Hunter

NCVS

National Center for Voice and Speech

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

Status and Progress Report

Volume 11/May 1997

NCVS Status and Progress Report
Volume 11/May 1997

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

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

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

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

The University of Illinois
Department of Speech and Hearing Science

NCVS Personnel

Administration

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

Area Coordinators
Research - Ingo Titze
Training - Patricia Zebrowski
Continuing Education - Julie Ostrem
Information Dissemination - Cynthia Kintigh

Advisory Board

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

Investigators, Affiliates and Support Staff

Fariborz Alipour, Ph.D.
Kristin Baker, Ph.D.
Bridget Berning, M.S.
David Berry, Ph.D.
Florence Blager, Ph.D.
Diane Bless, Ph.D.
James Brandenburg, M.D.
Myrna Burt
John Butler, M.D.
John Canady, M.D.
Geron Coale, M.A.
Stefanie Countryman, M.A.
Linda D'Antonio, Ph.D.
Charles Davis, Ph.D.
Wendy Edwards, B.A.
Jeffrey Fields, B.M.
John Folkins, Ph.D.
Charles Ford, M.D.
Curt Freed, M.D.
Amy Furness
Steven Gray, M.D.
Elizabeth Hammond, M.D.
Marilyn Hetzel, Ph.D.
Margaret Hoehn, M.D.
Henry Hoffman, M.D.
Richard Hurtig, Ph.D.
Bruce Jafek, M.D.
Joel Kahane, Ph.D.
Michael Karnell, Ph.D.
Judith King, Ph.D.
Cynthia Kintigh, M.A.
David Kuehn, Ph.D.
Melda Kündük, M.S.
Jennifer Lehnherr
Julie Lemke, A.A.
Russel Long, M.S.
Erich Luschei, Ph.D.
Kathryn Maes, Ph.D.
Martin Milder, B.S.
Paul Milenkovic, Ph.D.
Jerald Moon, Ph.D.
Lea Ann Moriarty, B.A.
Carrie Nachtwey, B.S.
John Nichols, B.A.
Chris O'Brien, M.D.
Lorraine Olson Ramig, Ph.D.
Julie Ostrem, B.S.
Namrata Patil, M.D.
Annette Pawlas, M.A.
Kathe Perez, M.A.
Donald Robin, Ph.D.
Robin Samlan, M.S.
Richard Schmidt, Ph.D.
Suzanne Segal, Ph.D.
Elaine Smith, Ph.D.
Marshall Smith, M.D.
Jennifer Spielman, M.A.
Elaine Stathopoulos, Ph.D.
Brad Story, Ph.D.
Linda Suckow, B.A.
Edie Swift, M.S.
James Teslow, Ph.D.
Susan Thibeault, M.S.
Laetitia Thompson, Ph.D.
Sue Thompson, Ph.D.
Ingo Titze, Ph.D.
Lou Tomes, Ph.D.
Jane VanDeWiel
Katherine Verdolini, Ph.D.
Jennifer Waldron, B.S.
Patricia Ward
Barbara Williams
Darrell Wong, Ph.D.
Raymond Wood, M.D.
George Woodworth, Ph.D.
Patricia Zebrowski, Ph.D.
Jane Zernicke, B.S.
Lynn Zimba, Ph.D.

Doctoral Students

Todd Brennan, M.S.
Roger Chan, B.S.
Eileen Finnegan, M.A.
Elisa Mordue, M.S.
John Nelson, M. Aud.
Phyllis Palmer, M.A.
Annie Ramos, M.S.
Nelson Roy, M.S.
Helen Sharp, M.S.

Postdoctoral Fellows

Michael Edgerton, D.M.A.
Katsuhide Inagi, M.D.
Aliaa Khidr, Ph.D.

Visiting Scholars

Hanspeter Herzel, Ph.D., Germany
Matthias Frölich, Dipl. Physiks., Germany
Dirk Michaelis, Dipl. Physiks, Germany

Contents

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

Part II. Tutorial reports

Forward

This issue of our NCVS Status and Progress Report emphasizes the clinical arm of our Center. In previous volumes, we have usually begun with the basic science papers and gradually proceeded toward the more applied papers. This is not necessarily our way of thinking. The clinic drives the laboratory, and the laboratory drives the clinic. The studies on laryngeal massage and other intervention methods (i.e., the Lee Silverman technique) should serve as a springboard to drive deeper into the underlying mechanisms of phonatory, respiratory, and articulatory function.

Ingo R. Titze, Director
May, 1997

Part I

Research papers submitted for peer review in archival journals

NCVS Status and Progress Report - 11
May 1997, 1-8

Psychological Correlates of Functional Dysphonia: An Investigation Using the Minnesota Multiphasic Personality Inventory

Nelson Roy, M.Cl.Sc.
Department of Otolaryngology-Head and Neck Surgery, The University of Wisconsin-Madison
Department of Communicative Disorders, The University of Wisconsin-Madison
Jay J. McGrory, Ph.D.
Psychologist, Private Practice, Windsor, Ontario, Canada
Stephen M. Tasko, M.H.Sc.
Department of Communicative Disorders, The University of Wisconsin-Madison
Waisman Center, The University of Wisconsin-Madison
Diane M. Bless, Ph.D.
Department of Otolaryngology-Head and Neck Surgery, The University of Wisconsin-Madison
Department of Communicative Disorders, The University of Wisconsin-Madison
Dennis Heisey, Ph.D.
Department of Surgery, The University of Wisconsin-Madison
Charles N. Ford, M.D.
Division of Otolaryngology-Head and Neck Surgery, The University of Wisconsin-Madison

Abstract

Abnormal psychological factors have been implicated in the development of functional dysphonia (FD). This investigation describes the personality/psychological characteristics of twenty-five female patients who had received the diagnosis of FD. All subjects experienced symptom resolution following voice therapy. While vocally asymptomatic, these remitted FD subjects completed the Minnesota Multiphasic Personality Inventory (MMPI), an objective personality questionnaire. When compared to a medical out-patient control group, the results showed that FD subjects scored significantly higher on 7 of 10 clinical scales, therefore suggesting an elevated degree of emotional maladjustment. A stepwise logistic discriminant analysis identified two clinical scales which provided valuable discriminatory power between the two groups. Scale one (Hs-Hypochondriasis) which measures the number and type of reported somatic complaints, and scale seven (Pt-Psychaesthenia), a measure of diffuse anxiety, discriminated the groups with 88% sensitivity and 89% specificity. The results suggested that in spite of symptom improvement after voice therapy, the FD subjects continued to exhibit poor levels of adaptive functioning, which may represent trait-like vulnerability. The clinical implications of these results for voice practitioners are discussed.

Introduction

The human voice is regarded by many as the "valve of emotion" and a "window to personality", implying that it is a conduit for the expression of emotions and temperament. Consequently, when voice becomes disturbed, abnormal psychological processes are frequently offered as potential causal mechanisms, as is often the case in functional dysphonia (FD), where no visible structural or neurological pathology exists to explain the voice disturbance (FD) (1,2). FD, which may account for more than 10% of cases referred to multidisciplinary voice clinics (3), occurs predominantly in women, and is frequently transient. It is often preceded by upper respiratory infection symptoms and varies in its response to treatment (6-18). Although the voice literature is replete with speculations regarding the personality-psychological characteristics as-

sociated with FD, the precise nature of the voice-psychology relationship remains unspecified (19). This investigation aims to describe the psychological characteristics of a group of patients with FD.

Defining Functional Dysphonia

Controversy surrounds the term "functional dysphonia", and some clinicians object to the label because of its etiologic and symptomatologic ambiguity (1,8,16). "*Functional* " implies a disturbance of physiological function rather than anatomical structure. *Functional* is usually contrasted with *organic* and often carries the added meaning of psychogenic. Stress and psychological conflict are frequently presumed to cause or exacerbate functional symptoms (20-22). Confusion exists because FD is often used as the general descriptive term for many disorders, and is sometimes broadly synonymous with "hysterical", "psychogenic", "conversion","psychosomatic", "muscle misuse", or "tension" dysphonia. Certain authors prefer the labels "psychogenic" or "conversion" voice disorder to describe the presumed psychological origins of the dysphonia (6,8). Theorists differ, however, concerning the relative contribution of psychological factors to the formation of functional voice disorders (23). There is currently no clear evidence of whether psychological processes should be considered causal, correlational or consequential.

Functional *dysphonia* and *aphonia* are sometimes regarded as disorders represented on a continuum of severity, and are believed to share a common etiology (12). In aphonia, patients lose their voice completely and articulate in a whispered breath stream, whereas dysphonia suggests phonation is preserved, but disturbed in quality, pitch and/or loudness (24). Most studies investigating personality and/or psychological processes group both disorders under the designation "psychogenic voice disorder", reflecting the etiological supposition. Some authors warn that distinctions must be made between aphonia and dysphonia to prevent overestimation of the role of psychological factors in "dysphonia" (25).

Psychogenic Mechanisms in FD

Most clinicians acknowledge that factors contributing to a dysphonic voice are often a complex blend of organic, psychological and social elements, any one of which may be a predisposing, precipitating, or perpetuating agent (8,26). Rammage et al. (27) described several psychopathological processes that might be active in symptom formation. One such mechanism was "conversion reaction". In conversion, the voice loss is believed to represent a symbolic somatization of psychodynamic conflict. In short, patients convert psychic distress into a somatic symptom. In such disorders, the dysphonia is typically described in relation to primary and/or secondary gain. (8,28-30).

In addition to conversion, other psychological processes have been proposed to explain FD, including the combined interaction of organic and psychogenic mechanisms. One example of this interaction is the "specificity hypothesis" offered by Alexander (31,32). This theory suggests that a specific stimulus (emotional conflict) elicits a distinctive response, or illness, and the organ affected (larynx) is determined by a genetic weakness or vulnerability. Milutinovic (9) recognized the extensive etiologic overlapping of organic and functional voice disturbances and suggested that "genetic factors, the state of the endocrine and neurovegetative systems, and psychological factors are significant in the development of functional dysphonia." (p.179). He felt that psychogenic aphonia and dysphonia should be considered as "*phononeuroses*." Since over half of his "phononeurotic" patients had documented infection of the upper respiratory airways preceding the voice disturbance, he concluded that a direct connection existed between the pathological state of the mucosa and the development of FD. Milutinovic speculated that organic changes in the larynx, pharynx and nose facilitate the appearance of a functional voice problem; that is, they direct the somatization of the psychodynamic conflict.

Schalen and Andersson (3) also emphasized the interaction between psychological and physiological influences. Based on their study of subjects with "psychogenic dysphonia and aphonia", they concluded that the high number of reported allergy/asthma symptoms (37.5%) justified a more detailed examination of the interrelationship between psychological factors and respiratory and phonatory disorders.

In addition to acknowledging the conversion explanation for FD, Nichol and his colleagues (26) suggested that "tensional symptoms arise from the overactivity of autonomic and voluntary nervous systems in individuals who are unduly aroused and anxious (p.644)." He added that such overactivity leads to hypertonicity of the intrinsic and extrinsic laryngeal muscles, resulting in muscle tension dysphonias sometimes associated with adjustment and anxiety disorders, or with certain personality trait disturbances. This generalized laryngeal hypertonicity is also a recurrent theme in the writings of Aronson and others (8,24,27-30).

Finally, most authors have viewed psychological factors as strongly influential in the development of FD, and have virtually ignored the possiblility that such processes could be the consequence of coping with an incapacitating voice disorder. Depression, anxiety, and tension are frequent psychological concomitants of chronic illness (33-35). The notion that such sequelae could be considered outcomes of a severe voice disturbance, rather than causal agents, has received little attention.

Table 1.
Literature Review:
Description of Subjects, Test Instrument(s), and Major Findings of Investigations of the Voice-Psychology Relationship.

Authors	Subjects	Test Instrument(s)	Major Findings and Interpretations
Aronson et al. (12)	psychogenic aphonia/dysphonia N= 24 F; 3 M 1 "mute" 11 "continuous whisper" 7 "intermittent whisper-phonation" 8 "continuous phonation"	Minnesota Multiphasic Personality Inventory (MMPI) clinical psychiatric interview	A period of acute/chronic stress antedated the onset of dysphonia in 74% of the patients. 93% of patients were judged to have difficulty dealing with anger. 26% of patients reported excessive somatic complaints. No patient was in acute psychiatric distress. A clinical impression of hysteria was observed in less than half of patients and 30% exhibited a conversion "V" profile. The authors suggest that the entire group had a "hysterical flavor".
Pfau (36)	psychogenic aphonia/dysphonia N= 46 F; 8 M	German equivalent of MMPI	Results suggested neurosis in 35% of female patients, 20% of whom were considered a hysterical reaction type. The majority of individual profiles were either uninterpretable (37%), or considered within normal limits (28%).
Kinzl et al. (10)	"hyperfunctional" and "hypofunctional" aphonia N= 22 F	psychiatric evaluations social support network assessments life event inventories	Patients did not have particular personality traits in common, nor exposed to comparable conflict situations. "Personality structures and psychopathological symptoms ranged from mild impairment to severe neurosis (p.134)." Hysterical personality traits were frequent, but not always present. 75% of patients have other psychosomatic functional disturbances in their histories. Authors suggest that aphonia is a homogeneous clinical syndrome with heterogeneous personality structures and psychopathologies underlying its development.
Gerritsma (6)	psychogenic dysphonia/aphonia N= 75 F; 7 M	Wilde's Amsterdam Biographical Questionaire (ABV) Social Anxiety Scale (SAS) Wolpe-Lazarus Assertiveness Scale (WLAS)	42% of patients scored high on neuroticism (N), and neurotic somatization (NS) scales, a pattern consistent with conversion symptoms; and, 40% scored low on the extraversion (E) scale, suggesting a tendency toward introversion. 4% met DSM (III) criteria for hysterical personality. 65% of subjects were socially anxious, nonassertive or both. Author suggests dropping the term hysterical dysphonia in favor of one of "conversion, psychogenic or functional" aphonia.
Friedl et al. (7)	functional dysphonia/aphonia (20) organic dysphonia (14) normal control (20)	multiple measures of personality and anxiety	FD patients show a tendency toward restraint, and in stressful situations, the result is an intensified anxiety state. Life events may influence pathogenesis of FD.
Friedl et al. (25)	functional dysphonia/aphonia	"empirical-psychological procedure"	Psychological conditions were major etiologic factors in aphonia, but only partially relevant in patients with functional dysphonia. Authors warn that distinctions must be made between aphonia and dysphonia to prevent overestimation of the role of psychological factors in "dysphonia".
House & Andrews (37)	functional dysphonia/aphonia dysphonia (65) aphonia (4) spastic (2)	Present State Examination (PSE) Bedford College Life Events and Difficulties Interview	Authors failed to find an association between voice type and PSE score or psychiatric diagnosis. "The majority of patients were remarkable for the apparent normality of their premorbid psychological and social functioning. Major mental illness was infrequently diagnosed and minor states of tension and anxiety predominated (33%) (p.488)." FD is not usually found in markedly abnormal personalities and previous episodes of conversion disorder are rare.

Standardized Assessment of Psychological Processes

Although the previous review of potential psychological mechanisms represents engaging speculation, empirical evidence to support these explanations has seldom been provided. Only a handful of studies exist that have used standardized instruments to assess the voice-psychology relationship. Table 1 provides a review of the major findings and interpretations. Direct comparison of the results is restricted because of significant methodological differences. Some studies neglected to use control groups or normative data, while other researchers selected comparison groups with unmatched or unspecified qualities. Most investigators did not describe whether subjects were vocally asymptomatic at the time of testing. It is therefore difficult to judge whether these psychological attributes reflect long-term "trait-like" characteristics, or merely represent reaction to the voice disorder (i.e., "state" attributes). These problems are complicated by inclusion of heterogeneous

voice disorder types, and the failure of the investigators to distinguish aphonic from dysphonic subjects. This might partially explain the diverse results regarding the frequency and degree of hysterical personality traits (6,10,12), conversion reaction (6,10), and psychopathological symptoms (6,10,12,36,37). These obstacles make it exceedingly difficult to appraise the specific nature of the voice-psychology relationship. As Green (19) states, "until more adequate research is conducted, psychological variables must be considered possible etiological, consequential and therapeutic factors (p.34)."

The primary goal of the present study was to assess the psychological characteristics of patients who presented with *functional dysphonia.* In accordance with the recommendation of Freidl (25), this investigation was restricted to those subjects exhibiting dysphonia, not aphonia. By contrasting the personality profiles of "remitted" functional dysphonics with a medical outpatient control, we attempt to shed further light on the voice-psychology relationship.

Methods

Subjects

Two groups were selected to participate in this investigation. The voice disordered group consisted of twenty-five female patients (ages 18 to 69 years, mean 44.6 yrs, +/- 12.5) who attended a hospital-based speech pathology service. All of the subjects presented with a voice disturbance that ranged in duration from 4 days to 3 years, with a mean duration of 8.5 (+/- 11.6) months. The diagnostic label of functional dysphonia was offered following comprehensive perceptual, acoustic and videolaryngoscopic examinations by a speech pathologist and otolaryngologist. FD diagnostic inclusion criteria were as follows: (a) a voice disturbance in the absence of visible mucosal disease or structural pathology; (b) no neurological pathology (specifically, vocal fold paresis, paralysis or motor speech disturbance); (c) no previous laryngeal surgery; and (d) no coexisting upper respiratory infection symptoms at the time of examination. All of the subjects were treated by one of two speech pathologists (NR & ST). Each subject responded favorably to voice therapy techniques; that is, voice improved substantially or returned to normal based on the perceptual judgment of both the clinician and patient. The negative laryngeal findings, combined with the positive response to treatment, excluded other possible explanations, and confirmed the diagnosis of FD.[1]

The control group consisted of 19 age- and gender-matched non-voice disordered medical outpatients (mean age 45.8 yrs, +/- 14.6; range 28 to 74 yrs.) who were recruited by their family physician during a non-emergent visit. The "control" subjects received medical care for physical complaints unrelated to laryngeal function, and denied any previous history of voice related problems. Because control subjects sought help for medical concerns, it was felt that they would serve as an appropriate comparison group.

[1] The presence of dysphonia and its positive response to treatment was substantiated by auditory-perceptual evaluation by four judges. Randomized pre- and post-treatment audiorecordings of the middle sentence of the Rainbow Passage (38) were rated on a seven-point equal-appearing interval scale, where "1" indicated normal voice and "7" indicated a severe voice disorder. Results confirmed a significant reduction in mean severity ratings across all subjects. The interested reader should refer to Roy et al. (48) for a complete description of the rating procedure.

Procedures

Description of the MMPI

During a followup session (mean 3.96 mos. post treatment, +/- 5.91), each remitted FD subject completed the **Minnesota Multiphasic Personality Inventory (MMPI)** (39) in a quiet setting. The MMPI is the most widely used personality test in the United States (40). The relatively unambiguous stimuli and the structured response format qualify the MMPI for classification as an objective technique of personality assessment. It permits inferences regarding general level of adjustment and degree of psychopathology. The MMPI consists of 566 self-reference statements; subjects respond "true" or "false" to each statement depending on whether or not the subject views the statement as descriptive of her behavior.

Scoring of completed inventories was undertaken by a licensed psychologist (JJM). The scoring procedure produced scores for three validity scales and ten basic clinical or personality scales. The raw scores from the standard validity and clinical scales were transformed to linear T-scores (mean 50; SD=10). The T-scores were then plotted to construct a profile. This profile served as the basis for generating inferences regarding the examinee. The validity scales L, F and K not only provided validation measures of the subject's test taking attitude (i.e., defensiveness, or responding in a socially desirable manner), but also furnished useful clinical information.

The clinical scale section of the MMPI profile was composed of ten scales, each with a number, abbreviation, and formal name. The ten clinical scales yielded information regarding problem areas for the examinee. Elevation of scores across each of the clinical scales (1 through 10) was one indicator of adjustment. In general, as more of the clinical scales were elevated (and as the degree of elevation increased), the greater was the probability that some psychopathology and poor level of adaptive functioning was present (41). Elevations were considered to be in the moderate range when T= 60 through 70, and to be in the marked range when T= 70 or above. This division into categories is a generally-accepted convention and should not be taken as absolute.

Results

All test patterns/profiles were within admissable validity limits. Only 32% of FD subjects (8 of 25) displayed "normal" MMPI profiles, as evidenced by no clinical scale

score exceeding T= 70. This is compared to 90% of the medical control patients (17 of 19) whose profiles fell within the normal range using the same criterion.

The means and standard deviations for the validity and clinical scale variables appear in Table 2. Asterisks indicate significant differences between groups based on t-test comparisons of mean T-scores. P-values for each of the comparisons are also provided. Inspection of these data revealed that the functional dysphonia group did not differ from the medical controls on any of the three validity scales. However, FD subjects differed significantly from the controls on 7 of the 10 clinical scales, specifically scales 1 Hs- Hypochondriasis, 2 D- Depression, 3 Hy- Hysteria (which constitute the "neurotic triad"), 6 Pa-Paranoia, 7 Pt- Psychasthenia, 8 Sc- Schizophrenia, and 0 Si- Social Introversion.

Figure 1 illustrates composite MMPI profiles generated from the mean T-score values for the FD and medical control group. The cursory scale descriptions that follow were derived from several sources and serve only to familiarize the reader (41-47). *Scale 1 Hs* measures the number of bodily complaints claimed by the examinee. Individuals who score high on this scale are characterized by denial of good health and the admission of a variety of vague somatic symptoms. This scale reflects "dispositional" attributes, suggesting that elevations tend to reflect long-standing behavior. *Scale 2 D* is a mood scale which provides an index of the examinee's present degree of pessimism, dissatisfaction, and sadness. This scale tends to be sensitive to current states of mood and mood changes, and as such, represents a "state" index. Elevated scorers report feeling depressed, blue, unhappy or dysphoric. *Scale 3 Hy* measures tendencies toward denial. One method of avoiding facing difficulty and conflict is to deny such situations exist. This scale measures the amount and type of such denial. It is considered a character scale. High scorers acknowledge many physical problems, but deny that they are worried about them. Elevations may be suggestive of persons who when under stress, avoid responsibility by developing physical symptoms. Persons may be symptom-free most of the time, but under stress the symptoms appear suddenly. These symptoms are likely to disappear just as suddenly when the stress subsides. *Scale 6 Pa* suggests suspiciousness, interpersonal sensitivity, and a rigid adherence to ideas and attitudes. *Scale 7 Pt* is a measure of diffuse anxiety, usually of a long-term nature; thus, general characterological anxiety. Individuals with elevations have a low threshold for anxiety and tend to be worried, indecisive, tense, and unable to concentrate. When scales 1-2-3 are combined with Scale 7, individuals will show increased anxiety and fearfulness in addition to the physical ailments suggested by the triad (1-2-3) profile. *Scale 8 Sc* measures mental confusion: the higher the elevation, the more confused the person is. Elements of this scale deal with social alienation, isolation, bizarre feelings, influence of external agents, peculiar bodily dysfunctions and general dissatisfaction. *Scale 0 Si* provides an index of a person's preference for being alone (high 0) or being with others (low 0). High scorers tend to be withdrawn, socially insecure, and anxious when in contact with people. All of these scales, with the exception of Scale 2 D are viewed as assessments of character, not mood. All but scale 2D represent relatively enduring personality structures (41-47).

Table 2.
Means, Standard Deviations, and Significance Levels for MMPI Validity and Clinical Scales for Functional Dysphonia and Medical Control Groups.

VALIDITY AND CLINICAL SCALES	FUNCTIONAL DYSPHONIA		MEDICAL CONTROL		p-value
	Mean	SD	Mean	SD	
L Scale	56.96	11.32	56.52	8.62	p=.8901
F Scale	59.88	13.21	53.42	11.39	p=.0959
K Scale	49.24	10.26	54.15	9.91	p=.1175
1-Hs: Hypochondriasis	65.84	10.11	51.47	8.68	*p<.0001
2-D: Depression	62.40	10.43	50.74	6.70	*p<.0001
3-Hy: Hysteria	63.88	13.03	52.42	11.39	*p<.0040
4-Pd: Psychopathic Dev.	56.44	9.63	52.32	8.29	p=.1431
5-Mf: Masc. Femininity	50.12	9.75	53.58	11.25	p=.2817
6-Pa: Paranoia	57.08	10.15	51.12	8.09	*p<.0430
7-Pt: Psychasthenia	61.24	7.39	47.16	9.09	*p<.0001
8-Sc: Schizophrenia	61.68	8.49	48.37	10.90	*p<.0001
9-Ma: Mania	53.40	10.71	51.89	10.80	p=.6479
0-Si: Social Introversion	57.84	6.32	53.05	5.05	*p<.0098

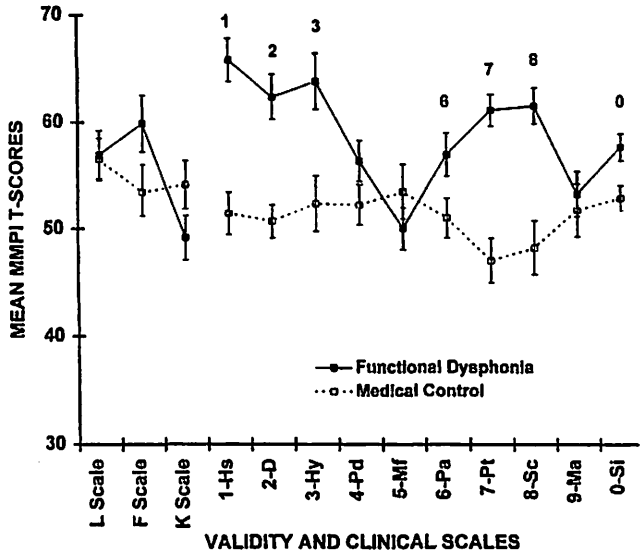


Figure 1. Composite MMPI profiles for functional dysphonia and medical control groups using mean T-score values (+/- standard error). Numbers above select scales indicate significant differences between the groups, based on t-test comparisons of mean T-scores.

A stepwise logistic discriminant analysis was performed to identify clinical scale scores that best distinguished the voice-disordered group from the controls. The clinical score variables were entered in a stepwise fashion, in the order of their discriminatory ability. Variable entry stopped when the remaining variables contributed no further significant discriminatory information. All variables remaining in the model were significant at the .05 critical level. Two scales sufficiently discriminated the FD subjects from the controls: Clinical Scale 1 (Hs), a general index of frequency of somatic complaints; and Scale 7 (Pt), a measure of diffuse anxiety. Figure 2 graphically illustrates a scatterplot of individual FD and medical control subjects using the two clinical scales. Subjects' T-scores are plotted for Scale 1 (Hs), which is represented on the y-axis, and Scale 7 (Pt), which is represented on the x-axis. This scatterplot of the distribution of scores for the FD subjects versus the controls clearly illustrates the discriminatory value of the two clinical scale variables.

For each subject, a predicted value was obtained from the estimated logistic regression equation. This predicted value is the probability that the subject is a control. If a probability of .4 or greater was used as the cutoff to predict a control, the estimated sensitivity was 88%, and the estimated specificity was 89%. That is, a functional dysphonia subject would be correctly identified as such about 88% of the time, and a control subject would be correctly identified as such about 89% of the time. These rates are remarkably high.

Discussion

Functional dysphonia is a disorder characterized by heterogeneous psychological and personality functioning. As a group, FD subjects exhibited significantly more

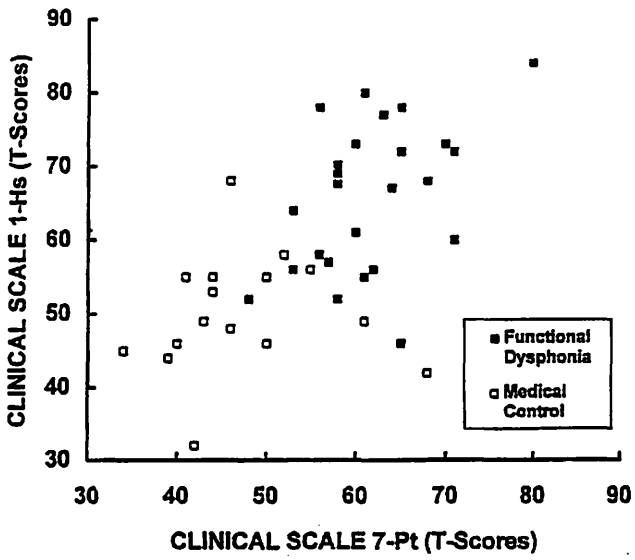


Figure 2. Scatterplot of T-scores obtained from clinical scale 1-Hs (y-axis) and clinical scale 7-Pt (x-axis) for individual FD and medical control subjects.

emotional adjustment problems and poorer levels of adaptive functioning than a medical outpatient control group. This array of problems included multiple somatic complaints, diffuse anxiety, and dysphoria.

Despite significant differences between FD subjects and the medical controls in seven of the clinical scales, it is also important to recognize that almost one-third of the FD subjects displayed profiles considered within the normal range. Inventories in most cases were administered several weeks following successful treatment, and therefore, one might argue that these results may actually underestimate levels of emotional maladjustment, especially in this "normal" subgroup. Because we did not administer the MMPI both before and after treatment, it is impossible to directly assess changes in psychological functioning associated with voice recovery. These "normal" subjects may represent an etiologically distinct subgroup, underscoring the heterogeneity of patients with FD. Clearly, more precise longitudinal methodologies are required to identify changes in psychological functioning following successful voice management.

These results are in general agreement with previous studies identifying elevated levels of anxiety, somatic complaints and introversion in this patient population (6,7,10, 12, 37). Although these data do not directly answer the question of whether psychological factors should be considered causal or a consequence of the disordered voice, several inferences can be made. Recognizing that all of the inventories were collected while the subjects were vocally asymptomatic, and assuming that treatment did not create emotional problems, it appears that many remitted FD patients continue to display emotional adjustment difficulties after their voice symptoms have abated. Thus, the poor levels of adaptive functioning, as demonstrated by elevated MMPI scores, can be interpreted as either representing (1) residual psychological effects; that is, the "emotional scar" of the voice problem, or (2) persistent personality characteristics. Because most of the elevated scores occurred on clinical scales with trait-like stability, the "emotional scar" argument seems less defensible. These findings suggest that a majority of FD patients display certain psychological/personality dispositions, including bodily preoccupation and anxiety, that are relatively enduring and seem to represent trait-like qualities. Interestingly, this may constitute a persistent vulnerability (diathesis) for the development of tensional or somatic symptoms when under conditions of psychological distress.

In the absence of a psychiatric or psychosocial interview, the question of whether functional dysphonia can be explained by a single mechanism, such as conversion reaction, is impossible to answer directly with this research methodology. Some qualified evidence from individual MMPI profiles however, does not appear to support conversion as the sole explanation. Specifically, elevations on Clinical Scales 1 and 3 of the MMPI, with an intervening valley of 10 or more T points at Scale 2, form what has come

to be called the "conversion V" (41-47). The general meaning of the 1-3 or 3-1 pattern is that these individuals convert psychological stress and difficulties into physical complaints. It is felt that the association of bland emotional unconcern/ affect (low Scale 2) with numerous physical complaints (contributing to elevation of Scale 1) and denial of anxieties and fears (raising Scale 3) is characteristic of conversion reactions. Although six FD subjects approximated the distinctive "V" profile, only three FD subjects met the above accepted standard for the "conversion V" configuration.

Additional evidence against a prototypical conversion profile is found in the presence of other clinical scale elevations. In particular, high scores on Scale 2 combined with an elevated Scale 7 suggest a pattern contrary to the celebrated "la belle indifference" sometimes felt to be characteristic of "hysterical" conversion patients. Rather than displaying bland indifference, these patients endorse many items consistent with dysphoria and cognitive distress. This is further complicated by a general tendency toward introversion (Scale 0 elevation), which is in direct contrast to the extraverted, histrionic style of the hysterical personality. Thus, if conversion is the primary mechanism underlying functional dysphonia, which has been suggested by several authors, then the overwhelming majority of these subjects seemingly do not present with the prototypical "conversion profile" as measured by this instrument.

The finding of only 12% of patients with the classic "conversion V" profile is in contrast to previous studies by Aronson (12) and Pfau (36), who identified this pattern in 30% and 20% of their patients, respectively. This may reflect differences related to subject selection. Their inclusion of aphonic subjects (and their relative proportions) might partly account for the discrepancy. Unfortunately, any further comparison between this research and the Aronson et al. study is impeded by the failure of those researchers to provide completed individual or group MMPI data.

In addition to conversion, other mechanisms need to be considered in order to explain functional dysphonia. Within the diagnostic category of functional dysphonia, it is reasonable to speculate that several subtypes exist with single or multiple etiologic factors. Other psychological mechanisms, including the interaction between trigger factors, voice use, anxiety and tensional states, require further exploration.

Clinical Implications

The discriminatory value of Scales 1 and 7 merits further discussion. These two scales may be potentially useful as a clinical tool to distinguish FD patients from other voice disorder types, for example, spasmodic dysphonia. The development of an abbreviated, but reliable screening device could raise the index of suspicion of one disorder type over another. At the very least, information regarding psychological functioning could assist in the assessment and treatment process.

From these results, it is clear that despite successful voice management, a substantial proportion of patients with FD continue to display emotional adjustment problems. Voice therapy may represent the first step in rehabilitation; however, relapse could be a serious problem in the absence of adjunctive counseling (48). The dysphonia in some patients might be a symptom of underlying psychological distress and tension. Failure to recognize such fundamental problems may limit long-term success in the treatment of functional dysphonia.

Acknowledgements

Grateful acknowledgement is made to Drs. David Crookston and Frank Ianni who generously recruited subjects to serve as medical controls. The authors express gratitude to the Vocal Function Laboratory staff (Waisman Center), especially Amy Mast for her valuable contributions at various stages during this project. Parts of this manuscript were presented at the American Speech and Hearing Association Convention, New Orleans, 1994.

This work was supported by the National Center for Voice and Speech through Grant P60 00976 for the National Institute on Deafness and Other Communication Disorders.

References

1. Morrison MD, Rammage L. Muscle misuse voice disorders: description and classification. Acta Otolaryngol (Stockh). 1993;113:428-434.

2. Colton RH, Casper JK, Hirano M. Understanding voice problems: A physiological perspective for diagnosis and treatment. Baltimore: Williams & Wilkins, 1990.

3. Schalen L, Andersson K. Differential diagnosis and treatment of psychogenic voice disorder. Clin Otolaryngol. 1992;17(3):225-230.

4. Bauer H. Die Bedeutung der Anamnese fur die Therapie von funktionellen Stimmstorungen. Sprache Stimme Gehor. 1980;4:93-99.

5. Schulte-Holthausen G. Das klinische Bild der psychogenen Dysphonie im Wandel der Zeit. Sprache Stimme Gehor. 1981;5:108-110.

6. Gerritsma, EJ. An investigation into some personality characteristics of patients with psychogenic aphonia and dysphonia. Folia Phoniatr. 1991;43:13-20.

7. Friedl W, Friedrich G, Egger J. Personality and coping with stress in patients suffering from functional dysphonia. Folia Phoniatr. 1990;42:144-149.

8. Aronson AE. Clinical Voice Disorders: An Interdisciplinary Approach. 3rd ed. New York:Thieme-Stratton; 1990.

9. Milutinovic Z. Inflammatory changes as a risk factor in the development of phononeurosis. Folia Phoniatr. 1991;43:177-180.

10. Kinzl J, Biebl W, Rauchegger H. Functional aphonia: Psychosomatic aspects of diagnosis and therapy. Folia Phoniatr. 1988;40:131-137.

11. Koufman JA, Blalock PD. Classification and approach to patients with functional voice disorders. Ann Otol Rhinol Laryngol. 1982;91:372-377.

12. Aronson AE, Peterson HW, Litin EM. Psychiatric symptomatology in functional dysphonia and aphonia. J Speech Hear Dis. 1966;31:115-127.

13. Fex F, Fex S, Shiromoto O, Hirano M. Acoustic analysis of functional dysphonia: Before and after voice therapy (Accent Method). J Voice. 1994;8:163-167.

14. Roy N, Leeper HA. Effects of the manual laryngeal musculoskeletal tension reduction technique as a treatment for functional voice disorders: Perceptual and acoustic measures. J Voice 1993; 7:242-249.

15. Bridger MM, Epstein R. Functional voice disorders: A review of 109 patients. J Laryngol Otol. 1989;97:1145-1148.

16. Pahn J, Friemaert K. Differential diagnostische und terminologische erwagungen bei sogenannten; funktionellen storungen im neuropsychiatrischen und phoniatrischen. Fachgebiet. 2. Phoniatrischer Aspeckt. Folia Phoniatr. 1988;40, 168-174.

17. Mans EJ. Psychotherapeutic treatment of patients with functional voice disorders. Folia Phoniatr. 1994;46:1-8.

18. Butcher P, Elias A, Raven R, Yeatman J, Littlejohns D. Psychogenic voice disorder unresponsive to speech therapy: psychological characteristics and
cognitive-behaviour therapy. British J Dis Comm. 1987;22:81-92.

19. Green G. The inter-relationship between vocal and psychological characteristics: A literature review. Austral J Hum Comm Dis. 1988;16:31-43.

