brain magnetic resonance imaging subcortical lesions(104). Evidence for disordered motor control based on disrupted interconnections between cortex,

subcortex, and basal ganglia have been postulated.(94,105) Laryngeal sensory and reflex functions may also play a role. It has been recognized that laryngeal reflex suppression is required for vocalization.(106) Laryngeal adductor reflex response to SLN stimulation has been reliably elicited from awake humans.(107) Abnormalities in inhibitory modulation of this reflex have been measured in patients with spasmodic dysphonia.(108) It has been hypothesized that abnormalities in laryngeal reflex suppression during vocalization may be related to the mechanism of phonatory disruption in laryngeal motor control disorders, including spasmodic dysphonia, vocal tremor, Parkinson's disease, and multiple systems atrophy.(5)

The perceptual voice signs and symptoms of adductor spasmodic dysphonia are characterized by a strain and struggle to talk, with accompanying voice breaks or stoppages.(20) Patients may present with a variety of voice qualities, including hoarseness, harshness, or tremor. Voice and pitch breaks may be present. A creaking, fry-type phonation pattern may be seen. Voice symptoms may vary in the individual, based on activity, stress level, and other factors. There is often seen "task specificity" in this voice disorder, as in other dystonias.(96,97,109) Emotional vocalizations such as laughing, crying, or singing are not associated with the symptoms seen in voluntary speech tasks. The difficulty in verbal expression experienced by these patients creates a great deal of stress, anxiety and depression.(110,111)

Objective measures of phonation in SD were reviewed in a recent article by Woodson et al.(112) Flexible fiberoptic laryngoscopy was found to be useful to identify the nature of vocal obstruction in SD.(112-114) Acoustic measures have not been diagnostic by themselves; however, when taken together provide a profile of vocal dysfunction shown by abnormal signal-to-noise ratio, perturbation measures, standard deviation of fundamental frequency, and voice break factor (number of voice breaks divided by maximum phonation time).(112) Aerodynamic measures have generally been characterized by low to normal ranges of mean glottal flow. (112,116,117) They were found to be helpful in documenting the effect of botulinum toxin laryngeal injections.(116,117)

The treatment of adductor SD has been a subject of multiple approaches(118), including behavioral(119), surgical(120,121), and, recently, laryngeal injection therapy(122,123) modalities. The general goal of all therapies is to reduce laryngeal hyperadduction. A comprehensive discussion of the treatment of adductor laryngeal dystonia and spasmodic dysphonia is found elsewhere in this book.

Malabduction

Malabduction or "wrong abduction" refers to the inappropriate or mistimed abduction of the vocal folds during phonation. The paired abductor muscles of the intrinsic laryngeal musculature are the posterior cricoarytenoid (PCA) muscles. While the PCA has primarily been considered for its function in respiration, it has also been shown to have an important role in speech and voice. Many speech sounds (eg. stop consonants, fricatives) require the absence of vocal fold vibration. Ludlow et al (124) reported that PCA is active during two specific phonatory tasks during the onset and offset of speech; abduction in the transition from closed glottis to phonation, and abduction in the transition from phonation to open glottis. Malabduction occurs when the vocal folds open inappropriately and allow excess escape of air through the glottis during speech.

Abductor Spasmodic Dysphonia

Malabduction has generally been associated with the disorder of abductor spasmodic dysphonia. Abductor spasmodic dysphonia is characterized by interruptions of breathy or whispered (unphonated) segments upon a normal or hoarse voice.(20) This type of voice disorder does not necessarily occur in isolation but its characteristic breathy voice breaks may also be a feature associated with other neurologic-based voice disorders such as adductor spasmodic dysphonia, severe vocal tremor, or myoclonus.(7,90) It may be seen as a compensatory behavior to alleviate adductory laryngeal spasms in adductor spasmodic dysphonia.(7,83,112,115) The incidence of abductor spasmodic dysphonia is much less than the adductor variety. In a series of 100 consecutive SD patients, 4 had "pure" abductor SD and 11 others had malabduction in combination with adductor SD or tremor.(125)

Several studies have reported on objective voice measures in patients with abductor SD. Fiberoptic laryngoscopy demonstrated synchronous and untimely abduction of both true vocal folds, exposing a wide glottic chink,(20,116) and bowing of the vocal folds.(126) Acoustic recordings documented sudden drops in fundamental frequency.(127) Aerodynamic measures found much higher than normal mean airflow rates.(127) Spectrograms demonstrated prolongation of the interval from initial stop consonant (voiced or unvoiced) to vowel, and less energy in higher formants,(128,129) or more high frequency spectral energy during voice breaks.(130) Surprisingly, electromyographic studies of both adductor and abductor SD have failed to demonstrate differentiation of subjects by symptom type, instead showing normal levels and patterns of intrinsic laryngeal muscle activation(131), with bursts of activity superimposed on these.(132) Another unsuspected finding has been the measurement of abnormal bursts of cricothyroid muscle activity in some abductor SD subjects.(133)

Recent advances in the medical treatment of abductor SD have involved botulinum toxin injection therapy.(133-135) These are discussed elsewhere in this book.

Phonatory Instability

Vocal fold vibration is generally not perfectly periodic. A number of factors combine to introduce perturbations (small changes in system behavior) and/or fluctuations (more severe deviations from the norm) into the phonatory mechanism.(136,137) This yields aperiodicity, or minor deviations from periodicity sometimes termed quasi-periodic.(21) Sources of aperiodicity include biomechanical asymmetries(66), variations in neuromuscular control and motor unit firing(138,139), airflow instabilities(140), mucus on the vocal folds(137), glottal source-vocal tract interactions(141), pulsatile blood flow(142,143), irregularities in the respiratory system(134), and chaotic motion resulting from non-linearities in the laryngeal mechanism(144). These factors, all present in the normal phonatory mechanism, may be exaggerated or amplified in the patient with a neurologic voice disorder. This creates phonatory instability. For purposes of classifying the phonatory effects of instability in neurologic voice disorders it is convenient to separate them into two categories; short-term and long-term.(21,145)

Short-term Phonatory Instability

Short-term instabilities are usually described as the cycle-to-cycle variations in frequency (jitter) and amplitude (shimmer). The effect of short-term instability on speech intelligibility may be associated with the perception of "hoarse-rough" voice quality.(146) Measurements of these variables are affected by physiological condition(147), sex(148), and vowel(149,150). Short-term phonatory instabilities have been identified in many neurologic voice disorders(151). Unfortunately, studies have yielded mixed results in distinguishing normal from abnormal voice-disordered populations or differentiating between neurologic voice disorders of various causes.(152,153) Stability measures may have utility in tracking the disease progression of neurologic disorders(151) or documenting efficacy of voice treatment.

Another class of short-term instabilities have also been described.(144) Waveforms seen in both normal and pathologic voice populations (including neurologic voice disorders such as vocal fold paralysis and spasmodic dysphonia) have been found to contain subharmonics, period doubling-tripling-quadrupling etc, or alternating aperiodic/periodic segments(144,154,155), consistent with phenomena seen in low dimensional chaotic dynamical systems(156). The theory of non-linear dynamics and chaos is being applied to investigate disordered voices in these terms.(144,157) Since many neurologically impaired voices are not amenable to analysis by traditional perturbation measures(145), non-linear dynamics may have increasing application in the analysis of neurologic voice disorders.

Long-term Phonatory Instability Accompanying Vocal Tremor

Long-term phonatory instability affects vocal fold vibration over a longer time interval than individual glottal cycles. These effects are seen to be superimposed on vocal fold vibration throughout the duration of phonatory effort, often lasting seconds. Long-term instabilities are created by modulation of the primary signal.(136) The periodic glottal airflow signal is superimposed on by a much slower modulation in frequency (FM) and/or amplitude (AM). This long-term phonatory instability may be normal (physiologic) or abnormal. Vibrato is an excellent example of physiologic vocal modulation(136,158,159). In neurological diseases, long-term instability is seen associated with vocal tremor.

Tremor is defined as "any involuntary, approximately rhythmic, and roughly sinusoidal movement".(14) These fluctuations in movement may be observed in various parts of the body. When these movements affect any or several portions of the speech system (respiratory, phonatory, articulatory) they may create vocal tremor. Tremors have been classified by etiology, clinical appearance, and described by frequency, amplitude, body distribution, and exacerbating and relieving factors (160). Though all individuals have physiologic tremor(14) that occurs with a frequency of 8 - 12 Hz, abnormal (pathologic) tremor has been reported on the range of 3 to 7 Hz. The tremor frequency is influenced by many factors, including body location, disease type. Tremors are divided into resting tremor and action tremor. Action tremors may be postural (holding a position against gravity), contraction (produced by isometric voluntary contraction), or kinetic/intention (during activity or goal directed movement).(160)

The voice characteristics in vocal tremor have generally been studied from acoustic recordings of sustained vowel phonation. Observation of oscillographic recordings gave information on the amplitude changes of the output.(161-164) With more sophisticated analysis techniques the acoustic signal can demodulated in amplitude and frequency components.(158,165) This data suggests that in patients with vocal tremor complex oscillations involving both frequency and amplitude, and variations in these exist(158,165-167). These patterns may differ in tremors of various neurological diseases.(88,168) Vocal tremor is a feature of several neurological diseases, including essential tremor, Parkinson disease, ataxic dysphonia of cerebellar lesions, amyotrophic lateral sclerosis, and others.

Essential Tremor

Essential tremor is the most common movement disorder, with a reported prevalence of 4 to 60 per 1000 people.(14) It is found in all age groups, but generally increases with advancing age. The voice is affected in 4 to

20% of cases(14), but more frequently the hands or head are involved. Voice problems may be the first symptom for which medical help is sought, as was the case in 13 of 31 patients with essential vocal tremor reported by Brown and Simonson.(161) Tremor symptoms at various body sites frequently co-exist. A hereditary-familial history of tremor is found in about 50% of cases, and reflects autosomal dominant transmission.(169) The cause of essential tremor is unknown,(14) however recently neurophysiological research using positron emission tomography suggests that olivocerebellar tracts are abnormal in this condition(170,171).

Vocal tremor has been described perceptually as quavering or tremulous speech(20,44,161) The symptoms are most noticeable on vowel prolongation. Vocal tremor of essential tremor may be confused with spasmodic dysphonia.(20,172) This is due to pitch breaks and voice arrests from large amplitude tremor of laryngeal structures that interrupt airflow(20) or allow excess air escape(173). On laryngeal examination the larynx is usually reported to move both at rest (quiet breathing) with increased movement during phonation.(20) Vertical oscillations in the neck may be visible. Electromyographic study of eight patients with essential vocal tremor found predominant involvement of the thyroarytenoid muscles as well as other extrinsic laryngeal muscles.(174)

The treatment of essential vocal tremor has involved pharmacotherapy (eg. propranolol, primidone, acetazolamide, alprazolam, phenobarbital) with mixed results.(83) Recent reports of treatment with laryngeal botulinum toxin injections have been promising.(83,175)

Vocal Tremor in Parkinson Disease

In addition to the other voice abnormalities seen in Parkinson disease involving hypoadduction, Logemann et al(13) reported that 13.5% of a group of patients with Parkinson disease had tremulousness of speech. In a similar recent study, perceptual ratings of 45 patients with idiopathic PD found that 20% were rated to have tremor on passage reading (Ramig LO, unpublished data). Thirty-five percent (7/20) of IPD patients were observed to have tremor at rest and 71% during sustained /i/ phonation based upon videolaryngoscopic views of the glottal and supraglottal region.(176) The frequency of vocal tremor in Parkinson disease has been reported in the range of 5 to 7 Hz, including both frequency and amplitude modulations.(27)

Vocal Tremor in Spasmodic Dysphonia

Several studies have documented the presence of vocal tremor in patients with spasmodic dysphonia. This vocal tremor is differentiated from essential tremor in that it is present during phonation but not at rest, other body structures are not generally involved, and there is no family history of tremor.(96,125) Blitzer et al reported that 25% of patients with SD were found to have an irregular tremor.(91) Koufman and Blalock observed tremor in 45% of 100 consecutive SD patients of all varieties (adductor, abductor, mixed, etc).(125) They described "dystonic tremor" as task-specific with crescendo onset just prior to and during phonation, then decrescendo with offset of voicing.

Vocal Tremor of Palatopharyngolaryngeal Myoclonus

An unusual form of myoclonus affects the soft palate, termed palatal myoclonus. It is distinguished from the irregular jerks of limb myoclonus by its periodicity, but felt not to be a tremor because of nonsinusoidal kinematics.(14) Aronson feels this is a slow form of tremor (usually 1 to 4 Hz) and technically may be a misnomer.(20) When the symptoms extend to involve the pharynx, larynx, diaphragm, and possibly the eye muscles it is termed palatopharyngolaryngeal myoclonus.(20) It affects speech similarly to essential vocal tremor, causing a quavering voice. When the amplitude of movement is large, voice arrests may appear, resembling spasmodic dysphonia.(20,177) The patient also complains of a clicking or popping sound in the ear, from movement of the palate and pharyngeal muscles that open and close the eustachian tube. Laryngeal dysfunction affecting airway(178) or swallowing(179) may occur. On examination, the rhythmic movement of the palate, pharynx, and larynx (usually bilateral but sometimes unilateral) is seen both at rest and during phonation. The neurologic lesions associated with this disorder interrupt the central tegmental tract, the olivo-dentate pathways or contralateral dentate nucleus, and connections via the red nucleus, leading to secondary vacuolar degeneration and hypertrophy of the inferior olive.(180) They are usually due to brainstem infarction or idiopathic degeneration. Though usually resistant to treatment(180), symptoms of palatal myoclonus have responded, with varying degrees of success, to treatment with serotonin precursors, carbamazepine, clonazepam, and various surgical approaches.(83)

Mixed Neurological Voice Disorders

Various combinations of ataxic (cerebellar), upper (spastic), and lower (flaccid) motor neuron findings may be found in the "mixed" dysphonias seen in several neurologic diseases. These include cerebellar lesions, amyotrophic lateral sclerosis (ALS), multiple sclerosis (MS), and Parkinson plus syndromes. These disorders are characterized by a mixture of laryngeal physical pathologies including both adduction and stability problems.

Ataxic Dysphonia: Cerebellar Lesions

The function of the cerebellum involves coordination and regulation of skilled movements. The effects of speech and voice from cerebellar lesions are felt to be a

breakdown in integration and alterations in motor programming (ataxia).(181) Causes of cerebellar lesions include cerebellar degeneration, infarcts, hemorrhage, or neoplasm. The characteristics of speech and voice in ataxic dysarthria surveyed by Aronson et al (N = 30) included harsh voice (21), monopitch (20), monoloudness (18), poor pitch control or loudness control (bursts of loudness) (24), and vocal tremor (5).(182) The tremor observed in cerebellar disorders is irregular and slow, approximately 3 to 8 Hz.(14) Kent et al (181) reported on the acoustic characteristics of speech in five cerebellar patients and found abnormally long vowel and segment durations; they suggested problems with control of fundamental frequency. In another study, long-term phonatory instability (tremor) reflected in pitch fluctuations was found via acoustic analysis of sustained vowels in 6 of 11 subjects with cerebellar atrophy, and 5 of 9 subjects with olivopontocerebellar atrophy.(183) Though the pathophysiology of this disorder suggests reduced laryngeal coordination in abduction/adduction movements,(44) there have been no studies of laryngoscopy or laryngostroboscopy in verification of this.

Amyotrophic Lateral Sclerosis

ALS is a progressive degenerative neurologic disease that affects both upper and lower motor neurons. Its cause is unknown, but findings suggest several etiologies, including autoimmune disorder (antibodies to GMI gangliosides), multifocal conduction block gammopathy, malignancy (eg. lymphoma), and familial/genetic forms(20,44). The incidence is 0.4 to 1.8 per 100,000. The male to female ratio is about 1.5:1.0. The average of diagnosis is 58 years, the disease is uncommon under 40 years of age. The median survival is about 17 months, and 20% of patients with ALS are alive at 5 years. The initial manifestations usually involve muscle weakness, cramps, and fasciculations.(184) One series (N = 123) found that 28% of patients with ALS presented with symptoms in the head, neck, larynx or voice. 68% exhibited slurred speech, 14% hoarseness, and 13% dysphagia as a presenting symptom.(184)

Aronson described the dysphonia of mixed flaccid-spastic dysarthria associated with ALS as a harsh, strain-strangle sound with degrees of breathiness, reduced loudness, audible inhalation, and "wet hoarseness"(20). The wet, gurgling sound is due to pooling of secretions in the pyriform sinuses and around the glottis. Carrow et al conducted a subjective perceptual voice assessment of 87 patients with motor neuron disease, including ALS.(185) 80% had harsh voice quality, 75% hypernasality, 65% breathy voice quality, 63% voice tremor, 60% strain-strangle voice, 38% high pitch, 8 percent low pitch.

The acoustic characteristics of ALS speech have been the object of recent study. In two studies, Ramig et al documented the progressive neurological degeneration of voice in ALS with increasing acoustic perturbation measures and decreasing harmonics-to-noise ratio.(11,27) Another study by Strand et al found wide individual variation in acoustic measures of voice in four women with ALS, pointing to the variability in site of and degree of neural involvement in these patients.(186) Aronson identified a distinctive rapid tremor or "flutter" in the voice of ALS patients.(20) Acoustic analysis of this tremor identified multiple combinations of levels of amplitude and frequency modulations when compared to control subjects.(168) The source of these productions remains obscure.