20. Morrison MD, Nichol H, Rammage LA. Diagnostic criteria in functional dysphonia. Laryngoscope. 1986;94:1-8.

21. Bass CM, ed. Somatization: Physical Symptoms & Psychological Illness. London: Blackwell Scientific Publications; 1990.

22. Kirmayer LJ, Robbins JM, eds. Current concepts of Somatization: Research and Clinical Perspectives. Washington: American Psychiatric Press Inc; 1991.

23. Cooper M. ModernTechniques of Vocal Rehabilitation. Springfield: Charles C. Thomas; 1973.

24. Boone DR, McFarlane S. The Voice and Voice Therapy. 4th ed. Englewood Cliffs, NJ: Prentice-Hall; 1988.

25. Friedl W, Friedrich G, Egger J, Fitzek I. Psychogenic aspects of functional dysphonia. Folia Phoniatr. 1993;45:10-13.

26. Nichol H, Morrison MD, Rammage LA. Interdisciplinary approach to functional voice disorders: The psychiatrist's role. Otolaryngol Head Neck Surg. 1993;108: 643-647.

27. Rammage LA, Nichol H, Morrison MD. The psychopathology of voice disorders, Hum Comm Can. 1987;11:21-25.

28. Greene MC, Mathieson L. The Voice and its Disorders. 5th ed. London: Whurr Publishers; 1989.

29. Stemple JC. Clinical Voice Pathology: Theory and Management. Columbus: Charles E. Merrill; 1984.

30. Stemple JC. Voice Therapy: Clinical Studies. St. Louis: Mosby Year Book; 1993.

31. Alexander F. Psychosomatic Medicine: Its Principles and Applications. New York: WW Norton and Company; 1950.

32. Alexander F, French TM, Pollock GH. Psychosomatic Specificity, Vol 1: Experimental Study and Results. Chicago: University of Chicago Press; 1968.

33. Dubovsky SL, Weissberg MP. Reactions to illness. In Clinical Psychiatry in Primary Care, 2nd edn. Baltimore: Williams & Wilkins; 1982.

34. Nemiah JC. Psychological complications of physical illness. In Foundations of Psychopathology. New York: Oxford University Press; 1961.

35. Reiser DE. Reactions to illness. In Reiser DE, Schroder AK. Patient Interviewing: The Human Dimension. Baltimore: Williams & Wilkins; 1985.

36. Pfau EM. Psychologische Untersuchungsergegnisse zur atiologie der psychogenen Dysphonien. Folia Phoniatr. 1975;25:298-306.

37. House A, Andrews HB. The psychiatric and social characteristics of patients with functional dysphonia. J Psychosom Res,1987;3:483-490.

38. Fairbanks G. Voice and Articulation Drillbook. New York: Harper and Row, 1960.

39. Hathaway SR, McKinley JC. The Minnesota Multiphasic Personality Inventory. New York: Psychological Corporation; 1972.

40. Lubin B, Larsen RM, Matarazzo JD, & Seever M. (1985). Psychological test usage patterns in five professional settings. American Psychologist. 1985; 40: 857-861.

41. Graham JR. The MMPI: A Practical Guide. 2nd ed.New York: Oxford University Press; 1987.

42. Graham JR. MMPI-2: Assesssing Personality and Psychopathology. New York: Oxford University Press; 1990.

43. Newmark CS, ed. MMPI Clinical and Research Trends. New York: Praeger; 1979.

44. Butcher JN, Dahlstrom WG, Graham JR, Tellegen A, Kaemmer B. MMPI-2 (Minnesota Multiphasic Personality Inventory-2): Manual of administration and scoring. Minneapolis: University of Minnesota Press; 1989.

45. Duckworth JC, Anderson WP. MMPI and MMPI-2: Interpretation Manual for Counselors and Clinicians. Philadelphia:Taylor and Francis Group; 1995.

46. Greene RL. The MMPI: An Interpretive Manual. 2nd ed. New York: Grune & Stratton; 1989.

47. Marks PA, Seeman W, Haller DL. The Actuarial Use of the MMPI with Adolescents and Adults. Baltimore: Williams & Wilkins Company; 1974.

48. Roy N, Bless DM, Heisey D, Ford CF. Manual circumlaryngeal therapy for functional dysphonia: An evaluation of short- and long-term treatment outcomes. J Voice (in press).

NCVS Status and Progress Report - 11
May 1997, 9-19

Manual Circumlaryngeal Therapy for Functional Dysphonia: An Evaluation of Short- and Long-Term Treatment Outcomes

Nelson Roy, M.Cl.Sc.
Diane M. Bless, Ph.D.
Department of Otolaryngology- Head and Neck Surgery, The University of Wisconsin-Madison
Department of Communicative Disorders, The University of Wisconsin-Madison
Dennis Heisey, Ph.D.
Department of Surgery, The University of Wisconsin-Madison
Charles N. Ford, M.D.
Department of Otolaryngology- Head and Neck Surgery, The University of Wisconsin-Madison

Abstract

Manual circumlaryngeal therapy (manual laryngeal musculoskeletal tension reduction) was used to treat 25 consecutive functional dysphonia patients. Pre- and post-treatment audiorecordings of connected speech and sustained vowel samples were submitted to auditory-perceptual and acoustical analysis to assess the immediate and long-term effects of a single treatment session. To complement audiorecordings, subjects were interviewed in follow-up regarding the stability of treatment effects. Pre- and post-treatment comparisons demonstrated significant voice improvements. No significant differences were observed between post-treatment measures, suggesting that vocal gains were maintained. Interviews revealed 68% of subjects reported occasional partial recurrences, typically less than 4 days in duration which resolved spontaneously. These results replicate and extend previous research suggesting the utility of manual circumlaryngeal therapy for functional voice disorders.

Introduction

Voice problems in the absence of identifiable structural or neurological pathology are commonly labeled "functional dysphonia" or "functional aphonia". These functional voice problems may account for more than 10% of cases referred to multidisciplinary voice clinics (1). In spite of near-consensus on the use of voice therapy as the primary approach with functional voice disorders, little objective evidence exists to support its efficacy (2). Several clinical investigators have reported success using diverse treatment strategies. In general, most of these reports are anecdotal and rely on single clinician judgments, patient report, chart review, and/or other less controlled means of determining improvement (3-6). Subject inclusion criteria are often deficient, and treatment techniques are seldom described satisfactorily to permit replication. Furthermore, insufficient follow-up data have made it difficult to evaluate maintenance of reported changes. This scarcity of objective data, combined with myriad treatment techniques, results in uncertainty surrounding therapy technique selection for functional dysphonia (FD). Assessment of individual voice therapies using objective outcome measures is imperative for rational, effective, and efficient patient management. The present investigation was undertaken to assess the immediate and long-term effects of manual circumlaryngeal therapy (manual laryngeal musculoskeletal tension reduction) on FD.

For clarity, it is necessary to distinguish functional *dysphonia* from *aphonia*. In aphonia, patients lose the voice completely and articulate in a whispered breath stream, whereas dysphonia suggests phonation is preserved, but disturbed in quality, pitch and/or loudness (7). The term *functional* implies a disturbance of physiological function rather than anatomical structure. In clinical circles, *func-*

tional is frequently contrasted with *organic* and often carries the added meaning of psychogenic. Stress and psychological conflict are frequently assumed to cause and/or exacerbate functional symptoms.

Some clinicians object to the use of the term functional dysphonia because of its etiologic imprecision and symptomatologic ambiguity. Certain authors favour the labels "psychogenic" or "conversion" voice disorder to describe the *presumed* psychological origins of dysphonia(3,8). In conversion, the voice disorder is felt to represent a symbolic somatization of psychodynamic conflict. Other clinicians prefer labels such as "muscle tension or muscle misuse" dysphonia, "vocal hyperfunction" to emphasize the role of poorly regulated laryngeal muscle tension (7). All such labels collectively reflect etiological presuppositions. Clear thresholds or discrete boundaries to separate these various diagnostic categories are lacking. Although psychological factors have been widely implicated, the precise mechanism(s) underlying the development and maintenance of functional dysphonia has not been explicitly elucidated (9).

In general, it is felt that a disordered voice in the context of a structurally normal larynx is the product of a complex blend of psychological, social, and physiological factors (10). Elevated laryngeal musculoskeletal tension is repeatedly cited as contributing to the development and maintenance of this form of dysphonia. Some authorities speculate that this state of tension originates from increased environmental stress and/or from personality structures that tend to induce tension (3,6,7). Therefore, many voice therapy techniques including manual circumlaryngeal therapy, yawn-sigh, Accent Method, and progressive relaxation aim to reduce or eliminate excess muscle tension in the laryngeal region, postulating that it will precipitate a proportional improvement in overall vocal function. Aronson (3) suggested that all patients with voice disorders, regardless of etiology, should be assessed for excess laryngeal musculoskeletal tension, either as a primary or a secondary cause of the persisting dysphonia. He described and advocated manual circumlaryngeal therapy as the primary approach for patients with musculoskeletal tension disorders. Aronson claimed that non-manual tension reduction techniques share the same objective as his manual approach, but these methods often fail because of the stubborn nature of excess musculoskeletal tension. Unfortunately, few data exist to substantiate or reject Aronson's contentions.

Of the existing literature on treatment outcomes, several investigations with the FD population merit review. Bridger and Epstein (11) reviewed 109 cases and suggested that the results of treatment for functional voice disorders were disappointing. Fewer than half of their patients could be considered "cured" (i.e., patients' voices returned to their premorbid state) after 6 months of voice therapy. No attempts were made by the authors to objectify their judgments of success or to describe the treatment(s) employed. Furthermore, no follow-up data were provided to assess the stability of the improvements observed.

More recently Fex and co-workers (12) reported acoustic changes following treatment using the Accent Method of voice therapy with ten patients with mixed diagnoses (seven normal vocal folds, and three with moderate vocal nodules). Pitch and amplitude perturbation quotients, normalized noise energy for 1 to 4 kHz, and fundamental frequency showed significant improvement following ten therapy sessions. Improvement in voice quality was based on the nonquantified opinions of the patient and the perceptual evaluation of the speech pathologist who administered the treatments. No follow-up data were provided to assess the durability of the observed changes.

Koufman and Blalock (13) reported the results of treatment on a group of patients with "laryngeal tension-fatigue syndrome". This disorder is characterized by fluctuating hoarseness and vocal fatigue. The authors described impressive results following a single treatment session using progressive relaxation, breath support, and pitch elevation. While these investigators relied on perceptual measures to document improvement, they failed to control adequately for reliability of listener judgments.

Roy and Leeper (14) evaluated the short-term effects of the manual laryngeal muscle tension reduction approach (as described by Aronson) with 17 FD patients. This technique involves kneading the laryngeal musculature in specific locations while observing changes in voice quality. Using acoustic and perceptual measures to assess the immediate results of the procedure, the authors reported that 93% of the subjects achieved an approximation of normal voice following a single treatment session. However, they did not provide any long-term maintenance data and relied solely on one- week post-treatment telephone contacts to judge recurrence.

The extant literature reveals variable results surrounding the effectiveness of treatment for FD. Few data exist evaluating long-term treatment effects . Although Aronson strongly advocates the use of his manual technique as a primary approach for patients with FD, he also concedes that little is known regarding the long-term fate of such treatment (p.315). This investigation, therefore, evaluates the short- and long-term consequences of manual circumlaryngeal therapy as a treatment for FD. As such, it represents an attempt to replicate and extend the results of Roy and Leeper (14).

Table 1.
Characteristics of Subjects Submitted to Manual Circumlaryngeal Therapy

DESCRIPTOR	MEAN (+/- SD)
Age (yr)	40.9 (13.03)
History of previous voice problems	84% (7.33)
History of environmental stress	76% (8.54)
Reports pain when palpated	88% (6.50)
History of excessive voice use	32% (9.33)
Smoker	12% (6.50)
Allergy	36% (9.60)
URI Symptoms prior to onset	48% (9.99)
Time elapsed (mo) to first follow-up (post 2-) audio and interview	3.58 (5.50)

Methods

Subjects

Twenty-five consecutive patients (mean age 40.9 (+/- 13.0) yrs.) referred to a hospital based speech-language pathology service, participated in this study. All subjects presented with vocal symptoms that ranged in duration from 5 days to 4 years with a mean duration of 8.8 (+/- 12.2) months. The diagnosis of functional dysphonia was offered following videolaryngoscopic examination by an otolaryngologist and a speech-language pathologist. Diagnostic inclusion criteria were: (a) a voice disturbance in the absence of any visible mucosal disease or structural pathology, (b) no neurological pathology, specifically vocal fold paresis, paralysis, or motor speech disturbance, (c) no previous laryngeal surgery, and (d) no coexisting upper respiratory infection symptoms at the time of examination. All subjects were assessed (Table 1) and treated by two speech pathologists (NR and ST) using the same management protocol.

Procedures

Assessment and Treatment Protocol: The assessment protocol, developed according to guidelines provided by Aronson (3), was designed to assess the voice in a traditional manner, but also to determine the presence, nature, and degree of emotional factors and the extent to which they contributed to the development and maintenance of the voice disorder (Appendix A). After completing the case history, voice evaluation, and muscle tension assessment (i.e., palpation), clinicians undertook the treatment protocol (Appendix B) with each subject. Briefly, this treatment involved kneading the extralaryngeal musculature in an anterior-posterior direction at specific locations while exerting a downward pull on the larynx. Target voice stimuli were presented concurrently while noting changes in voice quality. The assessment and treatment protocol was completed in a single session that ranged in duration from 50 minutes to 3 hours.

All 25 subjects' voices were recorded in a quiet setting on a high-quality audio-cassette tape recorder (Marantz model PMD201) with a professional-grade windshielded dynamic microphone (Shure Model 10A) maintaining a constant mouth-to-microphone distance of 1 cm. "The Rainbow Passage" (15) and sustained productions of three vowels /a, i, u/ were recorded prior to ("pre-") and immediately following the assessment and therapy session ("post 1-"; same session), and in a follow-up appointment ("post 2"). Because patients frequently traveled considerable distances at significant personal expense for voice assessment and treatment, regularly scheduled follow-up visits at pre-determined intervals were difficult. Where possible, every effort was made to schedule return visits within a 3-month window of the initial session. The mean length of time from initial treatment to first follow-up contact ("post 2-") was 3.6 (+/-5.5) months. Sixteen subjects were available for a second follow-up evaluation ("post 3-"). Nine subjects (36%) were lost to long-term follow-up because of relocation or travel distance. The mean length of time from initial contact to final contact ("post 3-"; N=16) was 16.5 (+/- 11.4) months.

In addition to audiorecordings, subjects were interviewed on each occasion regarding the stability of their post-treatment voices. Specific issues raised during the interview regarding relapse included (1) whether the patient's present voice quality was representative of her typical "out of clinic" voice, (2) whether she had experienced partial or total recurrence of dysphonic symptoms, (3) whether this recurrence was more or less severe than the original voice disturbance, (4) what the duration and frequency of recurrences were (i.e., how long and how often), (5) whether recurrences were preceded or accompanied by upper respiratory infection symptoms or elevated life stress, (6) whether the subject attempted self-treatment, and (7) what patterns, if any, observed by the subject could explain fluctuations in voice performance. When subjects were lost to follow-up (i.e., post 3- interview results were unavailable), post 2- interview reports were included as part of the "final contact"

data. No other voice therapies (e.g., vocal hygiene, relaxation exercises, etc.) were provided following the single management session. Subjects were instructed, however, to contact the attending clinician should they experience partial or complete recurrence of dysphonic symptoms. Only two subjects were submitted to a repeat treatment following the initial management session (employing the same technique) at 3 weeks and 36 months post-therapy respectively.

Auditory-Perceptual Evaluation: The middle sentence of the first paragraph of "The Rainbow Passage" (i.e., "these take the shape of a long round arch with its path high above and its two ends apparently beyond the horizon") (15) was extracted from the original pre-, post 1-, post 2-, and post 3-treatment audiorecordings and randomized onto a master tape for the purpose of auditory-perceptual evaluation. A similar procedure was used for evaluation of pre-, post 1-, post 2-, and post 3-treatment recordings of the sustained vowel / a /. The central 3 second segment of the original vowel recordings was copied in random order onto a second master tape, adjusting the signal level to prevent overload distortion. To permit an estimate of the intra-judge reliability of the perceptual ratings, approximately 10% of the samples on each experimental tape were repeated toward the end of the rating session.

To obtain severity ratings, four judges were instructed to rate each voice sample on a seven-point equal-appearing interval scale, where "1" indicated normal voice and "7" indicated a severe voice disorder. Four speech pathology graduate students employed in a university vocal function laboratory, who had received graduate level training in voice pathology, served as listeners. Ten practice samples were presented to familiarize the judges with the listening task and the range of severity of vocal dysfunction. These practice stimuli were representative of the experimental samples, but were not the same as those included in the experimental session. The selection of the practice samples was predicated on the agreement of the investigators. Following each practice sample, the judges compared and discussed their ratings until agreement on all practice items was achieved.

For the experimental session, judges were not permitted to discuss their ratings. Voice stimuli were presented at a comfortable listening level in a quiet setting. Listeners were required to rate connected speech and sustained vowel samples separately. All judgments were made during a single listening session.

Listener Reliability: Several estimates of inter- and intra-judge reliability and agreement were calculated (16) (see Table 2). Intraclass correlation coefficients (ICC(2,1)) (17) for listener ratings of sustained vowel and connected speech samples were employed to estimate the overall cohesiveness of the raters (interobserver reliability). Variance components used to compute the ICCs were estimated from the expected mean squares using a random effects analysis of variance (ANOVA). ICC = 0.92 and ICC = 0.88 were obtained for connected speech and sustained vowels, respectively. The 95% confidence intervals for the ICCs were also calculated (18). These values suggest that the true ICCs were within the range of 0.90 to 0.94 for connected speech and of 0.85 to 0.92 for sustained vowels . These estimates indicate acceptable between-rater reliability. The judges' overall means of 3.1 ranged from 3.0 to 3.2 for connected speech, and from 2.8 to 3.3 for sustained vowel samples.

Intra-judge agreement represents the extent to which the individual judges tended to make the same judgments at the second presentation. Intra-judge agreement was calculated separately for both connected speech and sustained vowel ratings using two criteria, i.e., pooled exact agreement, and pooled agreement within +/- 1 scale value. Pooled exact agreement across all judges was 54% and 60% for connected speech and sustained vowel voice samples, respectively. Intra-judge agreement calculated using the within +/- 1 scale value criterion, was 90% and 93% for connected speech and sustained vowel samples, respectively, suggesting acceptable concordance levels. Analysis comparing the first and second ratings suggested that listeners' ratings drifted significantly within the listening session (Table 2). A two-tailed sign test revealed a consistent trend across the repeated samples for connected speech (p=0.019). Inspection of the data demonstrated that severity ratings tended to increase at the second presentation, indicating that voices were rated as more severe later in the listening session. A similar drift trend was also detected for repeated sustained vowel ratings, but this tendency was non-significant. This drift phenomenon may partially account for the reduced exact agreement noted previously. Spearman's rho was also calculated to determine intra-judge reliability. Spearman correlation coefficients of 0.93 and 0.91 were obtained for connected speech and sustained vowel ratings.

Acoustic Evaluation: A signal processing package (Cspeech, Paul Milenkovic 1991; University of Wisconsin-Madison) was employed to measure fundamental frequency (Hz), % jitter, % shimmer, and signal-to-noise ratio (SNR-

Table 2.
Estimates of Inter- and Intra-Judge Reliability and Agreement for Listener Ratings of Severity

STATISTIC	CONNECTED SPEECH	SUSTAINED VOWEL
ICC (95% Confidence Interval)	0.92 (0.90 to 0.94)	0.88 (0.85 to 0.92)
Mean severity (Range)	3.1 (3.0 to 3.2)	3.1 (2.8 to 3.3)
Pooled exact agreement (Range)	54% (42% to 58%)	60% (47% to 65%)
Pooled agreement +/- 1 scale value (Range)	90% (83% to 92%)	93% (88% to 100%)
Increases / decreases (Sign Test: p value)	15 / 4 (0.019)	17 / 10 (0.165)
Spearman's rho (Range)	0.93 (0.87 to 0.98)	0.91 (0.90 to 0.92)

Table 3.
Acoustic Measurement Reliability for Connected Speech and Sustained Vowel Stimuli; Spearman Corelation Coefficients (* all p-values were less than 0.0001)

VOICE CONTEXT	Frequency	Jitter	Shimmer	SNR
CONNECTED SPEECH	0.99	0.91	0.85	0.92
SUSTAINED VOWEL	0.99	0.96	0.84	0.85

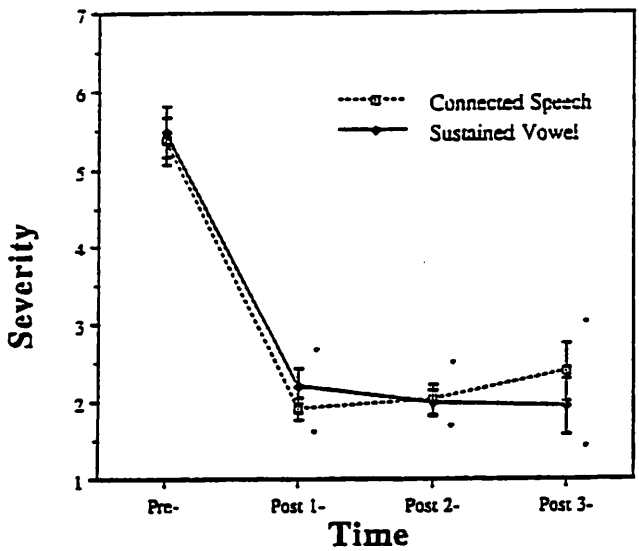


Figure 1. Means and standard errors (severity) for connected speech and sustained vowel voice contexts obtained at four stages during the treatment process. Asterisk (*) indicates significant difference (p<0.0001) when compared to pre-treatment measures.

dB) of the pre-, post-1, post 2-, and post 3- treatment voice samples. Contextual speech and sustained vowel productions were analyzed. The central 100 ms portion of the neutralized stressed vowel /ʌ/ in the word "above" from the middle sentence of "The Rainbow Passage" was chosen for analysis because of its voiced CVC environment. The central 1000 ms segment of the sustained vowel /a/ from the pre-, post-1 , post 2-, and post 3- treatment samples was also selected for analysis. The speech waveform was sampled at a rate of 22 kHz and filtered at 8 kHz.

Acoustic Measurement Reliability: Remeasurement of 15% of the voice samples from each voice context was undertaken to assess the test-retest reliability of each acoustic measure. Spearman rho correlations were calculated for each acoustic parameter and the results were correlation coefficients that ranged from 0.99 to 0.85, and 0.98 to 0.84 for connected speech and sustained vowels respectively (Table 3). All p- values were less than 0.0001, indicating acceptable remeasurement reliability. Similar trends were observed for Pearson correlations.

Table 4.
Means (+/- Standard Deviation) of Severity Measures Before and Following Treatment

VOICE CONTEXT	Pre-	Post 1-	Post 2-	Post 3-
	Mean (+/- SD)	Mean (+/- SD)	Mean (+/- SD)	Mean (+/- SD)
CONNECTED SPEECH	5.37 (1.50)	1.91 (0.74)	2.03 (0.98)	2.38 (0.51)
SUSTAINED VOWEL	5.48 (1.66)	2.20 (1.16)	1.98 (0.79)	1.94 (1.41)

Results

Repeated-measures ANOVA was used to evaluate time trends within the data. The dependent measures were first log-transformed to better meet the assumptions of homogeneous variances required by ANOVA. Fisher's protected least significant difference (LSD) procedure was used to compare the means.

Auditory-Perceptual Evaluation of Severity

Group means and standard errors for both voice contexts at each time interval are represented graphically in Figure 1. Group means and standard deviations data for severity measures before and following treatment are listed in Table 4. Because trends and significance patterns for connected speech and sustained vowel analyses corresponded, only the connected speech results are reviewed here in detail.

Following the single treatment session, a significant reduction in mean severity measures was observed. These improvements were maintained over time based on the repeated measures analysis. Pre-treatment severity ratings for connected speech samples were reduced from a mean of 5.37 to 1.91 immediately following the management session (p=0.0001). Comparison of pre-treatment mean severity measures with each subsequent post-treatment followup mean revealed significant differences (p<0.0004). No significant differences were identified when comparing among post-treatment means (post 1- vs. post 2-; post 1- vs. post 3-; and post 2- vs. post 3-). These results would indicate that subjects maintained the improvements in voice acquired in the single treatment session.

Inspection of individual subject data revealed the following relatively consistent patterns, regardless of the voice context selected for study (Table 5; following page). Twenty-four subjects (96%) were rated by listeners as improved on the basis of perceptual judgments of severity. That is, mean post-treatment severity scores were at least one scale value less severe when compared to pre-treatment measures. Twenty subjects (80%) improved at least two scale values. Furthermore, 64% of subjects (16/25) were considered normal or only mildly dysphonic immediately

Table 5.
Responses to Treatment Using Criterion Measures for Impovement/Relapse Derived From the Severity Rating Scale

CRITERION MEASURES	CONNECTED SPEECH	SUSTAINED VOWEL
Improvement (post 1- at least 1 scale value less severe than pre-treatment rating)	96% (24/25)	92% (23/25)
Improvement (post 1- at least 2 scale values less severe than pre-treatment rating)	80% (20/25)	76% (19/25)
Normal or mildly dysphonic following treatment (post 1- < or = 2 on severity rating scale)	64% (16/25)	68% (17/25)
Relapse (post 2- or post 3- at least 1 scale value more severe than post 1- rating)	25% (6/24)	4% (1/23)
Further Improvement following discharge (post 2- or post 3- at least 1 scale value less severe than post 1- rating)	17% (4/24)	17% (4/23)

following the treatment session (i.e., mean post-treatment severity measures were less than or equal to 2 on the severity rating scale of 1 to 7).

Relapse was assessed on an individual basis by comparing mean severity ratings at several points during followup. Some degree of relapse was assumed to have occurred if a subjects' post 2- or post 3- treatment mean severity score was at least one scale value more severe than the immediate post-treatment (post-1) value. Using this criterion, 28% of subjects were judged to have experienced partial recurrence. Only one subject however, experienced complete relapse based on perceptual ratngs. Several subjects (17%) were rated as superior to their immediate post-treatment voices (post 1-) suggesting further improvement after the initial management session.

Acoustic Evaluation

Time trends were tested for with repeated-measures ANOVA. With the exception of fundamental frequency, all acoustic measures were significantly improved immediately following the treatment session. Again, results and significance trends for connected speech and sustained vowels were comparable. Significant reductions in percent jitter (cycle-per-cycle variations in frequency) and percent shimmer (a measure of amplitude variability) were detected. Signal-to-noise ratio, a ratio of total energy to noise in the speech waveform, was observed to increase appreciably following the management session (Figure 2; following page). No significant changes were observed in vocal fundamental frequency.

Each subsequent post-treatment mean was also shown to be significantly different from the pre-treatment mean, regardless of the acoustic parameter assessed ($p<.0001$) (Table 6). Based on comparisons among post-treatment mean acoustic measures (post 1-, post 2- and post 3-), improvements observed after the single treatment session were maintained during the followup period.

Table 6.
Means (+/- Standard Deviations) for Acoustic Measures of Connected Speech and Sustained Vowel Samples Obtained at 4 Stages During the Treatment Process

ACOUSTIC MEASURES	Pre-		Post 1-		Post 2-		Post 3-	
	Mean (+/- SD)		Mean (+/- SD)		Mean (+/- SD)		Mean (+/- SD)	
CONNECTED SPEECH								
Frequency (Hz)	202.37	(99.95)	165.39	(32.45)	163.61	(29.78)	164.59	(91.82)
Jitter (%)	4.22	(4.45)	1.51	(1.25)	1.09	(1.28)	2.1	(3.93)
Shimmer (%)	28.4	(22.26)	7.74	(8.68)	7.03	(8.51)	9.62	(11.90)
SNR (dB)	8.19	(6.16)	14.63	(4.58)	15.69	(5.38)	15.04	(5.48)
SUSTAINED VOWEL								
Frequency (Hz)	241.87	(141.99)	200.86	(42.21)	195.72	(27.39)	185.45	(39.11)
Jitter (%)	3.78	(4.35)	0.57	(0.46)	0.39	(0.15)	0.69	(1.37)
Shimmer (%)	24.55	(20.72)	3.76	(2.94)	3.22	(1.54)	6.74	(17.77)
SNR (dB)	9.37	(6.47)	20.19	(4.82)	20.06	(4.44)	21.34	(5.75)

Interview Results

Follow-up interviews of each subject (Table 7) provided broad assessment of long-term results. Given the inherent constraints associated with self-report methodologies, including memory limitations of subjects, appropriate caution should be exercised during their interpretation. Because insufficient time had lapsed to accurately evaluate the presence or absence of recurrence in five subjects (final contact was less than 1 month post-treatment) and one subject was judged to have derived minimal benefit from the treatment, results are based on self-report data from 19 subjects. Subject attrition during long-term followup, therefore, may actually contribute to an underestimation of the frequency of recurrence.

Thirteen of the 19 subjects (68%) who responded to manual therapy reported episode(s) of recurrence during the follow-up period. Inspection of these data indicated that within 2 months following treatment, 77% (10/13) of these subjects reported experiencing episodes of relapse (mean =1.4 (+/- 0.99) months, range= 1 hour to 3 months). In the majority of cases, the total number of recurrences was low (less than three episodes) and was partially related to the duration of the follow-up period. However, individual variability was substantial, with one subject (#10) who reported experiencing relapse on a daily basis, as compared to other subjects (#3 and #17) who reported single incidents. Although the duration of relapse varied within subjects from several hours to several weeks, it was reported to be short-lived and typically persisted for less than 4 days. Furthermore, 62% (8/13) of subjects reported relapse that was "partial" rather than complete, indicating that the symptoms were judged less severe than the original dysphonia. Twenty-three percent of subjects judged the recurrent episode(s) to be as severe as the original voice problem. In 70% (9/13)

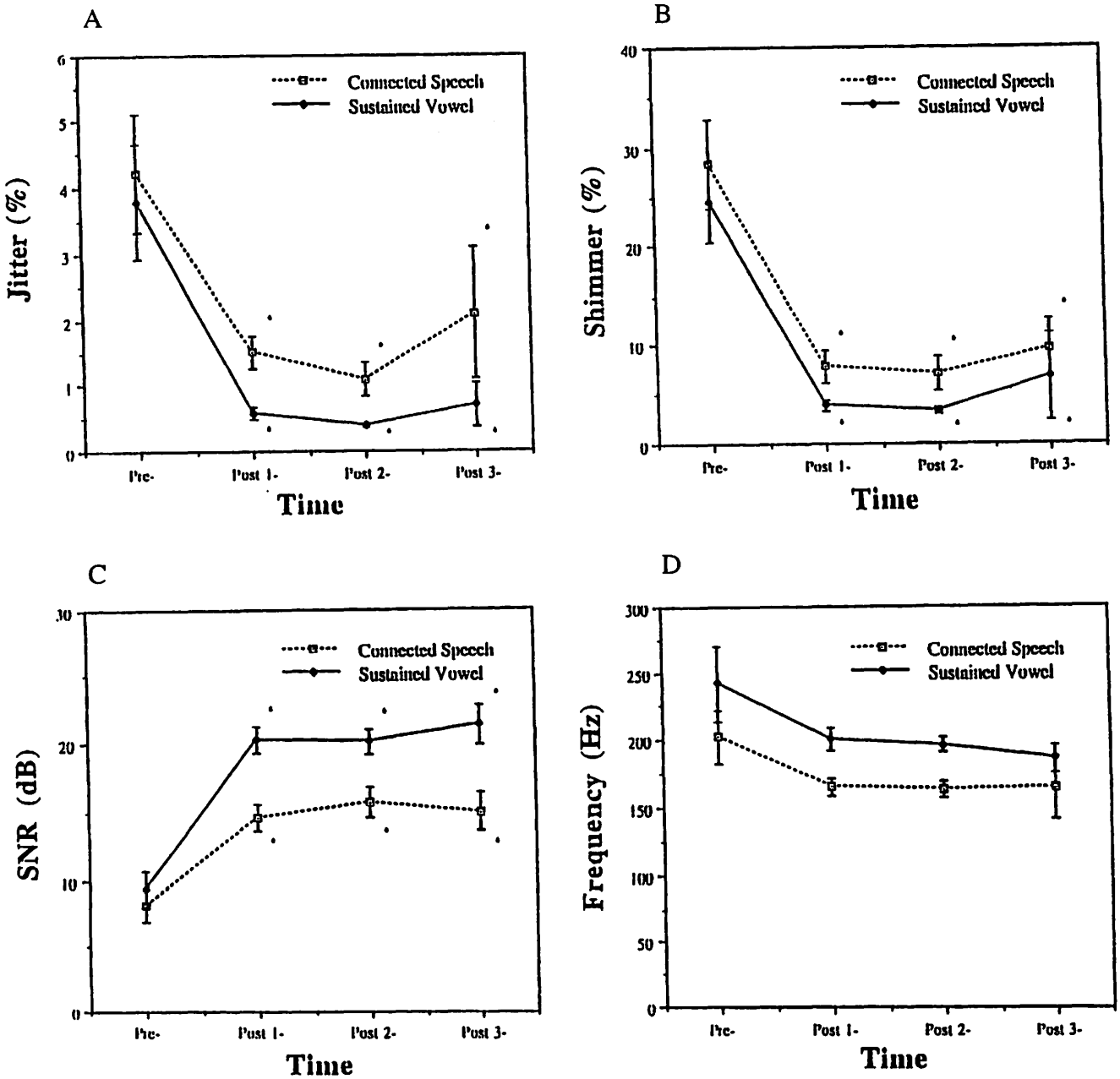


Figure 2. Means and standard errors of acoustic measures for connected speech and sustained vowel analysis obtained at four stages during the treatment process. Asterisk (*) indicates significant different (p<.0004) when compared to pre-treatment measures.

of subjects, the recurrence resolved spontaneously, whereas other subjects attempted "self laryngeal massage" with mixed results. Patients associated increases in life stress, mood disturbances, or cold-like symptoms as preceding or accompanying the return of dysphonic symptoms. No relationship was observed between pre-treatment duration of dysphonic symptoms and the frequency, severity, or duration of relapse. In spite of a high rate of reported recurrence, only two patients contacted the attending clinician for additional consultation/treatment. One patient (#6) experienced an episode of persistent relapse approximately 36 months following initial treatment and responded to a second treatment using the same procedure. The second patient (#17) experienced a minor recurrence within 21 days, sought additional treatment, and responded favorably.

Table 7.
Description of Relapse Characteristics Based on Follow-up Interviews with Individual Subjects

Subject	Time to Final Contact (mo post-Rx)	Relapse (+/-)	Time to First Relapse (mo)	Frequency of Relapses	Mean Duration of Relapse	Severity of Relapse </> Original Dysphonia	Resolution of Relapse	Self Massage	Psychological Counseling Recommended	URI or Evaluation in Life Stress Preceding Relapse
1	24	+	1.5	2	4 days	less	spontaneous	no	no	respiratory
2	3	No Rx effect	N/A	N/A	N/A	N/A	N/A	N/A	yes	depression
3	24	+	1	1	4 days	less	spontaneous	no	yes	stress
4	27	+	1	2	2 days	same	spontaneous	yes	no	stress
5	12	+	3	3	6 wk	same	spontaneous	no	yes	depression
6	36	+	3	7	4 days	same	self-massage, spont., and	yes	yes	stress
7	10	+	0.0014	>20	2 wk	variable	spontaneous	no	yes	depression
8	20	+	1.5	2	2 days	less	self-massage	yes	yes	stress
9	24	+	0.5	24	5 days	variable	spontaneous	no; pain	yes	respiratory
10	36	+	0.25	DAILY	30 min	less	spont., self-mass.	no	yes	no
11	6	-	N/A	N/A	N/A	N/A	N/A	yes	no	N/A
12	24	-	N/A	N/A	N/A	N/A	N/A	no	yes	N/A
13	18	-	N/A	N/A	N/A	N/A	N/A	no	no	N/A
14	18	-	N/A	N/A	N/A	N/A	N/A	no	no	N/A
15	0.10	No relapse based on post 2-, but patient lost to long-term follow-up								
16	18	+	1	2	5 hr	less	spontaneous	no	no	no
17	6	+	0.75	1	3 days	less	repeat MCT	yes	no	no
18	12	+	2	12	3 days	less	spontaneous	no	no	respiratory
19	6	-	N/A	N/A	N/A	N/A	N/A	yes	no	N/A
20	6	+	2.5	3	2 hr	less	spontaneous	no	no	stress
21	4	-	N/A	N/A	N/A	N/A	N/A	no	no	N/A
22	0.25	No relapse based on post 2-, but patient lost to long-term follow-up								
23	0.25	No relapse based on post 2-, but patient lost to long-term follow-up								
24	0.17	No relapse based on post 2-, but patient lost to long-term follow-up								
25	0.33	No relapse based on post 2-, but patient lost to long-term follow-up								

Discussion

The purpose of this investigation was to examine the immediate and long-term effects of manual circumlaryngeal therapy for patients with functional dysphonia (FD). Perceptual, acoustic, and interview techniques were used to assess vocal function before and after a single treatment session. Subjects demonstrated consistent improvement across perceptual and acoustic indices of vocal function immediately after treatment and during the follow-up period. The results of data gathered during post 2- and post 3-treatment interviews suggested that although perceptual and acoustic measures supported short-term maintenance of voice improvements, the majority of patients did report infrequent, partial, self-limiting, recurrences early in the follow-up phase.