Multiple Sclerosis

Multiple sclerosis (MS) is a disease of demyelinization of the white matter of the brain, brainstem, and spinal cord. It can disrupt neural function at any level. Symptoms vary in severity and in approximately two-thirds of the cases they wax and wane (exacerbation and remission), and in the other third, symptoms are progressive. MS can often have otolaryngological manifestations.(187) Impairment in speech ("scanning speech"; where every syllable is produced slowly and hesitantly with a pause after each syllable) has been considered a hallmark symptom of MS (part of "Charcot's triad"). Charcot (188) also reported a paroxysmal dysarthria accompanying MS that varied from severely unintelligible to temporarily intelligible speech production. The reports of the incidence of dysarthria and dysphonia accompanying multiple sclerosis are varied. Jensen (189) reported that the speech of 78% of a group of multiple sclerosis patients was rated defective. Beukelman et al(190) found that only 23% of 656 MS patients surveyed reported speech or other communication problems. In the series reported by Darley, et al(191) most MS patients had normal speech (59%) or minimal speech impairment (29%). When the voice was affected, the main perceptual findings were impaired loudness control and harshness, described as a "spastic-ataxic dysarthria".(191) Less frequently observed were impaired pitch control, inappropriate pitch level, breathiness, and hypernasality. This variety of vocal symptoms suggests that MS patients with voice abnormalities may be found to have problems in both laryngeal adduction (hyperadduction or hypoadduction) and phonatory stability. Using spectrographic and motion picture sound track analyses, Zemlin (192) reported extreme variability in sustained phonation produced by nineteen of thirty three MS patients he studied. Morrison and Rammage (193) reported the case of a patient who presented with 'spasmodic dysphonia' in which multiple sclerosis developed two years later. The voice eventually improved, though other neurologic symptoms progressed. Neurologic factors were felt to interact with psychological factors to cause the dysphonia. Speech therapy is felt to benefit patients with MS.(194)

Parkinson Plus Syndromes:(Shy-Drager Syndrome, Progressive Supranuclear Palsy)

There are a number of forms of Parkinsonism. The most common is idiopathic Parkinson disease (80%). The second largest group (12.2%) comprises Parkinsonism plus syndromes.(195) Patients in this group have symptoms of Parkinson disease and additional neurological disorders. Two of these include Shy-Drager syndrome and Progressive Supranuclear palsy (PSP). Shy-Drager is a rare, progressive disorder of the autonomic nervous system.(196) Its syndrome is described by orthostatic hypotension, urinary and rectal incontinence, impotence, loss of sweating, and other autonomic failure. Since the original report several studies have described respiratory, swallowing, and voice problems in conjunction with bilateral vocal fold paresis/paralysis in these patients.(196-198) The voice is described as monopitch, monoloud, with both strain/strangle and breathy qualities,(44,197) with functional evidence of laryngeal hypo- and/or hyperfunction as well as short and/or long-term phonatory instability.(199) EMG of laryngeal muscles in Shy-Drager syndrome have found denervation of posterior cricoarytenoid and interarytenoid muscles in these patients.(200)

Progressive supranuclear palsy (PSP) is a brainstem neurologic disease of progressive Parkinsonism and ocular motility disturbance.(201) Progressive gait and balance impairment is an early and disabling symptom. Speech difficulties in PSP are reported as a spastic, hypernasal, monotonous, low-pitched dysarthria.(202,203) Imprecise articulation may be the most prominent disordered speech feature of PSP.(203) Laryngoscopic assessment of one PSP patient revealed bowing of the vocal folds and false fold hyperadduction. Acoustic measures showed decreased vocal intensity as a primary factor reducing speech intelligibility. Application of the Lee Silverman Voice Treatment to this patient, and two additional selected patients with Parkinsonism plus syndromes, demonstrated significant improvement pre- to post-treatment as well as maintenance of these changes up to six months.(199)

Miscellaneous Neurolaryngeal Disorders

Apraxic Dysphonia

Lesions of the cortex and subcortical regions, such as those caused by stroke, head injury, or tumor, can produce apraxia of phonation, in conjunction with apraxia of respiration and articulation.(20) Phonatory characteristics can vary from mutism to trial-and-error nonspeech to whispered speech, with or without airflow.

Involuntary Vocalization

Verbal and non-verbal vocalizations may be encountered in a variety of neurological disorders. The most well-known of these is Gilles de la Tourette syndrome (TS). TS is a rare, chronic disorder of involuntary motor tics that begins in childhood and persists in adulthood.(204) It is seen predominantly in men and may have an autosomal dominant mode of inheritance. The involuntary vocalizations in TS range from unintelligible non-verbal noises to verbal tics including coprolalia (repetition of obscene words), echolalia (repetition of the last syllable, word, or sentence spoken by others), or palilalia (repeating by the patient of the last word or sentence in a phrase). The cause of TS is unknown. The symptoms of TS respond to haloperidol, as well as dopamine antagonist drugs. This suggests a possible role of dopaminergic hyperactivity in TS.(204)

Neurolaryngological and Voice Evaluation

It has been suggested that the laryngeal system can be regarded as a microcosm of the entire speech mechanism(2) and therefore the larynx may reflect neurological impairments before other speech components. Neurologic voice disorders can be isolated problems or occur in the context of extensive systemic disease. Because voice disorders have been reported as initial symptoms of various neurological disorders, such as Parkinson disease(13) and myasthenia gravis(30-33), phonatory evaluation should be considered for its contribution to neurological assessment(21).

Voice characteristics may contribute to neurological differential diagnosis and various intervention procedures may significantly improve quality of life, even in a patient otherwise suffering from a terminal illness. Therefore, the laryngologist, speech pathologist and neurologist should recognize the disorders for which voice and speech problems and complaints may be the initial presentation, and the disorders for which medical, surgical, and/or behavioral intervention may improve laryngeal function, in the context of the individual's general physical condition.

The neurolaryngological examination involves several components. These include medical history, phonation assessment, general otolaryngology/head and neck examination, laryngeal imaging examination, directed neurological examination, and, when appropriate, special diagnostic testing.

A careful history of both medical and neurologic symptoms provides the basis for evaluation of the patient with a neurological voice disorder. The voice history includes the nature of the vocal complaint. In general, voice complaint symptoms may be described as hoarseness, vocal fatigue, breathiness, reduced phonational range, aphonia, pitch breaks or inappropriately high pitch, strain/strangle voice, and tremor.(44) One or (frequently) several of these complaints may be given by a patient with a neurological voice disorder. Voice problems in conjunction with other speech production difficulties such as unclear speech, slurring, or unintelligible speech suggest underlying neurologic disease. The onset and progression of symptoms is elicited, as well as what improves and worsens the voice. Symptoms

may vary to some degree in a patient with a neurological voice disorder, however, speech and/or voice is usually not totally normal.(205) Stress worsens most neurologic voice disorders, so this should not be taken as an indication of purely functional etiology of symptoms.(205) Other symptoms of laryngeal dysfunction, ie. swallowing or respiratory complaints are reviewed; if these also occur, a neurologic disease may be suspected. The patient's complete medication history is discussed. A review of systems pays particular attention to respiratory, gastrointestinal, endocrine, as well as neurologic areas. Questions about fatigue, balance, gait disturbance, tremor, changes in handwriting, sensation, and weakness all pertain to the neurologic system review.

The next portion of the examination comprises the phonation assessment. Ideally, it should be conducted jointly by the voice/speech pathologist and laryngologist. Separately or together, they assess the quality of the patient's voice, speech, resonance, articulation and breath support. Such examination begins while listening to the patient relate their medical history. In addition, sustained vowel phonation allows evaluation of phonatory stability, tremor, and movement of laryngeal supporting structures. If the patient is a singer and has complaints related to singing, then observation of the patient while singing is mandatory so that the problem may be demonstrated. Professional vocalists may develop neurological voice disorders; problems during the vocally demanding tasks of singing may be an initial symptom.(206) Testing of diadochokinetic rates of the articulators (tongue, lip, jaw) can be done by having the patient rapidly repeat /pa/,/ta/,/ka/. Lack of precision or crispness suggests lower motor neuron or muscular compromise; irregularity in timing or stress may suggest cerebellar involvement. Sustained phonation of the vowels "ah" and "ee" while alternately pinching and releasing the nares together with reading sentences with high pressure and nasal sounds such as "Suzy stayed all summer" can be screening procedures for velopharyngeal incompetence. Increased nasal resonance or nasal air emission during these tasks suggests lack of adequate velopharyngeal closure. Comprehensive assessment of the laryngeal mechanism in the context of motor speech control is summarized by Yorkston and Beukelman.(207)

The general otolaryngologic/head and neck examination should be conducted, including assessment of ears, nasal passages, facial structures, oral cavity and oropharynx, nasopharynx, and neck inspection and palpation.

Laryngeal imaging examination is a crucial component of the laryngologist's evaluation. The instrument of choice for examination is the flexible fiberoptic laryngoscope.(65) When no lesion is visible in the larynx in a patient with dysphonia, functional, neurogenic, or psychogenic causes should be investigated. These require visualization of the structures of the speech mechanism dynamically. The flexible fiberoptic laryngoscope allows examination of the larynx, pharynx, palate, and velopharyngeal sphincter at rest and during a variety of tasks, including connected speech, sustained phonation, coughing, singing, and swallowing. This capability is extremely helpful in the assessment of neurologic voice disorders. The rigid laryngeal telescope is also helpful, but not useful in every case. It provides superior image quality for view of the larynx at rest and during sustained vowel phonation, and for viewing the vocal fold mucosal waves under stroboscopic lighting.(10) Videodocumentation for review and patient education is also an important aspect of the laryngeal imaging examination.

The patient with a neurological voice disorder may have other neurological examination findings that aid the physician in diagnosis. These have been reviewed in an excellent summary by Rosenfield.(205) While this neurological assessment may not replace a complete evaluation by a neurologist, it helps the laryngologist to understand the nature of the illness of which the voice may only be a part. It also facilitates communication with the neurologist about the patient's problem. Components include a cranial nerve examination, assessment of muscle strength and tone (motor and extrapyramidal), and coordination (cerebellar). Though cranial nerve testing is familiar to the otolaryngologist, other aspects of neurologic examination are not. However, these should be routinely performed in cases of suspected neurologic voice disorders. In the elderly patient, a screening examination for tone (resistance to passive movement) may be more revealing than an assessment of muscle strength.(84) Increased tone on passive movement of the limbs may be an upper motor neuron sign, rigidity of wrist or elbow movement an early sign of Parkinson's disease. Coordination and cerebellar function are assessed by observation of gait and Romberg sign. Following the Romberg test the coordination of the upper extremities is evaluated by finger-nose testing. Observation of the outstretched limbs allows evaluation of postural (cerebellar) tremor. Resting tremor that improves with movement is seen in Parkinson disease. During the examination the patient is observed for adventitious (unintended or involuntary) movements (eg. tremor of head or limbs, circumoral twitches, blepharospasm, dyskinesias, or dystonias). Observation of the patient's handwriting may be helpful in identifying ataxia or tremor.

Special diagnostic tests are indicated in selected cases. Laryngeal electromyography has been found to be helpful in the diagnostic assessment of vocal fold paralysis, differentiating paralysis from fixation, and in determining prognosis for recovery.(59,60,208-210) The workup of vocal fold paralysis, in the absence of identifiable etiology, involves imaging studies along the entire course of the nerve on the affected side.(38) Regarding spasmodic dysphonia, in the absence of secondary causes for dystonia,(82,211) Rosenfield et al (88) recommended additional medical evalu-

ation with a thyroid stimulating hormone level (TSH) to identify concomitant hypothyroidism, and Swenson et al(211) suggested a complete blood count and sedimentation rate to rule out systemic vasculitis. In the patient with suspected myasthenia gravis a Tensilon® test should be performed. Those with presumed underlying central neurologic lesions are evaluated with MRI imaging.

Treatment Issues Related to Neurological Disorders of the Larynx

This chapter has organized neurological disorders of the larynx according to the phonatory dysfunction dimensions of hypoadduction, hyperadduction, malabduction, and short and long-term instabilities, and mixed abnormalities. This classification system lends itself to treatment planning that spans the various etiological factors.

When planning treatment for neurological disorders of the larynx it is important to consider that patients often have disorders of multiple speech subsystems (e.g., respiration, articulation and velopharyngeal). Enhancing laryngeal function must be considered in relation to the impact of each subsystem on functional speech production. Recent findings support the key role of phonation in enhancing oral communication in neurological disorders.(212) Treatment goals must also consider the role of language or cognitive disorders (aphasia, dementia) and the progressive nature of many neurological diseases when planning treatment to maximize functional communication.

There are various forms of treatment which may effect neurological disorders of the larynx. Some treatments (neuropharmacological or neurosurgical) are designed to treat the neurological disorder and may have co-occurring effects on laryngeal or speech function. Other treatments (behavioral, laryngeal-surgical) are designed to directly treat the phonatory dysfunction and improve the voice. These are ideally conducted by an interdisciplinary, collaborative group. Thus, the combined efforts of the speech pathologist, neurologist and otolaryngologist can work to provide the patient with a neurological disorder of the larynx optimal speech intelligibility(21).

Acknowledgments

This work was supported by NIH Grants R01 DC0-1150 and P60 DC 00976 from the National Institute of Deafness and Communication Disorders. This chapter is dedicated with deep appreciation to the memory of Dr. Wilbur James Gould.

References

1. Blitzer A, Brin MF, Sasaki CT, et al. Eds. *Neurologic Disorders of the Larynx*. New York, Thieme Medical Publishers, 1992.

2. Barlow SM, Netsell R, Hunker CJ: Phonatory disorders associated with CNS lesions. In Cummings CW et al Eds: *Otolaryngology-Head and Neck Surgery (1st Ed)*, St Louis, CV Mosby, 1986, p 2087-2093.

3. Larson CR: Brain mechanisms involved in the control of vocalization. *J Voice* 2:301-311, 1988.

4. Gacek RR, Malmgren LT: Laryngeal motor innervation - central. In Blitzer A, et al Eds. *Neurologic Disorders of the Larynx*. New York, Thieme Medical Publishers, 1992, p 29-35.

5. Larson CR, Yoshida Y, Sessile BJ, Ludlow CR: Higher level motor and sensory integration. In Titze IR Ed: *Vocal Fold Physiology: Frontiers in Basic Science*, San Diego, Singular Publishing Group, 1993,p 227-275.

6. Ward PH, Berci G, Calcaterra TC: Advances in endoscopic examination of the respiratory system. *Ann Otol Rhinol Laryngol* 83:754-760, 1974.

7. Hanson DG: Neuromuscular disorders of the larynx. *Otolaryngol Clin North Am* 24:1035-1051,1991.

8. Yanagisawa E: Physical examination of the larynx and videolaryngoscopy. In Blitzer A, et al Eds. *Neurologic Disorders of the Larynx*. New York, Thieme Medical Publishers, 1992, p 82-97.

9. Baken R: *Clinical Measurement of Speech and Voice*. Boston, Little, Brown and Company, 1987.

10. Bless DM: Measurement of vocal function. *Otolaryngol Clin North Am* 24:1023-1033, 1991.

11. Ramig LO, Scherer RC, Klasner ER, et al: Acoustic analysis of voice in amyotrophic lateral sclerosis: a longitudinal case study. *J Speech Hear Disord* 55:2-14, 1990.

12. Countryman S, Ramig LA: Effects of intensive voice therapy on voice deficits associated with bilateral thalamotomy in Parkinson's disease: a case study. *J Med Speech-Lang Pathology* 1:233-249,1993.

13. Logemann JA, Fisher HB, Boshes B et al: Frequency and occurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson's patients. *J Speech Hear Dis* 42:47-57, 1978.

14. Elble RJ, Koller WC: *Tremor*. Baltimore, MD. Johns Hopkins University Press, 1990.

15. Koufman JA, Isaacson G: The spectrum of vocal dysfunction. *Otolaryngol Clin North Am* 24:985-988,1991.

16. Darley F, Aronson A, Brown J: Differential diagnostic patterns of dysarthria. *J Speech Hear Res* 12:246-269, 1969.

17. Darley F, Aronson A, Brown J: Clusters of deviant speech dimensions in the dysarthrias. *J Speech Hear Res* 12:462-496, 1969.

18. Ward PH, Hanson DG, Berci G: Photographic studies of the larynx in central laryngeal paresis and paralysis. *Acta Otolaryngol (Stockh)* 91:353-367, 1981.

19. Ward PH, Hanson DG, Berci G: Observations on central neurologic etiology for laryngeal dysfunction. *Ann Otol Rhinol Laryngol* 90:430-441, 1981.

20. Aronson AE: *Clinical Voice Disorders*, 3rd Ed., New York, Thieme, 1990.

21. Ramig LA, Scherer RC: Speech therapy for neurological disorders of the larynx. In Blitzer A et al Eds. *Neurologic Disorders of the Larynx*. New York, Thieme Medical Publishers, 1992, p 163-181.

22. Hanson DG, Gerratt BR, Ward PH: Cinegraphic observations of vocal pathology in Parkinson's disease. *Laryngoscope* 92:348-353,1984.

23. Smith ME, Ramig LO, Dromey C, et al: Intensive voice treatment in Parkinson's disease: laryngostroboscopic findings. *J Voice*, 1994, in press.

24. McHenry MA, Wilson RL, Minton JT: Management of multiple physiologic system deficits following traumatic brain injury. J Med Speech-Lang Path 2:59-74, 1994.

25. Younger DS, Lange DJ, Lovelace RE, et al: Neuromuscular disorders of the larynx. In Blitzer A, et al Eds. *Neurologic Disorders of the Larynx*. New York, Thieme Medical Publishers, 1992, p 246.

26. Mastaglia FL: Genetic myopathies. In Swash M, Oxbury J Eds: *Clinical Neurology*, New York, Churchill Livingstone, 1991, p 1286.

27. Ramig LA, Scherer RC, Titze IR et al: Acoustic analysis of voices of patients with neurological disease: rationale and preliminary data. *Ann Otol Rhinol Laryngol* 97:164-172, 1988.

28. Fenchiel GM: Clinical syndromes of myasthenia in infancy and childhood. *Arch Neurol* 35:97-103, 1978.

29. Scadding GK, Harvard CWH. Pathogenesis and treatment of myasthenia gravis. *Br Med J* 283:1008-1019, 1981.

30. Carpenter RJ, McDonald TJ, Howard FM: The otolaryngologic presentation of myasthenia gravis. *Laryngoscope* 89:922-928, 1979.

31. Dhillon RS, Brookes GB: Myasthenia gravis in otolaryngological practice. *Clin Otolaryngol* 9:27-34, 1984.

32. Calcaterra TC, Stern F, Herrmann C, et al: The otolaryngologist's role in myasthenia gravis. *Trans Am Acad Ophthalmol Otolaryngol* 76:308-312, 1972.

33. Neiman RF, Mountjoy JR, Allen EL: Myasthenia gravis focal to the larynx: report of a case. *Arch Otolaryngol* 101:569-570, 1975.