These results extending the work of Roy and Leeper (14) confirm the clinical utility and efficiency of manual techniques for patients with FD. This investigation revealed that short-term results are impressive, but on the basis of patient report, long-term results are less robust. Thus, manual circumlaryngeal therapy may best be viewed as an incomplete remedy for most patients, whereas for some it may provide permanent relief of symptoms. Unfortunately, the present study provides little objective evidence upon which to base predictions of who is vulnerable to relapse. It has been speculated by several authors that FD patients who suffer chronic, unresolved, psychological conflict respond incompletely to voice therapy, and/or are at higher risk for experiencing frequent and prolonged relapses. Because of the small number of subjects and reliance on self-report methodologies to assess emotional adjustment, we have insufficient data to confirm these suppositions. Research employing empirical methodologies is required to assess the relationship between psychological factors and long-term relief of vocal symptoms.

The finding that patients responded to a single treatment session is encouraging, but a serious question lingers as to whether improvements in voice are a consequence of reduced laryngeal musculoskeletal tension or due to factor(s) unrelated to the manual technique. It is evident that positive treatment effects cannot logically prove the

case for etiology of the voice disorder in question. The presence, absence, and reduction of excess laryngeal tension was not objectively evaluated. The absence of a non-treatment or alternative treatment control group leaves open the possibility of numerous alternative explanations for the observed treatment effects, including placebo effects, the clinician's instructions, expectations, experience, and confidence. Because there exists a complex interaction between psychological and physiological components in functional voice disorders, and because the reactive psychological effects of clinical interaction (e.g. the psychosocial interview) could not be controlled, it would be presumptuous to attribute changes in voice exclusively to a reduction in laryngeal muscle tension.

These data do not *directly* substantiate Aronson's (3) claim that improvements in voice quality are proportional to a reduction in laryngeal musculoskeletal tension and lowering of the position of the larynx in the neck (p.314). Two observations however, may furnish *indirect* evidence to bolster Aronson's contention concerning the relationship between voice improvement and laryngeal tension reduction. Prior to treatment, 88% of subjects reported pain and/or tenderness in the laryngeal region during palpation (Table 1). During and immediately following the manual procedure, most subjects reported gradual reduction and then amelioration of laryngeal pain. This finding is consistent with Aronson's clinical descriptions (p.314). Whether this confirms his hypothesis that laryngeal pain is related to muscle cramping and that kneading of such muscles results in reduced tension and pain remains open to speculation. During the treatment phase, subjects typically progressed through stages of decreasing dysphonia until symptoms gradually remitted, consistent with Aronson's accounts (p.315). Whether this represents a steady reduction in laryngeal tension, as Aronson maintains, also remains open for debate. No such pattern of gradual remission occurred in two patients (#4 and #10) whose symptoms instantaneously resolved within seconds following initiation of the procedure. This finding raises the question as to whether tension could have been reduced sufficiently in such a brief period to account for the sudden relief of symptoms. One might logically interpret this finding as indicating that factors other than tension reduction are responsible for the observed voice change. However, manually impeding elevation of the larynx may have prevented the subject from assuming an unnatural laryngeal posture, and may represent a plausible explanation for the sudden change.

Given that both psychological and physiological factors have been implicated in the pathogenesis of FD, no single explanation and interpretation of the results is possible. It is reasonable to consider that several clusters exist within the diagnostic category of FD. We suggest that excess laryngeal muscle tension plays a variable role depending on which factors precipate and maintain the voice disorder. Anecdotally, subject #4, who responded extremely rapidly, denied pain during laryngeal palpation. This lack of laryngeal pain combined with rapid amelioration may represent a separate and distinct pathophysiology. This example may serve to illustrate the etiologic heterogeneity of the patients included in the broad diagnostic category of FD.

This investigation also provides helpful information regarding the character and likelihood of recurrence using this approach. If, as theorized by numerous investigators, psychological and personality factors contribute to the development, maintenance, and recurrence of functional dysphonia, then careful research is needed regarding the kind of stressor that can trigger a dysphonic episode in a predisposed individual. Maintenance may be influenced by environmental stimuli that can negatively or positively influence the recovering voice patient. Behavioral and psychological approaches might need to be integrated more consistently for long-term success. For some patients superior results may be found when manual laryngeal techniques are combined with supportive counseling and/or more frequent clinical support. This treatment technique, along with most other voice therapies, focuses on the overt disorder of phonation; until more is known of the etiologic factors, it may be unrealistic to expect great advances in long-term "cure" rates. For some patients, controlled or short relapses may be a viable outcome even when total cure was the original goal.

In summary, manual circumlaryngeal therapy appears to be an effective and efficient treatment for functional dysphonia. Because so little information exists regarding the long-term consequences of other behavioral approaches, it is difficult to evaluate these results against other treatments. Many clinicians have anecdotally reported rapid success with patients using approaches that do not involve manual techniques. Because no controlled comparison of alternative treatment techniques was performed, these results do not support the superiority of manual approaches over any other behavioral technique. Rather, they do support the notion that manual circumlaryngeal therapy should be considered early in the treatment selection process. Further research must be undertaken to compare the short- and long-term consequences of other behavioral techniques used with the FD patient population to place these results in context. Future investigations should also aim to isolate the causative factors that precipitate significant voice improvement. Vocal, psychological, physiological, and clinician characteristics must be better defined to separate the critical variables contributing to successful management of functional dysphonia.

Acknowledgments

Grateful acknowledgment is made to Stephen M. Tasko, who provided treatment to four of the subjects. The authors express gratitude to the Vocal Function Laboratory staff (Waisman Center), especially Amy Mast and Dawn Salyards for their valuable contributions at various stages during this project. The authors also thank Paul Bergamin, Stephanie Harvey and Vanessa Sharbure for their assistance with data analysis.

Parts of this manuscript were presented at the American Speech and Hearing Association Convention, New Orleans, 1994.

This work was supported by the National Center for Voice and Speech through Grant P60 00976 for the National Institute on Deafness and Other Communication Disorders.

Appendix A: Voice Assessment and Interview Protocol

(i) Baseline audiorecordings of (a) "The Rainbow Passage" and (b) sustained vowel productions / a, i, u /.

(ii) Familiarization period.

(iii) Onset of the voice disorder: Conditions surrounding the onset of the dysphonia were explored including when and how the voice disturbance began.

(iv) History of previous periods of partial or total voice loss.

(v) Course of the voice disorder: consistent or variable?

(vi) Events associated with onset: The major emphasis was to determine if the voice disorder began in association with emotional problems (i.e., stress, anxiety, conflict, and communication breakdown).

(vii) Traditional evaluation of vocal function, inclucing MPT, and pitch and loudness ranges.

(viii) Musculoskeletal tension evaluation: This was used to determine subjectively (a) the extent of laryngeal elevation, (b) pain in response to pressure in the region of the larynx, and (c) sites of focal nodularity or taut bands. By encircling the larynx with the thumb and index finger in the region of the thyrohyoid space, it was established whether the space had been narrowed by laryngeal elevation. The presence or absence of pain was detected by manually palpating in the area of the hyoid bone and thyrohyoid space bilaterally.

(ix) Post-treatment audiorecordings of (a) "The Rainbow Passage" and (b) sustained vowels / a, i, u /.

(x) Post 2- treatment audiorecordings and interview.

(xi) Post 3- treatment audiorecording and interview.

Appendix B: Treatment Protocol

(i) Review of the laryngologist's and videolaryngoscopic findings with reassurance that there was no evidence of lesions or disease.

(ii) Discussion and explanation of the patient's voice in relation to the effects of emotion on muscle tension and phonatory control.

(iii) Discussion of the therapy plan when the patient appeared to understand the mechanism of the voice disorder. The therapy approach was explained and the outlook for recovery was established.

(iv) Manual circumlaryngeal therapy (the manual laryngeal musculoskeletal tension reduction technique) was undertaken following the description of Aronson (3). First, (a) the hyoid bone was encircled with the thumb and index finger, which were worked posteriorly until the tips of the major horns were felt; (b) light pressure was exerted with the fingers in a circular motion over the tips of the hyoid bone; (c) the procedure was repeated beginning from the thyroid notch and working posteriorly; (d) the posterior borders of the thyroid cartilage just medial to the sternocleidomastoid muscles were located and the procedure was repeated; (e) with the fingers over the superior borders of the thyroid cartilage, the larynx was worked downward, and moved laterally at times; and (f) the patient was asked to hum or prolong vowels during the above procedures, while changes in vocal quality were noted. Improvement in voice was immediatedly reinforced. The improved voice was progressively shaped from vowels and words (usually automatic serial speech, i.e., counting, days of the week) to short phrases (e.g., "many men in the moon", "one monday morning") to sentences, and finally conversation.

(v) After completion of the procedure, the results of the therapy approach were discussed with the patient in terms of factors that could be contributing to the voice disorder and whether further psychological assessment or counseling was necessary.

(vi) All patients were encouraged to telephone a friend, relative, or spouse while in the clinic to stabilize the voice.

(vii) Arrangements were made with each patient for a follow-up session to assess the effects of the treatment.

References

1. Schalen L, Andersson K. Differential diagnosis and treatment of psychogenic voice disorders. **Clin Otolaryngol** 1992; 17:225-30.

2. Reed CG. Voice therapy: A need for research. **J Speech Hear Dis** 1980; XLV: 157-69.

3. Aronson AE. **Clinical Voice Disorders: An Interdisciplinary Approach.** (3rd edn.) New York: Thieme Stratton, 1990.

4. Cooper M. Spectrographic analyses of fundamental frequency and hoarseness before and after vocal rehabilitation. **J Speech Hear Dis** 1974;39: 276-87.

5. Milutinovic Z. Results of vocal therapy for phononeurosis: Behavior approach. **Folia Phoniatr** 1990; 42, 173-177.

6. Stemple JC. **Voice Therapy: Clinical Studies.** St. Louis: Mosby Year Book, 1993.

7. Boone DR, McFarlane S. **The Voice and Voice therapy.** (4th edn.). Englewoood Cliffs, NJ: Prentice-Hall 1988.

8. Pahn J, Friemaert K. Differential diagnostische und terminologische Erwagungen bei sogenannten funktionellen Storungen im neuropsychiatrischen und phoniatrischen Fachgebiet. 2. Phoniatrischer Aspekt. **Folia Phoniatr** 1988; 40, 168-74.

9. Green G. The inter-relationship between vocal and psychological characteristics: A literature review. **Austral J Hum Comm Dis** 1988; 16: 31-41.

10. Rammage LA, Nichol H, Morrison MD. The psychopathology of voice disorders. **Hum Commun Can** 1987; 11: 21-25.

11. Bridger MM, Epstein R. Functional voice disorders: A review of 109 patients. **J. Laryngol Otol** 1989; 97: 1145-48.

12. Fex F, Fex S, Shiromoto O, Hirano M. Acoustic analysis of functional dysphonia: Before and after voice therapy (Accent Method) **J Voice** 1994; 8: 163-67.

13. Koufman JA, Blalock PD. Vocal fatique and dysphonia in the professional voice user: Bogart-Bacall syndrome. **Laryngoscope** 1988; 98: 493-99.

14. Roy N, Leeper HA. Effects of the manual laryngeal musculoskeletal tension reduction technique as a treatment for functional voice disorders: Perceptual and acoustic Measures. **J Voice** 1993; 7: 242-49.

15. Fairbanks, G. **Voice and Articulation Drillbook.** New York: Harper and Row, 1960.

16. Kreiman J, Gerratt BR, Kempster GB, Erman A, Berke GS. Perceptual evaluation of voice quality: Review, tutorial, and a framework for future research. **J Speech Hear Res** 1993; 36: 21-40.

17. Ebel R. Estimation of the reliability of ratings. **Psychometrica** 1951; 16: 407-24.

18. Snedecor GW, Cochran WG **Statistical Methods.** (7th edn.) Ames: Iowa State University Press, 1980.

NCVS Status and Progress Report - 11
May 1997, 21-35

Speech Rehabilitation and Short-term Changes in Satisfaction of Employment in Parkinson Disease: A Preliminary Multiple Case Study

Stefanie Countryman M.A., CCC-SLP
Wilbur James Gould Voice Research Center, The Denver Center For The Performing Arts
Lorraine Olson Ramig Ph.D., CCC-SLP
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
Annette Arnone Pawlas M.A., CCC-SLP
Wilbur James Gould Voice Research Center, The Denver Center For The Performing Arts
Laetitia L. Thompson Ph.D.
Neuropsychology Laboratory, The University of Colorado School of Medicine

Abstract

Three employed individuals with speech and voice deficits as a result of Parkinson disease are presented. Each participant completed an intensive 1 month course of the Lee Silverman Voice Treatment. Measures of speech and voice, employment satisfaction, and quality of life were administered pre- and immediately post-treatment to determine the effects of intensive speech rehabilitation on employment satisfaction and performance. All three participants improved on the measures of speech, employment, and quality of life outcomes immediately post-treatment from the employee (participant), employer of the participants, and a "blind" rater perspective. The implications of these findings in relation to the role of the speech-language pathologist in vocational rehabilitation as well as future research are discussed.

Parkinson disease (PD) is a progressive neurological disorder which effects well over one million Americans (Duvoisin, 1984). At least 75% of these individuals have speech disabilities (Aronson, 1990; Logemann, Fisher, Boshes, & Blonsky, 1978; Streifler & Hofmann, 1984) as well as other physical problems that can limit employment and overall functional participation in society (Mutch, Strudwick, Roy, & Downie, 1986; Oxtoby, 1982). While individuals with Parkinson disease report their speech disabilities negatively effect their employment and quality of life (King, Ramig, & Lemke, 1996; Oxtoby, 1982), the relationship between speech rehabilitation, employability and job satisfaction in Parkinson disease has not been addressed. The Americans with Disabilities Act of 1990 guarantees equal employment opportunities for individuals with disabilities, including speech. Therefore, it is important to evaluate whether speech-language pathology services can help individuals with Parkinson disease improve job satisfaction or maintain employment after onset of speech disabilities.

Recently, Ramig and colleagues developed an efficacious program for speech rehabilitation in Parkinson disease (Ramig, Bonitati, Lemke, & Horii, 1994; Ramig, Countryman, O'Brien, Hoehn, & Thompson, 1996; Ramig, Countryman, Thompson, & Horii, 1995). The Lee Silverman Voice Treatment (LSVT) is an intensive voice treatment program designed specifically for individuals with Parkinson disease. Individuals receive treatment 4 times a week, one hour a day for four weeks (16 sessions). After receiving the LSVT, at least 80% of individuals with PD reported on to date demonstrated statistically significant, quantifiable improvements in objective measures of speech and voice (Ramig et al., 1994;Ramig et al., 1995). In addition, individuals with PD receiving the LSVT have demonstrated long-term maintenance of treatment effects (Ramig et al., 1996).

Contrary to common perception, PD is not a disease only of the elderly. At least 35% of individuals with PD are diagnosed before the age of 40 (Parkinson's Education Program Exchange, 1984). Threat of job loss or interference from the disease in job duties has been documented in this population (King et al., 1996; LaPlante, 1988; Oxtoby, 1982; Shoulson, 1989). For example, Shoulson (1989) found that 40% of 400 individuals with PD surveyed indi-

cated their disease threatened their full-time employment. King et al., (1996) reported that individuals with Parkinson disease experienced a reduction of communication in their job settings as a direct result of their disease. In addition, some individuals reported that difficulty with their speech was the most debilitating aspect of their disease (King et al., 1996).

Although the degenerative progression of Parkinson disease is not reversible at this time, early and effective speech rehabilitation may significantly impact the natural course of speech deterioration (Ramig et al., 1995; Ramig et al., 1996) and encourage individuals with PD to remain actively employed. The cost of maintaining disabled individuals on long-term disability can be substantial (Peters, Large, Elkind, 1992; Richardson, 1994; Straaton, Maisiak, Wrigley, & Fine, 1995). Successful speech rehabilitation (LSVT) may deter some of this health care expense by helping individuals withParkinson disease improve overall communication skills to compete in the work place longer and more effectively, improving the quality of life for individuals with PD.

The purpose of this preliminary study was to investigate the short-term impact (pre- to post-treatment) of speech rehabilitation (LSVT) on employment satisfaction and performance in three individuals with Parkinson disease. It was hypothesized that speech rehabilitation could potentially provide individuals with PD the necessary skills to improve overall job performance and satisfaction and communication skills within the work setting.

Method

Participants

Three participants with idiopathic Parkinson disease (PD), spanning a range of employment circumstances, participated in this study. Participant characteristics and employment descriptions are presented in Table 1. During the initial interviews, communication-related employment difficulties were reported by all three participants. Participant 1 reported he had lost his job as a lead mechanic at the airport because he could not communicate effectively with the pilots of the airplanes. Participant 2 reported that he had to repeat himself numerous times on the phone while at work and felt his voice was hoarse and fatigued by the end of the day. Participant 3 reported that she was limited to small group discussions and was not asked to do vocal tasks at work. From pre- to immediately post speechtreatment, the participants did not partake in other forms of treatment (physical therapy, occupational therapy etc.), were considered stable on their anti-Parkinson medication, and did not change anti-Parkinson medication.

Because cognitive and memory impairment in Parkinson disease (Karayanidis, 1989; Levin, Llabre & Weiner, 1989; Mohr, Juncos, Cox, Litvan, Fedio, & Chase, 1990) may interact with the impact of speech rehabilitation on sustained employment and job satisfaction in these participants, neuropsychological testing was completed by all three participants. The pre-treatment cognitive test battery focused on a combination of tests that have been found to differentiate neurologically diseased patients who have been able to maintain employment from those who have not (Heaton, Chelune, & Lehman, 1978; Newnan, Heaton, & Lehman, 1978). The computerized neuropsychological screening battery was administered by a neuropsychologist within the month prior to speech therapy. Scores were obtained in seven cognitive domains: attention/mental control, memory, reasoning/calculation, spatial processing, informational processing speed, informational processing accuracy, and reaction time. Overall functioning and proficiency scores were also calculated. An individual's scores were compared with the general population and individuals who are in the same age range and have similar levels of education. Reliability coefficients are satisfactory and the screening battery has been shown accurate for separating demented participants from controls (Powell, Kaplan, Whitla, Weintraub, Catlin, & Funkenstein, 1993).

Participant pre-treatment cognitive scores are reported in Table 2. Participant 2 and 3s' scores were within normal limits in all areas. Participant 1's scores were within normal limits except for the category "information processing speed." However, this score was only mildly impaired and the neuropsychologist stated that this deficit should not effect the participant's success in the speech therapy program.

Table 1.
Pre-treatment descriptive characteristics of Participants 1, 2, and 3.

	Gender	Age (years)	Hoehn and Yahr Stage (1967)*	Years since diagnosis	Speech severity rating**
Participant 1 airplane mechanic	M	58	3	10	mod-sev
Participant 2 technical staff manager (phone company)	M	55	2	3	mild
Participant 3 volunteer (elementary school)	F	41	2	1.5	mild-mod

* Stages on the Hoehn and Yahr range from 1 - 5, with 1 being the least severe and 5 the most severe.
** Rated by the attending speech-language pathologist.

Table 2.
Scores from the computerized neuropsychological screening battery from the seven cognitive domains tested for Participants 1, 2, and 3. Scores were obtained pre-treatment and are reported in percent.

Cognitive Domain	Participant 1	Participant 2	Participant 3
Attention/Mental control	42	47	39
Memory	63	82	63
Reasoning/calculation	55	63	25
Spatial Processing	68	75	70
Information Processing speed	13	82	73
Information processing accuracy	84	42	18
Reaction time	61	68	53
General cognitive functioning	47	66	42
General cognitive proficiency	21	47	27

Note: Scores below the 16th percentile are considered impaired, scores between the 16th and 30th percentile are below average.

All three participants completed a pre-treatment otolaryngological evaluation including video-laryngostroboscopy to rule out any contraindications to voice therapy as well as assess vocal fold functioning. Pre-treatment, Participants 1 and 3 had mild incomplete closure of the glottis during normal pitch, normal loudness phonation. Pre-treatment, Participant 2 had mild incomplete closure during the same task as well as a small posterior glottal gap (chink). No vocal fold pathology that would contradict voice treatment was observed in the participants pre-treatment.

Treatment

The Lee Silverman Voice Treatment ([LVST] Ramig et al., 1994; Ramig et al., 1995; Ramig et al., 1996) was administered to these participants. Treatment techniques focused on increasing vocal loudness and efficiency by targeting the hypothesized underlying laryngeal pathophysiology and optimizing phonatory and respiratory effort and coordination across the motor speech system using the global variable "loud" (Dromey, Ramig, & Johnson, 1995; Schulman, 1989). The LSVT wasspecifically designed to maximize participant motivation for the speech tasks and facilitate immediate carryover of increased vocal loudness into functional communication. Specific techniques included optimizing vocal fold closure through vocal isometric exercises and improving respiratory support by breathing deeply and frequently while speaking loud ("thinking loud") and speaking on "top of the breath."

A typical session included a high number of successive repetitions of the following activities: maximum sustained vowel phonation (/a/), generating the highest and lowest fundamental frequencies the participant could obtain, and speech production tasks using the same phonatory and respiratory techniques used in sustained phonation. The high number of repetitions during the treatment as well as the intensiveness of therapy (4 times a week for 4 weeks) are consistent with theories of motor learning and skill acquisition (Schmidt, 1975; 1988) and muscle training (Saxon & Schneider, 1995). Treatment techniques also enhance the participant's overall "knowledge of results," which is critical in skill acquisition (Salmoni, Schmidt, & Walter, 1984). A tape recorder was used to provide feedback on vocal quality and loudness to the participant during each task. In addition, during the four week program, participants are trained in self-monitoring of loudness (sensory calibration) and habituation of the increased loudness. The LSVT method has been described in further detail elsewhere (Ramig et al., 1995; Ramig, Pawlas, & Countryman, 1995)

Rationale for Measures of Speech, Employment, and Quality of Life Outcomes

The measurements of speech, employment and quality of life outcomes used in this study were chosen through interdisciplinary consultation with speech, vocational rehabilitation experts, human resource, and neuropsychological personnel.

Measures of Speech Outcomes To obtain pre- to post-treatment objective data on the participants' speech, the following measures were collected.

Sound Pressure Level. Because of the high frequency of occurrence of reduced loudness in individuals with Parkinson disease (Aronson, 1985; Canter, 1965; Critchley, 1981; Darley, Aronson, & Brown, 1969a, b; Fox & Ramig, 1996) and the impact of reduced loudness on speech intelligibility (Ramig, 1992), sound pressure level (SPL) during sustained phonation, reading, and monologue was measured.

Forced Vital Capacity. To gauge respiratory volume of the participants, forced vital capacity was measured.

Self-Perceptual Ratings. To evaluate self-perception of changes in speech pre- to immediately post-treatment, participants completed a visual analogue scale (Kempster, 1984) for the variables loudness, overall intelligibility, participation in conversation, and initiation of conversation. Loudness was chosen to correspond with the acoustic measure of intensity, while overall intelligibility, participation in conversation and initiation of conversation were all considered measures of improved intelligibility and confidence in speaking.

Naive Listener Perceptual Ratings. To determine the extent of pre- to immediately post-treatment changes in speech and voice characteristics, individuals unfamiliar with the participants completed a perceptual listening task. Four naive listeners completed a visual analogue scale for the variables loudness, monotone, slurs or mumbles, overall intelligibility, and speaks in a confident manner, while listening to a recording of the participants reading the "Rainbow Passage" pre- and post-treatment. Loudness was chosen for perceptual analysis because of the correspondence to the acoustic variable intensity as well as its effect on overall speech intelligibility (Ramig, 1992). Monotone, slurs, overall intelligibility and confidence in speaking were chosen because these variables could potentially reflect overall communication skills as well as effect perception of communication skills within an employment setting.

Measures of Employment Outcomes. Three aspects of employment were quantified in this study: job satisfaction from the employer perspective, job satisfaction from the employee perspective, and communication in the individual's work setting from an employee, employer, and "blind" rater perspective. The following measures were used to evaluate pre- to immediately post-treatment changes in these three areas: The Minnesota Satisfactoriness Scale (MSS) (Gibson, Weiss, Dawis, & Lofquist, 1970), The Job Satisfaction Scale (Brayfield & Rothe, 1951), the first section of the Communication Profile for Speakers with Motor Speech Disorders Questionnaire (Yorkston, Bombardier, & Hammen, 1994), employer perceptual ratings of the participants' voice and

speech, and an on-site (place of employment) observation by a "blinded" certified speech-language pathologist.

Measure of job satisfaction-employer perspective. The Minnesota Satisfactoriness Scale (MSS) is a 28 item questionnaire completed by an individual's employer or supervisor. The scale assesses the following 5 job skill areas: general, performance, personal adjustment, dependability, and conformance. This scale was chosen because it is an effective tool for measuring employer satisfaction with employee performance (Gibson et al., 1970).

Measure of job satisfaction-employee perspective. The Job Satisfaction Scale is a 6 item subscale of *An Index of Job Satisfaction Scale* (Brayfield & Rothe, 1951) with 5 answer choices that range between strongly agree to strongly disagree. This scale was chosen as a measure of the employee's satisfaction with their job and job performance. The reliability of this subscale has been reported (Cronbach's = 0.88-0.91, [Price & Mueller, 1986]).

Measures of communication within job-setting-employee perspective. The first section of the Communication Profile for Speakers with Motor Speech Disorders (Yorkston et al., 1994) was used to evaluate participant's self-evaluation of communication. This qualitative measurement helps gauge relationships between communication and employment as well as gain a perspective of functional communication and its impact on the participant.

Measures of communication within job-setting-employer perspective. Participant employers perceptually rated the participants' communication using the same visual analogue rating scale as the participant self-rating visual analogue scales. The same variables of loudness, overall intelligibility, initiation of conversation and participation in conversation were rated.

Measures of communication within job-setting-"blind" rater perspective. On-site observations of the employee's communication within the work setting were completed by a "blind" rater to obtain a non-biased opinion of the participant's communication skills. The rater was a certified speech-language pathologist who was unaware whether or not the participant was receiving speech treatment. She was aware that the participants had Parkinson disease.

Measures of Quality of Life Outcomes. Psychosocial measures were also administered to determine quality of life in relation to disease, employment and communication. Five standardized assessments were completed by the participants pre- and immediately post-treatment. These included: the Sickness Impact Profile (SIP; Bergner, Bobbit, Carter, & Gilson, 1981), the Patient Assessment of Own Functioning Inventory (PAOFI; Chelune, Heaton, & Lehman, 1986), the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961), and the Well Being Scale (Ryff, 1989). All psychosocial scales were completed once pre-treatment and once post-treatment.

Sickness Impact Profile. The SIP was chosen for assessment because it provides information about how people perceive their disease affects other aspects of their lives (e.g., employability, social, or recreational functioning). It is one of the best validated scales available (McDowell & Newell, 1987) and contains a section specificallyon communication functioning.

Patient Assessment of Own Functioning Inventory. The PAOFI was chosen for assessment because it effectively assesses participant perception of mental and language functioning, which could potentially influence job performance and satisfaction.

Beck Depression Inventory. The BDI assesses participant perception of depression and was chosen given the important relationship between depression and perceived quality of life (Beck et al., 1961; Karayandis, 1989).

Well Being Scale. The Well Being Scale assesses positive psychological functioning such as self-acceptance, purpose in life, personal growth, positive relations with others, autonomy and environmental mastery, all of which may potentially effect perceived job satisfaction and performance.

Data Collection and Analysis of Measures of Speech, Employment, and Quality of Life Outcomes

Measures of Speech Outcomes. Sound pressure level (SPL) and acoustic data were collected 3 times pre-treatment and 2 times immediately post-treatment in an IAC sound-treated booth to establish a stable baseline and account for potential performance variability observed in individuals with Parkinson disease (Canter, 1965; King, Ramig, Lemke, & Horii, 1994). To measure SPL, a Bruel and Kjaer 2230 sound level meter was placed 30cm from the participants' lips. To record acoustic data, a head mount microphone (AKG C410) was placed 8 cm from the lips. Both signals were recorded onto a Sony PC-108M eight channel digital audiotape recorder. Sound pressure level and acoustic data were collected while participants performed the following tasks: six maximum duration sustained "ah" vowel phonations, one reading of the "Rainbow Passage," and 30 seconds of conversational monologue.

Sound Pressure Level. For analysis of SPL, the SLM signal was pre-amplified and then digitized at 5000 samples per second into a VAX system computer through a 16-bit resolution DSC-200 A/D converter. The SLM signal was then analyzed using a custom built software analysis program which displayed the mean SPL of the signal in decibels. To avoid inclusion of nonvoiced or pausal segments in the analysis, data points below the noise floor 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 nonvoiced or pausal segments occurred. The percentage was input into the analysis program and the signal was displayed with the percentage cutoff line. This enabled the analyzer to deter-

mine if the correct percentage waschosen. After the correct percentage was determined, the program displayed the mean SPL of the signal.

Forced Vital Capacity. Forced vital capacity (FVC) was collected using a Collins wet spirometer (Model RS 2785). Participants were asked to take their deepest breath and blow out "as hard, fast and as long as you can." This task was repeated three times at the beginning and once at the end of each recording session and the best performance was taken as forced vital capacity.

Self-perceptual Ratings. Participant self-rated perceptual scales were completed by the participant immediately preceding the first two pre-treatment and first two post-treatment data collection sessions. Standard procedures for analysis of visual analogue scales (Boeckstyns & Backer, 1989) were used to obtain participant self-rated perceptual data.

Naive Listener Perceptual Ratings. To prepare tapes for the naive listener perceptual task, participant readings of the "Rainbow Passage" from the first pre-treatment recording and the first post-treatment recording were dubbed from the eight channel digital audio tape onto a separate digital audio tape. The acoustic signal from the eight channelaudio tape was input to a 2 channel Panasonic SV-3700 professional digital audio tape recorder. The signal from the 2 channel recorder was then input to a Hewlett-Packard 3466A digital multimeter. The signal was played through the multimeter and the voltage was adjusted using the record level on the 2 channel audio tape player to correspond to the decibel level of the calibration tone on the original 8 channel tape. The standard formula for converting SPL to volts was used to determine voltage level for the SPL levels. This procedure allowed each participant's passage to be recorded onto the 2 channel tape at the actual SPL level recorded at 30 cm during the data collection session.

The pre- and post-treatment readings of the "Rainbow Passage" of each participant were blocked (paired together) and recorded onto the tape in random order. A voice marker was put on the tape for each reading sample that allowed the listeners to know which participant and sample they were rating. For example, "Participant 1 sample 1" or "Participant 1 sample 2." The naive listeners did not know if the samples were pre-treatment or post-treatment readings. One participant was repeated on the tape to determine intra-listener reliability. This "blocking" procedure was used so the listeners could compare each participants' voice and speech pre- to post-treatment. It was thought that perceivable changes in the voice and speech of each participant could be representative of communication skill changes in his/her work environment.

To complete the perceptual study by naive listeners, the 2 channel tape was played to each individuallistener. The playback level was set for the actual SPL levels at 30cm. The listeners were seated approximately 3 feet from the loud speaker in an IAC sound treated booth. Each listener was seated at the same exact distance throughout the listening task. Listeners were instructed to listen to each sample of reading and rate the characteristics presented on the visual analogue scale. Explanation of the visual analogue scale was given until the listeners felt comfortable with the instructions and task. The listeners were allowed to listen to each reading sample as many times as needed to complete the rating scale accurately. Standard procedures for measuring visual analogue scales were used to analyze the naive listener task.

Measurement of Employment Outcomes. Measurement scales completed by the employer of each participant were sent to them in the month prior to speech therapy and within the first week after treatment. All other measurement scales and assessments were completed within the month prior to speech therapy and within the week after treatment. All were completed once both pre-treatment and immediately post-treatment.

Measure of Job Satisfaction-employer perspective. Each section of the Minnesota Satisfactoriness Scale calculates a raw score which is then converted to a percentile score based upon the type of labor or employment category for the participant (Gibson et al.,1970). A higher score on the MSS indicates greater satisfaction in job performance. Scores in the 75th percentile and above are considered in the highly satisfactory range, scores between 26 and 74 are considered average and scores below 25 are considered poor. Confidence bands are also reported for this scale which are equal to one standard deviation from the mean. The general portion of the scale is considered an overall score for job performance (Gibson et al., 1970).

Measure of job satisfaction-employee perspective. The answer choices on the Job satisfaction scale range between strongly agree to strongly disagree. The choice "strongly agree" indicates greatest job satisfaction. If the participant marked a choice on the scale post-treatment that was closer to the strongly agree answer, it was considered an improvement in job satisfaction for that question. If the participant marked a choice on the scale post-treatment that was closer to "strongly disagree," it was considered a reduction in job satisfaction. (Brayfield & Rothe, 1951; Price & Mueller, 1986).

Measure of communication within job-setting - employee perspective. The first portion of the Communication Profile for Speakers with Motor Speech Disorders was used to determine the participants' perception of difficulty in relation to communication. This section of the scale consists of 30 questions that specifically ask if the participant is having difficulty withcommunication in particular situations. For example, question 4 states "I find it difficult when I am explaining a new project to someone at work." The participant can respond by checking the agree, disagree or doesn't apply column. For all questions in this section of the questionnaire an agree response indicates difficulty. The

results of this scale were quantified by counting the number of agree, disagree, and doesn't apply responses in this first 30 item section. A reduction in the number of agree responses post-treatment would indicate less communication difficulty (Yorkston et al., 1994).

Measure of communication within job-setting - employer perspective. Standard procedures for measuring visual analogue scales were used to analyze the employer perceptual ratings of the participants (Kempster, 1984).

Measure of communication within job-setting - "blind" rater perspective. For the on-site observations a job site interaction data sheet was created by an employment consultant for the observations. The observer rated aspects of the participants' communication ability during a 45-60 minute period. The number of interactions between the participant and a coworker or customer/client, duration of the interactions, whether the interactions were work or non-work related and the quality of the speech were noted. The overall quality of the speech during the observation was rated by the observer on a scale of 1-5 (1=mild, 2=mild-moderate, 3=moderate, 4=moderate-severe, and 5=severe). In addition, using a scale adapted from Frey (1984), which assesses dysarthria on multiple levels, the participant's communication ability within the work site was rated as normal, impairment, disability, or handicap. A participant would be rated "impaired" if the speech demonstrated any impairment, "disability" if the speech impairment created reduced speech intelligibility, and "handicapped" if speech impairment reduced functional communication in situations such as vocational settings (Frey, 1984). Criteria of the on-site observations were reviewed with the rater until the rater was comfortable with the rating forms and techniques. Participant comments by the rater were also solicited on the participants' communication ability throughout each observation. The rater observed the participants one time pre-treatment and one time post-treatment. The rater did not know if the participant received treatment or not within the four week period between observations.

Measures of Quality of Life. Sickness Impact Profile Sickness Impact Profile assesses the extent to which the participants perceive their illness effects 12 different areas of their lives. For the purpose of this paper the areas of work, social interaction, communication, and overall score will be reported. Higher scores on the SIP indicate a greater impact of illness on the reported area (Bergner et al., 1981).

Patient Assessment of Own Function Inventory. Scoring for the Patient Assessment of Own Functioning Inventory is qualitative and results are reported anecdotally (Chelune et al., 1981).

Beck Depression Inventory. Scores for the Beck Depression Inventory are reported. A higher score indicates greater depression (Beck et al., 1961).

Well Being Scale. The Well Being Scale assesses 6 areas: autonomy, environmental mastery, personal growth, positive relations, purpose in life, and self acceptance. Scores are reported for the Well Being scale. Higher scores on the Well Being Scale indicate more normal functioning (Ryff, 1989).

Reliability. Twenty percent of the SPL and perceptual rating scale data were reanalyzed to determine intra-examiner reliability. Reliability ranged between 0.99 and 1.0 for all measures for all variables and participants. Intraparticipant reliability for participant self perceptual ratings ranged between 0.88 and 0.93 for all variables.

For the perceptual ratings from the naive listeners intra as well as inter-listener reliability were completed. Data from one of the 4 listeners were omitted from the study due to poor intralistener reliability. Intralistener reliability for the remaining three listeners ranged between 0.81 and 0.99. Interlistener reliability for these three listeners ranged between 0.82 and 0.92. Because reliability was high for these listeners, all three ratings were combined for one mean and standard deviation score for each of the rated variables for each participant.