34. Warren WR, Gutmann L, Cody RC, et al: Stapedius reflex decay in myasthenia gravis. *Arch Neurol* 34:496-497, 1977.

35. Rontal M, Rontal E, Leuchter W, et al: Voice spectrography in the evaluation of myasthenia gravis of the larynx. *Ann Otol Rhinol Laryngol* 87:722-728, 1978.

36. Tyler HR: Neurology of the larynx. *Otolaryngol Clin North Am* 17:75-79, 1984.

37. Oosterhuis JHGH: Myasthenia gravis and other myasthenic syndromes. In Swash M, Oxbury J Eds: *Clinical Neurology*, New York, Churchill Livingstone, 1991, p 1368.

38. Terris DJ, Arnstein DP, Nguyen HH: Contemporary evaluation of unilateral vocal cord paralysis. *Otolaryngol Head Neck Surg* 107:84-90, 1992.

39. Annino DJ, MacArthur CJ, Friedman EM. Vincristine-induced recurrent laryngeal nerve paralysis. *Laryngoscope* 102:1260-1262, 1992.

40. Cannon S, Ritter FN: Vocal cord paralysis in postpoliomyelitis syndrome. *Laryngoscope* 97:981-983, 1987.

41. Nugent KM: Vocal cord paresis and glottic stenosis: a late complication of poliomyelitis. *South Med J* 80:1594-1595, 1987.

42. Blau JN, Kapadia R: Idiopathic palsy fo the recurrent laryngeal nerve: a transient cranial mononeuropathy. *Br Med J* 4:259-261, 1972.

43. Ward PH, Berci G: Observations on so-called idiopathic vocal cord paralysis. *Ann Otol Rhinol Laryngol* 91:558-563, 1982.

44. Colton RH, Casper JK: *Understanding Voice Problems: A Physiological Perspective for Diagnosis and Treatment.* Baltimore, Williams & Wilkins, 1990.

45. Gerratt BR, Hanson DG, Berke GS: Glottographic measures of laryngeal function in individuals with abnormal motor control. In Baer T, Sasaki C, Harris K Eds. *Laryngeal Function in Phonation and Respiration*, Boston, Little, Brown, and Co., 1987, p 521-534.

46. Hanson DG, Gerratt BR, Karin R, et al: Glottographic measures of vocal fold vibration: laryngeal paralysis. *Laryngoscope* 98:348-353, 1988.

47. Woo P, Colton R, Shangold L: Phonatory airflow analysis in patients with laryngeal disease. *Ann Otol Rhinol Laryngol* 96:549-555, 1987.

48. Hirano M, Kurita S, Yukizane K, et al: Asymmetry of the laryngeal framework: a morphological study of cadaver larynges. *Ann Otol Rhinol Laryngol* 98:135-140, 1989.

49. Sanders I, Wu BL, Mu L, et al: The innervation of the human larynx. *Arch Otolaryngol Head Neck Surg* 119:934-939, 1993.

50. Crumley R, McCabe B. Regeneration of the recurrent laryngeal nerve. *Otolaryngol Head Neck Surg* 92:442-447, 1982.

51. Kirchner JA. Atrophy of laryngeal muscles in vagal paralysis. *Laryngoscope* 76;1753-1765, 1966.

52. Quiney RE, Michaels L. Histopathology of vocal cord palsy from recurrent laryngeal nerve damage. *J Otolaryngol* 19:237-241, 1990.

53. Smith ME, Berke GS, Gerratt BR, et al: Laryngeal paralyses: theoretical considerations and effects of laryngeal vibration. *J Speech Hear Res* 35:545-554, 1992.

54. Hirano M, Bless DM. *Videostroboscopic Examination of the Larynx.* San Diego, Singular Publishing Group, 1993, p 167-169.

55. Isshiki N, Ishikawa T. Diagnostic value of tomography in unilateral vocal cord paralysis. *Laryngoscope* 86:1573-1578, 1976.

56. Dedo HH. The paralyzed larynx: an electromyographic study in dogs and humans. *Laryngoscope* 80:1455-1517, 1970.

57. Woodson GE: Configuration of the glottis in laryngeal paralysis. II: animal experiments. *Laryngoscope* 103:1235-1241, 1993.

58. Woodson GE: Configuration of the glottis in laryngeal paralysis. I: clinical study. *Laryngoscope* 103:1227-1234, 1993.

59. Hirano M, Nozoe I, Shin T, Maeyama T: Electromyography for laryngeal paralysis. In Hirano, M, Kirchner JA, Bless DM, Eds. *Neurolaryngology: Recent Advances*, College-Hill, Boston, 1987, pg. 232-248.

60. Koufman J: Management of vocal fold paralysis. *The Visible Voice*, 2:50, 1993.

61. Ward PH, Berci G, Calcaterra TC: Superior laryngeal nerve paralysis: an often overlooked entity. *Trans Am Acad Ophthalmol Otolaryngol* 84:78-89,1977.

62. Adour KK, Schneider GD, Hilsinger RL: Acute superior laryngeal nerve palsy: analysis of 78 cases. *Otolaryngol Head Neck Surg* 88:418-424, 1980.

63. Tanaka S, Hirano M, Umeno H. Laryngeal behavior in unilateral superior laryngeal nerve paralysis. *Ann Otol Rhinol Laryngol* 103:93-97, 1994.

64. Sataloff RT: Structural and neurological disorders and surgery of the voice. In Sataloff RT ed. *Professional Voice: The Science and Art of Clinical Care*, New York, Raven Press, 1991, p 267-299.

65. . Koufman JA: Approach to the patient with a voice disorder. *Otolaryngol Clin North Am* 24:989-998, 1991.

66. Isshiki N, Tanabe M, Ishizaka K, et al: Clinical significance of asymmetrical vocal cord tension. *Ann Otol Rhinol Laryngol* 86:58-66, 1977.

67. Berke G, Moore D, Hantke D, et al: Laryngeal modeling: theoretical, in-vitro, in-vivo. *Laryngoscope* 97:871-881, 1987.

68. Hornykiewicz O. Metabolism of brain dopamine in human parkinsonism: neurochemical and clinical aspects. In E Costa, L Cote and M Yahr (Eds), New York Raven Press, 1966.

69. Canter GJ. Speech characteristics of patients with Parkinson's disease I. intensity, pitch and duration. *J Speech Hear Disord* 28:221-229, 1963.

70. Aronson AE, Brown JR, Litin EM et al. Spastic dysphonia II. Comparison with essential (voice) tremor and other neurologic and psychogenic dysphonias. *J Speech Hear Disord* 33:220-231, 1968.

71. Hirose H, Joshita Y: Laryngeal behavior in patients with disorders of the central nervous system. In Hirano M, Kirchner JA, Bless DM Eds: *Neurolaryngology: Recent Advances*, Boston, Little, Brown, and Co., 1987, p 258-266.

72. Ramig LO, Countryman S, Thompson L, et al.: Efficacy of voice therapy for patients with Parkinson disease. J Speech Hear Res, 1994, in submission.

73. Gracco C, Marek K: Laryngeal manifestations of early Parkinson's disease: data characterizing stage of disease and severity of symptoms. Presented at the Conference of Motor Speech, Sedona, AZ, March, 1994.

74. Larson K, Ramig L, Scherer R: The on-off effect of medication on voice and speech in Parkinson disease. *J Med Speech Path*, 1994, in press.

75. Freed C, Breeze R, Rosenberg N, et al: Survival of implanted fetal dopamine cells and neurologic improvement 12 to 46 months after transplantation for Parkinson disease. *New Engl J Med* 327:1549-1555, 1993.

76. Larson K, Ramig L, Freed C, et al: Fetal cell transplant: effects on acoustic characteristics of voice and speech. Presented at the Motor Speech Conference, Sedona, AZ, March, 1994.

77. Sarno MT: Speech impairment in Parkinson's disease. *Arch Phys Med Rehab* 49:269-275, 1968.

78. Weiner WJ, Singer C: Parkinson's disease and nonpharmacologic treatment programs. *J Am Geriatr Soc* 37:359-363, 1989.

79. Ramig LO, Bonitati C, Lemke J, et al: Voice treatment for patients with Parkinson disease: development of an approach and preliminary efficacy data. *J Med Speech-Lang Path*, 1994, in press.

80. Dromey C, Ramig LO, Johnson A: The control of vocal intensity in Parkinson disease: a longitudinal case study of the effects of intensive voice treatment. *J Speech Hear Res*, 1994, in submission.

81. Ramig LO, Dromey C: Aerodynamic characteristics of speech in patients with Parkinson disease: pre and post intensive voice therapy. *J Speech Hear Res*, 1994, in submission.

82. Ramig LO: Speech therapy for Parkinson's disease. In Koller W, Paulson G (Eds), *Therapy of Parkinson's Disease*, New York, Marcel Dekker, Inc. (in press).

83. Brin MF, Fahn S, Blitzer A, et al: Movement disorders of the larynx. In Blitzer A, Brin MF, Sasaki CT, et al. Eds. *Neurologic Disorders of the Larynx*. New York, Thieme Medical Publishers, 1992, p 248-278.