Statistical Design. Due to the variety of measurements used to determine pre- to post-treatment changes, results are reported using various methods. To determine pre- to post speech treatment changes in the measures of SPL, participant self-perceptual ratings, naive listener ratings, and the MSS, a comparison of the means of each participant's results were made. This method of statistical comparison is considered appropriategiven the small and heterogeneous sample and the irreversible nature of the speech treatment program (Kratochwill & Levin, 1992). Differences in the pre- to post-treatment means were considered noteworthy if the mean of the two post-treatment sessions exceeded the mean of the three pre-treatment sessions, where applicable, by $\pm$ 1 standard deviation. This criteria is considered meaningful and acceptable and is analogous to a statistical measurement of a large size effect (Cohen, 1988). Application of inferential statistical techniques is arguable for case studies (Barlow & Hersen, 1984; McReynolds & Kearn, 1983) given the limited sample size.

The Communication Profile for Speakers with Motor Speech Disorders, the Beck Depression Inventory, and the Well Being Scale provide a single number result. The employer perceptual ratings and the Sickness Impact Profile provide a pre- to post-treatment percentage point. Standard deviations for these 5 measures could not be calculated because the measurements were administered only 1 time pre-treatment and 1 time post-treatment. The pre- to post-treatment directional changes for these measurements are reported.

The Job Satisfaction Scale, portions of the On-site observation, and the PAOFI all have subjective results. Pre- to post-treatment subjective changes are reported on these measurements.

Table 3.
Means and standard deviations (in parentheses) for the speech variables sound pressure level (SPL), forced vital capacity (FVC), and participant self perceptual ratings. Means are listed for each participant for the 3 pre-treatment sessions, for the 3 pre-treatment sessions combined (total mean), the two post-treatment sessions and the 2 post-treatment sessions combined (total mean).

Participants	Pretreatment 1	Pretreatment 2	Pretreatment 3	Pretreatment total mean (sd)	Post-treatment 1	Post-treatment 2	Post-treatment total mean (sd)
Participant 1							
SPL (dB) 30 cm							
Sustained Phonation	70.28 (3.30)	69.24 (2.19)	69.60 (2.08)	69.71 (0.53)	81.86 (1.40)	81.23 (1.80)	81.55 (0.45)
Rainbow Passage	69.88 (2.41)	70.12 (2.08)	71.58 (3.02)	70.53 (0.92)	76.93 (2.88)	80.62 (2.83)	78.78 (2.61)
Monologue	69.07 (2.94)	69.88 (1.80)	70.14 (3.86)	69.70 (0.56)	76.13 (2.92)	78.54 (2.02)	77.34 (1.70)
FVC (L)	3.3	3.3	3.2	3.27 (0.06)	3.6	3.4	3.5 (0.14)
Self perceptual rating scale (%)							
Loudness	27	29	DNT*	28.00 (1.41)	65	79	72.00 (9.90)
Overall intelligibility	26	30		28.00 (2.83)	83	82	82.50 (0.71)
Participation in conversation	54	71		62.50 (12.02)	93	89	91.00 (2.83)
Initiation of conversation	37	60		48.50 (16.26)	80	86	83.00 (4.24)
Participant 2							
SPL (dB) 30cm							
Sustained Phonation	74.52 (1.20)	75.62 (1.30)	74.65 (0.36)	74.93 (0.60)	83.42 (0.67)	81.15 (2.16)	82.29 (1.61)
Rainbow Passage	67.95 (2.34)	69.88 (2.79)	71.40 (3.09)	69.74 (1.73)	81.72 (3.77)	81.07 (3.29)	81.40 (0.46)
Monologue	69.28 (3.21)	71.06 (4.39)	74.61 (3.05)	71.65 (2.71)	78.64 (2.69)	80.10 (2.56)	79.37 (1.03)
FVC (L)	4.3	4.2	4.1	4.20 (0.10)	4.1	4.3	4.20 (0.14)
Self perceptual rating scale (%)							
Loudness	30	39	DNT	34.50 (6.36)	70	72	71.00 (1.41)
Overall intelligibility	37	47		42.00 (7.07)	76	60	68.00 (11.31)
Participation in conversation	36	35		35.50 (0.71)	79	62	70.50 (12.02)
Initiation of conversation	37	36		36.50 (0.71)	80	64	72.00 (11.31)
Participant 3							
SPL (dB) 30cm							
Sustained Phonation	62.55 (1.33)	61.29 (1.38)	61.27 (0.76)	61.70 (0.73)	82.04 (0.82)	81.36 (1.14)	81.70 (0.48)
Rainbow Passage	63.75 (2.36)	61.88 (2.76)	61.79 (2.83)	62.47 (1.11)	72.74 (4.05)	74.49 (2.26)	73.62 (1.24)
Monologue	60.19 (3.35)	58.76 (4.67)	60.26 (2.37)	59.74 (0.85)	68.85 (3.70)	72.29 (3.33)	70.57 (2.43)
FVC (L)	2.9	2.8	2.8	2.83 (0.06)	2.7	2.8	2.75 (0.07)
Self perceptual rating scale (%)							
Loudness	46	51	DNT	48.50 (3.54)	70	66	68.00 (2.82)
Overall intelligibility	58	69		63.50 (7.78)	78	77	77.50 (0.71)
Participation in conversation	51	65		58.00 (9.90)	77	83	80.00 (4.20)
Initiation of conversation	37	59		48.00 (15.56)	65	70	67.50 (3.54)

* Did not test - this scale was not administered pre-treatment 3.

Results

Measures of Speech Outcomes. Sound Pressure Level. As shown in Table 3, all three participants improved on measures of SPL during sustained vowel phonation, reading, and the 30 second monologue pre- to post-treatment. Participant 1 increased mean SPL pre- to post-treatment 11.84 dB for sustained phonation, 8.25 dB for reading and 7.64 dB during conversational monologue. Participant 2 increased mean SPL pre- to post-treatment 7.36 dB for sustained phonation, 11.66 dB for reading and 7.72 dB during conversational monologue. Participant 3 increased mean SPL pre- to post-treatment 20.00 dB for sustained phonation, 11.15 dB for reading, and 10.83 during conversational monologue. All post-treatment increases in SPL for all participants exceeded the pre-treatment standard deviation range.

Forced Vital Capacity. Changes in measures of forced vital capacity for all the participants pre- to post-treatment were negligible (Table 3).

Self-Perceptual Ratings. As shown in Table 3, all participants self-rated improvement pre- to post-treatment on the variables ofloudness, overall intelligibility, participation in conversation, and initiation of conversation. Participant 1 rated a mean improvement from the 2 pre-treatment sessions to the 2 post-treatment sessions of 44 (loudness), 54.5 (overall intelligibility), 28.5 (participation in conversation), and 34.5 (initiation of conversation) percentage points. Participant 2 rated a mean improvement of 36.5 (loudness), 26 (overall intelligibility), 35 (participation), and 35.5 (initiation) percentage points. Participant 3 rated a mean improvement of 19.5 (loudness), 14 (overall intelligibility), 22 (participation), and 19.5 (initiation) percentage points. All post-treatment rated improvements for all participants

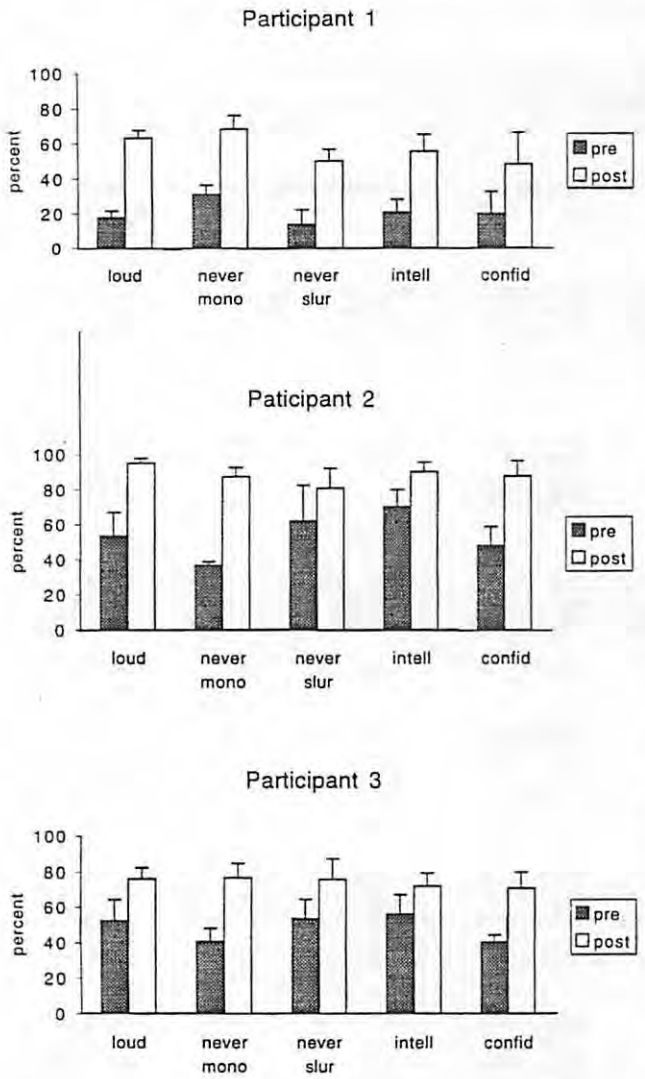


Figure 1. Naive listener ratings of always loud, never monotone, never slur, overall intelligibility, and confidence in speaking for Participants 1, 2, and 3.

exceeded the pre-treatment standard deviation range for each perceptual variable.

Naive Listener Perceptual Ratings. As seen in Figure 1, naive listeners rated all three participants improved on the variables always loud enough, never monotone, never slurs, overall intelligibility, and confidence in speaking. Listeners rated a mean improvement for Participant 1 of 45.33 (loudness), 37 (monotone), 36 (never slur), 34.33 (overall intelligibility), and 27.99 (perceived confidence in speaking) percentage points. Participant 2 was rated a mean improvement of 41.67 (loudness), 50.66 (never monotone), 18.67 (never slur), 19.67 (overall intelligibility), 39.67 (perceived confidence in speaking) percentage points and Participant 3 was rated a mean improvement of 23.67 (loudness), 36 (never monotone), 22 (never slur), 16 (overall intelligibility), and 30.34 (perceived confidence in speaking) percentage points. All post-treatment mean improvements across all variables for all participants exceeded the pre-treatment standard deviationranges (Figure 1) except for the variable "never slur" for Participant 2.

Table 4.
Percentiles and confidence bands (in parentheses) for the Minnesota Satisfactoriness Scale as rated by the employers of Participants 1 and 3 pre- and post-treatment. Higher scores indicate higher employer satisfaction.

Scale category	Participant 1 Pre-treatment	Participant 1 Post-treatment	Participant 3 Pre-treatment	Participant 3 Post-treatment
Performance	55 (42-63)	75 (66-85)	80 (69-87)	85 (76-96)
Conformance	99 (90-99)	99 (90-99)	80 (74-86)	99 (84-99)
Dependability	99 (70-99)	99 (70-99)	99 (71-99)	99 (71-99)
Personal adjustment	75 (65-85)	85 (74-99)	50 (37-56)	80 (74-99)
General	80 (73-87)	91 (86-94)	73 (65-79)	94 (91-98)

Measures of Employment Outcomes. Measure of Job Satisfaction - employer perspective - MSS. The employers for Participants 1 and 3 completed this scale pre- and post-treatment. Participant 2 did not want his employer to partake in this portion of the study. As shown in Table 4, the employer of Participant 1 rated pre- to post-treatment improvement from the 55th to the 75th percentile for the category performance, rated stability in the 99th percentile for conformance and dependability, rated improvement from the 75th to the 85th percentile for personal adjustment, and improvement from the 81st to the 91st percentile for the "general" category pre- to post-treatment. The employer of Participant 3 rated pre- to post-treatment improvement from the 80th to 85th percentile for performance, 80th to 99th percentile for conformance, rated stability at the 99th percentile for dependability, improvement from the 50th to the 80th percentile for personal adjustment and improvement from the 73rd to the 94th percentile for the "general" category. The improvements generally exceeded the standard deviation (see confidence bands) for each category (see Table 4).

Measure of Job Satisfaction - employee perspective - Job Satisfaction Scale. Table 5 indicates pre- to post-treatment changes the participants rated for each category on this scale. All participants rated some improvement on this scale. Participant 1 rated improvement in the categories "enthusiasm about job" and "satisfaction with job", a decrease in the category "would not consider taking another job" and no change in the categories "enjoyment in job," "like job more than the average worker," and "seldom bored with job." Participant 2 rated an improvement in "enjoyment in job," "enthusiasm about job," and "would not consider taking another job" and no change in the categories "satisfaction with job," "like job more than the average

Table 5.
Reported changes pre- to post-treatment on the Job Satisfaction Scale for Participants 1, 2, and 3. A plus sign indicates improvement, a minus sign indicates a reduction, and a blank indicates no change for each category of job satisfaction.

Category	Participant 1	Participant 2	Participant 3
Enjoyment in job		+	+
Enthusiasm about job	+	+	+
Satisfaction with job	+		+
Like job more than the average worker			+
Seldom bored with job			+
Would not consider taking another job	-	+	+

worker," and "seldom bored with job." Participant 3 rated improvement post-treatment in all categories: "enjoyment in job," "enthusiasm about job," "satisfaction with job," "like job more than the average worker," "seldom bored with job," and "would not consider taking another job."

Measure of the role of communication - Communication Profile for Speakers with Motor Speech Disorders. Pre-treatment, Participant 1 reported 30/30 agree responses, indicating difficulty with all areas presented on this portion of the Communication Profile. Post-treatment this participant indicated 0/30 agree responses, indicating no difficulty with the areas presented on the scale. Participant 2 reported a total of 18/30 agree responses pre-treatment and 4/30 agree responses post-treatment. Participant 3 reported 22/30 agree responses pre-treatment and 9/30 agree responses post-treatment. These changes suggest that all participants had less difficulty communicating in a variety of situations, including employment, post-treatment. For example, Participant 1 agreed he had difficulty talking to a coworker in the officepre-treatment but reported no difficulty with this type of communication situation post-treatment. Participant 2 reported difficulty while at a meeting talking with several other people pre-treatment but no difficulty with this type of communication situation post-treatment and Participant 3 reported difficulty explaining a new project with someone at work pre-treatment, but no difficulty with this type of communication post-treatment.

Employer perceptual ratings. Employer perceptual ratings were completed for Participants 1 and 3. Participant 2 did not want his employer to partake in this apect of the study. The employer of Participant 1 rated an improvement of 47 (loudness), 47 (overall intelligibility), 64 (participation in conversation), and 41 (initiation of conversation) percentage points pre- to post-treatment. The employer of Participant 3 rated a change of 3 (loudness), -2 (overall intelligibility), -10 (participation in conversation) and -31 (initiation of conversation) percentage points pre- to post-treatment.

On-site observations. Pre-treatment, Participant 1 was noted anecdotally by the "blind" rater to be "very difficult to understand on phone" when she called to set-up the pre-treatment observation. The rater noted that effective communication with Participant 1, even after he repeated himself 3-4 times, was difficult. The pre-treatment on-site observation was conducted on 4 interactions for a total of 3.5 minutes ofcontinuous speech. The participant's speech severity was rated a 5 (severe) during this observation. In addition, his speech was classified as "handicap" (Frey, 1984). Post-treatment, the "blind" rater noted a very noisy environment on the day of the observation. The observation was conducted on 7 interactions for total of 6.33 minutes of continuous speech. Speech severity was rated between 2 and 3 (mild-moderate) on the 5 point scale and classified as "handicap" based on the participant's continued loss of his original job position. During the post-treatment observation, the "blind" rater commented the participant was "much easier to understand" than during the first observation.

The pre-treatment observation for Participant 2 was conducted on a speech sample of 13 interactions for total of 20 minutes of continuous speech. Pre-treatment the speech severity for Participant 2 was rated 1 (mild). His speech was rated "impaired" (Frey, 1984). The post-treatment observation was conducted on 5 interactions for total of 9.25 minutes of continuous speech. His speech severity was rated 1 (mild) and his speech was rated "impaired" (Frey, 1984). The rater reported anecdotally improvement in the participant's overall loudness, but stated the participant was still having difficulty with voice quality.

The pre-treatment observation for Participant 3 was conducted on a speech sample of 21 interactions for total of 28 minutes of continuous speech. Her speech severity was rated 2 (mild-moderate) and her speech was rated "impaired" (Frey, 1984). Pre-treatment, the rater noted the participant's work environment was very noisy. The post-treatment observation was conducted on 16 interactions for atotal of 19.5 minutes of continuous speech. The rater noted again that the participant's work environment was very noisy. Her speech severity was rated a 1 (mild) while her speech was rated within normal limits (Frey, 1984). Post-treatment the rater noted increased loudness in the voice of Participant 3 as well as the ability to self-monitor and correct moments of reduced loudness. The rater noted, anecdotally, she "did not recognize the participant's voice on the phone" when she called to set-up the post-treatment observation due to the loudness changes in the speech of Participant 3 post-treatment.

Measures of Quality of Life. Sickness Impact Profile. Scores for the Sickness Impact Profile (SIP) are reported pre- and post-treatment in Table 6 (following page). Both pre- and post-treatment, Participant 1 reported a slight impact of illness on his life overall, with the most affected areas as communication and social interaction. However, post-treatment he reported his illness had less impact in these areas (32.14% pre-treatment; 9.66% post-treatment). In addition, pre-treatment the participant re-

Table 6.
Pre- to post-treatment scores for the measures Sickness Impact Profile (SIP), Beck Depression Inventory, and Well Being Scale for Participants 1, 2, and 3.

Participants	Pre-treatment	Post-treatment
Participant 1		
SIP*		
work	20.39	10.68
social interaction	32.14	9.66
communication	32.14	9.66
overall score	8.28	4.92
Beck Depression Inventory**	5	6
Well Being Scale***		
Autonomy	64 (good)	65 (good)
Environmental Mastery	64 (normal)	71 (good)
Personal Growth	55 (below average)	65 (normal)
Positive Relations	56 (below average)	66 (normal)
Purpose in Life	66 (normal)	73 (good)
Self Acceptance	62 (normal)	67 (normal)
Participant 2		
SIP		
work	0	0
social interaction	0	0
communication	0	0
overall score	0	0
Beck Depression Inventory	0	0
Well Being Scale		
Autonomy	65 (normal)	68 (normal)
Environmental Mastery	71 (good)	79 (good)
Personal Growth	71 (good)	73 (good)
Positive Relations	70 (good)	76 (good)
Purpose in Life	75 (good)	82 (high)
Self Acceptance	71 (good)	82 (high)
Participant 3		
SIP		
work	0.00	0.00
social interaction	24.34	2.48
communication	18.48	18.48
overall score	15.04	8.77
Beck Depression Inventory	13	11
Well Being Scale		
Autonomy	58 (normal)	69 (normal)
Environmental Mastery	68 (normal)	64 (normal)
Personal Growth	69 (normal)	75 (normal)
Positive Relations	69 (normal)	71 (normal)
Purpose in Life	67 (high)	61 (normal)
Self Acceptance	66 (normal)	72 (normal)

* Higher scores indicate a greater impact of illness.
** Higher scores indicate greater depression.
*** Higher scores indicate more normal functioning.
General descriptions of functioning scores are in parentheses.

ported an impact of his illness on the area of "work" with less impact post-treatment (20.39% pre-treatment; 10.68% post-treatment). Pre-treatment as well as post-treatment, Participant 2 reported no significant impact of his illness on any aspects of his life. Participant 3 reported her illness impacted her communication and social interactions pre-treatment. Post-treatment she reported less impact of her disease on social interaction (24.34% pre-treatment; 2.48% post-treatment), but no change in other areas .

Patient Assessment of Own Functioning Inventory. For the Patient Assessment of Own Functioning (PAOFI) scale, Participant 1 reported, anecdotally, significant difficulty with pronouncing words and writing pre-treatment. Post-treatment, Participant 1 reported the same writing problems but less difficulty with pronunciation. Participant 2 reported no significant problems pre- or post-treatment on the PAOFI. Participant 3 reported difficulties in the areas of writing, spelling, and performing tasks with her left hand pre-treatment on the PAOFI. This participant reported no change in these areas post-treatment.

Beck Depression Inventory. On the Beck Depression Inventory (BDI) the scores of both Participant 1 and 2 were within normal limits pre- and post-treatment. Participant 3 indicated mild depression pre- and post-treatment on the BDI (Table 6).

Well Being Scale. Scores for the Well Being Scale are reported in Table 6. Pre-treatment, the scores for Participant 1 were within normal limits for all areas of this scale except personal growth (55; below average) and relations with others (56; below average). Post-treatment, the scores for Participant 1 were within normal limits for all areas. Participant 2 and 3 scored within normal limits to the upper end of normal limits for allareas of this scale pre-treatment and remained in the same range post-treatment.

Discussion

The purpose of this study was to determine the effects of intensive voice treatment (Lee Silverman Voice Treatment [LSVT]) on employment satisfaction and performance in individuals with Parkinson disease. The findings reported here are an initial step in determining the effects of speech rehabilitation on employment outcomes for individuals with Parkinson disease.

All three participants improved pre- to post-treatment for the variables SPL during sustained phonation, reading and conversational monologue, and self-rated improvement on the perceptual variables loudness, overall intelligibility, participation and initiation of conversation. In addition, pre- to post-treatment changes for all three participants were documented by three naive listeners for the perceptual variables loudness, monotonicity, slurring, overall intelligibility and confidence in speaking. These data suggest that all participants improved their overall communication functioning pre- to post-treatment. These findings are consistent with previous reports of the effectiveness of the LSVT (Ramig et al., 1995; Ramig et al., 1996).

In addition to the post-treatment improvements documented in the measures of speech outcomes, allthree participants demonstrated some pre- to post-treatment improvement on the measures of employment outcomes from the employer, employee, or "blind" rater perspectives. The employer of Participant 1 rated him improved on the Minnesota Satisfactoriness Scale pre- to post-treatment for the categories performance, personal adjustment, and general. In addition, the employer of Participant 1 rated him improved

post-treatment on the perceptual variables of always loud enough, overall intelligibility, initiation of conversation, and participation in conversation. This finding suggests this employer was more satisfied with the overall performance and communication abilities of the participant at the employment setting post-treatment. The employer of Participant 3 rated her improved after treatment on the Minnesota Satisfactoriness Scale for the categories of performance, conformance, personal adjustment, and general, but rated relatively no improvement post-treatment on the perceptual variables. The SPL data, self-perceptual ratings and naive listener data contradict the employer's perceptual ratings as does Participant 3's report that her employer requested she read stories out loud to the class more often after treatment. It is difficult to determine why the employer rated her worse post-treatment on certain perceptual variables. It may be that completing the perceptual scales pre- and post-treatment increased the employer's awareness of the overall communicative ability of the participant, thus making the employer a more critical evaluator post-treatment.

Job satisfaction results from the employee perspective suggested all participants improved in this area post-treatment. Participant 3 had the greatest amount of improvement post-treatment in reported jobsatisfaction, while Participant 1 had the least. Participant 1 may have had the least amount of change in job satisfaction due to his continued loss of original job position post-treatment and his desire to regain that position. During the course of therapy it became evident that Participant 1 was upset over losing his original job as lead airplane mechanic. However, it also became apparent that communication difficulties were not the only reason for loss of this position. The participant was having considerable amount of physical involvement from his PD, by his own report, which prevented him from completing the duties of his original job. These findings suggest that improved job satisfaction post-treatment may vary among individuals and can be associated with other aspects of Parkinson disease.

For communication ability within the job setting (Communication Profile), all three participants rated improvement after treatment. In addition, all participants reported anecdotally that increased communicative effectiveness at the employment setting improved interactions with coworkers, customers/clients, and friends at work as well as "feeling more confident when speaking" post-treatment. For example, Participant 1 reported contemplating reapplying for his original job position because of his improved communication skills and increased confidence in his communication skills. Participant 2 reported his voice was less fatigued at the end of the day, thus enabling him to complete more difficult and complex phone calls anytime of the day rather than only in the morning. Participant 3 reported more confidence speaking in large groups. These data suggest that improved communication ability increased the confidence of each participant and their ability to perform necessary communication tasks effectively within their respective employment sites.

Finally, the speech and voice skills of all three participants were rated improved by the "blind" rater during the post-treatment on-site observations. Although the "blind" rater continued to rate the speech of Participant 1 "handicapped" post-treatment, she stated that the speech of Participant 1 was "significantly improved and much more intelligible" during the second on-site observation. He was still rated "handicapped" due to his continued loss of original job position. After treatment, the speech of Participant 2 remained rated "impaired" due to voice quality issues, although the rater noted that the speech of Participant 2 was louder after treatment. The speech of Participant 3 was rated improved to within normal limits post-treatment. These findings suggest that each participant was able to carry-over and utilize effectively, treatment gains into the employment setting.

The quality of life measures revealed that Participants 1 and 3 reported less impact of their illness on their social interactions after treatment. In addition, Participant 1 reported less impact on his communication and work as well as improved pronunciation. It is difficult to determine why Participant 2 and 3 did not rate an impact of their illness on their work on the Sickness Impact Profile (SIP) pre-treatment even though they reported during the pre-treatment interviews that their illness effected their ability to communicate while at work. One explanation is that the questions that address work on this profile are geared towards a more severe disability than was apparent in these participants (e.g., "I work for only short periods of time or take frequent rests"). Since this scale is a yes or no scale, it may not be sensitive enough for an individual with mild changes in communication or employment as a result of their disease.

No other changes in quality of life measures were noted for these participants post-treatment. As with the SIP, it could be that the measures used in this study were not sensitive enough to psychosocial changes immediately following speech therapy or it could be that not enough time had passed for improved communication skills to have impacted psychosocial functioning significantly. Long-term follow-up of changes in psychosocial measures in relation to speech rehabilitation and employment is currently being examined to further investigate this issue. However, the findings from the SIP suggest that improved post-treatment psychosocial functioning may be attributed to improved communicative functioning. Further research is needed in this area.

It has been demonstrated that difficulty with communication can be a barrier to employment (McCann, 1992). However, documenting the effects of speech treatment on employment outcomes can be challenging. Until now, no studies have been conducted that examine post speech treatment employer ratings on job satisfaction and commu-

nication skills in individuals with Parkinson disease. One study that included employer ratings investigated the effects of stuttering therapy on employment. The authors reported that employer perceptions of the employees (individuals who stutter) were enhancedpost-treatment (Craig & Calver, 1991). Our findings are consistent with those of Craig and Calver (1991). Improving employer perceptions of employees who have speech/voice difficulties should increase the ability of the individual to compete with nondisabled individuals in their vocations or other vocations that require effective communication skills. Furthermore, having employers rate employees is an initial step to include employers in the rehabilitative process. Including the employer in this process should improve employers' understanding of communication disabilities associated with PD as well as other chronic neurological conditions.

Collecting non-biased objective and subjective data within the employment setting following speech treatment is difficult. In the current study the employers were aware that the employees were receiving speech treatment and may have been biased to the participants' changes. Further investigation is needed that includes "blinded" employers. Moreover, because each participant was aware of the "blinded" on-site observer, speech performance may have been enhanced during the observation time. Regardless, the on-site observations in this study were an initial attempt to determine effects within the employment setting and are an important aspect of this study. These on-site observations provided a baseline of each participant's speech and voice impairment and its effect on employability as well as an accurate assessment of each individual's speech and voice functioning within the employment setting. The on-site observations in this study suggested that each participant was communicating more effectively post-treatment within their respective employment settings.

Finally, the bias resulting from the feelings of each participant surrounding speech rehabilitation is another challenge for determining employment outcomes following speech rehabilitation. Because each participant invests a significant amount of time in therapy, they may rate their speech "better" regardless of actual outcome (placebo effect). If the participant reports a feeling of greater confidence with his/her speech, this may enhance or improve employee performance or confidence of performance. However, in this study the objective data (SPL) supported the perceived improvement of each participant, suggesting overall improvement of communicative effectiveness within the job setting. To address and understand these biases more clearly, our ongoing research is examining a LSVT treated group vs. a placebo treatment group.

Measuring quality of life is extremely important when determining speech therapy effects on employability, since chronic disease can adversely effect psychosocial functioning (Swanson, Cronin-Stubbs, & Sheldon, 1989) and occupational and financial status adversely effect life satisfaction (Ben-Sira, 1986). Although psychosocial functioning is important to the success of employability (Rucker & Metzler, 1995), few vocational rehabilitation studies have addressed this issue pre- to post vocational treatment (Cook, 1983; Riipinen, Hurri, & Alaranta, 1994). If psychosocial functioning is negatively affected it can disrupt work leading to increased stress levels which leads to lower self-esteem and sense of mastery (Pearlin, Menaghan, Lieberman, & Mullan, 1981). By examining quality of life ineach participant, this study attempted to determine the effect of communicative impairment on overall disability of an individual within their employment setting. The results suggest that increased effectiveness of communication and improved satisfaction with employment can positively effect quality of life.

Many individuals with PD report retiring early as a result of their disease (Mutch et al., 1986). Because individuals with Parkinson disease can live for an extended period of time after onset of disease, it is important to help the individuals of employment age maintain their employment status. Maintaining their employment status could potentially save thousands of dollars annually in social security and disability payments. Currently, Social Security Disability Insurance (SSDI) and Supplemental Security Income (SSI) cost the government an estimated 30 - 50 billion dollars every year (Richardson, 1994; Straaton et al., 1995). Keeping individuals with PD employed as long as possible is an economic gain for society (Teasdale, Christensen, & Pinner, 1993), not only in taxpayer dollars, but in the individual himself paying taxes and maintaining financial independence. Moreover, it has been shown that individuals who begin to receive SSDI and SSI benefits are more difficult to rehabilitate back into employment (Berkowitz, 1981; Better, Fine, Simison, Doss, Walls, & McLaughlin, 1979; Kunce, Cope, Miller, & Lesowitz, 1972; Straaton et al., 1995; US general accounting office, 1987; Walls, Maason, & Werner, 1977). In addition, because the positive effects of the LSVT have been shown to last for up to 2 years following treatment (Ramig, Countryman, Pawlas, Dromey, & Fox, 1996), it could potentiallyextend the employment of individuals with PD. Overall, vocational rehabilitative programs are less expensive to the government than disability maintenance programs (Hill & Wehman, 1983; Peters, Large, & Elkind, 1992; Straaton et al., 1995). Thus, it is extremely important to begin speech therapy (LSVT) for individuals with PD as soon as possible. Ongoing research in our laboratory is addressing this issue of early intervention.

Adequate employer support, when administering speech treatment to an employed individual, is important for a vocational rehabilitation program to be the most effective (Galvin, 1986). If an employer seems unwilling to make the necessary adjustments so an employee may receive this

treatment, the program may lose some of its effectiveness. However, the ADA provides modification of work schedules for medical appointments (Walk, Ahn, Lampkin, Nabizadeh, & Edlich, 1993), including speech therapy appointments. The employer may need to be informed that disability as a result of speech impairment is included under the ADA (Walk et al., 1993), therefore ensuring the right of the employee to receive speech therapy to help maintain employability. In addition, the potential gains to the employer should outweigh the potential inconvenience of the speech therapy program. The two employers involved in this study were very supportive of their employees' attempts to increase communicative effectiveness. Both participants were allowed the necessary time to complete the 1 month speech treatment program.

Speech rehabilitation and changes in satisfaction of employment and/or employability is an important topic within the realm of speech-language pathology (ASHA, 1992). Research on the effects of speech treatment on employability in individuals with chronic neurological disease is lacking in the literature. Given the importance of functional outcome measurements now being addressed in this field, measuring employment gains before and after speech therapy is an excellent measure of functional outcomes. However, Storey and colleagues (1995) reported that SLPs rarely participate as members of the vocational rehabilitation support team. Speech-language pathologists should be aware that companies who hire or employ speech disabled individuals may be willing to shoulder some cost of rehabilitation. They may provide inservice training to promote and increase communication strategies and awareness of disabilities in the general work population (Storey, Ezell, & Lengyel, 1995), to ensure effective employee performance. Generalization of successful speech rehabilitation to an employment setting could potentially empower the disabled employee as well as increase the value of the speech-language pathologist in the employment setting.

In conclusion, these data suggest that the LSVT had an impact immediately post-treatment on the satisfaction of employment and job performance of each participant. Improving effective communication can positively impact satisfaction with employment from the employer and employee perspective. Typically, in the speech and vocational rehabilitation literature, outcomes of therapy are predicted by job vs. no job (Riipinen et al., 1994). However, this type of measurement is not appropriateor sensitive enough for the individual that is currently employed. This study provides positive preliminary, functional evidence of the effects of speech rehabilitation on employment in employed individuals with Parkinson disease. Given the high rate of speech impairment observed in individuals with PD as well as the high number of employed individuals with PD, this is an important area of investigation and report.

Acknowledgments

The authors would like to acknowledge the participants in this study, for their courage and perseverance. The research presented here was supported in part by OE-NIDRR Grant No. 8133G40108.

References

Aronson, A. E. (1985). Clinical Voice Disorders. New York: Thieme Inc.

Aronson, A. E. (1990). Clinical Voice Disorders. New York: Thieme-Stratton.

American Speech-Language Hearing Association (1992). Let's Talk: Employment rights of people with communication disabilities. ASHA, 34, 57-58.

Barlow, D. H., & Hersen, M. (1984). Single case experimental designs: strategies for behavior change. New York: Pergamon Press.

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.

Ben-Sira, Z. (1986). Disability, stress, and readjustment: the function of the professional's latent goals and affective behavior in rehabilitation. Social Science and Medicine, 23, 43-55.

Bergner, M., Bobbitt, 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.

Berkowitz, M. (1981). Disincentives and the rehabilitation of disabled persons. Annual Review of Rehabilitation, 2, 40-54.

Better, S. R., Fine, P. R., Simison, D., Doss, G. H., Walls, R. T., & McLaughlin, D. E. (1979). Disability benefits as disincentives to rehabilitation. Milbank Q, 57, 348-363.

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.

Brayfield, A. H., & Rothe, H. F. (1951). An Index of Job Satisfaction. Journal of Applied Psychology, 35, 307-311.

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.

Chelune, G. J., Heaton, R. K., & Lehman, R. A. W. (1986). Neuropsychological and personality correlates of patients' complaints of disability. In G. Goldstein and R.E. Tartar (eds.) Advances in Clinical Neuropsychology, 3, 95-126, New York: Plenum Press.

Cohen, J. (1988). Statistical power analysis for behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Assoc.

Cook, D. W. (1983). Postservice adjustment of former rehabilitation center clients: a longitudinal analysis. Rehabilitative Literature, 44, 194-200.

Craig, A. R., & Calver, P. (1991). Following up on treated stutterers: studies of perceptions of fluency and job status. Journal of Speech and Hearing Research, 34, 279-284.

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

Darley, F. L., Aronson, A. E., & Brown, J. R. (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, 462-496.

Dromey, C., Ramig, L. O., & Johnson, A. B. (1995). Phonatory and articulatory changes associated with increase vocal intensity in Parkinson disease: a case study. Journal of Speech and Hearing Research, 38 (4), 751-764.

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

Frey, W. D. (1984). Functional assessment in the '80s: A conceptual enigma, a technical challenge. In A. S. Halpern & M. J. Fuhrer (Eds.), Functional Assessment in Rehabilitation. Baltimore: Paul H. Brookes.

Fox, C., & Ramig, L. O. (1996) Sound pressure level in males and females with Parkinson disease. in review. American Journal of Speech-Language Pathology.

Galvin, D. E. (1986). Employer-based disability management and rehabilitation programs. Annual Review of Rehabilitation, 5, 173-215.

Gibson, D. L., Weiss, D. J., Dawis, R. V., & Lofquist, L. H. (1970). Manual for the Minnesota Satisfactoriness Scales. Minnesota Studies in Vocational Rehabilitation, 27.

Heaton, R. K., Chelune, G. J., & Lehman, R. A. (1978). Using neuropsychological and personality testsfo assess the likelihood of patient employment. Journal of Nerve Men Disorders, 166, 408-416.

Hill, M., & Wehman, P. (1983). Cost benefit analysis of placing moderately and severely handicapped individuals into competitive employment. Journal of the Association of the Severely Handicapped, 8, 30-32.

Karayanidis, F. (1989). Parkinson's disease: A conceptualization of neuropsychological deficits within an information-processing framework. Biological Psychology, 29,149-179.

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

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 (1), 29-42.

King, J., Ramig, L. O., & Lemke, J. (1996). Communication ability in Parkinson disease in relation to employment satisfaction: a survey. unpublished manuscript.

Kratochwill, T., & Levin, J. R. (1992). Single-case research design and analysis: New directions for psychology and education. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kunce, J. T., Cope, C. S., Miller, D. E., & Lesowitz, N. (1972). Rehabilitation outcomes and the public assistance recipient. Rehabilitation Literature, 33, 204-206.

LaPlante, M. P. (1988). Data on disability from the National Health Interview Survey 1983-1985. An InfoUse Report. Washington DC. US National Institute on Disability and RehabilitationResearch.

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

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, 43, 47-57.

McCann, J. (1992). Neuro-rehabilitation and employment. Occupational Health, 44 (9), 260-262.