84. Glick TH: *Neurologic Skills: Examination and Diagnosis*. Boston, Blackwell Scientific Publications, 1993.

85. Aring CD: Supranuclear (pseudobulbar) palsy. *Arch Int Med* 115;19, 1965.

86. Penney JB, Young AB, Shoulson I, et al: Huntington's disease in Venezuela: 7 years of follow-up on symptomatic and asymptomatic individuals. *Movement Disord* 5:93-99, 1990.

87. Ramig LA: Acoustic analyses of phonation in patients with Huntington's disease: preliminary report. *Ann Otol Rhinol Laryngol* 95:288-293, 1986.

88. Rosenfield DB, Donovan DT, Sulek M, et al: Neurologic aspects of spasmodic dysphonia. *J Otolaryngol* 19:231-236, 1990.

89. Izdebski K: Symptomatology of adductor spasmodic dysphonia: a physiologic model. *J Voice* 6:306-319, 1992.

90. Cannito M, Johnson H: Spastic dysphonia: a continuum disorder. *J Commun Disord* 14:215-223, 1981.

91. Blitzer A, Lovelace RE, Brin MF, et al: Electromyographic findings in focal laryngeal dystonia (spastic dysphonia). *Ann Otol Rhinol Laryngol* 94:591-594, 1985.

92. Blitzer A, Brin MF, Fahn S, et al: Clinical and laboratory characteristics of focal laryngeal dystonia: study of 110 cases. *Laryngoscope* 98:636-640, 1988.

93. Freeman FJ, Cannito M, Finitzo-Hieber T: Classification of spasmodic dysphonia be visual perceptual and acoustic means. In Gates GA, ed: *Spastic Dysphonia: State of the Art*. New York, The Voice Foundation, 1985, p 5-18.

94. Finitzo T, Freeman F: Spasmodic dysphonia: whether and where: results of seven years of research. *J Speech Hear Res* 32:541-555, 1989.

95. Blitzer A, Brin MF: The dystonic larynx. *J Voice* 6:294-297, 1992.

96. Hanson DG, Logemann JA, Hain T: Differential diagnosis of spasmodic dysphonia: a kinematic perspective. *J Voice* 6:325-337, 1992.

97. Cannito MP, Kondraske GV, Johns DF: Oral-facial sensorimotor function in spasmodic dysphonia. In Moore CA, Yorkston KM (Eds.), *Dysarthria and Apraxia of Speech: Perspectives on Management*, Paul H. Brookes Publishing Co., Baltimore, 1991, p 205-225.

98. Jacome DE, Yanez GF: Spastic dysphonia and Meigs disease [letter]. *Neurology* 30:349, 1980.

99. Robe E, Brumlik J, Moore P: A study of spastic dysphonia (neurologic and electroencephalic abnormalities). *Laryngoscope* 70:219-245, 1960.

100. Finitzo-Hieber T, Freeman FJ, Gerling IJ, et al: Auditory brainstem response abnormalities in adductor spasmodic dysphonia. *Am J Otolaryngol* 3:26-30, 1982.

101. Tolosa E, Montserrat L, Bayes A: Blink reflex studies in focal dystonias: enhanced excitability of brainstem interneurons in cranial dystonia and spasmodic torticollis. *Movement Disorders* 3:61-69, 1988.

102. Cohen LG, Ludlow CL, Warden BS, et al: Blink reflex excitability recovery curves in patients with spasmodic dysphonia. *Neurology* 39:572-577, 1989.

103. Devous MD, Pool KD, Finitzo T, et al: Evidence for cortical dysfunction in spasmodic dysphonia: regional cerebral blood flow and quantitative electrophysiology. *Brain Lang* 39:331-344, 1990.

104. Schaefer SD, Freeman FJ, Finitzo T, et al: Magnetic resonance imaging findings and correlations in spasmodic dysphonia patients. *Ann Otol Rhinol Laryngol* 94:595-601, 1985.

105. Pool KD, Freeman FJ, Finitzo T, et al: Heterogeneity in spasmodic dysphonia: neurologic and voice findings. *Arch Neurol* 48:305-309, 1991.

106. Davis PJ, Bartlett D, Luschei E, Berke GS: Coordination of the respiratory and laryngeal systems in breathing and vocalization. In Titze IR Ed. *Vocal Fold Physiology: Frontiers in Basic Science*, San Diego, Singular Publishing Group, 1993, pp 189-226.

107. Ludlow CL, VanPelt F, Koda J: Characteristics of late responses to superior laryngeal nerve stimulation in humans. *Ann Otol Rhinol Laryngol* 101:127-134, 1992.

108. Ludlow CL: Evaluation of methods fordifferentialdiagnosis of the spasmodic dysphonias. Presented at the Fifth Annual Pacific Voice Conference, San Francisco, CA, October, 1992.

109. Bloch CS, Hirano M, Gould WJ: Symptom improvement of spastic dysphonia in response to phonatory tasks. *Ann Otol Rhinol Laryngol* 94:51-54, 1985.

110. Ginsberg BI, Wallack JJ, Srain JJ, et al: Defining the psychiatric role in spastic dysphonia. *Gen Hosp Psychiatr* 10:132-137, 1988.

111. Murry T, Cannito MP, Woodson GE: Spasmodic dysphonia: emotional status and botulinum toxin treatment. *Arch Otolaryngol Head Neck Surg* 120:310-316, 1994.

112. Woodson GE, Zwirner P, Murry T, et al: Functional assessment of patients with spasmodic dysphonia. *J Voice* 6:338-343, 1992.

113. Parnes SM, Lavarato AB, Myers EN: Study of spastic dysphonia using videofiberoptic laryngoscopy. *Ann Otol Rhinol Laryngol* 87:322-326, 1978.

114. Davis PJ, Boone DR, Carroll RL, et al: Adductor spastic dysphonia: heterogeneity of physiologic and phonatory characteristics. *Ann Otol Rhinol Laryngol* 97:179-185, 1988.

115. Woodson GE, Zwirner P, Murry T, et al: Use of flexible fiberoptic laryngoscopy to assess patient with spasmodic dysphonia. *J Voice* 5:85-91, 1991.

116. Zwirner P, Murry T, Swenson M, et al: Effects of botulinum toxin therapy in patients with adductor spasmodic dysphonia: acoustic, aerodynamic, and videoendoscopic findings. *Laryngoscope* 102:400-406, 1992.

117. Woo P, Colton R, Casper J, et al: Analysis of spasmodic dysphonia by aerodynamic and laryngostroboscopic measurements. *J Voice* 6:344-351, 1992.

118. Miller RH, Woodson GE: Treatment options in spasmodic dysphonia. *Otolaryngol Clin North Am* 24:1227-1237, 1991.

119. Heuer RJ: Behavioral therapy for spasmodic dysphonia. *J Voice* 6:352-354, 1992.

120. Dedo HH: Recurrent laryngeal nerve section for spasmodic dysphonia. *Ann Otol Rhinol Laryngol* 85:451-459, 1976.

121. Tucker HM: Laryngeal framework surgery in the management of spasmodic dysphonia. *Ann Otol Rhinol Laryngol* 98:52-54, 1989.

122. Ludlow CL: Treatment of speech and voice disorders with botulinum toxin. *JAMA* 264:2671-2675, 1990.

123. Brin MF, Blitzer A, Stewart C, Fahn S: Treatment of spasmodic dysphonia (laryngeal dystonia) with local injections of botulinum toxin: review and technical aspects. In Blitzer A, et al Eds. *Neurologic Disorders of the Larynx.* New York, Thieme Medical Publishers, 1992,p 214.

124. Ludlow CL, Sedory SE, Fujita M: Neurophysiological control of vocal fold adduction and abduction for phonation onset and offset during speech. In Gauffin J and Hammarberg B Eds. *Vocal Fold Physiology: Acoustic, Perceptual, and Physiological Aspects of Voice Mechanisms.* San Diego, Singular Publishing Group, Inc., 1991, p 197-206.

125. Koufman JA, Blalock PD: Classification of laryngeal dystonias. *The Visible Voice.* 1(3):4-5,19-23, 1992.

126. Hartman DE, Aronson AE: Clinical investigations of intermittent breathy dysphonia. *J Speech Hear Disord* 46:428-432, 1981.

127. Merson RM, Ginsberg AP: Spasmodic dysphonia: abductor type. a clinical report of acoustic, aerodynamic, and perceptual characteristics. *Laryngoscope* 89:129-139, 1979.

128. Wolfe VI, Bacon M: Spectrographic comparison of two types of spastic dysphonia. *J Speech Hear Disord* 41:325-332, 1976.

129. Zwitman DH: Bilateral cord dysfunctions: abductor type spastic dysphonia. *J Speech Hear Disord* 44:373-378, 1979.

130. Koufman JA, Blalock PD: Diagnosis of spasmodic dysphonia by spectral analysis. *The Visible Voice.* 1(3):6-7,15-18, 1992.

131. Watson BC, Schaefer SD, Freeman FJ, et al: Laryngeal electromyographic activity in adductor and abductor spasmodic dysphonia. *J Speech Hear Res* 34:473-482, 1991.

132. Van Pelt F, Ludlow CL, Smith PJ: Comparison of muscle activation patterns in adductor and abductor spasmodic dysphonia. *Ann Otol Rhinol Laryngol* 103:192-200, 1994.