McDowell, I., & Newell, C. (1987). Measuring Health: a guide to rating scales and questionnaires. New York: Oxford University Press.

McReynolds, L. V., & Kearns, K. P. (1983). Single-subject experimental designs in communicative disorders. Baltimore: University Park Press.

Mohr, E., Juncos, Cox, C., Litvan, I., Fedio, P., & Chase, T. N. (1990). Selective deficits in cognition and memory in high-functioning Parkinsonian patients. Journal of Neurology, Neurosurgery and Psychiatry, 53, 603-606.

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.

Newnan, O.S., Heaton, R.K., & Lehman, R. A. W. (1978). Neuropsychological and MMPI correlates of patients' future employment characteristics. Perceptual and Motor Skills, 46, 635-642.

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

Parkinson's Education Program (PEP) Exchange (1984). IV:7.

Pearlin, L. I., Menaghan, E. G., Lieberman, M. A., & Mullan, J. T. (1981). The stress process. Journal of Health and Social Behavior, 22, 337-356.

Peters, J., Large, R. G., & Elkind, G. (1992). Follow-up results from a randomized controlled trial evaluating in- and outpatient pain management programmes. Pain, 51, 41-50.

Powell, D.H., Kaplan, E. F., Whitla, D., Weintraub, S., Catlin, R., & Funkenstein, H. H. (1993). MicroCog: Assessment of Cognitive Functioning Manual. San Antonio, TX: Psychological Corporation.

Price, J. L. & Mueller, C. W. (1986). Absenteeism and Turnover of Hospital Employees. Connecticut: Jai Press Inc.

Ramig, L. O. (1992). The role of phonation in speech intelligibility: A review and preliminary data from patients with Parkinson's disease. Intelligibility in Speech Disorders: Theory, Measurement and Management. Amsterdam: John Benjamin.

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., O'Brien, C., Hoehn, M., & Thompson, L. (1996). Intensive speech treatment for patients with Parkinson disease: short- and long-term comparison of two techniques. Neurology. in press.

Ramig, L. O., Countryman, S., Pawlas, A. A., Dromey, C., & Fox, C. (1996). Voice treatment in Parkinson disease: 2-year follow-up. A paper presented at the American Speech-Language Hearing Association annual convention. Seattle, WA.

Ramig, L. O., Countryman, S., Thompson, L., & Horii, Y. (1995). Comparison of two forms of intensive speech treatment for Parkinson disease. Journal of Speech and Hearing Research, 38, 1232-1251.

Ramig, L. O., Pawlas, A. A., & Countryman, S. (1995). The Lee Silverman Voice Treatment (LSVT): A practical guide for treating the voice and speech disorders in Parkinson disease. Iowa City (IA): National Center for Voice and Speech, University of Iowa.

Richardson, M. (1994). The impact of the American with disabilities act on employment opportunity for people with disabilities. Annual Review of Public Health, 15, 91-105.

Riipinen, M., Hurri, H., & Alaranta, H. (1994). Evaluating the outcome of vocational rehabilitation. Scandinavian Journal of Rehabilitative Medicine, 26, 103-112.

Rucker, K. S., & Metzler, H. M. (1995). Predicting subsequent employment status of SSA disability applicants with chronic pain. The Clinical Journal of Pain, 11, 22-35.

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. Journal of Personality and Social Psychology, 57 (6), 1069-1081.

Salmoni, A. W., Schmidt, R. A., & Walter, C. B. (1984). Knowledge of results and motor learning: a review and critical reappraisal. Psychological Bulletin, 95, 335-386.

Saxon, K. G., & Schneider, C. M. (1995). Vocal Exercise Physiology. San Diego, CA: Singular Publishing.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. Psychological Review, 82, 225-260.

Schmidt, R. A. (1988). Motor Control and Learning. Champaign, IL: Human Kinetic Publishers.

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

Shoulson, I. (1989). "Effect of Deprenyl on the progression of disability in early Parkinson's disease." New England Journal of Medicine, 321, 1364-1371.

Storey, K., Ezell, H., & Lengyel, L. (1995). Communication strategies for increasing the integration of persons in supported employment: a review. American Journal of Speech-Language Pathology, 4 (2), 45-54.

Straaton, K. V., Maisiak, R., Wrigley, J. M., & Fine, P. R. (1995) Musculoskeletal disability, employment, and rehabilitation. The Journal of Rheumatology, 22 (3), 505-513.

Streifler, M., & Hofman, S. (1984). Disorders of verbal expression in Parkinsonism. In: Hassler RG, Christ JF, eds. Advances in Neurology, 40, 385-393.

Swanson, B., Cronin-Stubbs, D., & Sheldon, J. A. (1989). The impact of psychosocial factors on adapting to physical disability: a review of the research literature. Rehabilitation Nursing, 14 (2), 64-68.

Teasdale, T. W., Christensen, A., & Pinner, E. M. (1993). Psychosocial rehabilitation of cranial trauma and stroke patients. Brain Injury, 7 (6), 535-542.

United States General Accounting Office. (1987). Social security: little success achieved in rehabilitating disabled beneficiaries. GAO/HRD, 88-11.

Walk, E. E., Ahn, H. C., Lampkin, P. M., Nabizadeh, S. A., & Edlich, R. F. (1993). Americans with disabilities act. Journal of Burn Care and Rehabilitation, 14 (1), 92-98.

Walls, R. T., Maason, C., & Werner, T. J. (1977). Negative incentives to vocational rehabilitation. Rehabilitation Literature, 38, 143-150.

Yorkston, K. M., Bombardier, C., & Hammen, V. L.(1994) Dysarthria from the viewpoint of individuals with dysarthria. Motor Speech disorders: Advances in Assessment and Treatment. J. A. Till, K. M. Yorkston, & D. R. Beukelman (eds). Baltimore: Paul Brooks Publishing. 19-36.

NCVS Status and Progress Report - 11
May 1997, 37-47

Supraglottal Hyperadduction in an Individual with Parkinson Disease: A Clinical Note

Jennifer Hicks M.A.
Bill Wilkerson Center Nashville, Tennessee
Stefanie Countryman M.A., CCC-SLP
Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts
Lorraine Olson Ramig Ph.D., CCC-SP
Department of Communication Disorders and Speech Sciences, The University of Colorado Boulder
Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts
Marshall E. Smith M.D.
The University of Colorado Health Sciences Center
Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts

Abstract

Recent treatment for voice problems associated with idiopathic Parkinson disease has primarily focused on increasing reduced vocal loudness and improving true vocal fold hypoadduction; common voice deficits observed in these individuals. This study presents an individual with reduced vocal loudness and supraglottic hyperadduction accompanying Parkinson disease and the outcome following a course of the Lee Silverman Voice Treatment (LSVT). Posttreatment observations included increased vocal loudness, decreased supraglottic hyperadduction, and improved intonation and overall voice quality when compared with pretreatment observations. These results suggest that in this individual, supraglottic hyperadduction was due to a secondary compensatory behavior resulting from mild true vocal fold hypoadduction that responded positively to adduction therapy (LSVT). This study also demonstrates the use of a continuum of observations ranging from functional ratings to physiological measures to evaluate the impact of intensive voice treatment and identify mechanisms underlying treatment-related change in an individual with Parkinson disease.

Introduction

Common neurological disorders, such as Parkinson disease, are often characterized by laryngeal abnormalities that result in voice disorders (Aronson, 1990; Brin, Fahn, Blitzer, Ramig, & Stewart, 1992; Darley, Aronson, & Brown, 1975; Logemann, Fisher, Boshes, & Blonsky, 1978). These disorders may contribute to reductions in vocal loudness, vocal quality and overall speech intelligibility (Aronson, 1990; Kent, Kent, Weismer, Sufit, Rosenbek, et al., 1990; Pawlas, Ramig, Countryman, O'Brien, & Hoehn, in submission; Ramig, 1992). Recently, Ramig and Scherer (1992) and Smith and Ramig (1995) developed an approach to planning behavioral treatment for voice disorders associated with neurological disease. This approach focuses on the disordered laryngeal function in combination with the neurological etiology and is designed to maximize improvement in overall intelligibility as efficiently as possible.

In this approach, the disordered voice is considered in relation to problems of adduction or phonatory instability (Ramig & Scherer, 1992). Disorders of hypoadduction are distinguished from those of hyperadduction. Disorders of hypoadduction are characterized by inadequate true vocal fold closure, reduced loudness, and breathy, hoarse vocal quality. Hypoadduction may accompany a variety of neurological disorders, but is typically associated with lower motor neuron (flaccid) disorders, multiple sclerosis (Beukelman, Kraft, & Freal, 1985; Darley, Brown, & Goldstein, 1979; Jensen, 1960), some cases of closed head injury (Vogel & Von Cramon, 1982; Von Cramon, 1981), and Parkinson disease (Hanson, Gerratt, & Ward, 1984; Perez, Ramig, Smith, & Dromey, 1996; Smith, Ramig, Dromey, Perez, Samandarii, 1995). The type and extent of vocal fold hypoadduction may be associated with the site and extent of the related neurological damage (Aronson, 1990).

Individuals with Parkinson disease may have mild to severe bowing of the true vocal folds or other forms of vocal fold incompetence (Aronson, 1990; Perez et al., 1996; Smith, Ramig, Dromey et al., 1995; Ward, Berci, & Calcaterra, 1977) during speech. As a result, many individuals with PD present with reduced loudness, a weak, breathy, or hoarse vocal quality, in addition to monotone of pitch and loudness, shortened breath groups and vocal tremor (Brin, et al., 1992; Darley et al., 1975; Pawlas et al., in submission; Ramig & Gould, 1986). Furthermore, individuals with PD may exhibit reduced respiratory volumes for speech resulting from limited thoracic excursion and difficulty coordinating the respiratory and laryngeal systems (Critchley, 1981; Dromey, Ramig, & Johnson, 1995; Ramig & Dromey, 1996; Solomon & Hixon, 1993).

In contrast, disorders of hyperadduction involve excessive true vocal fold adduction and are not typically characteristic of PD. Laryngeal hyperadduction may be found in all varieties of hyperkinetic movement disorders (Brin et al., 1992), but is commonly associated with upper motor neuron lesions, Huntington's disease, and dystonias (Aronson, 1990; Darley, Aronson, & Brown, 1969a, b). Hyperadduction occurs in different forms; it can be constant, as in the strain-strangled phonation of pseudobulbar palsy, or variable, as in the alternating adductor laryngospasm in spasmodic dysphonia (Aronson, 1990; Darley et al., 1969b). In some cases hyperadduction may be so severe as to result in complete termination of phonation or aphonia.

A second form of hyperadduction involves adduction of the ventricular (false) vocal folds (Arnold & Pinto, 1960; Aronson, 1990). The exact etiology of supraglottic hyperadduction has been debated (Arnold & Pinto, 1960; Jackson & Jackson, 1935; Roy, 1994; Von Hake, Ganzman, & Mauer, 1989), though it has been associated with laryngeal paralysis, cerebral disease, and cerebellar lesions (Arnold & Pinto, 1960; Von Hake et al., 1989). Supraglottal hyperadduction may also be a means of compensation for true vocal fold hypoadduction (Arnold & Pinto, 1960; Feinstein, Hilger, Szachowicz, & Stimson, 1987; Jackson & Jackson, 1935; Roy, 1994; Von Doersten, Izdebski, Ross, & Cruz, 1992; Woo, Casper, Colton, & Brewer, 1992; 1994).

For example, individuals with Parkinson disease may occasionally present with supraglottic hyperadduction (i.e. anterior-posterior foreshortening and ventricular hyperadduction or phonation) (Arnold & Pinto, 1960; Hanson et al., 1984; Von Hake et al., 1989; Smith, Ramig, Dromey, et al., 1995) and true vocal fold hypoadduction. Such observations are consistent with the opinion that ventricular fold adduction may result as a secondary compensatory mechanism in individuals with vocal fold hypoadduction. The combination of supraglottal hyperadduction and glottal hypoadduction in individuals with PD creates a clinical dilemma for the speech-language pathologist, who would typically treat individuals with Parkinson disease for a hypoadducted voice disorder.

However, if supraglottic hyperadduction is secondary compensation for inadequate vocal fold closure, therapy designed to improve true vocal fold adduction should be effective in reducing supraglottic hyperadduction and improving the voices of these individuals. An initial study by Smith, Ramig, Dromey and colleagues (1995) noted improved true vocal fold adduction with reduction in mild supraglottic hyperadduction following a course of the Lee Silverman Voice Treatment (LSVT) in individuals with Parkinson disease. This suggests that the mild supraglottic hyperadduction observed in these individuals may have been secondary compensation for an inadequate glottal source.

The LSVT was developed by Ramig and colleagues for remediation of voice and speech disorders associated with Parkinson disease (Countryman & Ramig, 1993; Countryman, Ramig, & Pawlas, 1994; Ramig, Bonitati, Lemke, & Horii, 1994; Ramig, Countryman, O'Brien, Hoehn, & Thompson, 1996; Ramig, Countryman, Thompson, & Horii, 1995) and other neurological disorders (Ramig, Countryman, Pawlas, & Fox, in press). The Lee Silverman Voice Treatment (LSVT) program is based upon the underlying vocal fold hypoadduction (Hanson et al., 1984; Smith, Ramig, Dromey, et al., 1995), reduced range of motion in laryngeal musculature, and poor respiratory drive typically observed in individuals with Parkinson disease. Therapy is designed to maximize overall speech intelligibility by focusing on increased or efficient true vocal fold adduction and improved coordination of the respiratory and laryngeal systems (Ramig & Dromey, 1996). Following conventional wisdom, the speech-language pathologist may determine this method of voice treatment inappropriate for an individual with PD with reduced loudness accompanied by moderate to severe supraglottal hyperadduction. There may be a concern that the adduction exercises would exacerbate the hyperfunctioning system. Furthermore, not all individuals with PD with reduced loudness have apparent vocal fold hypoadduction on laryngeal stroboscopic examination (Smith, Ramig, Dromey, et al., 1995). Again, conventional wisdom considers true vocal fold adduction exercises unsuitable for individuals with adequate vocal fold closure.

The purpose of this paper is to report speech and voice data from one individual with Parkinson disease who had a mild-moderately soft, hoarse voice, severe supraglottic hyperadduction, including ventricular fold vibration during soft, normal, and loud phonation and true vocal fold hypoadduction during soft phonation pretreatment. It was hypothesized that the LSVT would improve the individual's primary deficit (true vocal fold hypoadduction), reduce the need for secondary compensatory behavior (supraglottic hyperadduction) and result in improved loudness, intonation, and overall vocal quality posttreatment.

Method

Participant

The participant was diagnosed with idiopathic Parkinson Disease (IPD), and determined stable on his anti-Parkinson medication by a neurologist prior to entering this study. His medications did not change nor did he receive other therapies (i.e. physical, occupational etc.) during the study period. Furthermore, the participant had not received previous speech or voice treatment prior to the start of this program.

The participant was a 60 year-old male first diagnosed with IPD in December 1994. At the time of enrollment (April 1995) in the LSVT program, he was classified as stage III (moderate-severe) Parkinson disease (Hoehn & Yahr, 1967). His score on the motor section of the *Unified Parkinson Disease Rating Scale* ([UPDRS] Fahn, Elton, & members of the UPDRS development committee, 1987) was 33, indicating mild impairment (higher scores on the UPDRS indicate greater impairment; scores range from 0-108 for the motor section). Antiparkinson medication included selegiline hydrochloride (Eldyprel®). An initial speech mechanism exam revealed structure and function of speech musculature within normal limits. Perceptual speech and voice characteristics, as rated by the attending speech-language pathologist, included mild to moderately reduced loudness, monotone pitch, a hoarse vocal quality and abnormally low pitch during sustained phonation and conversational speech. The attending speech-language pathologist rated overall speech and voice impairment as mild to moderate. Pretreatment the participant and a family member reported his speech and voice deficits impacted the functionality and effectiveness of his communication. For example, during telephone and in-person conversations, the participant reported he had to repeat himself numerous times for the listener to hear and understand him. At the time of treatment, the participant was self-employed in a small consulting firm and typically spent approximately 1-2 hours speaking per day, usually on the phone. The participant reported no problems with chewing or swallowing.

Treatment

The LVST was administered to this participant. Treatment techniques focused on increasing vocal loudness and phonatory efficiency by targeting the hypothesized underlying laryngeal pathophysiology and optimizing phonatory and respiratory effort and coordination across the motor speech system using the global variable "loud" (Dromey et al., 1995; Schulman, 1989). The LSVT was specifically designed to maximize participant motivation for the speech tasks and facilitate immediate carryover of increased vocal loudness and vocal quality into functional communication. Specific techniques included optimizing vocal fold closure and efficiency through vocal isometric exercises and improving respiratory support by instructing the participant to inspire deeply and frequently before speaking loud ("thinking loud") and to speak on "top of the breath."

A typical session included a high number of successive repetitions of the following activities: maximum sustained vowel phonation (/a/), generating the highest and lowest fundamental frequencies the participant could obtain, and speech production tasks using the same phonatory and respiratory techniques used in sustained phonation. The high number of repetitions during the treatment as well as the intensiveness of therapy (4 times a week for 4 weeks) are consistent with theories of motor learning and skill acquisition (Schmidt, 1975; 1988) and muscle training (Saxon & Schneider, 1995). Treatment techniques also enhance the participant's overall "knowledge of results," which is critical in skill acquisition (Salmoni, Schmidt, & Walter, 1984). A tape recorder was used to provide feedback on vocal quality and loudness to the participant during each task. In addition, during the four week program, the participant was trained in self-monitoring of loudness and vocal quality (sensory calibration) and habituation of the increased loudness and improved vocal quality into conversational speech. At no point during therapy was the participant taught to sustain phonation or speak in a voice that was pressed or strained. The LSVT method has been described in further detail elsewhere (Ramig, Pawlas, & Countryman, 1995).

Measurement Variables and Rationale

Variables that assess a continuum of glottic and supraglottic functioning as well as objective and perceptual characteristics of the participant's speech and voice were chosen to document pre to posttreatment changes. Sound pressure level, mean fundamental frequency and its variability, maximum duration of sustained vowel phonation, electroglottographic data, perceptual ratings of the participant's speech and voice by experienced practitioners and videolaryngostroboscopic data are presented.

Sound Pressure Level. Because of the high frequency of occurrence of reduced loudness in individuals with Parkinson disease (Aronson, 1985; Canter, 1965; Critchley, 1981; Darley et al., 1969a, b; Fox & Ramig, in press), the impact of reduced loudness on speech intelligibility (Ramig, 1992), and its close relationship to vocal fold adduction and function, sound pressure level (SPL) during sustained phonation, reading, and monologue was measured.

Maximum Duration of Sustained Vowel Phonation. Maximum duration of sustained vowel phonation was chosen for measurement because of its relationship to laryngeal (Yanagihara, Koike, & Leden, 1966) and respiratory functioning (Boone, 1977). Stimulation of these two mechanisms is important for training respiratory and laryngeal coordination and improving overall vocal loudness and functioning in individuals with PD (Ramig, 1992).

Mean Fundamental Frequency and Fundamental Frequency Variability. Mean fundamental frequency and its variability (semitone standard deviation [stsd]) were chosen for measurement to evaluate whether changes in phonatory and respiratory effort and vocal intensity as well as the treatment exercises generalized to other aspects of the participant's speech. Changes in fundamental frequency often accompany changes in vocal intensity (Jacob, 1968; Linville & Korabic, 1987). Positive changes in these two variables may indicate better intonation and overall vocal quality in this participant. Changes in the measures of fundamental frequency variability and intensity have been previously documented in IPD patients following a course of the LSVT (Ramig, et al., 1994; Ramig et al, 1995; Ramig et al., 1996).

Electroglottographic Width 50% (EGGW50). The electroglottographic width 50% was chosen for analysis due to its theoretical relationship to true vocal fold adduction (Scherer & Vail, 1988; Scherer, Vail, & Rockwell, 1995; Titze, 1984). EGGW50 provides information about the duration of glottal closure relative to the duration of complete glottal cycle. EGGW50 is a measure of glottal adduction and is defined as the width of the EGG signal at 50% of the wave's amplitude divided by the period where the width is defined between the positive-going (glottal closing) and negative-going (glottal opening) portions of the EGG waveform (Scherer et al., 1995). It is highly correlated with other measures of glottal adduction, such as the abduction quotient (Titze, 1984) and frame by frame videoendoscopic analysis of vocal process adduction.

Expert Listener Perceptual Ratings. To determine the extent of pre to immediately posttreatment changes in speech and voice characteristics, expert listeners unfamiliar with the participant completed a perceptual listening task based upon a tape recording of the participant reading the "Rainbow Passage" (Fairbanks, 1960). The variables rated by the listeners were "strained" voice quality, "hoarse or rough" voice quality, and "strong" voice. Hoarseness, roughness and weak voice are all symptoms of PD and would be expected to decrease following a course of the LSVT. Because of the participant's pretreatment supraglottic hyperadduction and the treatment focus on true vocal fold adduction, "strained" voice quality was included in the ratings.

Videolaryngostroboscopic Examination and Ratings. After completion of a otolaryngological history and examination, a laryngostroboscopic examination was done to evaluate laryngeal and vocal fold functioning. To obtain a non-biased description of the participant's laryngeal functioning pre to immediately posttreatment, experienced raters completed a "blinded" perceptual study. Variables assessed by perceptual videostroboscopic ratings included supraglottal functioning (anterior-posterior compression and false vocal fold adduction) and degree of glottal competence (i.e. true vocal fold adduction).

Data Collection

Sound pressure level (SPL), mean fundamental frequency and its variability, and electroglottographic data were collected 3 times pre-treatment and 2 times immediately posttreatment to establish a baseline and account for potential variability in speech production associated with Parkinson disease (King, Ramig, Lemke, & Horii, 1994). All data collection sessions were scheduled at approximately the same time of day to minimize effects of medication fluctuation and were conducted by the same experimenter.

During each data collection session the participant was seated in an IAC sound-treated booth with a headset microphone (AKG-410) positioned 8 cm in front of his lips (Titze & Winholtz, 1994). After preamplification through an ATI-1000, the microphone signal was recorded onto a Sony Digital PC-108M (DAT) eight-channel recorder. The microphone signal was used to collect speech samples for the perceptual listening task as well as mean fundamental frequency and its variability. To collect sound pressure level data a Bruel and Kjaer 2230 sound level meter was placed 30 cm from the participant's mouth. During each speaking and voicing task the experimenter hand recorded the peak vocal SPL measures that were continuously displayed at 1 second intervals from the digital output of the sound level meter. To collect electroglottographic data the electrodes of a Synchrovoice Inc. Research Electroglottograph (EGG) were placed on the participant's neck over the thyroid lamina during each recording session. After amplification, (Tektronix Amplifier 502 TM 506), the EGG signal was recorded onto the eight-channel DAT recorder.

Measures of SPL, mean fundamental frequency, and fundamental frequency variability for reading and speaking were collected while the participant read the "Rainbow Passage" and spoke for 30 seconds on a topic of interest (monologue) at a comfortable rate and loudness. SPL, maximum duration, and electroglottographic data of sustained vowel phonation were collected while the participant sustained phonation of the vowel /a/ for as long as possible. A timer with a second hand was placed within the participant's view to encourage him to monitor his performance and sustain phonations longer with each repetition. Four to six maximally sustained vowels were collected during each recording session.

The speech samples for the perceptual listening task were collected by using the recorded microphone signal from the participant's second pretreatment and first posttreatment data collection session reading of the "Rainbow Passage." These passages were dubbed in random order with loudness normalized, onto a "master" rating tape that contained several samples of individuals with PD reading the "Rainbow Passage. Twenty percent of the samples were repeated to determine intrarater reliability.

For the videolaryngostroboscopic data, an examination was completed one time within the week preceding

treatment and one time within the week immediately following treatment. To collect the videolaryngostroboscopic data, the nasal passage was anesthetized with 4% lidocaine spray. Endoscopic examination was conducted with both an Olympus ENF-P3 fiberscope and Nagashima SFT-70 degree rigid laryngoscope using well-described techniques (Bless, Hirano, & Feder, 1987). Images were recorded with a CCD camera, using a 35 mm lens for the fiberscope and a 60 mm lens for the rigid telescope and a SVHS tape recorder. Examinations were conducted using both a rigid endoscope and a flexible fiberscope due to documented differences between views (Shaik & Bless, 1986; Södersten & Lindstad, 1992). Flexible fiberscope views have been reported as more representative of natural speech functioning (Shaik & Bless, 1986: Södersten & Lindstad, 1992; Smith, Ramig, Dromey et al., 1995). The abnormal posture, i.e. elevation of the larynx, and extension of the tongue, which occurs during a rigid endoscopic examination, may result in incomplete glottal closure during phonation and an abnormally high pitch (Södersten & Lindstad, 1992). In addition, soft phonation may cause differences in closure to become more apparent, making the glottis appear much more open during phonation when viewed under the rigid endoscope as compared with the flexible (Södersten & Lindstad, 1992).

The endoscopic examination protocol is briefly summarized as follows: with the fiberscope, the larynx was visualized under constant light during quiet respiration, sustained phonation of the vowel /i/, and counting for ten seconds. The strobe light source was then used to visualize the larynx and vocal folds during phonation of sustained vowel /i/ under several conditions, including (1) normal pitch, normal loudness, (2) normal pitch, soft phonation, and (3) normal pitch, loud phonation. Trials were repeated until adequate samples were obtained as judged by the otolaryngologist completing the examination. The fiberscope was withdrawn and the stroboscopic examination of the larynx was repeated with a rigid telescope for the sustained-vowel tasks only. For the purposes of this paper, only the stroboscopic conditions will be reported.

Two "master" study videotapes were created that included normal, loud, and soft phonation segments of the participant, during stroboscopic examination, before and after treatment. The participant's samples were included on this tape as a part of a larger study. Each phonation segment was shown as viewed through both a flexible fiberscope and a rigid endoscope. The order of samples was randomized and 20% of the recording samples were repeated to assess intrarater reliability. The audio signal from each individual on the tape was removed to eliminate auditory perceptual cues. Instead, a "prompter" was overdubbed onto the audio track to provide the information that identified the participant number and tasks performed during each sample to allow raters to identify them.

Data Analysis

Sound Pressure Level. SPL during sustained vowel phonation, "The Rainbow Passage," and conversational monologue were calculated using the continuously hand recorded peak vocal SPL that was displayed at 1 second intervals from the digital output of the sound level meter. Because peak vocal SPL can only be recorded from the sound level meter during speech output, pauses and hesitations were not included in the analysis. Using this method, mean vocal SPL measures have been reported accurate when compared to a custom built soft-ware program for measuring SPL from the sound level meter signal (Countryman & Ramig, 1993; Ramig et al., 1995). The mean and standard deviation of each set of output data for each task, i.e. sustained phonation, reading, and monologue, were calculated.

Maximum Duration of Sustained Vowel Phonation. To calculate duration measures of maximally sustained vowel phonation, each phonation (4 - 6 tokens for each recording session) was input into a Hewlett Packard Model 54503A MHz digitizing oscilloscope at a sampling rate of 10 samples per second. Cursors were hand-positioned to mark the monitored-displayed zero crossing preceding the first negative-going peak at the onset and the zero crossing following the final positive-going peak at the offset of each vowel. The four to six tokens were then measured, averaged, and reported for each session in seconds.

Mean Fundamental Frequency and Its Variability. To determine measures of mean fundamental frequency and fundamental frequency variability (stsd) during reading and conversational monologue, the microphone signal was digitized at 5000 samples per second and analyzed on a 486 computer using C-Speech software (Milenkovic, 1987). The mean fundamental frequency and hertz standard deviation were then calculated and displayed by the program. The hertz standard deviation was then converted using a standard formula to express frequency variability in semitones (stsd).

Electroglottographic Data (EGGW50). The electroglottic signal was low passed filtered at 10 kHz and digitized at a sampling rate of 20 kHz onto a VAX computer system. To calculate EGGW50, in-house software averaged 13-20 consecutive EGG cycles from the temporal midpoint of each vowel obtained during each voice recording session. The mean of four tokens from each pretreatment and post-treatment session were then averaged.

Perceptual Data - Expert Listeners. Three trained speech-language pathologists, with 3 or more years experience, served as raters for the listening task. The listeners were individually seated in an IAC sound-treated booth approximately 3 feet from the loudspeaker. The tape was played at approximately 75dB at 30 cm for all samples. The listeners were given a visual analog (VA) scale and asked to rate the perceptual variables of "strained" voice quality,

"hoarse or rough" voice quality, and "strong" voice quality during participant taped readings of the "Rainbow Passage." On one side of the VA scale was the word "always" and the other side "never," e.g. always a "strong" voice, never a "strong" voice. The listeners were asked to make a mark along the line that would best represent their impression of the participant's voice during reading. Due to high intralistener reliability (Pearson product correlation coefficients > 0.90), data from all three listeners were used in this study. Standard procedures for analysis of visual analog scales (Boeckstyns & Backer, 1989) were used to analyze and obtain perceptual data. The data from each expert listener are reported individually.

Videostroboscopic Data. The variables assessed for the study from the laryngostroboscopic examinations included glottal incompetence and two supraglottal hyperadduction variables - anterior/posterior compression and false fold movement, seen under six observation conditions. These were (1) flexible telescope view of vowel /i/ produced at the participant's self-judged normal pitch and normal loudness (NPNL), (2) flexible telescope view of vowel /i/ produced at the participant's loud phonation effort, (3) flexible telescope view of vowel /i/ produced at the participant's soft phonation effort, (4) rigid telescope view of vowel /i/ produced at the participant's self-judged NPNL, (5) rigid telescope view of vowel /i/ produced at the participant's loud phonation effort, and (6) rigid telescope view of vowel /i/ produced at the participant's soft phonation effort. Five raters with experience in videostroboscopic ratings assessed the recordings. Four were speech-language pathologists, and one was an otolaryngologist. They each independently viewed and rated the "master" tape. Of the five raters, two had intrarater reliability greater than 0.70 (Pearson product correlation coefficients). Data from these two raters were used in this study and are reported individually.

A modification of a study rating form from Bless (1991) which was proven effective in a previous study (Smith, Ramig, Dromey et al., 1995) was used to record and interpret the videostroboscopic rating data in relation to treatment-related changes. The raters' judgments were standardized for this rating system by prior consensus on interpretation of the scale. Degree of glottal incompetence was rated during the most closed portion of the glottal cycle on stroboscopy. A 5 point scale was used: 1= no incompetence, 2= mild incompetence - folds just not touching, 3=moderate incompetence - 50% of the length of the folds not touching with 1-2 mm gap, 4= severe incompetence, and 5= extreme incompetence - no vocal fold contact throughout the length of the folds and a large 3-4 mm gap. Supraglottal hyperadduction was rated for both anterior-posterior compression and false-fold compression on a 5 point scale as well. For anterior-posterior compression 1= no shortening of glottal length compared with resting, nonphonating state, 2=mild, glottal length shortened by 25%, 3=moderate, glottal length shortened by 50%, 4=severe, glottal length shortened by 75% and 5= extreme, arytenoids touch laryngeal surface of epiglottis, obscuring glottal view. For false fold movement 1= no false fold overclosure, laryngeal ventricles easily seen, 2= mild, one or both false folds obscure laryngeal ventricles, 3= moderate, one or both false folds obscure ventricles and a portion of the true folds, 4=severe, true folds barely visible and 5= extreme, false folds touch and cover entire glottis and may interfere with glottal vibration.

Reliability

For the variables SPL, maximum duration of sustained vowel phonation, and fundamental frequency and its variability, intraexaminer measurement reliability using Pearson product correlation coefficients ranged between 0.99 and 1.0. Interexaminer reliability for the SPL measurement method has been shown reliable (0.93-0.99 [Fox & Ramig, in press]). For the expert listener perceptual rating task intralistener reliability ranged between 0.92 and 0.97. For the two videolaryngostroboscopy raters, intrarater reliability was 0 .79 and 0.90. Interlistener and interrater reliability are not reported for the perceptual listening and videostroboscopy rating tasks because the results from each rater are reported and discussed individually.

Statistical Design

To evaluate pre to post speech treatment changes in the measures of SPL, fundamental frequency and its variability, and maximum duration, a comparison of the means of the participant's results were made. This method of statistical comparison is considered appropriate given the small sample size (Barlow & Hersen, 1984; McReynolds & Kearn, 1983) and the irreversible nature of the speech treatment program (Kratochwill & Levin, 1992). Differences in the pre to posttreatment means were considered noteworthy if the mean of the two posttreatment sessions exceeded the mean of the three pretreatment sessions, where applicable, by $\pm$ 1 standard deviation. This criteria is considered meaningful and acceptable and is analogous to a statistical measurement of a large size effect (Cohen, 1988). The expert listener and videolaryngostroboscopic perceptual ratings are reported separately for each rater.

Results

Sound Pressure Level

As shown in Table 1, the participant improved on measures of SPL during sustained vowel phonation reading and the 30 second monologue pre to posttreatment. The participant increased mean SPL pre to posttreatment 8.32 dB for sustained phonation, 13.09 dB for reading, and 12.75 dB for conversational monologue. The posttreatment increases in SPL exceeded the pretreatment standard deviation range.

Table 1.
Means and standard deviations (in parenthesis) for the speech variables sound pressure level (SPL), maximum duration of sustained vowel phonation, mean fundamental frequency, fundamental frequency variability (STSD), and EGGW50. Means are listed for the participant for the 3 pretreatment sessions, for the 3 pretreatment sessions combined (mean pre), for the 2 posttreatment sessions, and the 2 posttreatment sessions combined (mean post).

Variable	PRE 1	PRE 2	PRE 3	Mean Pre (sd)	POST 1	POST 2	Mean Post (sd)
SPL (dB) 30 cm							
Sustained vowel	78.80 (0.95)	76.50 (1.70)	80.90 (0.58)	78.73 (2.20)	86.30 (0.82)	87.80 (0.38)	87.05 (1.06)
Rainbow	73.50 (2.68)	71.15 (2.54)	74.11 (2.78)	72.90 (1.59)	87.30 (2.83)	84.68 (2.37)	85.99 (1.85)
Monologue	72.50 (2.42)	68.56 (3.35)	71.94 (2.76)	71.00 (2.13)	86.00 (2.34)	81.50 (3.14)	83.75 (3.18)
Max duration (seconds)	34.55 (3.06)	35.75 (3.21)	31.15 (7.01)	33.82 (2.39)	32.97 (3.26)	34.11 (3.05)	33.54 (0.81)
Mean Fo (Hertz)							
Rainbow	86.97 (10.90)	92.66 (12.26)	94.92 (14.17)	91.52 (4.10)	126.39 (26.08)	118.54 (24.27)	122.47 (5.55)
Monologue	92.09 (12.55)	86.68 (9.53)	86.75 (9.96)	88.51 (3.10)	123.45 (21.14)	105.05 (14.04)	114.25 (13.01)
STSD							
Rainbow	2.17	2.29	2.59	2.35 (0.22)	3.59	3.56	3.58 (0.02)
Monologue	2.36	1.90	1.99	2.08 (0.24)	2.97	2.32	2.65 (0.46)
EGGW50	0.671 (0.020)	0.629 (0.010)	0.684 (0.020)	0.661 (0.029)	0.671 (0.010)	0.688 (0.020)	0.680 (0.012)

Maximum Duration of Sustained Vowel Phonation (seconds)

Maximum duration of sustained vowel phonation remained relatively stable pre to posttreatment (Table 1).

Mean Fundamental Frequency

The participant increased mean fundamental frequency pre to posttreatment 30.95 Hz during reading and 27.50 Hz during conversational monologue (Table 1). These posttreatment increases exceeded pretreatment standard deviation ranges.

Fundamental Frequency Variability (STSD)

The participant increased semitone standard deviation (stsd) pre to posttreatment 1.23 stsd during reading and 0.57 stsd during conversational monologue (Table 1). These changes exceeded pretreatment standard deviation ranges for these tasks.

EGGW50

As shown in Table 1, the EGGW50 data remained relatively stable pre to posttreatment. This stability suggests no noticeable decrease or increase of true vocal fold adduction pre to posttreatment during sustained vowel phonation. Pre and posttreatment, these values are within the high range of normal phonation (i.e. not characteristic of a breathy or pressed voice) for this measure (Scherer et al., 1993).

Perceptual Measurements

For the variable never "hoarse or rough," raters 1 and 2 rated a positive change of 6 (85% to 91%) and 25 (68% to 93%) percentage points respectively pre to posttreatment. Rater 3 rated relatively no change (-1 percentage point [97% to 96%]) pre to posttreatment for this variable. For the variable always "strong" raters 1, 2, and 3 noted a positive change of 15 (73% to 88%), 36 (64% to 100%), and 2 (95% to 97%) percentage points respectively pre to posttreatment. The participant was rated never "strained" 99% and 100% of the time both pre and posttreatment by raters 1 and 2 respectively. Rater 3 rated a change of -6 percentage points (98 to 92) for the participant pre to posttreatment for this variable.

Laryngostroboscopic Measurements

Flexible Scope. All laryngostroboscopic ratings for the participant are summarized in Table 2 (following page). For glottal incompetence during soft phonation, both raters noted mild incompetence pretreatment and none posttreatment. For anterior/posterior (A/P) compression, both raters noted mild A/P compression during normal and loud phonation pretreatment and none posttreatment. Both raters noted severe false fold movement pretreatment during normal and loud phonation. In addition, rater 2 assessed mild false fold movement for soft phonation pre-treatment. Posttreatment, both raters assessed no false fold movement for the same tasks.

Rigid Scope. For glottal incompetence, rater 1 indicated mild incompetence during loud phonation pretreatment and none posttreatment. No incompetence was noted during soft or normal loudness phonation pre or

Table 2.
Perceptual ratings of videolaryngostroboscopic samples of the participant pre and posttreatment. Ratings include glottal incompetence, anterior/posterior compression, and false fold movement for both the flexible and rigid views during normal, loud, and soft phonation.

Variable	Pre	Post	Pre	Post
Glottal Incompetence*				
Flexible Scope				
normal phonation	1	1	1	1
loud phonation	1	1	1	1
soft phonation	2	1	2	1
Rigid Scope				
normal phonation	1	1	1	1
loud phonation	2	1	1	1
soft phonation	n/a	n/a	n/a	n/a
Anterior/Posterior Compression *				
Flexible Scope				
normal phonation	2	1	2	1
loud phonation	2	1	2	1
soft phonation	1	1	1	1
Rigid Scope				
normal phonation	1	1	1	1
loud phonation	1	1	1	1
soft phonation	n/a	n/a	n/a	n/a
False fold movement *				
Flexible				
normal phonation	3	1	3	1
loud phonation	4	1	5	1
soft phonation	1	1	2	1
Rigid				
normal phonation	1	1	1	1
loud phonation	1	1	2	1
soft phonation	n/a	n/a	n/a	n/a

* Ratings for flexible and rigid scope: 1=none, 2=mild, 3=moderate, 4=severe, 5=extreme

posttreatment by either rater. For anterior/posterior compression, both raters assessed no compression for all conditions pre and posttreatment. For false fold movement, rater 2 assessed mild hyperadduction during loud phonation pretreatment and none posttreatment. No false fold movement was noted during soft or normal loudness phonation pre or posttreatment by either rater.

Discussion

The purpose of this study was to evaluate the effect of the LSVT on an individual with idiopathic Parkinson disease (IPD) who had soft, hoarse, monotone voice, severe supraglottic hyperadduction, and mild true vocal fold hypoadduction pre to posttreatment. It was hypothesized that the Lee Silverman Voice Treatment (LSVT) program, which focuses on true vocal fold adduction, would improve the participant's true vocal fold hypoadduction, decrease his supraglottic hyperadduction and have positive effects on his soft, hoarse, monotone voice. The findings reported here suggest that the participant responded positively to the LSVT in all of these domains.

Pre to posttreatment the participant increased SPL levels during sustained phonation as well as increased SPL and fundamental frequency and its variability (stsd) during reading and monologue. These findings are consistent with previous reports of LSVT treatment related changes in IPD (Dromey et al., 1995; Ramig & Dromey, 1996; Ramig et al., 1994; Ramig et al., 1995; Ramig et al., 1996). In addition, the expert listeners and raters indicated improved voice quality and true vocal fold adduction without straining as well as complete reduction in supraglottal hyperadduction in this participant posttreatment. Although SPL in sustained phonation increased and the raters noted no true vocal fold hypoadduction posttreatment, little change was observed in measures of maximum duration of sustained vowel phonation and EGGW50 (measure of adduction) posttreatment. These last two findings are inconsistent with previous post LSVT reports (Brosovic, 1994; Ramig et al., 1994; Ramig et al., 1995; Ramig et al., 1996).

Normally, the mechanism responsible for posttreatment increases in SPL and fundamental frequency and its variability in an individual with IPD is attributed to improved true vocal fold adduction and respiratory support (Dromey et al., 1995; Ramig & Dromey, 1996; Smith, Ramig, Dromey, et al., 1995). We suggest that through training optimal vocal fold adduction (i.e. adduction with adequate respiratory support, loudness and quality), the participant increased his SPL and eliminated the need for secondary compensatory supraglottal hyperadduction posttreatment. Reducing supraglottal hyperadduction or constriction in the vocal tract can increase SPL by alleviating a damping effect on the acoustic signal (Sundberg & Gauffin, 1979). Enlarging the vocal tract will increase SPL as well as improve overall vocal quality. This mechanism together with probable increases in subglottal air pressure are likely responsible for increased SPL posttreatment. Although subglottal air pressure was not measured here, it has been reported to increase following a course of the LSVT in a group of individuals with idiopathic Parkinson disease (Ramig & Dromey, 1996). The change in SPL suggests the participant improved efficiency of his true vocal fold adduction, reduced constriction in the supraglottal area, and learned to manage his respiratory support at the level of the larynx more effectively posttreatment.

The changes observed posttreatment in the participant's supraglottal hyperadduction are not surprising when considering the theory for treating voice disorders suggested by Ramig and Scherer (1992) and Smith and Ramig (1995). The underlying physiological mechanism

contributing to the participant's voice disorder and supraglottal hyperadduction appears to have been inadequate true vocal fold closure. By focusing on this primary deficit rather than the resulting secondary compensatory behavior, overall speech and voice improvement was achieved in a simple and efficient manner. The simple focus of the adduction therapy program eliminated true vocal fold hypoadduction and supraglottal hyperadduction and increased loudness, intonation, and vocal quality in this participant. These changes were supported by the objective and perceptual measures as well as the participant's report that the treatment increased his ability to communicate effectively with coworkers, customers, family, and friends.

A more traditional approach to treating this individual may have been to eliminate the supraglottal hyperadduction through relaxation techniques and then focus on the decreased loudness and intonation and hoarse voice. We believe this more traditional approach to treatment would not have been as effective or as efficient given the probable origin of the supraglottal hyperadduction and reported difficulty individuals with Parkinson disease have with complex instructions and material. These findings correspond with past reports that treatment designed to restore true vocal fold adduction or repair true vocal fold pathology has reduced supraglottal hyperadduction (Feinstein et al., 1987; Smith, Ramig, Dromey et al., 1995; Von Doersten et al., 1992; Von Hake et al., 1989).

It should be noted, however, that the changes observed in this individual with Parkinson disease may not be indicative of all individuals with PD presenting with supraglottal hyperadduction. We advocate treating these individuals on a case-by-case basis. Stimulability testing is recommended during pretreatment screening to determine if the individual is capable of producing a louder voice easily and without straining. If straining is noted during the screening or extreme cases of supraglottal hyperadduction are observed, a trial period of the LSVT is recommended to determine appropriateness of the individual for the treatment.

This study demonstrates the use of a continuum of measures ranging from functional ratings to physiological measures to evaluate the impact of intensive voice treatment and identify mechanisms underlying treatment-related change in an individual with Parkinson disease. Some of the measures used here (SPL, fundamental frequency, maximum duration, expert listener ratings, videolaryngostroboscopy) are available to the practicing clinician and are important for documentation of treatment-related change and reimbursement. Physiological measures such as the EGGW50 provide a better understanding of the underlying mechanism contributing to treatment-related changes. A wide spectrum of measures is important to obtain a comprehensive and accurate picture of treatment-related change from functional impact to physiological bases.

In conclusion, the case presented here is an initial step towards increasing the professional's knowledge of the origin and treatment of supraglottal hyperadduction in individuals with idiopathic Parkinson disease. The study demonstrates that adduction therapy, if completed properly and efficiently, can enhance the speech and voice in an individual with PD with mild vocal fold hypoadduction pretreatment. Clearly, the findings are limited to one case of supraglottal hyperadduction in Parkinson disease. Further research on this topic is needed to generalize treatment effects on supraglottal hyperadduction in individuals with idiopathic Parkinson disease, resulting in effective speech treatment for a wide variety of individuals with IPD.

Author Note

Lorraine Olson Ramig, Department of Speech-Language and Hearing Sciences; Marshall Smith, Department of Otolaryngology.

The authors wish to express deep gratitude to the individual who participated in this study. This study was supported in part by grants from the Office of Education - National Institute for Disability and Rehabilitative Research (8133G40108) and from the National Institutes of Health - National Institute for Deafness and Communicative Disability (DC01150) and the Hearst Foundation.

Correspondence concerning this article should be addressed to Stefanie Countryman, at the Wilbur James Gould Voice Research Center, 1245 Champa St., Denver, Colorado 80204, (303) 893-4000 ext. 2622. Electronic mail may be sent via internet to <Stefanie@star.dcpa.org>.

References

Aronson, A. E. (1985). Clinical Voice Disorders. New York: Thieme Inc.Aronson, A. (1990). Clinical Voice Disorders. New York: Thieme.

Arnold, G., & Pinto, S. (1960). Ventricular dysphonia: New interpretation of an old observation. Laryngoscope, 70, 1608-1622.Barlow, D. H., & Hersen, M. (1984). Single Case Experimental Designs: Strategies for Behavior Change. New York: Pergamon Press.

Beukelman, D.R., Kraft, G.H., & Freal, J. (1985). Expressive communication disorders in persons with multiple sclerosis: A survey. Archives of Physical Medicine and Rehabilitation, 66, 675-677.

Bless, D.M. (1991). Assessment of laryngeal function. In C. Ford & D. Bless (Eds.). Phonosurgery: Assessment and Surgical Management of Voice Disorders. New York: Raven Press.

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

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.

Boone, D.R. (1977). The Voice and Voice Therapy. Englewood Cliffs, NJ: Prentice-Hall.

Brin, M., Fahn, S., Blitzer, A., Ramig, L. O., & Stewart, C. (1992) Movement disorders of the larynx. In A. Blitzer, M. Brin, C. Sasaki, S. Fahn, & K. Harris. (Eds.). Neurologic Disorders of the Larynx. New York: Thieme Medical Publishers.

Brosovic, G. (1994). Electroglottographic measures of vocal fold adduction pre to posttreatment in patient's with Parkinson disease. an unpublished mater's thesis. University of Colorado - Boulder.

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.

Cohen, J. (1988). Statistical Power Analysis for Behavioral Sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Assoc.

Countryman, S. & Ramig, L.O. (1993). Effects of intensive voice therapy on speech deficits associated with bilateral thalamotomy in Parkinson disease. A case study. Journal of Medical Speech-Language Pathology, 1(4), 233-249.

Countryman, S., Ramig, L.O., & Pawlas, A.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. (1981). Speech disorders of Parkinsonism: A review. Journal of Neurology, Neurosurgery, and Psychiatry, 44, 751-758.

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

Darley, F., Aronson, A., & Brown, J. (1969b). Clusters of deviant speech dimensions in the dysarthrias. Journal of Speech and Hearing Research, 12, 462-496.

Darley, F., Aronson, A., & Brown, J. (1975). Motor Speech Disorders. Philadelphia: W.B. Saunders.

Darley, F., Brown, J., & Goldstein, N. (1979). Dysarthria in multiple sclerosis. Journal of Speech and Hearing Research, 15, 229-245.

Dromey, C., Ramig, L. & Johnson, A. (1995). Phonatory and articulatory changes associated with increased vocal intensity in Parkinson disease. Journal of Speech and Hearing Research, 38, 751-764.

Fahn, S., Elton, R. L., & members of the UPDRS development committee. (1987). Unified Parkinson's disease rating scale. In S. Fahn, C.D. Mardsen, M. Goldstein, & D. B. Caine (eds). Recent Developments in Parkinson Disease. 2, 153-163, New York: Macmillan Press.

Fairbanks, G.(1960). Voice and Articulation Drill Book. New York: Harper & Brothers.

Feinstein, I., Hilger, P., Szachowicz, E., & Stimson, B. (1987). Laser therapy of plica ventricularis. Annals of Otology, Rhinology, & Laryngology, 96, 56-57.

Fox, C. & Ramig, L. O. (in press). Sound pressure level and self-perception of speech and voice in men and women with idiopathic Parkinson disease. American Journal of Speech-Language Pathology.

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

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

Jackson, C., & Jackson, C.L. (1935). Dysphonia plicae ventricularis. Archives of Otolaryngology, 21, 157-167.

Jacob, L.A. (1968). A normative study of laryngeal jitter. Unpublished master's thesis, University of Kansas, Lawrence.

Jensen, J.R. (1960). A study of certain motor-speech aspects of the speech of multiple sclerosis patients. Unpublished doctoral dissertation. University of Wisconsin, Madison.

Kent, R., Kent, J., Weismer, G., Sufit, R., Rosenbek, J., Martin, R., & Brooks, B. (1990). Impairment of speech intelligibility in men with amyotrophic lateral sclerosis. Journal of Speech and Hearing Disorders, 55, 721-728.

King, J., Ramig, L., Lemke, J. & Horii, Y. (1994). Longitudinal changes in acoustic parameters of phonation. Journal of Medical Speech-Language Pathology, 2(1), 29-42.

Kratochwill, T., & Levin, J. R. (1992). Single-Case Research Design and Analysis: New Directions for Psychology and Education. Hillsdale, NJ: Lawrence Erlbaum Associates.

Linville, S.E. & Korabic, E. (1987). Fundamental frequency characteristics of elderly women's voices. Journal of the Acoustical Society of America, 81, 1196-1199.

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, 43, 47-57.

McReynolds, L.V., & Kearns, K.P. (1983). Single-Subject Experimental Designs in Communicative Disorders. Baltimore: University Park Press.

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

Pawlas, A. A., Ramig, L. O., Countryman, S., O'Brien, C., & Hoehn, M. (in submission). Speech and voice characteristics in idiopathic Parkinson disease.

Perez, K., Ramig, L.O., Smith, M.E. & Dromey, C. (1994). The Parkinson larynx: tremor and videostroboscopic findings. Journal of Voice, 10(4), 354-361.

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. Philadelphia: John Benjamins Publishing Co.

Ramig, L.O., Bonitati, C., Lemke, J. & Horii, Y. (1994). Voice treatment for patients with Parkinson disease: Development of an approach and preliminary efficacy data. Journal of Medical Speech-Language Pathology, 2(3), 191-209.

Ramig, L. O., Countryman, S., O'Brien, C., Hoehn, M., & Thompson, L. (1996). Intensive speech treatment for patients with Parkinson disease: short- and long-term comparison of two techniques. Neurology, 47(6), 1496-1504.

Ramig, L. O., Countryman, S., Pawlas, A. A., & Fox, C. (in press). The Lee Silverman Voice Treatment: application to multiple sclerosis, closed head injury, and stroke. Phonoscope.

Ramig, L.O., Countryman, S., Thompson, L., & Horii, Y. (1995). A comparison of two forms of intensive speech treatment for Parkinson disease. Journal of Speech and Hearing Research, 38(6), 1232-1251.

Ramig, L.O., & Dromey, C. (1996). Aerodynamic mechanisms underlying treatment-related changes in vocal intensity in patients with Parkinson disease. Journal of Speech and Hearing Research, 39(4), 798-807.

Ramig, L.O., & Gould, W. (1986). Speech characteristics in Parkinson's disease. Neurologic Consultant, 4(1), 1-8.

Ramig, L.O., Pawlas, A. A., & Countryman, S. (1995). The LSVT: A Practical Guide for Treating the Voice and Speech Disorders in Parkinson Disease. Iowa City, IA: The National Center for Voice and Speech.

Ramig, L.O. & Scherer, R.C. (1992). Speech therapy for neurologic disorders of the larynx. In A. Blitzer, C. Sasaki, M. Brin., & K. Harris (Eds.). Neurologic Disorders of the Larynx. New York: Thieme Medical Publishers.

Roy, N. (1994). Ventricular dysphonia following long-term endotracheal intubation: A case study. Journal of Otolaryngology, 23(3), 189-193.

Salmoni, A. W., Schmidt, R. A., & Walter, C. B. (1984). Knowledge of results and motor learning: a review and critical reappraisal. Psychological Bulletin, 95, 335-386.

Saxon, K. G., & Schneider, C. M. (1995). Vocal Exercise Physiology. San Diego, CA: Singular Publishing.

Scherer, R.C., & Vail, V.J. (1988). Measures of laryngeal adduction. Journal of the Acoustical Society of America, suppl 84 (1), S81 (A).

Scherer, R.C., Vail, V.J., & Rockwell, B. (1995). Examination of the Laryngeal Adduction Measure EGGW. In: F. Bell-Berti & L. Raphael (Eds.). Producing Speech: Contemporary Issues. New York: American Institute of Physics.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. Psychological Review, 82, 225-260.

Schmidt, R. A. (1988). Motor Control and Learning. Champaign, IL: Human Kinetic Publishers.

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

Shaik, A., & Bless, D. (1986). Preliminary comparison of vocal fold vibration with strobotelescopic and strobofiberscopic examinations. In: S. Hibi, M. Hirano, & D., Bless (Eds.). Proceedings of International Conference on Voice, 2:9-14, Kerume, Japan.

Smith, M.E. & Ramig, L.O. (1995). Neurological disorders and the voice. In W. Gould, J. Rubin, G. Korovin, & W. Sataloff. (Eds.). Diagnosis and Treatment of Voice Disorders. New York: Igaku-Shoin.

Smith, M.E., Ramig, L.O, Dromey, C., Perez, K. & Samandarii, R. (1995). Intensive voice treatment in Parkinson disease: Laryngostroboscopic findings. Journal of Voice, 9(4), 453-459.

Södersten, M. & Lindestad, P.Å. (1992). A comparison of vocal fold closure in rigid telescopic and flexible fiberoptic laryngostroboscopy. Acta Otolaryngologica, 112, 144-150.

Solomon, N.P. & Hixon, T.J. (1993). Speech breathing in Parkinson's disease. Journal of Speech and Hearing Research, 36(2), 294-310.

Sundberg, J., & Gauffin, J. (1979). Waveform and spectrum of the glottal voice source. In Lindblom and Ohman (Eds.), Frontier of Speech Communication Research, pp. 301-320, Orlando, Florida: Academic Press.

Titze, I. (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., & Winholtz, W. (1994). Effect of microphone type and placement on voice perturbation measurements. Journal of Speech and Hearing Research, 36, 1177-1190.

Vogel, M. & Von Cramon, D. (1982). Dysphonia after traumatic midbrain damage: a follow-up study. Folia Phoniatrica, 34, 150-159.

Von Cramon, D. (1981). Traumatic mutism and the subsequent reorganization of speech functions. Neuropsychologica, 19, 801-805.

Von Doersten, P., Izdebski, K., Ross, J., & Cruz, R. (1992). Ventricular dysphonia: A profile of 40 cases. Laryngoscope, 102, 1296-1301.

Von Hake, C., Ganzman, I. & Mauer, T. (1989). Diagnosis and management of ventricular dysphonia. Journal of the American Osteopathic Association, 89(2), 181-183.

Ward, P.H., Berci, G., & Calcaterra, T.C. (1977). Superior laryngeal nerve paralysis: an often overlooked entity. Trans-American Academy of Opthalmology & Otolaryngology, 84, 78-89.

Woo, P., Casper, J., Colton, R. & Brewer, D. (1992). Dysphonia in the aging: Physiology v. disease. Laryngoscope, 102, 139-144.

Woo, P., Casper, J., Colton, R. & Brewer, D. (1994). Diagnosis and treatment of persistent dysphonia after laryngeal surgery: a retrospective analysis of 62 patients. Laryngoscope, 104, 1084-1091.

Yanagihara, N., Koike, Y., & von Leden, H. (1966). Phonation and respiration. Function study in normal subjects. Folia Phoniatrica, 18, 323-340.

NCVS Status and Progress Report - 11
May 1997, 49-57

Correlation Between Vocal Functions and Glottal Measurements in Unilateral Vocal Fold Paralysis Patients

Katsuhide Inagi, M.D.
Aliaa A. Khidr, M.D., Ph.D.
Charles N. Ford, M.D.
Department of Surgery, The University of Wisconsin Clinical Science Center
Diane M. Bless, Ph.D.
Department of Surgery, The University of Wisconsin Clinical Science Center
Department of Communicative Disorders, The University of Wisconsin Clinical Science Center
Dennis M. Heisey, Ph.D.
Department of Surgery, The University of Wisconsin Clinical Science Center

Abstract

Observations and analysis of glottal characteristics are critical in choosing the best modality for surgery in patients with unilateral vocal fold paralysis (UVP). This study suggests that multiple glottal characteristics influence the vocal product in patients with UVP. In addition to the horizontal position of the paralyzed vocal fold (deviation from the mid-line), the glottal area, degree of bowing of the paralyzed and contralateral vocal folds, maximum separation between vocal folds, compensatory glottal maneuvers, and the vertical glottic closure plane significantly influenced the quality of the voice. Clinicians should be aware of these observations to facilitate treatment planning and assessing the results of surgical procedures used to improve voice quality in cases of UVP.

Introduction

Unilateral vocal fold paralysis (UVP) due to injury of the recurrent laryngeal nerve can result in various degrees of change in voice quality and compromise the patient's ability to communicate. In spite of the recent phonosurgical developments offering a plethora of surgical options for treating UVP (1-6), there are few guidelines for matching procedure and patients. This is due, in part, to a lack of information about the glottal parameters critical to changes in the voice quality of these patients.

Although numerous neuropathological studies (7,8) have provided detailed descriptions of the horizontal position of the paralyzed fold relative to midline. Evaluation of the effect of changes in the position on vocal fold vibration has been limited to theoretical models of phonation (9) or in vivo human larynges, where control of anatomical variables is not possible (10). Recently, the authors have used excised human larynges to measure different glottal parameters to standardize and evaluate glottal measurements specifically as they relate to the results of thyroplasty type I. These studies indicated that the anatomical configurations resulting from UVP were complex, and varying these resulted in different voice qualities.

For several decades, clinicians have been trying to predict the vocal function (loudness, pitch, and quality) of the different patterns of unilateral vocal fold paralysis by using the position of the paralyzed vocal fold (deviation from mid-line) as a means of classification or reference (11). This traditional position classification schema has several inherent problems: It is subjective and does not describe other glottal characteristics such as bowing, vertical differences in vocal fold level, and compensatory glottal maneuvers. Moreover, quantitative analysis of the paralyzed vocal fold position (median, paramedian, intermediate, and lateral) and its effect on the resultant voice quality have not been fully described. Woodson et al. (12) reported that laryngeal paralysis results in a complex alteration of the

glottis rather than simple limitation of the movement of the paralyzed vocal fold, accordingly, the horizontal position of the paralyzed vocal fold by itself has no rational value in predicting the changes in the voice quality. Furthermore, because standard guidelines for classification do not exist, different clinicians use different references in describing the same classification terms; some clinicians use maximum posterior separation (7), while others prefer using a description of the subjective relationship between the paralyzed vocal fold and mid-glottal planes (11). Differences in description make inter-institution communication difficult and comparison between institutions nearly impossible.

This study was designed to investigate four questions related to the position of the vocal folds in UVP and vocal outcome. (In the context of this study, outcome is defined as voice as measured by subjective and objective vocal function assessments.): 1) Is there any correlation between traditional horizontal position classification of the paralyzed vocal fold and vocal function? 2) Is it possible to predict vocal function using traditional position classification with other subjective visual parameters. 3) Which single objective measurement correlates best with vocal function? 4) Which combination, if any, of objective measurements predicts vocal function?

Materials and Methods

Patients

A retrospective study was done on 43 patients, 20 males and 23 females, with unilateral vocal fold paralysis, selected from the University of Wisconsin Clinical Science Center database. Selection criteria included no previous surgical treatment, no vocal fold lesions, no neurological or respiratory diseases, and availability of good videostroboscopic, aerodynamic and acoustic recordings. Male patients' ages ranged from 22 to 92 years (mean 50), female patients' ages ranged from 20 to 68 years (mean 43.2). The etiologies of paralysis for the 43 patients were subclassified as iatrogenic (n=26), idiopathic (n=15), or trauma without laryngeal structural damage (n=2).

Voice Assessments

Each patient had undergone a battery of tests of vocal function. Tests included indirect laryngeal examination and videostroboscopic recording using the strobe system and 70 degree rigid laryngoscope (Kay, Lincoln Park, NJ). Laryngeal videorecordings were obtained during modal phonation (defined as the patient's most comfortable habitual phonatory mode), maximum voluntary inspiration and rest breathing. This was followed by aerodynamic and acoustic recordings using a phonatory function analyzer (Nagashima PS-77, Tokyo, Japan), and a DAT audio recorder (in a sound isolated audio booth).

Objective Voice Measures

Measures obtained from each patient included maximum phonation time(MPT), frequency range (Fr-range), intensity range (sound pressure level (SPL) - range), and mean air-flow rate (MFR). and Acoustic measurements included jitter, shimmer, and signal-to-noise ratio (SNR). All acoustic data were analyzed using a locally developed hardware and software program, C-Speech.

Subjective Measures

Auditory perceptual voice rating was done on the GRBAS scale that assesses grade, roughness, breathiness, asthenia, and strain. Visual-perceptual parameters of the laryngeal video recordings were made on a rating form (Figure 1). Both auditory and visual recordings were played from a randomized master tape to two judges with extensive experience in the field of voice disorders (an otolaryngologist and a phoniatrician). A four-point equal-appearing interval rating scale of 0 to 3 was used for the subjective measurements. Grade 0 indicates no symptoms, grade 1 = mild symptoms, and grade 3 = severe symptoms. Consensus rating was done by two raters with extensive experience in voice disorders. Both audio and video recordings were selected based on their quality. Selection criteria for video images included sharp focus, image centralization, no tilting, and clear view of the anterior commissure , vocal processes and posterior glottis. Video recordings that did not meet these criteria were not used.

The horizontal position of the paralyzed vocal fold was classified as follows: median position, in which the paralyzed vocal fold is positioned at the anatomical midline; lateral position, in which the vocal fold is fixed in an abducted position (comparable to the position of the healthy vocal fold in maximum abduction in maximal voluntary

Rating Form

Subject Number _____

Tape _____ Counter _____

Videostroboscopic Recording				
Paralyzed Vocal Fold	Right	Left		
Paralyzed VF Horizontal Position	Median	Paramedian	Intermediate	Lateral
Paralyzed VF Vertical Position	On Glottal Plane		Off Glottal Plane	
Glottal Gap Position	Anterior	Middle	Posterior	All
Glottal Gap Shape	Longitudinal	Spindle	Hourglass	Triangle
Glottal Gap Size	Absent(0)	Small(1)	Medium(2)	Large(3)
Vocal Fold Bowing	Absent(0)	Mild(1)	Moderate(2)	Severe(3)
Vocal Fold Mucosal Wave Magnitude	Absent(0)	Mod. Red.(1)	Mild Red.(2)	Normal(3)
Vocal Fold Movement Amplitude	Absent(0)	Mod. Red.(1)	Mild Red.(2)	Normal(3)
Predominant Open Phase	Normal(0)	Mild(1)	Moderate(2)	Severe(3)
Predominant Asymmetry	Normal(0)	Mild(1)	Moderate(2)	Severe(3)
Latero-Medial Compression	Normal(0)	Mild(1)	Moderate(2)	Severe(3)
Antero-Posterior Compression	Normal(0)	Mild(1)	Moderate(2)	Severe(3)
Auditory Perceptual Rating				
Grade	Normal(0)	Mild(1)	Moderate(2)	Severe(3)
Roughness	Normal(0)	Mild(1)	Moderate(2)	Severe(3)
Breathiness	Normal(0)	Mild(1)	Moderate(2)	Severe(3)
Asthenia	Normal(0)	Mild(1)	Moderate(2)	Severe(3)
Strain	Normal(0)	Mild(1)	Moderate(2)	Severe(3)

Figure 1. This form was used for rating the visual-perceptual parameters of the laryngeal video recordings and perceptual voice rating.

inspiration); intermediate (cadaveric) position, in which the vocal fold is positioned at the mid-plane between median and lateral position; and paramedian position, in which the vocal fold is positioned at the mid-plane between the median position and intermediate position.

The vertical position of the paralyzed vocal fold was evaluated using slow phase motion video playback and was classified as follows: on glottal plane, when the edge of the normal vocal fold was seen to meet or approximate that of the paralyzed vocal fold at the same plane during vibration; off glottal plane, when the edge of the normal vocal fold was seen to meet or approximate the paralyzed vocal fold at a higher or lower plane than its glottal edge. It should be noted that all samples were previously subjectively judged to be central to avoid the optical illusion of closure plane differences that are due to asymmetric placement of the endoscope.

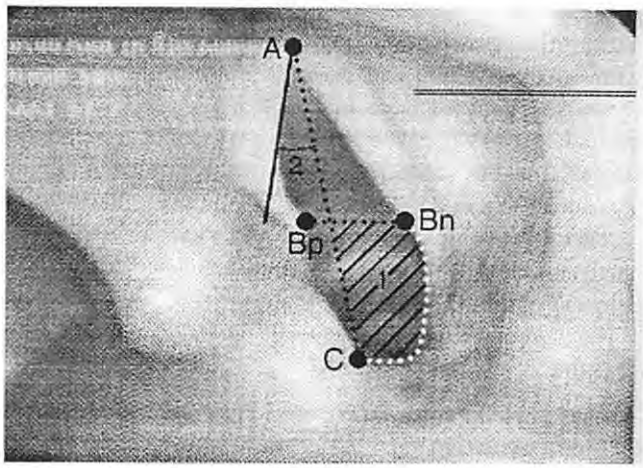


Figure 2a. Digitized glottal images of videostrobscopic recording were used for measurements. 1: Normal side cartilaginous area (GA-nor-car)(2650pixels). 2: Paralyzed anterior vocal fold angle(ant-VFA-par)(24.7°).

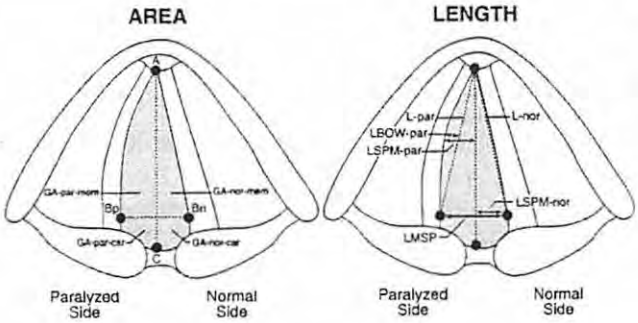


Figure 2b. -left- Glottal area (GA) measurements included paralyzed side membranous area (GA-par-mem), and the normal side membranous area (GA-nor-mem); paralyzed side cartilaginous area (GA-par-car), and normal side cartilaginous area (GA-nor-car); total membranous glottal area (GA-all-mem=GA-par-mem+GA-nor-mem); total cartilaginous glottal area (GA-all-car=GA-par-car+GA-nor-car); total paralyzed side glottal area (GA-all-pos=GA-par-mem+GA-par-car); total normal side glottal area (GA-all-nor=GA-nor-mem+GA-nor-car); and total glottal area (GA-all=GA-par-mem+GA-nor-mem +GA-par-car+GA-nor-car). -right- Length (L) measurements included length of the paralyzed vocal fold (L-par) and the normal vocal fold (L-nor); degree of bowing in the paralyzed vocal fold (LBOW-par) and normal vocal fold (LBOW-nor); distance of the paralyzed vocal fold from midline (LSPM-par) and the normal vocal fold (LSPM-nor); and maximum separation between both vocal folds during phonation (LMSP).

Objective Measurements of Glottal Images

Objective measurements were obtained from digitized glottal images from videostrobscopic recordings (Figure 2a). Three glottal frames were selected for each patient; full adduction during modal phonation (-- /AD), full abduction during maximum voluntary inspiration (-- /AB), and at resting position (-- /R). Four reference points were determined for each glottal frame: anterior commissure (point A), paralyzed side vocal process (point Bp), normal side vocal process (point Bn), and mid-interarytenoid point (point C). The mid-interarytenoid point was carefully determined only after reviewing vocal fold movement patterns so as not to be biased by the position of the paralyzed vocal fold. The maximum medial movement of the normal vocal fold during modal phonation was used to identify the midline. A marker was then put at the site of the normal vocal process (during maximum adduction) and a horizontal line was extended posteriorly to meet the posterior glottis at point (C). These reference points were saved on the digitized image to facilitate standardization of measurements of glottal area, length and angle.

Glottal Area (GA) (Figure 2b) was divided by line Bp-Bn and by line A-C to designate the paralyzed side membranous area (GA-par-mem), and normal side membranous area (GA-nor-mem), and paralyzed side cartilaginous area (GA-par-car) and normal side cartilaginous area

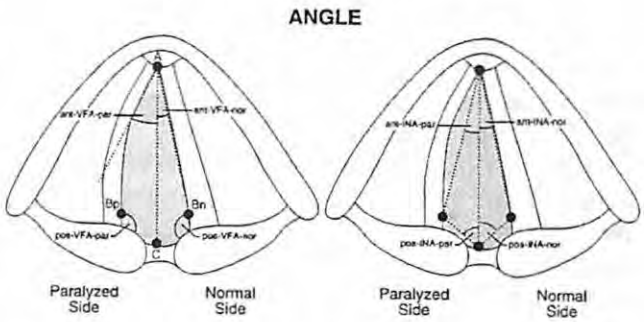


Figure 2c. Angle (A) measurements included angle of paralyzed anterior vocal fold angle from the midline (ant-VFA-par) and normal anterior vocal fold angle (ant-VFA-nor); the anterior inner angle of the paralyzed vocal fold (ant-INA-par) and normal anterior inner angle (ant-INA-nor); the posterior inner angle of the paralyzed vocal fold (pos-INA-par) and normal posterior inner angle (pos-INA-nor); the posterior angle of the paralyzed vocal fold from the arytenoid (pos-VFA-par) and normal posterior vocal fold angle (pos-VFA-nor); . total anterior vocal fold angle (ant-VFA-all=ant-VFA-par+ant-VFA-nor); and total inner angle (ant-INA-all=ant-INA-par+ ant-INA-nor).

(GA-nor-car). Five additional measurements were calculated from the each of four areas: Membranous glottal area (GA-all-mem=GA-par-mem+GA-nor-mem), cartilaginous glottal area (GA-all-car=GA-par-car+GA-nor-car), paralyzed side glottal area (GA-all-par=GA-par-mem+GA-par-car), normal side glottal area (GA-all-nor=GA-nor-mem+GA-nor-car), and total glottal area (GA-all=GA-par-mem+GA-nor-mem +GA-par-car+GA-nor-car).

In order to avoid the effect of the changing vocal fold length, the Length (L) measurements (Figure 2b) were calculated for three glottal conditions (adduction, abduction and at rest). Length measures included : the length of the paralyzed vocal fold (L-par) measured between point A and Bp, the length of the normal vocal fold (L-nor) measured between point A and Bn, and the degree of bowing in the paralyzed (LBOW-par) and normal (LBOW-nor) vocal folds measured between the point of maximum concavity and line ABp or ABn (connecting the anterior commissure and the corresponding vocal process). In addition, the separation of the paralyzed (LSPM-par) vocal fold from midline , the separation of the normal (LSPM-nor) vocal fold from midline, and maximum separation between the vocal folds during phonation (LMSP). The greatest distance between the vocal folds was measured from the full adduction image. Midline was defined as the connecting line between the anterior commissure and the interarytenoid point (A-C).

For standardization of the glottal measurements, the length measurements were divided by the average of both L-par and L-nor in each of the glottal conditions and area measurements were divided by the square of the average of both L-par and L-nor in each of the glottal conditions.

Angle (A) measurements (Figure 2c) were calculated from the full adduction laryngeal images. They included; Anterior angle of the paralyzed (ant-VFA-par) and normal (ant-VFA-nor) vocal folds, measured between lines drawn tangential to the anterior of the inner surface of each vocal fold and the midline; anterior inner angle on the paralyzed side (ant-INA-par) and normal (ant-INA-nor) side measured between line A-Bn(Bp) and line A-C; posterior inner angle on the paralyzed side (pos-INA-par) and normal (pos-INA-nor) side measured between line Bn(Bp)-C and line C-A; posterior angle of the paralyzed (pos-VFA-par) and normal angle (pos-VFA-nor) vocal folds, measured between lines drawn tangential to the inner surface of the posterior of each vocal fold and tangential to the inner surface of each arytenoid; and the total anterior vocal fold angle (ant-VFA-all) and total inner angle (ant-INA-all), measured for both paralyzed and normal sides.

All measurements were performed using a NIH Image 1.44 software (USDHSS-National Institutes of Health, Bethesda, MD) and an Apple Macintosh II FX (Cupertino, CA). Each measurement was outlined by hand and subsequently computed.

Statistical Analysis

Measures of vocal function were determined from statistical analysis of aerodynamic, acoustic and auditory perceptual voice ratings. Ranked Spearman correlation analysis was used to examine the relation between the aerodynamic, acoustic, and auditory perceptual vocal measures and subjective (videostroboscopic) and objective (glottal) measures. The correlation (R^2) was considered significant at p-value of less than 0.05. Variables that showed significant correlation with the measures of vocal function were then used in a forward stepwise regression analysis to evaluate the combination of parameters that were most influential in predicting the measures of vocal function. Aerodynamic and acoustic variables used in this stepwise regression were adjusted for age and sex.