133. Ludlow CL, Naunton RF, Terada S, et al: Successful treatment of selected cases of abductor spasmodic dysphonia using botulinum toxin. *Otolaryngol Head Neck Surg* 104:849-855, 1991.

134. Rontal M, Rontal E, Rolnick M, et al: A method for the treatment of abductor spasmodic dysphonia with botulinum toxin injections: a preliminary report. *Laryngoscope* 101:911-914, 1991.

135. Blitzer A, Brin MF, Stewart C, et al: Abductor laryngeal dystonia: a series treated with botulinum toxin. *Laryngoscope* 102:163-167, 1992.

136. Titze IR: *Principles of Voice Production.* Prentice Hall, 1994.

137. Titze IR: Measurements for the assessment of voice disorders. In Cooper JA Ed. *Assessment of Speech and Voice Production: Reasearch and Clinical Applications*, NIDCD Monograph, Volume 1:1991, pg 42-49.

138. Larson C, Kempster G, Kistler M: Changes in voice fundamental frequency following discharge of single motor units in cricothyroid and thyroarytenoid muscles. *J Speech Hear Res* 30:552-558, 1987.

139. Titze IR: A model for neurologic sources of aperiodicity in vocal fold vibration. *J Speech Hear Res* 34:460-472, 1991.

140. Davies P, McGowan R, Shadel C, Scherer R: Practical flow duct acoustics applied to the vocal tract. In Titze IR, Ed. *Vocal Fold Physiology: Frontiers in Basic Science*, San Diego, Singular Publishing Group, Inc, 1993, p. 93-142.

141. Rothenberg M: Acoustic interaction between the glottal source and the vocal tract. In Stevens KN, Hirano M, Eds. *Vocal Fold Physiology.* University of Tokyo Press, Tokyo, 1979, pg 305-328.

142. Orlikoff RF, Baken RJ: The effect of the heartbeat on fundamental frequency perturbation. *J Speech Hear Res* 32:576-583, 1989.

143. Orlikoff RF: Vowel amplitude variation associated with the heart cycle. *J Acoust Soc Am* 88:2091-2098, 1990.

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

145. Scherer RC, Gould WJ, Titze IR, et al: Preliminary evaluation of selected acoustic and glottographic measures for clinical phonatory function analysis. *J Voice* 2:230-244, 1988.

146. Heiberger VL, Horii Y: Jitter and shimmer in sustained phonation. In Lass NJ Ed. *Speech and Language: Advances in Basic Research and Practice*, Vol. 7, Academic Press, New York, 1982, pg 299-332.

147. Ramig LA, Ringel RL: The effects of physiological aging on selected acoustic characteristics of voice. *J Speech Hear Res* 26:22-30, 1983.

148. Linville SE: Intraspeaker variability in fundamental frequency stability: an age-related phenomenon? *J Acoust Soc Am* 83:741-745, 1988.

149. Wilcox K, Horii Y: Age and changes in vocal jitter. *J Gerontol* 35:194-198, 1980.

150. Horii Y. Jitter and shimmer differences among sustained vowel phonations. *J Speech Hear Res* 25:12-14, 1982.

151. Ramig LA, Scherer RC, Titze IR, et al: Acoustic analysis of voices of patients with neurologic disease: rationale and preliminary data. *Ann Otol Rhinol Laryngol* 97:164-172, 1988.

152. Ludlow C, Coulter D, Gentges F: The differential sensitivity of frequency perturbation to laryngeal neoplasms and neuropathologies. In Bless DM, Abbs JH, Eds. *Vocal Fold Physiology: Contemporary Research and Clinical Issues*. San Diego, College-Hill Press, 1983, pg 381-392.

153. Zwirner P, Murry T, Woodson GE: Phonatory function in neurologically impaired patients. *J Commun Disord* 24:287-300, 1991.

154. Baken RJ: Irregularity of vocal period and amplitude: a first approach to the fractal analysis of voice. *J Voice* 4:185-197, 1990.

155. Herzel H, Steinecke I, Mende W, et al: Chaos and bifurcations during voiced speech. In Mosekilde E. Ed. *Complexity, Chaos and Biological Evolution*, Plenum Press, New York, 1991, pg 41-50.

156. Gleick J: *Chaos: Making a New Science*. New York, Viking Penguin, 1987.

157. Berry DA, Titze IR: Analysis of rough voice from a nonlinear dynamics point of view (Abs). *J Acoust Soc Am* 94 (pt 2):1762, 1993.

158. Ramig LA, Shipp T: Comparative measures of vocal tremor and vocal vibrato. *J Voice* 2:162-167, 1987.

159. Horii Y: Acoustic analysis of vocal vibrato: a theoretical interpretation of data. *J Voice* 3:36-43, 1989.

160. Jankovic J: Tremors: pathophysiology, differential diagnosis and pharmacology. *Neuro Cons* 2:1-8, 1984.

161. Brown JR, Simonson J: Organic voice tremor: a tremor of phonation. *Neurology* 13:520-525, 1963.

162. Hachinski VC, Thomsen IV, Buch NH: The nature of primary vocal tremor. *Can J Neurol Sci* 2:195-197, 1975.

163. Hartman DE, Overholt SL, Vishwanat B: A case of vocal cord nodules making essential (vocal) tremor. *Arch Otolaryngol* 108:52-53, 1982.

164. Massey EW, Paulson GW: Essential vocal tremor: clinical characteristics and response to therapy. *South Med J* 78:316-317, 1985.

165. Ludlow CL, Bassich CJ, Connor NP, et al: Phonatory characteristics of vocal fold tremor. *J Phonetics* 14:509-515, 1986.

166. Philippbar SA, Robin DA, Luschei ES: Limb, jaw and vocal tremor in Parkinson's patients. In Yorkston K, Beukelman D, Eds. *Recent Advances in Clinical Dysarthria*. San Diego, College-Hill Press, 1991.

167. Hartman DE, Abbs JH, Vishwanat B: Clinical investigations of adductor spastic dysphonia. *Ann Otol Rhinol Laryngol* 97:247-252, 1988.

168. Aronson AE, Ramig LO, Winholtz WS, et al: Rapid voice tremor, or "flutter," in amyotrophic lateral sclerosis. *Ann Otol Rhinol Laryngol* 1992;101:511-518.

169. Murray TJ: Essential tremor. *Can Med Assoc J* 124:1559-1570, 1981.

170. Colebatch JG, Findley LJ, Frackowiak RS, et al: Preliminary report: activation of the cerebellum in essential tremor. *Lancet* 336:1028-1030, 1990.

171. Jenkins IH, Bain PG, Colebatch JG, et al: A positron emission tomography study of essential tremor: evidence for over activity of cerebellar connections. *Ann Neurol* 34:82-90, 1993.

172. Aronson AE, Hartman DE: Adductor spastic dysphonia as a sign of essential (voice) tremor. *J Speech Hear Disord* 46:52-58, 1981.

173. Lebrun Y, Devreux F, Jean-Jaques R, et al: Tremulous speech. *Folia Phoniat* 34:134-142, 1982.

174. Koda J, Ludlow CL: An evaluation of laryngeal muscle activation in patients with voice tremor. *Otolaryngol Head Neck Surg* 107:684-696, 1992.

175. Stager SV, Ludlow CL: Responses of stutterers and vocal tremor patients to treatment with botulinum toxin. In Jankovic J, Hallett M (Eds). *Therapy with Botulinum Toxin*, New York, Marcel Dekker.

176. Perez KS, Ramig LO, Smith ME, Dromey C : Videostroboscopic observations in patients with Parkinson's disease. Presented at the Voice Foundation Symposium, June, 1994.

177. Doody RS, Rosenfield DB: Spasmodic dysphonia associated with palatal myoclonus. *Ear Nose Throat J* 69:829-832, 1990.

178. Toland AD, Porubsky ES, Coker NJ, et al: Velopharyngo-laryngeal myoclonus: evaluation of objective tinnitus and extrathoracic airway obstruction. *Laryngoscope* 94:691-695, 1984.

179. Drysdale AJ, Ansell J, Adeley J: Palato-pharyngo-laryngeal myoclonus: an unusual cause of dysphagia and dysarthria. *J Laryngol Otol* 107:746-747, 1993.

180. Lapresle J: Palatal myoclonus. *Adv Neurol* 43:265-273, 1986.

181. Kent RD, Netsell R, Abbs JH: Acoustic characteristics of dysarthria associated with cerebellar disease. *J Speech Hear Res* 1979;22:627-648.

182. Aronson AE, Brown JR, Litin EM, et al: Spastic dysphonia II. Comparison with essential (voice) tremor and other neurologic and pychogenic dysphonias. *J Speech Hear Disord* 1968;33:220-231.

183. Ackermann H, Ziegler W: Acoustic analysis of vocal instability in cerebellar dysfunctions. *Ann Otol Rhinol Laryngol* 103:98-104, 1994.

184. Carpenter R, McDonald T, Howard F: The otolaryngologic presentation of amyotrophic lateral sclerosis. *Trans Am Acad Ophthalmol Otolaryngol* 1978;86:479-484.

185. Carrow E, Rivera V, Mauldin M, et al: Deviant speech characteristics in motor neuron disease. *Arch Otolaryngol* 1974;100;212-218.