ANOVA and chi square were used for comparative analysis of the subjective and objective vocal function measures to evaluate variations in patients with different vocal fold positions from midline and planarity. Statistical analysis was performed using SAS PROC GLM with the lsmeans option (13).

Results

Correlation Between Horizontal Position and Measures of Vocal Function

As shown in Table 1, the horizontal position of the paralyzed vocal fold showed no significant correlation with the any of the measures of vocal function except for breathiness and SNR in the vowels /a/ and /i/ ($p<0.05$).

Comparison Among Different Horizontal Positions and Measures of Vocal Function

Comparative analysis showed no significant differences among the paramedian, intermediate and lateral positions of the paralyzed vocal fold (Table 1). Observed differences were mainly found between median and other positions of the paralyzed vocal fold. Breathiness was significantly different between the patients with vocal fold in the median and intermediate positions. SNR for the tense vowel /i/ was significantly different in the patients with the median position compared to those with paramedian or intermediate position. SNR for the vowel /a/ showed significant differences between the patients with paralysis at the median and paramedian position paralysis.

Table 1.
Comparison* Between Measures of Vocal Functions and Paralyzed Vocal Fold Position

Paralyzed VF Position	Breathiness	SNR /a/	SNR /i/
Median	1.00±1.15 (13)	16.7± 4.79(8)	23.9± 5.27(9)
Paramedian	1.53± 1.13(13)	10.5 ± 3.27(8)	15.7± 4.89(8)
Intermediate	2.31± 0.95(13)	9.73± 5.45(2)	16.1± 7.54(3)
Lateral	2.33±1.15 (3)	16.5 (1)	15.6 (1)

* mean ± SD (n)

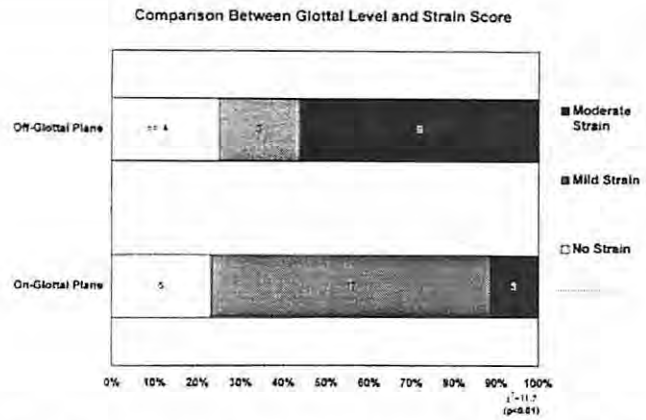


Figure 3. Significantly ($p<0.01$) more patients with off-plane glottal closure showed moderate strain than patients with on-plane closure.

Table 2.
Visual Subjective Predictors* for Measures of Vocal Functions

Voice Functions	n	Forced Variable	1st Predictor	2nd Predictor	R^2(%)
Grade	36	par-VF-position*	par-Bowing (0.443)	par-Amplitude (-0.454)	52
Breathiness	36	par-VF-position	par-Bowing (0.729)		42
Asthenia	36	par-VF-position	par-Bowing (0.616)	nor-Bowing (0.573)	44
MPT	33	par-VF-position	Gap Size (-4.15)		26
Jitter /i/	16	par-VF-position	Compensation (0.07)		80
SNR /i/	16	par-VF-position	Gap Size (-5.68)		64

par=paralysis side; nor=normal side;
( )=Regression coeffient
position=horizontal deviation from mid-line

Table 3a.
Correlation+ Between Subjective (GRBAS) Measures of Vocal Function and Glottal Measurements

Parameters \ Voice Functions	Grade	Roughness	Breathiness	Asthenic	Strain
GA-all	AD**	-	AD**	AD**	-
GA-all-mem	AD**	-	AD**	AD**	-
GA-all-par	AD*	-	AD** R*	AD*	-
GA-par-mem	AD**	-	AD** R**	AD**	-
GA-par-car	-	-	-	-	ab** r*
ant-VFA-all	AD**	-	AD**	AD**	-
ant-VFA-par	AD** AB*	-	AD** AB* R*	AD** AB*	-
ant-INA-all	-	-	AD*	-	-
ant-INA-par	-	-	AD** R*	-	-
pos-VFA-par	-	-	ad* r*	-	AD* R*
LSPM-par	-	-	AD** R**	AD*	-
LMSP	AD*	-	AD**	AD*	-
LBOW-par	AD*	-	-	AD**	-

†Capital letters=Positive correlation;
Lower case letters=Negative correlation;
*=$p<0.05$; **=$p<0.01$.

Correlation Between Vertical VF Position (Glottal Plane) and Measures of Vocal Function

Strain was significantly higher ($p < 0.01$) in patients with off-plane glottic closure (Figure 3). No other correlation between vocal fold planarity and any of the rest of the vocal outcome measures were significant.

Subjective Predictors of Measures of Vocal Function After Adjusting for Position of Paralyzed Vocal Fold

The most influential predictor variables selected by forward stepwise regression (after adjusting for paralyzed vocal fold position) are presented in Table 2. Fifty-two percent of the variation in grade was predicted by the combination of horizontal position of the paralyzed vocal fold, degree of vocal fold bowing of the paralyzed side, and severity of decreased amplitude of the paralyzed vocal fold. Forty-two percent of the variation in breathiness was predicted by the combination of horizontal position of the paralyzed vocal fold and the degree of bowing of the paralyzed vocal fold. Forty-four percent of the variation in asthenia was predicted by the combination of horizontal paralyzed vocal fold position and the degree of bowing of

Table 3b.
Correlation+ Between Objective Measures of Vocal Functions and Glottal Measurements

Voice Functions / Parameters	MPT	Fr-range	SPL-range	MFR /a,i,u/	Jitter /a,i,u/	Shimmer /a,i,u/	SNR /a,i,u/
GA-all	ad*	-	-	-	-	AD*	ad*
GA-all-mem	ad*	-	-	-	-	AD*	ad*
GA-all-car	-	-	-	AD*	-	AD*	-
GA-all-par	-	-	-	AB*	AB* R*	AD* AB** R*	ad* ab** r*
GA-all-nor	-	-	-	-	-	AD*	-
GA-par-mem	r*	-	-	AB*	AB** R*	AB** R*	ad* ab** r**
GA-par-car	-	-	-	AD*	-	AD*	ad* ab*
GA-nor-car	-	-	-	AD*	-	-	-
ant-VFA-all	-	-	-	-	-	AD*	ad*
ant-VFA-par	ab* r*	-	-	AB*	AB* R*	AD** AB** R*	ad* ab** r*
ant-INA-all	-	-	-	AD*	-	AD*	ad**
ant-INA-par	ad*	-	-	AD* AB*	AB* R*	AD* AB** R*	ad** ab* r*
pos-INA-par	-	-	-	AD* AB* R**	-	-	r*
pos-VFA-par	-	-	-	-	ad*	ad** r*	AD** R*
LSPM-par	-	-	-	AB*	AB*	AD* AB** R*	ad** ab** r**
LMSP	-	-	-	-	-	AD*	ad**
LBOW-par	-	AB**	-	AB*	AB** R*	ad** AB** R*	ab**
LBOW-nor	ad*	R*	-	-	AB** R**	AB* R*	ab** r**

+Capital letters=Positive correlation;
Lower case letters=Negative correlation;
*=p<0.05 ; **=p<0.01.

the paralyzed and normal vocal folds. The combination of the horizontal paralyzed vocal fold position and glottal gap size predicted 26% of the variability in MPT and 64% of the variation in SNR in the tense vowel /i/. The combination of the horizontal paralyzed vocal fold position and antero-posterior and/or medio-lateral compensation predicted 80% of the variability in jitter for the tense vowel /i/.

Correlation Between Objective Measures and Perceptual Measures of Vocal Function

Grade, breathiness and asthenia showed similar correlation patterns with measures of the paralyzed side glottal area, membranous glottal area, anterior vocal fold angle and maximum separation length during adduction (LMSP) (Table 3a; previous page). MPT was negatively correlated with total and membranous glottal area, degree of normal vocal fold bowing, and anterior inner angle during phonation (Table 3b). Jitter, shimmer, and SNR showed several significant correlations with the objective glottal measurements. Glottal area and anterior vocal fold angle were positively correlated with both jitter and shimmer and negatively correlated with SNR. MFR showed significant positive correlation with the normal cartilaginous glottal area during phonation and anterior inner angle of the paralyzed side at rest. Roughness and SPL range showed no significant correlation with any of the objective measurements.

Prediction of Vocal Outcome Using Objective Measures

Glottal measurements predicted a larger degree of the variability in the values of the objective measures of vocal function (MFR, jitter, shimmer, and SNR) in comparison to the subjective perceptual (GRBAS ratings) measures of vocal function (Table 4). The area measurements were more critical in predicting the subjective judgments of voice. Conversely, measurements of angles and degree of bowing were more critical in predicting the objective measures of vocal function. The total glottal membranous area during phonation predicted 35-45% of the variability in the perceptual ratings of grade, breathiness and asthenia. Both the posterior inner angle of the paralyzed side at rest and degree of bowing of the paralyzed vocal fold in adduction, regardless of the cartilaginous glottal area of the normal side during phonation, were sensitive to changes in the MFR values. The degree of bowing of both the paralyzed and normal vocal folds, regardless of anterior inner angle, were sensitive predictors for shimmer and SNR.

Table 4.
Objective Predictors for Measures of Vocal Functions

Voice Functions / Parameters	n	1st Predictor	2nd Predictor	3rd Predictor	4th Predictor	Percent R^2
Grade	40	GA-all-mem / AD (25.2)	GA-all / AD (-14.18)			35
Breathy	39	GA-all-mem / AD (16.1)				40
Asthenia	40	GA-all-mem / AD (39.1)	GA-all-par / AD (-33.1)	LSPM-par / AD (8.43)		45
Strain	39	GA-par-car / R (-7.84)				15
Fr-range	35	ant-VFA-par / R (0.29)				15
MFR /a/	32	pos-INA-par / R (3.82)	LBOW-par / AB (2370)			31
MFR /i/	29	pos-INA-par / R (4.23)	LBOW-par / AB (3363)	GA-nor-car / AD (24723)	pos-INA-par / AD (4.71)	69
MFR /u/	25	pos-INA-par / R (4.05)	GA-nor-car / AD (18716)			50
Jit /u/	9	LBOW-par / AB (9.43)				88
Shimmer /a/	17	LBOW-nor / AB (115)	ant-INA-all / AD (0.956)			63
Shimmer /i/	18	GA-all-nor / AD (293)	GA-par-mem / R (120)	LSPM-par / R (-49.3)		81
Shimmer /u/	9	LBOW-par / AB (606)				87
SNR /a/	17	LBOW-nor / R (-102)	ant-INA-par / AB (-0.315)			68
SNR /i/	19	LMSP / AD (-77.6)				43
SNR /u/	9	LBOW-par / AB (-301)				68

( ) : Regression coefficient

Discussion

Despite the limitations inherent in any retrospective study, the results reported here on 43 subjects with identical carefully collected data is worthy of discussion because it appears that horizontal position of the vocal fold is inadequate for either planning or evaluating treatment. Traditionally, clinicians have used a simple description of vocal fold position to make decisions concerning optimal surgical management of glottic incompetence secondary to UVP. Our retrospective data demonstrated that the horizontal position of the paralyzed vocal fold showed no significant correlation with most of the measures of vocal function. Consequently it should not be considered the sole basis of classification. The additional parameters of glottal area, degree of bowing of the paralyzed and contralateral vocal folds, maximum separation between vocal folds, compensatory glottal maneuvers, and the vertical glottic closure plane were found to significantly influence the quality of the voice in these patients.

Using this more descriptive schema of glottal configuration, it is easier to target problems associated with the severity of dysphonia in UVP patients and to select the most appropriate surgical option. For example, a severely dysphonic patient with a paralyzed vocal fold in the paramedian position might experience a poor vocal outcome even after phonosurgical correction of deviation from midline if the surgery did not address bowing or off-plane asymmetry of the vocal folds. In such a case, the use of injection augmentation or thyroplasty alone would likely not be as good as if it were combined with arytenoid adduction to correct the glottic plane problem.

Four specific questions were addressed in this study. The answer to the first question concerning traditional classification and voice product was consistent with literature on the importance of the subglottal shape and glottal configuration in vocal fold vibration. In other words, it is not just the vocal fold deviation from midline that should be taken into account, but the shape of the subglottis and the glottic plane of closure are also important in determining the convergence of the vocal folds during vibration. Given the same degree of deviation from the midline, the severity of dysphonias varied widely; however, when the plane of closure and bowing of the vocal folds were taken into consideration, the variability was reduced. Patients in

whom vocal fold closure was off plane exhibited evidence of hyperfunction and increased effort to compensate for the deficit. It appears that procedures to alter the glottic plane would help reduce phonatory effort and should be a focus for future research. It may also be that radiographic studies should be used to complement videostroboscopy to better determine which vertical dimensional variables contributing to voice would help in the selection of treatment. In short, horizontal deviation from midline as traditionally designed does not correlate well with 'vocal function' and other factors must be considered.

The second question addressed whether perception of horizontal vocal fold position combined with visual observation of bowing and plane of the vocal fold predicted vocal function. The results clearly support the notion that no single parameter provides the whole picture. Rather, the results suggest that a combination of observations yield a better description relating to the voice product. To the extent that these same variables are obtained post-treatment and that clinicians can predict how specific procedures affect each of the parameters (position, bowing, plane, etc.), making these observations routinely should help clinicians select appropriate treatment and predict vocal outcome. Since this study was made on individuals with UVP prior to treatment, it remains to be seen from future research if these same principles apply post-treatment.

The third question examined if there was a single measure that best correlated with vocal functions. The answer to this question is essentially, no. Many of the anatomical parameters correlated significantly to other measures of vocal functions except roughness, frequency range, and sound pressure level range. Combination of factors had similar predictive values. Some of the strange correlations, such as a negative correlation between paralyzed side cartilaginous glottal area and grade of strain, may have been due to overcrossing of the normal arytenoid and overhanging of the paralyzed side arytenoid during strained voice thus obscuring the posterior cartilagenous area on the paralyzed side. However, the majority of our results show reasonable correlations and suggest that the combinations could be used to better explain how the glottal configuration relates to voice production.

The final question addressed whether any combination of measures could predict vocal function. Our results suggest that a combination of objective parameters can predict voice product. But some of these combinations present difficulties in interpretation because of the co-existing positive and negative correlations. For example, in scoring perceptual grade of voice, membranous glottal area as the predominant predictor was positively correlated to the severity of the grade, but total glottal area as the second predictor was negatively correlated. Since the strength of the correlation of membranous glottal area was twice as great as that of the total glottal area, this combination suggested that the balance between membranous glottal area and cartilaginous glottal area during phonation was important. Interestingly a combination of positive and negative correlations was also shown as the best predictor of an asthenic voice. The strength of the correlation of membranous glottal area was nearly identical to that of the paralyzed glottal area in predicting the asthenia grade. This combination suggested that the ratio of normal glottal area compared with paralyzed cartilaginous area during phonation was positively correlated with the severity of asthenia.

The measure of the posterior inner angle represents how far the vocal process is located away from the midline. This parameter reflects the total area of the posterior third of the vocal fold and is the initially predominant predictor of the mean air-flow rate (MFR). MFR was also predicted by the normal cartilaginous glottal area. The combination of these parameters suggests the importance of the posterior valving function, regardless of the anterior glottal area, in significantly correlating with the MFR. Hirano (14) has described the cartilaginous glottis as the breathing glottis and the membranous as the phonatory glottis. In cases of UVP, when the vocal folds fail to approximate the breathing glottis may overlap the phonatory glottis causing friction and distorting the acoustic signal. When the vocal folds cannot be approximated, the delicate balance of tissue and aerodynamic forces necessary to set the vocal folds into oscillation is impeded. The present data support the notion that approximation of the vocal process plays an important role in setting the right conditions and should not be neglected in consideration of surgery.

Our study suggests that the membranous area is more accurately measured during phonation than at rest and the cartilaginous area is more accurately measured at rest than phonation. Moreover the measurement of the angle of the vocal fold at rest, rather than in the adduction or abduction phase was most powerful. At rest, the natural position of the paralyzed vocal fold is less affected by other factors; during phonation, the paralyzed vocal fold is stretched longitudinally by the normal vocal fold, reducing the appearance of the bowing. The measures made at rest were more reflective of the severity of the bowing and its functional consequences. Thus clinical protocols in the future would be enhanced by including at rest breathing tasks as well as phonation tasks.

The data presented here raise some interesting questions relative to the perceptual judgment of overall grade or severity of voice disturbance and perturbation measures. Less than 50% of the perceptual judgments on any parameter could be predicted by the measures. This is probably related to the complex nature of voice production and the fact that the superimposed physical strain may introduce more turbulence or other factors contributing to voice precepts. It also could be interpreted to support the

contention that listening alone is insufficient, even though perceptual judgments clearly need to be considered. It may be this mismatch that accounts, in part, for some surgical failures. Clinicians depending on auditory perceptual judgments and glottal horizontal position would be unable to predict the contribution of plane or hyperfunction to the voice problem or its impact on selection of surgical procedure, the need for augmentative behavioral management or its impact on subsequent vocal function.

The other question raised concerns perturbation values. Perturbation values have recently fallen out of favor because of the difficulty of accurately measuring jitter or shimmer in voices that exhibit more than 10% perturbation. It has even been suggested that perturbation measures be eliminated for clinical purposes. Individuals with vocal fold paralysis typically have high perturbation values, and it is this high perturbation value during the production of tense vowels that appears to have the greatest predictive value. This is interpreted to mean that despite their problems perturbation measures may have some value, and that while clinicians may regard the measure with caution they should not be totally disregarded.

Conclusion

In conclusion, this study suggests many glottal characteristics influence the voice product in patients with unilateral vocal fold paralysis. Clinicians must be aware of these observations in planning phonosurgical procedures to optimally improve voice quality. Assessment decisions based on horizontal position alone are too simplistic and may lead to suboptimal management decisions. Glottal measurements present themselves as useful objective means for the quantification of vocal process characteristics, and with further study may lead to more precise management decisions.

Acknowledgments

This work was supported by the National Center for Voice and Speech through Grant P60 00976 from the National Institute on Deafness and Other Communication Disorders.

The authors would like to thank Ms. Celeste Kirk for editorial assistance, Ms Joan Kozel for medical illustrations, and people working at the voice lab for their continuous help throughout this project.

Bibliography

1) Ford, C.N., Bless, D.M., Loftus, J.M.:Role of Injectable Collagen in the Treatment of Glottic Insufficiency: A Study of 119 Patients. Ann Otol Rhinol Laryngol, 101: 237-247, 1992.

2) Isshiki, N., Okamura, H., Ishikawa, T.:Thyroplasty Type I(Lateral Compression) for Dysphonia due to Vocal Cord Paralysis or Atrophy. Acta Otolaryngol (Stockh), 80: 465-473, 1975.

3) Isshiki, N., Tanabe, M., Sawada M.:Arytenoid Adduction for Unilateral Vocal Cord Paralysis. Arch Otolaryngol, 104: 555-558, 1978.

4) Crumley, R.L., Izdebski, K., Mcmicken, B.:Nerve Transfer versus Teflon® Injection for Vocal Cord Paralysis:a Comparison. Laryngoscope, 98: 1200-1204, 1988.

5) Tucker, H.M., Rusnov, M.:Laryngeal Reinnervation for Unilateral Vocal Cord Paralysis. Ann Otol Rhinol Laryngol, 90: 380-390, 1990.

6) Netterville, J.L., Stone, R.E., Luke, E.S., et al.: Silastic Medialization and Arytenoid Adduction: the Vanderbilt Experience. A Review of 116 Phonosurgical Procedures. Ann Otol Rhinol Laryngol, 102: 413-424, 1993.

7) Koufman, J.A., Walker, F.O., Joharji, G.M.: The Cricothyroid Muscle does not Influence Vocal Fold Position in Laryngeal Paralysis. Laryngoscope, 105: 368-372. 1995.

8) Woodson, G.E.,: Configuration of the Glottis in Laryngeal Paralysis. I: Clinical Study. Laryngoscope, 103: 1227-1234, 1993.

9) Smith, M.E., Berke, G.S., Gerratt, B.R., et al.: Laryngeal Paralyses: Theoretical Considerations and Effects on Laryngeal Vibration. J. Speech Hearing Res 35:545-554. 1992.

10) Sercarz, J.A., Berke, G.S.,Ming, Y.,et al.: Videostroboscopy of Human Vocal Fold Paralysis. Ann Otol Rhinol Laryngol, 101:567-577, 1992.

11) Kerr, A.G., Groves, J., Stell, P.M.:*Scott-Brown.s Otolaryngology (5th ed.).* Butterworths, London, pp179-185, 1987.

12) Woodson, G.E., Murray, T.: Glottic Configuration after Arytenoid Adduction.. Laryngoscope,104: 965-969, 1994.

13) SAS Institute Inc.: *SAS/STAT User's guide Version 6.* (4th ed.). Vol.2. Cary, NC, 1989.

14) Hirano, M.:*Clinical Examination of Voice.* Springer-Verlag Wein New York, pp4-5, 1981.

NCVS Status and Progress Report - 11
May 1997, 59-66

Efficacy as a Function of Timing in Repeated Botulinum Toxin Injections

Katsuhide Inagi, M.D.
Charles N. Ford. M.D.
Department of Surgery, The University of Wisconsin Medical School
Arthur A. Rodriquez, M.D.
Division of Otolaryngology Head and Neck Surgery, The University of Wisconsin Medical School
Departments of Rehabilitation Medicine and Neurology, The University of Wisconsin Medical School
Edward Schultz, Ph.D.
Department of Anatomy, The University of Wisconsin Medical School
Diane M. Bless, Ph.D.
Dennis M. Heisey, Ph.D
Department of Surgery, The University of Wisconsin Medical School

Abstract

This pilot study was designed to determine if the interval between repeated botulinum toxin injections influenced physiologic and histologic effects on laryngeal muscles in a rat laryngeal model The physiologic measurements included digitized videomicroscopic recording of vocal fold movement and electromyography. The histological measurements included muscle fiber size and digitized optical density of laryngeal muscles after glycogen depletion by electrical stimulation. The results demonstrated that the effect of timing of the second injection was strongly correlated to laryngeal changes. Most notable were results in the subjects that underwent injections 6 weeks apart. We hypothesize that these findings might be related to terminal axonal sprouting with reinnervation. The results from this study help confirm and expand the validity of the rat laryngeal model in understanding of the effect of botulinum toxin. Moreover, we believe that the data might be extrapolated to prove useful in predicting human responses to botulinum toxin treatment for functional dystonias such as spasmodic dysphonia.

Introduction

Spasmodic dysphonia (SD) is a focal dystonia characterized by dysrhythmic contraction of the laryngeal musculature that often results in severe voice disturbance and impaired communication. Since the etiology of this disorder remains in doubt, a definitive cure is unknown. Symptomatic control has been achieved with various medications, behavioral therapy, psychotherapy, recurrent nerve resection, thyroplasty techniques, chronic nerve stimulation, and partial thyroarytenoid myomectomy. Periodic injection of botulinum toxin (BT) has proven to be effective in temporarily relieving the symptoms of adductor SD. Unfortunately the therapeutic effect of BT is transient and the injection does not always correct the disturbance. The exact BT dosage, site of injection, and timing of repeat injections are decided on an empirical basis guided by patient demands and clinician preference, for there is little scientific basis for those decisions in the treatment of SD.

Clinical studies have demonstrated the efficacy of large (1,2) and small dosages (3-5), and have assayed whether bilateral (6,7) or unilateral (8,9) injection of BT produced the most sustained voice improvement. Our most recent study (10) suggests the importance of precise placement of injectate in the larynx for prolonging the effect. It remains unclear how the spacing of intervals between injections influences efficacy. There is also little objective data describing the measured effects of various treatments. One approach to addressing these issues would be an animal model in which physiological assays could be employed to determine efficacy as a function of repeated BT injections into the larynx.

Most of the studies of BT injection in different volumes and dosages have been limited to studying the effect of a single dosage in an animal model. Cohen et al. (11) worked out an effective dose based on dose-response curves from dog larynges. George et al. (12) mapped the affected muscle fibers using glycogen depletion techniques in the canine thyroarytenoid muscle. In our rat laryngeal model (13, 14) we demonstrated the relationships between effect and both dosage and volume using videomicroscopic and electromyographic response measurements; in these preliminary studies we established suitable dosages and volumes to inject in rat larynges. The use of a rat laryngeal model has greatly improved our ability to investigate the efficacy of BT therapy. This convenient animal model should prove helpful in developing a rational basis for treatments decisions in humans.

Materials and Methods

This pilot study was designed to determine if the interval between repeated BT injections influenced physiologic and histologic effects on laryngeal muscles. The physiologic measurements included digitized video-microscopic recording of vocal fold movement and electromyography. The histological measurements included muscle fiber size and digitized optical density of laryngeal muscle after glycogen depletion by electrical stimulation. These studies evaluated both morphological and functional changes induced by BT injection.

Rats

Nine adult male Sprague-Dawley rats, weighing between 425 and 534g (458.8 ± 31.1 g) were used for this study. Rats were anesthetized with ketamine (48 mg/kg) and xylazine (4 mg/kg) by intraperitoneal injection each experimental day (three times a week: Monday, Wednesday, and Friday). They were grouped by the time interval between the initial injection and the second injection; the groups were Control (no second injection, n = 1), 1-Week (7 days after initial injection, n = 2), 2-Week (n = 2), 4-Week (n = 2), and 6 weeks group (n = 2).

Botulinum Toxin Injection Methods

Freshly prepared crystalline complex (30 U/ng) BT was diluted with sterile saline to the dosages of 0.001 Unit in a volume of 0.4 µl. Diluted toxin was precisely injected unilaterally into the region of the thyroarytenoid muscle at the mid-portion of vocal fold using a 5-µl syringe with a 25-gauge 2-inch needle by a microlaryngoscopic technique. The side (left, right) of initial injection was randomly chosen. The second injection was given using the same methods and same side as the initial injection. For the control rat, 0.4 µl saline was injected in right vocal fold on the initial injection day only.

Physiological Measures

Laryngoscopy and Video-endoscopy . Each anesthetized rat was secured on a custom-designed operating platform in a nearly vertical position with the mouth secured open. The customized laryngoscope was inserted orally and adjusted to provide the best view of the larynx via an operating microscope (Zeiss, Thornwood, New York). The operating microscope was connected to the video equipment using a c-mount and micro-camera (CIRCON ZIMMER, Santa Barbara, California) which allowed us to view and record the larynx simultaneously. All rat vocal fold movements were recorded at the same magnification (21X). Recordings were made at the time of interval injection and objective measurements were obtained from digitized glottal images of videostrobscopic recordings (Fig. 1). Two glottal frames, selected for each rat, consisted of the image during full adduction and full abduction of respiration. Four reference points were determined for each glottal frame: posterior commissure (point P), absolutely vertical line (line H) at the point P, and edge of the right (point R) and left (point L) vocal fold at 2 mm from point P. Lateral angles between line H and line RP (angle H-RP) or LP (angle H-LP) in each glottal frame (adduction and abduction) were measured. Movement angles of the right and left vocal fold were calculated by measuring the difference between angle H-RP in adduction frame and angle H-RP in abduction frame, and between angle H-LP in adduction frame and angle H-LP in abduction frame respectively. All measurements were performed using Image 1.59 software (USDHSS-National Institutes of Health, Bethesda, MD) and a Macintosh II fx (Apple, Cupertino, CA). The number of days required to return to pre-injection movement was calculated.

Electromyography. Electromyographic sampling of the thyroarytenoid muscle was performed on a Nicollet Viking electromyograph (Mound, Minnesota) with a 50-mm, 26-gauge monopolar needle electrode and a monopolar subcutaneous reference electrode. Filter settings of 20 Hz

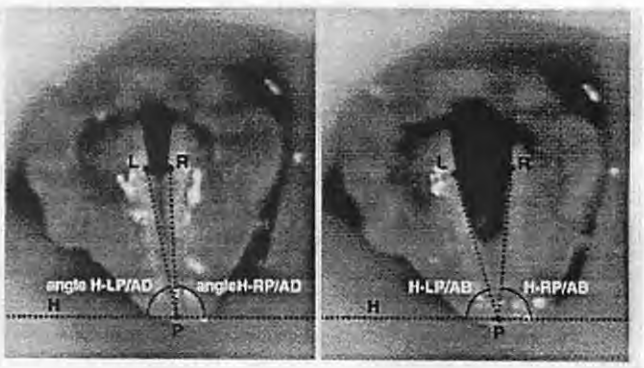


Figure 1. Digitized glottal images of videostrobscopic recordings. The angle H-RP/AD was 81.2, the angle H-LP/AD was 91.5, the angle H-RP/AB was 71.9, and the angle H-LP/AB was 85.6. Movement angles of the right vocal fold was 9.3, and left vocal fold was 5.9.

Table 1.
Rating Criteria of Interference Pattern

Grade	Interference pattern	
	Gap	Amplitude
0	no	0
1	many	much reduced
2	few	little reduced
3	no	good (equal to intact side)

Interference pattern was graded on a scale of 0 to 3.

to 10 kHz and a sweep speed of 10 msec/division were used. Display sensitivity of 50μV/division was used for interference pattern analysis. The most complete interference pattern from the thyroarytenoid muscle was obtained by manipulation of the electrode within the muscle to maximize high-frequency content and amplitude. Samples were taken before the first injection, then every 2-3 days until 70 days after the second injection. The time to return to pre-injection levels of interference was calculated. Interference pattern measurements were judged by two judges. One was blinded to the injected dosage and volume. Interference patterns were semi-quantitatively graded from 0 to 3 (Table 1). References for judgments in EMG measurements were obtained by comparison both to measurements made prior to injection and to the contralateral side. Greater interference response indicated greater muscle fiber activity.

Histological Measures

All surviving rats underwent surgery under anesthesia to isolate both recurrent laryngeal nerves on day 70 after the second injection. The isolated nerves were stimulated by bipolar stimuli for 20 min at 10 mA and 10 Hz, to achieve glycogen depletion on the intristic laryngeal muscles. Tracheotomy was performed to preserve ventilation during these procedure. All rats were then sacrificed by intracardiac injection of 7.5% chloral hydrate. Their larynges were immediately excised and frozen in liquid nitrogen. Frozen larynges were sectioned in the coronal plane continuously by cryostat at a thickness of 40 ìm from the anterior edge of the thyroid cartilage to the posterior surface of the cricoid cartilage. The frozen block was then remounted to slice the posterior cricoarytenoid muscle transversely. Sections were air-dried for 20 min. They were fixed in Histochoice (AMRESCO, Solon, Ohio) for 5 min, and washed in double-distilled water for 30 sec.

Muscle fiber size was determined as areas and perimeters of the lateral thyroarytenoid muscles on both the injected and uninjected side. The lateral thyroarytenoid muscle was divided into five parts; superior, median, inferior, lateral, and central. Measurement areas that were limited by the microscopic objective magnification (10 X) were randomly chosen in each part. All muscle fibers in the field were measured and their areas and perimeters digitized.

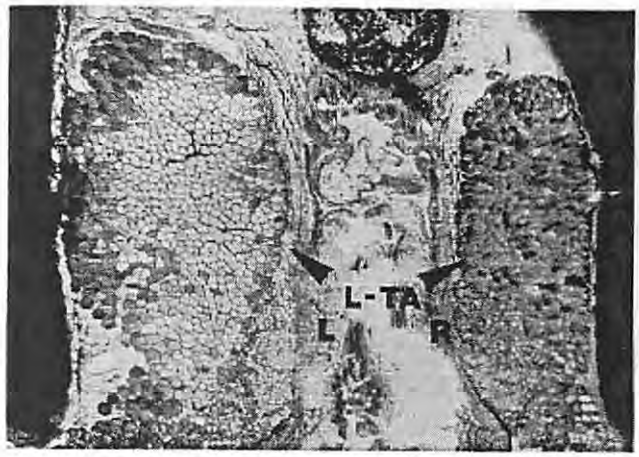


Figure 2. The mean actual optical density of the lateral thyroarytenoid (L-TA) muscle was 64.5 pixels in right (R) and 33.0 pixels in left (L) side after glycogen depletion and PAS staining in the 6-Week group. The size of the right lateral thyroarytenoid muscle was smaller than that of the left.

Optical density of muscles was determined on samples immediately adjacent to sections used to fiber size measure. Fibers were stained by the periodic acid Schiff (PAS) method. Sections washed of fixation solution were incubated in 1% periodic acid solution for 10 min. and then washed with double-distilled water for 5 min. with a water change after the first 2 min.. Then they were stained with Schiff reagent (Fisher Scientific, Fair Lawn, NJ) for 10 min and washed in 0.5% sodium metabisulfite for 2 min three times. Slides were dehydrated and covered. Optical densities as gray scales ranging from 0 to 256 pixels (Fig. 2) were digitized using a microscope (Leits Wetzlar, Germany) connected to the computer (Mac II fx) with software (Image 1.59) using a c-mount and CCD camera (CCD 72, Dage MTI, Michigan, IN). The whole area of the lateral thyroarytenoid muscle in coronal section at the level of the anterior third of the thyroarytenoid muscle was selected for measurements and each measurement was performed three times under the same magnification (2.5X) and same power of the light source. To minimize error due to the light source, we calibrated the source by absolute optical density using Photographic Step Tablet No. 2 (Kodak, Rochester, NY) before and after each measurement.

Data Analysis

Of the nine rats, one in the 4-Week group died on day 39 after the initial injection (day 11 after second injection), apparently from an anesthesia-related cause. The remaining eight rats were all observed for 70 days following second injection.

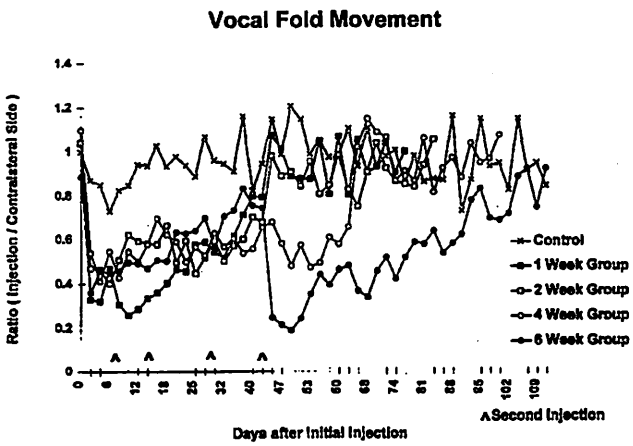


Figure 3. Longitudinal observation of mean vocal fold movement ratio in each subject group. The reductions of vocal fold movement induced by botulinum toxin with repeated injection at 1-Week or at 6-Week group were evident.

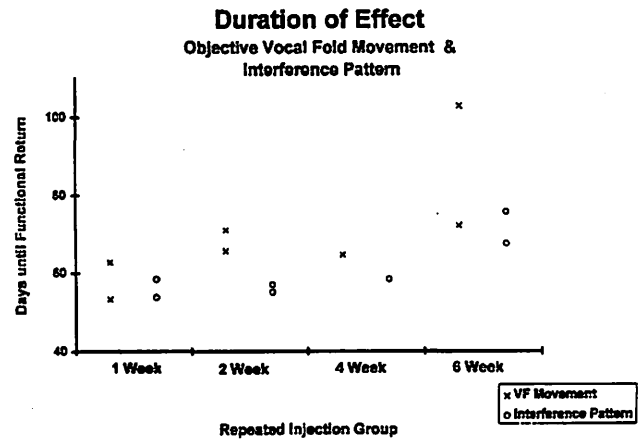


Figure 4. The subject groups were positively correlated with the time to recover function of vocal fold movement (p=0.012). Overall ANOVA was not quite significant (p=0.073). Significant pairwise differences were shown between the 6-Week group and the 1-Week group (p=0.030) and the 6-Week and 2-Week groups (p=0.030).

Measurements of vocal fold movement, muscle fiber area and perimeter, and optical density were analyzed as ratios of the injection side to the contralateral side. Use of these ratios reduced errors due to associated confounding factors such as anesthetic condition of rats, recurrent laryngeal nerve stimulation condition, histological stain variability, and artifacts.

The duration of the botulinum toxin effect was defined as the time required to recover 100% of the pre-injection status for the measurements of interference pattern and vocal fold excursion, as predicted from linear regression.

The cross-product ratio of change after the second injection (post-injection divided by pre-injection) to change after the initial injection (post-injection divided by pre-injection) of the vocal fold movement ratio was used for assessing effect of botulinum toxin. Pre-injection data for the initial injection was obtained on the day before injection. All data measured within 7 days before the second injection were used for the second pre-injection data. All measurements within 7 days after injection were used for both the initial and the second pre-injection data.

Spearman rank correlation analysis was used to examine the relation between all measurements and second injection groups. The correlation (r_s) was considered significant at $p<0.05$. If the Spearman rank correlation analysis for any group was significant, that group was further examined by comparing all pairs of means with Fisher's Protected Least Significant Difference Procedure (15). All testing was done at the nominal 5%, and the reported p-values control the comparison wide Type I error rate within each dependent variable. Inter-judge agreement on subjective assessments of change was measured.