186. Strand EA, Buder KM, Yorkston KM, et al: Differential phonatory characteristics of four women with amytrophic lateral sclerosis. *J Voice* 1994;in press.

187. Garfinkle TJ, Kimmelman CP: Neurologic disorders: amyotrophic lateral sclerosis, myasthenia gravis, multiple sclerosis, poliomyelitis. *Am J Otolaryngol* 1982;3:204-212.

188. Charcot, JM: Lectures on the diseases of the nervous system. Vol.1 New Sydenham Society: London, 1877.

189. Jensen, JR: A study of certain motor-speech aspects of the speech of multiple sclerosis patients. Ph.D. dissertation, University of Wisconsin, 1960.

190. Beukelman DR, Kraft GH, Freal J: Expressive communication disorders in persons with multiple sclerosis: a survey. *Arch Phys Med Rehab* 66:675-677, 1985.

191. Darley FL, Brown JR, Goldstein N: Dysarthria in multiple sclerosis. *J Speech Hear Res* 15:229-245, 1972.

192. Zemlin, WR: A comparison of the periodic function of vocal fold vibration in a multiple sclerosis and a normal population. Ph.D. dissertation. University of Minnesota, 1962.

193. Morrison M, Rammage L: *The Management of Voice Disorders*. San Diego, CA, Singular Publishing Group, 1994, pg 161-162.

194. Farmakides MN, Boone DK: Speech problems of patients with multiple sclerosis. *J Speech Hear Disord* 25:385-390, 1960.

195. Stacy M, Jankovic J: Differential diagnosis of Parkinson's disease and the parkinsonism plus syndromes. *Neurol Clin* 10:341-359, 1992.

196. Bawa R, Ramadan HH, Wetmore SJ: Bilateral vocal cord paralysis in Shy-Drager syndrome. *Otolaryngol Head Neck Surg* 109:911-914, 1993.

197. Hanson DG, Ludlow C, Bassich C: Vocal fold paresis in Shy-Drager syndrome. *Ann Otol Rhinol Laryngol* 92:85-90, 1984.

198. Williams AJ, Hanson D, Calne D: Vocal cord paralysis in Shy-Drager syndrome. *J Neurol Neurosurg Psychiatr* 42:151-153, 1979.

199. Countryman S, Ramig LO, Pawlas AA: Speech and voice deficits in parkinsonian plus syndromes: can they be treated? *J Med Speech-Lang Path*, 1994, in press.

200. Guidi AM, Bannister R, Gibson WPR, et al: Laryngeal electromyography in multiple system atrophy with autonomic failure. *J Neurol Neurosurg Psychiatr* 44:49-53, 1981.

201. Steele JC, Richardson JC, Olszewski J: Progressive supranuclear palsy. *Arch Neurol* 10:333-359, 1964.

202. Kluin KJ, Foster NL, Berent S, et at: Perceptual analysis of speech disorders in progressive supranuclear palsy. *Neurol* 43:563-566, 1993.

203. Metter EJ, Hanson WR: Dysarthria in progressive supranuclear palsy. In Moore CA, Yorkston KM (Eds.), *Dysarthria and Apraxia of Speech: Perspectives on Management*, Paul H. Brookes Publishing Co., Baltimore, 1991, pg 127-136.

204. Tolosa E, Peña J: Involuntary vocalizations in movement disorders. *Adv Neurol* 49:343-363, 1988.

205. Rosenfield DB: Neurolaryngology. *ENT J* 66:323-326, 1987.

206. Sataloff RT: The physical examination. In Sataloff RT ed. *Professional Voice: The Science and Art of Clinical Care*, New York, Raven Press, 1991, pp 91-100.

207. Yorkston KM, Beukelman DR, Bell KR: Laryngeal function. In *Clinical Management of Dysarthric Speakers*, Boston, College Hill Press, Little, Brown, and Co., 1988, pg 241-267.

208. Miller RH, Rosenfield DB: The role of electromyography in clinical laryngology. *Otolaryngol Head Neck Surg* 1984;92:287-291.

209. Parnes SM, Satya-Murti S: Predictive value of laryngeal electromyography in patients with vocal cord paralysis of neurogenic origin. *Laryngoscope* 1985:95:1323-1326.

210. Lovelace RE, Blitzer A, Ludlow CL: Clinical laryngeal electromyography. In Blitzer A, et al Eds. *Neurologic Disorders of the Larynx*. New York, Thieme Medical Publishers, 1992, pg. 66-81.

211. Swenson MR, Zwirner P, Murry T, et al: Medical evaluation of patients with spasmodic dysphonia. *J Voice* 6:320-324, 1992.

212. Ramig LO, Dromey C, Johnson A, et al: The roles of phonatory and articulatory effort in the treatment of patients with Parkinson disease. Presented at the Motor Speech Conference, Sedona, AZ, March, 1994.

NCVS Status and Progress Report - 7
December 1994, 229-230

Vocal Fold Scarring - Current Concepts in Management: The Role of Collagen Injection

Charles Ford, M.D.
Division of Otolaryngology-Head & Neck Surgery, Department of Surgery, The University of Wisconson-Madison

Improving the voice of patients with scarred vocal folds represents a challenge to the phonosurgeon. Scar tissue in the membranous vocal fold disrupts normal vibratory function by changing the physical properties of the tissues and destroying the body-cover interface. Such changes can result in glottic insufficiency and the symptoms can be even more pronounced when scarring involves deeper musculoskeletal laryngeal structures. Although there is no ideal method to resolve such problems, there are several treatment approaches that have proven helpful. Behavioral management with voice therapy and anti-reflux measures can be effective over time. Medical treatment with steroids and antibiotics occasionally plays a role in prophylaxis and early management of scarring. Aspects of the problem can be addressed with a variety of surgical techniques from scar lysis to tissue replacement using flaps and grafts. Injection of alloplastics can be used in some instances to augment deficient vocal fold tissue but the resultant fibrosis might aggravate the problem. Bioimplants such as fat and collagen are more suitable fillers for the scarred vocal fold and soluble bovine collagen has been demonstrated to soften scar tissue.

Encouraged by studies showing softening of facial scars with repeated injections of Zyderm Collagen Implant (Collagen Corporation, Palo Alto, Ca.)[1,2] we performed a series of experiments injecting canine larynges with bovine collagen preparations. These implants promoted minimal inflammatory response, seemed to be rapidly invaded by host fibroblasts and partially replaced by new host collagen[3,4] within 6-12 months. It was postulated that the metabolically active invading fibroblasts produced sufficient collagenase to affect the equilibrium between scar production and breakdown to amino acids. A series of 119 patients with glottic insufficiency were subsequently treated with bovine collagen injections, nearly half of whom had scarred larynges[5]. The voice results of collagen injection for patients with glottic insufficiency due to scarring were less dramatic than in groups without scarring, but overall these patients showed objective and subjective voice improvement without serious complications. These findings were perhaps more significant because of the selected patient population where other treatment options were limited, and collagen injection addressed both the glottic insufficiency and the physical properties of the damaged vocal fold tissues.

In a study limited to scarred larynges[6], three groups of patients were treated with collagen injection: (1) patients who had undergone cordectomy for cancer, (2) a group of failed Teflon injections with fibrosis status-post removal of Teflon granulomas, and (3) iatrogenic injuries secondary to aggressive vocal fold stripping procedures. The functional voice results were proportional to the amount of residual normal tissue; patients with stripped vocal folds did best and those who had extensive tissue loss withonly a residual scar tissue band after cordectomy had the poorest results. Multiple injections were employed in several patients. Good results were also associated with preservation of at least one intact vocal fold. It appeared that voice therapy played an important adjunctive role in improved results but there may have been other metabolic and functional factors that occurred naturally over time that might have favorably affected outcome.

We conclude that injectable collagen plays a role in the management of dysphonia secondary to vocal fold scarring. Although there are other options for filler substances to displace the vocal fold and correct glottic insufficiency, collagen injection appears to promote softening of existing scar tissue and should be considered as a primary or adjunctive measure in the management of scarred vocal folds. Bovine cross-linked collagen (Phonagel or Zyplast, Collagen Corporation, Palo Alto, Ca.) remains restricted due to FDA concern about possible immune reactions, but we are

currently investigating the use of autologous collagen as a safe substitute injectable bioimplant. It is important to recognize components of the pathology in every scarred larynx case and design treatments accordingly. Major musculoskeletal strictures might require open surgical procedures and framework surgery can play a role in cases of severe glottic insufficiency. That component of the dysphonia caused by stiffness of the vocal fold proper can be addressed by vocal fold collagen injection.

References

1. Knapp TR, Kaplan EN, Daniels JR. Injectable collagen for soft tissue augmentation. Plast Reconstr Surg 60:398-405, 1977.

2. Stegman SJ, Tromovitch TA. Implantation of collagen for depressed scars. J Dermatol Surg Oncol 6:450-453, 1980.

3. Ford CN, Martin DW, Warner TF. Injectable collagen in laryngeal rehabilitation. Laryngoscope 94:513-518, 1984.

4. Ford CN. Histologic study of injectable collagen in the canine larynx. Laryngoscope 96:1248-1257, 1986.

5. 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 101:237-247, 1992.

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