Table 2.
The cross-product ratio of the pre- and post-injection changes of vocal fold movement ratio

	Repeated Injection Group			
	1-Week	2-Week	4-Week	6-Week
Rat #1	0.505	0.385	0.374	1.23
Rat #2	0.972	0.470	0.463	2.00

Only the 6-Week group showed that the effect of botulinum toxin induced by second injection was stronger than the effect induced by the initial injection.

Results

Vocal Fold Movement

The longitudinal observations of mean vocal fold movement ratio in each group are shown in Figure 3. The length of time needed to recover pre-injection vocal fold movement ranged from 53.4 days in the 1-Week group to 102.7 days in the 6-Week group (Fig. 4). The time needed to recover function was positively correlated ($p<0.012$) with the groups. The 6-Week group needed over 19 days on average more than any other group to recover pre-injection vocal fold movement, but this is not significant. The cross-product ratio of the changes of vocal fold movement ratio between pre- and post-injection is shown in Table 2. Only the 6-Week group showed an effect of the second injection that was stronger than that of initial injection. Although overall ANOVA was not quite significant ($p=0.056$), significant pairwise differences were shown between the 6-Week group and 2-Week group ($p=0.020$) or 4-Week group ($p=0.020$).

Electromyography

Time needed to recover to pre-injection interference activity of muscle fibers ranged from 55.1 days in the 2-Week group to 75.8 days in the 6-Week group. Both rats in the 6-Week group needed longer to recover muscle fiber function than any other groups. Although overall ANOVA was not quite significant (p=0.073), significant pairwise differences were shown between the 6-Week group and 1-Week group (p=0.030) or 2-Week interval group (p=0.030) (Fig 4).

Muscle Fiber Size

The ratios of the mean muscle fiber area in lateral thyroarytenoid muscle ranged from 0.995 in the 4-Week group to 0.801 in the 6-Week group (Fig. 5). Mean muscle fiber area ratio was negatively correlated (p=0.046) to the intervals between injections. Mean muscle fiber perimeter ratios also showed a negative correlation; there was a stronger (p=0.01) correlation between the perimeter ratio and the subject groups than between the area ratio and the groups (Fig. 5). Although the overall ANOVA was not quite significant (p=0.11), some groups showed significant pairwise differences. The perimeter ratio of the 6-Week group was significantly smaller than that of the 1-Week (p=0.04) or 2-Week (p=0.04) groups.

Optical Density

The ratios of the injection side to the contralateral side for optical density measurements of the lateral thyroarytenoid muscle were significantly (p<0.01) correlated to intervals between injections (Fig. 6). The optical density ratios of both rats in the 6-Week group were 1.55 and 1.75, higher than for any other rats in any other groups. Although the overall ANOVA was not quite significant (p=0.099), some groups showed significant pairwise differences. The 6-Week group had significantly higher optical density ratios than the 1-Week group (p=0.037) or 2-Week group (p=0.030). Actual measurements of the mean optical density of the injection side revealed significant (p=0.044) differences between the groups. The mean optical density of the 6-Week group was 55.8, significantly higher than 38.1 in the Control (p=0.031), 41.5 in the 1-Week group (p=0.032), 41.2 in the 2-Week group (p=0.030), and 28.8 in the 4-Week group (p=0.010).

Inter-rater Agreement

Exact agreement was 88.6% (636/718), and agreement within one unit was 0.14% (1/718). Spearman rank correlation revealed significant inter-judge correlation (r_2=0.82, p=0.001). These results indicate a fairly high degree of inter-judge agreement.

Discussion

This study was designed to investigate one of the variables that appears to have an impact on the efficacy of botulinum toxin (BT) injections when used in the treatment of spasmodic dysphonia. Although BT injection provides an effective means of symptom control in many patients with spasmodic dysphonia, the results are transient and repeated injections are required to achieve adequate control of symptoms. Selection of appropriate dosage, volume, and placement of injectate are usually decided empirically, based on the clinician's experience and subjective patient responses. The timing of intervals between injections varies

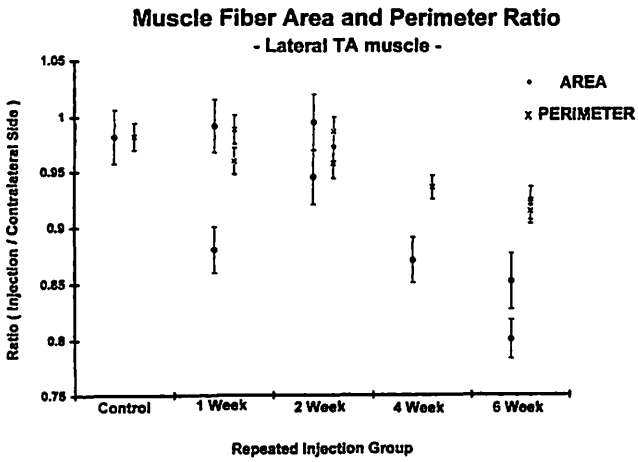


Figure 5. The intervals between repeated injections were significantly correlated to mean muscle fiber area ratios (p=0.046) and mean muscle fiber perimeter ratios (p=0.01). Although the overall ANOVA was not quite significant (p=0.11), the perimeter ratio of the 6-Week group was significantly smaller than that of the 1-Week (p=0.04) and 2-Week (p=0.04) groups.

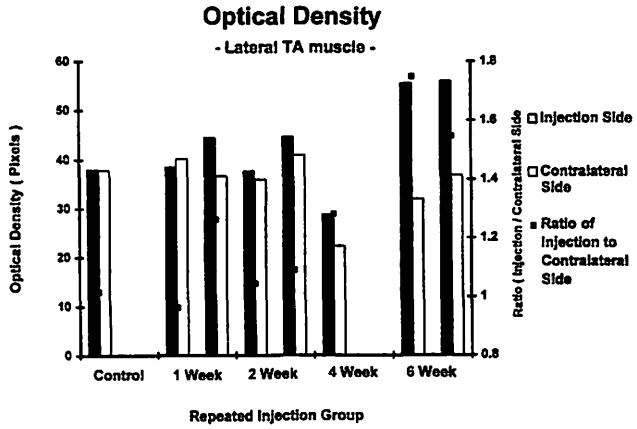


Figure 6. The ratios of the injection side to the contralateral side using the optical density measurements of lateral thyroarytenoid muscle were significantly (p<0.01) correlated to injection intervals of the subject groups. Although the overall ANOVA was not quite significant (p=0.099), the 6-Week group had significantly higher optical density ratios than the 1-Week group (p=0.037) and the 2-Week group (p=0.030). Actual measurements of the mean optical density of the injection side revealed significant (p=0.044) differences between the groups. The mean optical density of the 6-Week group was significantly higher than other that of groups.

but decisions are usually patient-driven, with some patients waiting until symptoms are severe and others seeking treatment pre-emptively before symptoms recur. To develop a scientific basis for prescribing BT injection parameters, there is a need for an animal model that can be used to demonstrate the physiological and morphological consequences of altering treatment variables. Although dogs have been used extensively for laryngeal studies, the canine larynx is anatomically quite different from the human, and, in our ongoing comparative studies of larynges from other species, there does not appear to be an ideal model. We have found the rat convenient and useful in preliminary studies (13,14) investigating the effect of dosage, volume, and placement of BT on the magnitude and duration of effect. In this study, the same model is used to establish pilot data on the effect of the intervals between injections on the physiologic and histologic changes induced by laryngeal injection of BT.

The methods employed are largely based on our prior BT studies (13,14) in the rat. The *dosage of 0.001 U in a volume of 0.4* μl was chosen because we had demonstrated that it was the lowest dosage and volume that produced the maximal duration of functional paresis without inducing complete paralysis or airway obstruction in the rat. The selection of the dose was also an attempt to simulate the incomplete paresis that is typically achieved in the human for effective symptom control and minimal side-effects. The *lateral thyroarytenoid muscle site for injection* was chosen because it is a relatively constant, a large muscle in the rat, and composed of type II fibers; this makes it analogous to the human wherein the thyroarytenoid is the muscle most typically injected for control of adductor spasmodic dysphonia. The *intervals between initial and repeated injection* were selected on the basis of linear regression analysis of our pilot data measuring the length of functional ablation from a single BT injection. We found that function was maximally reduced at 1 week post-injection; measurable recovery began to be noticed at 2 weeks, 50% return was noted at 4 weeks, and 75%-80% at 6 weeks.

Table 3.
Comparison between single and repeated injection in the time need to 100% functional return.

Functional Measure	Single Injection	Repeated Injection			
		1-Week	2-Week	4-Week	6-Week
Vocal Fold Movement	52.7	55.7	54.5	60.6	82.1
Interference Pattern	57.0	56.1	56.0	58.4	71.7

Unit : days

The time needed to recover pre-injection status for the measurements of interference pattern and vocal fold movement were compared between single botulinum toxin injection (preliminary study (14)) and subject groups.

Full return of vocal fold movement was evident at 52.7 days and complete return to pre-injection interference patterns was noted on EMGs at 57 days after initial injection (Table 3). We chose to study intervals in the period between maximal functional ablation and the recovery of 75-80% of function.

Our preliminary studies indicated a need to control for the mechanical effect of needle insertions in addition to other variables associated with the procedures in this experiment. It was apparent that repeated insertion of the EMG electrodes caused some scar formation in the laryngeal muscles of the rat. Longitudinal observations of vocal fold excursion revealed some decrease bilaterally that might be ascribed to such changes. To eliminate such factors from affecting the data in this study, we elected to employ the concept of ratios: The injected side was compared to the non-injected side and the difference expressed as a ratio. In this fashion the common events occurring due to repeated anesthesia inductions and bilateral needle insertions were accounted for by each rat serving as its own control.

The results of this study indicate a relationship between the interval between injections and the magnitude of BT-injection effect. Although there was little difference noted among the 1-week, 2-week, and 4-week groups, the 6-week interval group demonstrated a distinctly more prolonged pattern of recovery based on measurements of vocal fold movement and interference pattern. This pattern of physiologic alteration was corroborated by glycogen depletion studies demonstrating a prolonged absence of glycogen depletion capability evident by significantly greater optical densities of the PAS-stained 6-week group compared to all other groups and the controls. Histologic studies also showed a much greater degree of morphologic alteration in the 6-week group as reflected in reduced muscle fiber area and perimeter ratios.

It is not possible to extrapolate these data directly to provide a formula for treating humans with BT. These findings, however, provide a segment of the scientific matrix that might be used in developing a more rational approach to treatment. It appears that short intervals between injections (before evidence of 50% functional recovery) may not produce as prolonged a functional BT effect as somewhat longer intervals. However, as we did not study injection intervals longer than 6 weeks (at which time functional recovery was about 80%) in this model, it is possible that there is an optimal interval which could be identified if we studied longer intervals.

The relative ineffectiveness of short interinjection intervals to produce a prolonged functional ablation might be due to the relative efficacy of BT to block acetylcholine release at the neuromuscular junction of muscles that have progressed further along in the reinnervation process (16,17). The association of terminal axonal sprouting with reinnervation was evident in our studies (unpublished) using rat

gastrocnemius muscle, wherein we observed sprouting beginning two weeks after BT injection. More time is probably required for these sprouts to develop functional endplates and thus end the myoneural blockade of the muscle fiber. The resumption of myoneural junction function should be associated with a decrease in fibrillations and improvement of the interference pattern. In a previous study (14) we found that after a single BT injection of the rat thyroarytenoid muscle using the same dosage as in this study, fibrillations were maximally present within 1 week and were absent after 3 weeks. The interference pattern was maximally depressed at 1 week and returned to baseline by approximately 7 weeks. These data suggest that many axonal sprouts have established functional reconnection between 3 and 7 weeks post injection. It is possible that BT injected within a short interval after the initial injection might be metabolized or denatured without binding to the target axon because the binding sites are occupied. After the axonal sprouts have established functional reconnection or, alternatively, when the botulinum toxin is metabolized within the presynaptic axon terminal, the neuromuscular junctions might be more vulnerable to a second injection. Repeating the injection at shorter intervals might not create a greater effect than would be possible by simply combining the dosage of the two injections into a single injection.

In addition to the restoration of myoneural junction function, the re-establishment of muscle fiber diameter and strength may play a roll in determining optimal timing of repeated BT injection intervals. In order to maximally prolong the desired weakness in patients receiving multiple BT injections, it may be important to reinject the muscle after myoneural junction function has been restored, but before muscle atrophy has been reversed. In this study, muscle fiber circumference was smallest in the group reinjected at 6 weeks. Further study using longer interinjection intervals will be needed to determine the optimal reinjection interval.

One of the problems with this study is the small number of rats used in each experimental group. This is largely due to our inability to garner sufficient funding to support more comprehensive studies, for there is scientific skepticism about the use of the rat for such studies. Such skepticism is in part methodological but should be dispelled by this study and our other pilot experiments demonstrating the feasibility of assessing the effects of BT injection in the rat larynx using customized videoendoscopy, digitized measures of vocal fold excursion, EMG assessment of interference patterns, and quantification of glycogen depletion with PAS staining and optical density determinations. The other issue is whether this physiological and morphological information is valid in the human. Clearly there is not an anatomical correlation. That, however, can also be said of canine experiments, which have gained acceptance largely because of the convenience of working with a larger animal model. Increasing numbers of studies with the rat larynx have improved our understanding of the anatomy and physiology of the structure; some analogies in fiber types of the different muscle groups and functional correlates are intriguing. We propose that further use of this rat laryngeal model might shed light on basic questions about BT metabolic action in laryngeal muscles and provide an approach to correlating our knowledge about the role of sprouting and morphological changes with neuromuscular functional changes induced by intralaryngeal BT injections.

Conclusion

It appears that the interval between initial and subsequent botulinum toxin injection does affect the physiological and morphological changes induced by the toxin. These changes include decreased vocal fold excursion, loss of normal EMG interference pattern, and failure to deplete glycogen stores with prolonged stimulation as measured by the optical density of post-stimulation PAS-stained lateral thyroarytenoid muscle. There was little enhancement of effect when repeat injections were given at 1 to 4 weeks, but those rats receiving their second injection at 6 weeks exhibited a markedly heightened response. Further study with the rat laryngeal model would be useful in achieving a greater understanding of the mechanisms of action of botulinum toxin in the larynx and in developing methods to enhance laryngeal denervation and to delay undesirable reinnervation.

Acknowledgments

This work was supported by the National Center for Voice and Speech through Grant P6000976 from the National Institute on Deafness and Other Communication Disorders.

The authors would like to thank Ms. Celeste Kirk for editorial assistance, Ms. Joan Kozel for illustrations, and the people working at the voice lab, including Bridget Berning, Kelly Cavanaugh, Kelle Hasenberg, Darin Johnson, Carrie Nachtwey, Linda Suckow, Jennifer Waldron, and Jane Zernicke for their continuous help throughout this project.

References

1) Miller RH, Woodson GE, Jankovic J. Botulinum toxin injection of the vocal fold for spasmodic dysphonia. Arch Otolaryngol Head Neck Surg 1987;113:603-605.

2) Ludlow CL, Naunton RF, Sedory SE, Schulz GM, Hallet M. Effect of botulinum toxin injections on speech in adductor spasmodic dysphonia. Neurology 1988; 38 1220-1225.

3) Brin MF, Blitzer A, Fahn S, Gould W, Lovelace RE. Adductor laryngeal dystonia (spastic dysphonia): treatment with local injections of botulinum toxin (Botox ©). Movement Disorder 1989;4; 4:287-296.

4) Blitzer A, Brin MF, Fahn S, Lovelace RE. Localized injections of botulinum toxin for the treatment of focal laryngeal dystonia (spastic dysphonia). Laryngoscope 1988;98:193-197.

5) Truong DD, Rontal M, Rolnick M, Aronson AE, Mistura K. Double-blind controlled study of botulinum toxin in adductor spasmodic dysphonia. Laryngoscope 1991:101:630-634.

6) Ludlow CL, Bagley J, Yin SG, Koda J. A comparison of injection techniques using botulinum toxin injection for treatment of the spasmodic dysphonias. J Voice 1992;6(4): 380-386.

7) Maloney AP, Morrison MD. A comparison of the efficacy of unilateral versus bilateral botulinum toxin injections in the treatment of adductor spasmodic dysphonia. J Otolaryngol 1994;23(3):160-164.

8) Adams SG, Hunt EG, Charles DA, Lang AE. Unilateral versus bilateral botulinum toxin injections in spasmodic dysphonia: acoustic and perceptual results. J Otolaryngol 1993;22: 171-175.

9) Zwirner P, Murray T, Woodson GE. A comparison of bilateral and unilateral botulinum toxin treatments for spasmodic dysphonia. Eur Arch Otorhinolaryngol 1993;250:271-276.

10) Inagi K, Ford CN, Bless DM, Heisey D. An analysis of factors affecting botulinum toxin results in spasmodic dysphonia. J. Voice 1996;10(3):306-313.

11) Cohen SR, Thompson JW. Use of botulinum toxin to lateralize true vocal cords: A biochemical method to relieve bilateral abductor vocal cord paralysis. Ann Otol Rhinol Laryngol 1987;96: 534-541.

12) George EF, Zimbler M, Wu B, Biller HF, Sanders I. Quantitative mapping of the effect of botulinum toxin injections in the thyroarytenoid muscle. Ann Otol Rhinol Laryngol 1992;101:888-892.

13) Inagi K, Schultz E, Ford CN. An anatomical study of the rat larynx: Establishing the rat model for neuromuscular function. Otolaryngol Head and Neck Surgery 1995.(in press)

14) Inagi K, Rodriquez AA, Ford CN, Heisey DM. Transoral electromyographic recordings in botulinum toxin injected rat larynges. Ann Otol Rhinol Laryngol, 1995. (in press)

15) SAS Institute Inc. SAS/STAT User's Guide Version 6. 4th ed. Vol 2. Cary NC, 1989.

16) Borodic GE, Ferrante RJ, Pearce LB, Alderson K. Pharmacology and Histology of the Therapeutic Application of Botulinum Toxin, 1st ed. New York. Marcel Dekker, 1994:119-157.

17) Cichon JV, McCaffrey TV, Litchy WJ, Knops JL. The effect of botulinum toxin type A injection on compound muscle potential in an in vivo rat model. Laryngoscope 1995;105: 144-48.

NCVS Status and Progress Report - 11
May 1997, 67-74

An Anatomical Study of the Rat Larynx: Establishing the Rat Model for Neuromuscular Function

Katsuhide Inagi, M.D.
Division of Otolaryngology-Head and Neck Surgery, The University of Wisconsin Medical School
Edward Schultz, Ph.D.
Department of Anatomy, The University of Wisconsin Medical School
Charles N. Ford, M.D.
Division of Otolaryngology-Head and Neck Surgery, The University of Wisconsin Medical School

Abstract

The gross and microscopic anatomy of the rat larynx was studied with particular attention to myology and neuromuscular structures to further validate it as a model to evaluate morphological and functional changes induced by botulinum injection. A laryngeal alar cartilage (LAlC), alar cricoarytenoid (ACA) muscle, and a superior cricoarytenoid muscle (SCA) were identified as new structures not previously described. Two portions (medial and lateral) of the thyroarytenoid muscle were distinguished. The function of the ACA was suggested to be similar to the aryepiglottis muscle in the human and the function of the SCA was suggested to be similar to the human interarytenoid muscle. The predominant pattern of motor endplate (MEP) distribution in rat laryngeal muscles (PCA, LCA, CT, and SCA) was to have MEPs concentrated mostly at the mid-belly of muscle where they were distributed throughout the cross-sectional area of the mid-belly. The TA and ACA differed from this pattern. The lateral TA had MEPs concentrated at the anterior third of its belly and medial TA were located at mid-belly. MEPs in the ACA were located mostly at the posterior portion of muscle. Muscle fiber-typing showed subtle differences between intrinsic laryngeal muscles. Fast fibers were predominant in the rat laryngeal muscles. This study supports the expanded use of rats in studies on laryngeal neuromuscular function and disease in humans.

Introduction

Rats are popular laboratory animals for experimental models because of their low cost, availability, and ease of handling. Although studies using the rat larynx as experimental models are uncommon, rat larynges have been used recently in studies of inhalation toxicity (1-3), allograft rejection (4,5), and irradiation responses (6).

We have developed a rat laryngeal model for assessing the effects of botulinum toxin injections (7) to ameliorate certain spastic conditions. The model has been facilitated through the development of a novel technique to observe the rat larynx endoscopically during and after injection of toxin. This technique has greatly improved our ability to investigate the efficacy of botulinum toxin for extended periods after injection by measuring physiologic changes.

To maximize our ability to evaluate morphological and functional changes induced by botulinum injection in the rat model, we have undertaken a number of studies to provide a thorough understanding of the anatomy of rat larynx. Although a few researchers (8-10) have described laryngeal structure, especially related to intrinsic laryngeal musculature, photographs of these muscles have not been made available. Moreover, no studies have described the distribution of motor endplates (MEPs) or muscle fiber-types, both of which are of interest in neurolaryngologic phenomena such as the effects of botulinum toxin.

In this study, we present gross and microscopic rat laryngeal anatomy with particular attention to neuromuscular structures.

Materials and Methods

Twelve Sprague-Dawley male adult breeder rats, weighing between 450g and 550g, were used for this study. The rats were sacrificed by intracardiac injection of 7.5%

chloral hydrate after which their larynges were immediately excised. Larynges were used either whole or in sections for basic anatomy, motor-endplate distribution, and fiber-typing.

Anatomy of the Rat Larynx

The anatomy of each of the laryngeal muscles was studied in three rats by microdissection using Nikon (Chiyoda, Tokyo, Japan) and Zeiss (Thornwood, New York, USA) operating microscopes. Overlying connective tissue was removed from the laryngeal muscles after which photographs were taken. From these photographs the maximum length of each muscle was digitized using NIH Image 1.44 software (USDHSS-National Institute of Health, Bethesda, MD) and an Apple computer II fx (Cupertino, CA).

Motor-Endplate Distribution

Whole Mounts: The larynges of three rats were fixed in fixative solution (40% formaldehyde 20 ml, sodium chloride 4.25 g, acid sodium phosphate monohydrate 0.8 g, anhydrous disodium phosphate 1.3 g, and deionized water 180 ml) and prepared as whole mounts to visualize the distribution of motor endplates. The method for doing this is basically the same as that indicated below for sectioned material except that the staining times are protracted.

Sections: The larynges from an additional four animals were removed and immediately frozen in liquid nitrogen. Transverse and coronal sections were made of all muscles from two rats for each plane and placed on microscopic slides. Serial coronal sections were taken from the anterior part of the thyroid cartilage and serial transverse sections were taken from the tip of the epiglottis. All sections were numbered and placed in order on glass microscope slides. The slides were processed as indicated below to reveal the presence of motor endplates.

Staining Technique: Frozen larynges were sectioned continuously by cryostat at a thickness of 20 ìm (transverse and coronal planes). They were fixed on slides in acetone for 10 min, then air dried for 20 min. They were washed with phosphate buffered saline (PBS) for 5 min three times, then stained with cholinesterase stain (5-bromoindoxyl acetate 8 mg, ethanol 6 ml, potassium ferrocyanide 0.126 g, potassium ferricyanide 0.1 g, tris-HCL 0.084 g, tris-base 0.008 g, calcium chloride 0.066 g, and double-distilled water 60 ml) for 30 min until the blue reaction product appeared at the site of each motor endplate. Samples were washed in PBS for 5 min three times, and counterstained with eosin for 10 min.

Measurement: Numbered sections from the serial sets were selected at 60 ìm intervals for each muscle. This was to ensure that the number of stained MEPs counted in each muscle came from matched levels (area). Sections were selected every 60 ìm, and the number (numerical density) of stained motor endplates in each laryngeal muscle was counted. An average number of MEPs in matched sectioned levels was graded from 0 to 3 (0 = no MEPs, 1 = 1 to 20 MEPs, 2 = 21 to 40 MEPs, and 3 = over 41 MEPs). The graded numbers from transverse-sectioned series or coronal-sectioned series were plotted to describe the distribution of MEPs in each laryngeal muscle.

Myosin Fiber-Typing

Four animals were used to assess the distribution of fiber-types. The larynges from two of the four rats were excised, fixed in Histochoice (AMRESCO, Solon, Ohio, USA), then embedded in paraffin. Paraffin-embedded sections were cut at 10 ìm and placed on glass slides and immunostained for myosin. Frozen sections obtained from the larynges of the two remaining rats were used to determine motor-endplate distributions or both myosin and motor endplate distribution. The distribution of fast-myosin-containing myofibers was determined using paraffin and frozen sections by immunostaining using antibody MY-32 (Sigma Chemical Company, St. Louis, MO, USA), which stains all fast-myosin-containing fibers. Fibers containing no fast myosin remained unstain.

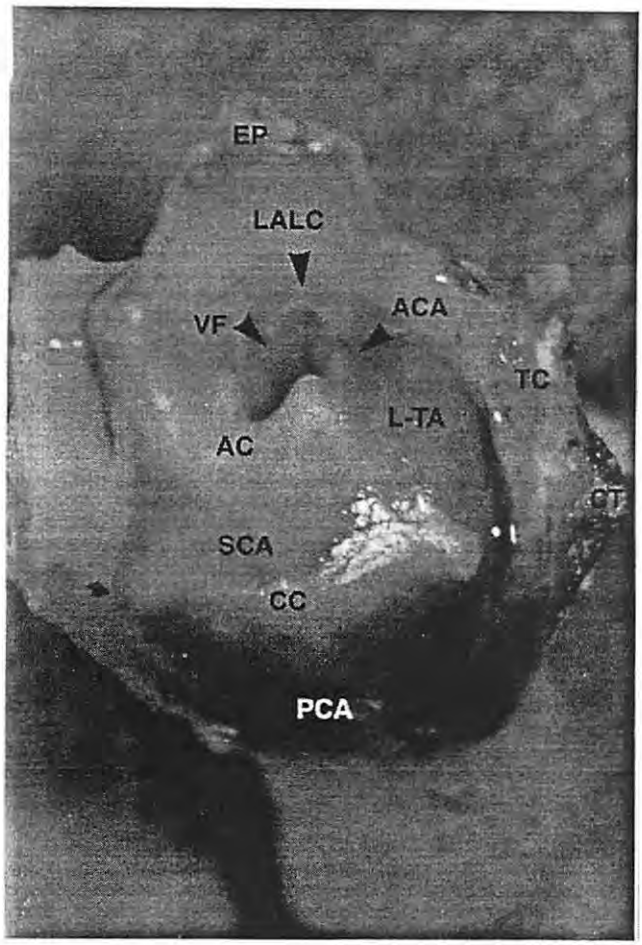
Sections are incubated in primary antibody (MY-32 diluted 1/600) overnight at room temperature, then washed with PBS for 5 min three times. They are incubated in 20% goat serum for 10 min to block non-specific reaction, and then incubated in a second antibody (mouse Ig-G, diluted 1/100) for 30 min. Sections are washed in PBS for 5 min three times, and incubated in diaminobenzidene (DAB) at room temperature until the reaction product appears. Slides are dehydrated and covered. Photographs of ten random fields for each laryngeal muscle were taken at 100 X magnification. The numbers of MY-32-positive (stained) and -negative (unstained) muscles were counted.

This study was performed in accordance with the PHS Policy on Human Care and Use of Laboratory Animals, the NIH Guide for Care and Use of Laboratory Animals, and the Animal Welfare Act (7 U.S.C. et seq.); the animal use protocol was approved by the Institutional Animal Care and Use Committee (IACUC) of University of Wisconsin.

Results

Anatomy of the Rat Larynx

The skeleton of the rat larynx consists of thyroid cartilage (TC), cricoid cartilage (CC), and arytenoid cartilage (AC). An extra cartilage, shaped like a wing, was observed at the anterior half of vocal fold near the base of the epiglottis. We named this cartilage the laryngeal alar cartilage (LAlC) because of its delta wing shape (Figure 1). This cartilage was connected to the base of the epiglottis by bundles of connective tissue and served as an attachment site for muscles described in Table 1. Figure 1 showed the posterior view of the superior aspect of the rat larynx.



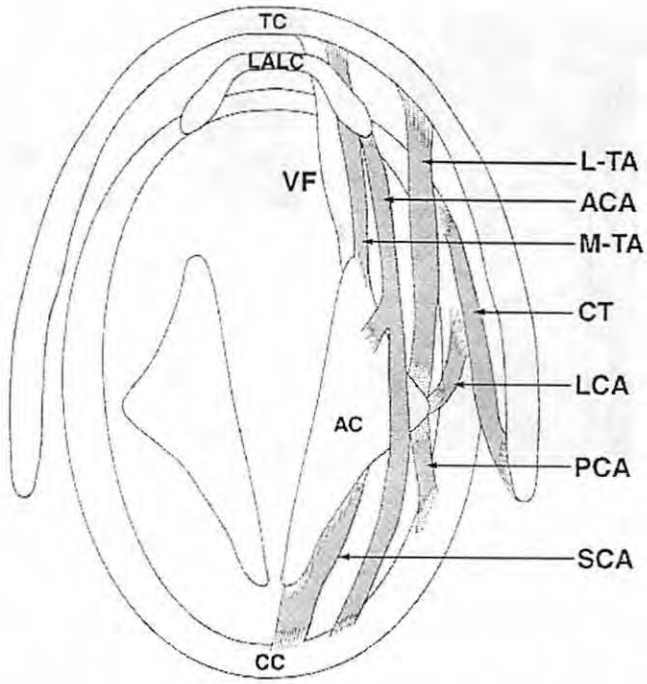


Figure 1. (A; left) Posterior view of the superior aspect of the rat larynx. The skeleton of the rat larynx consists of thyroid cartilage (TIC), cricoid cartilage (CC), arytenoid cartilage (AC), and laryngeal la cartilage (Laic). Myology of the rat larynx consists of thyroarytenoid muscle (TA; lateral TA[L-TA] is shown), cricothyroid muscle (CT), lateral cricoarytenoid muscle (LCA; not shown), posterior cricoarytenoid muscle (PCA), alar cricoarytenoid muscle (ACA), and superior cricoarytenoid muscle (SCA). LAIC, ACA, and SCA are structures that have not been described before. (EP: epiglottis). (B; above): The schematic drawing of the muscle insertion/ origin relative to the cartilages.

The thyroarytenoid muscle (TA), cricothyroid muscle (CT), lateral cricoarytenoid muscle (LCA), and posterior cricoarytenoid muscle (PCA) were located as in the human larynx (Table 1). The TA was apparently divided into a medial TA and a lateral TA by vertical septum (Fig 3a; following page). Two additional muscles were identified that are not seen in the human larynx. One, the longest laryngeal muscle (average length, 3.02 mm), was attached anteriorly to the LAlC. This muscle ran posteriorly in a position that was superior or cephalic to the TA. Posteriorly this muscle attached to the medial and lateral portions of the muscular process of the arytenoid. The lateral-most portion of the muscle passed beyond the arytenoid to attach to the CC. We named this muscle the alar cricoarytenoid muscle (ACA) (Figure 1). The second newly identified muscle ran posteriorly and medially from the arytenoid to a midline tubercle on the cricoid. We named this muscle the superior cricoarytenoid muscle (SCA) (Figure 1). The lateral part of SCA was positioned just inferior to the ACA during its posterior course.

Table 1.
Anatomy of the Rat Larynx and Myosin Fiber-Typing

Muscle	Attachment From - To	Average Length (mm)	(Range)	Fast Fiber (%)	(Range)
TA Lateral	AC - TC	2.93	(2.81-3.12)	100	
Medial	AC - TC		-	100	
LCA	AC - CC	1.77	(1.69-1.81)	94.9	(91.7-97.9)
CT	TC - CC	2.30	(2.20-2.49)	71.2	(68.0-73.8)
PCA	AC - CC	3.93	(3.81-4.09)	81.2	(76.1-85.9)
SCA	AC - CC	1.17	(1.09-1.21)	99.0	(98.3-99.5)
ACA	LAlC-both AC & CC	3.02	(2.89-3.12)	100	

Motor-Endplate Distribution

Whole Mounts: TA (lateral). Motor endplates (MEPs) were located in a transverse band on the lateral aspect of the lateral TA that continued onto the superior aspect this muscle. On the superior aspect of this muscle, this band appeared in a zigzag fashion (Figure 2a; following page). The medial portion of the TA could not be observed directly in whole mounts.

LCA. MEPs were located in a band that ran transversely across the midbelly of the LCA (Figure 2a).

ACA. MEPs were in a band that ran horizontally from lateral to medial at the posterior belly near the medial portion of muscular process of the ACA (Figure 2a).

SCA. MEPs were located in a band that ran across the lateral third of the belly in the posterior portion and midbelly in the medial portion of the SCA (Figure 2a).

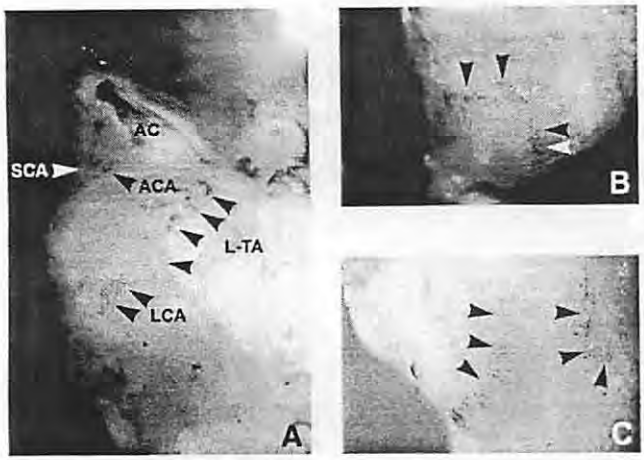


Figure 2. Whole mount of a rat larynx stained by cholinesterase stain to reveal the distribution of MEPs as black spots (A: lateral view after removal of TC, B: lateral view, C: posterior view). The MEP band of the lateral TA (A) runs in a zigzag fashion on the superior aspect this muscle. The MEP band of the LCA (A) runs transversely across the midbelly. MEPs of the CT (B) are located in an diagonally arched band that run transversely across the midbelly of the muscle. MEPs of the PCA (C) are located in a band that runs diagonally from medial in the superior portion of the muscle to lateral in the inferior portion.

CT. MEPs were located in a diagonally arched band that ran transversely across the midbelly of the CT muscle (Figure 2b). The band extended from the lateral-inferior aspect of the CT to the anterior-superior aspect of the muscle near the midline of the larynx.

PCA. MEPs were located in a band that went diagonally from medial in the superior portion of the PCA to lateral in the inferior portion (Figure 2c).

Microscopic Sections: For these studies rat larynges were cut in either the coronal or transverse plane and the distribution of MEPs is reported with respect to these planes. Each of muscles below was sectioned from anterior to posterior in the coronal plane, and from superior to inferior in the transverse plane.

Lateral TA

Coronal Sections: MEPs were concentrated in the superior portion of the anterior belly of the muscle (not illustrated), and in the inferior portion of the posterior belly (not illustrated). The greatest density of MEPs was in the anterior 1/3 of the muscle (Figure 3a) where they were scattered throughout the breadth of the muscle section.

Transverse Sections: As the muscle was sectioned from superior to inferior, MEPs were located from anterior to posterior (Figure 3a). The MEP band was widest and of the highest density in the mid-portion of the muscle belly (Figure 3f).

Medial TA

Coronal Sections: A similar pattern as that seen in the lateral TA was observed. MEPs were located in the superior portion of the anterior belly and in the inferior portion of the posterior belly. Unlike the lateral TA the greatest density was of MEPs was in the mid-portion of the muscle and were not as scattered (Figure 3a).

Transverse Sections: As the muscle was sectioned from superior to inferior, MEPs were located from the mid-belly of the muscle. In the inferior portion of the muscle, MEPs were located in the posterior belly. There was a paucity of MEPs in the anterior portion of the muscle belly. The greatest number of MEPs were located in sections from the muscle mid-portion where they were scattered throughout the muscle (Figure 3f).

LCA

Coronal Sections: MEPs were present in the superior portion of the anterior belly of the muscle. In the mid-belly of the muscle MEPs were at their highest density and were scattered throughout the breadth of the muscle (Figure 3b). In the posterior belly MEPs were located in a median position.

Transverse Sections: As the muscle was sectioned from superior to inferior MEPs were located first laterally in the anterior portion of the muscle and then in the median position in the posterior muscle belly. The highest density of MEPs was in the sections from the mid-portion of the muscle belly and were present in the entire breadth (Figure 3g).

CT

Coronal Sections: In the anterior belly of the muscle MEPs ran in a small band between the TC and CC. In sections from the mid-belly MEPs were located in a wide band toward the lateral portion of the muscle (Figure 3b). This pattern of distribution persisted in the posterior belly of the muscle, although the number of MEPs was decreased (Figure 3b).

Transverse Sections: As the muscle was sectioned form superior to inferior, MEPs were seen first in a band running through the mid-belly of the muscle (Figure 3g). In sections from the mid-belly, two MEP bands were observed (not illustrated). In the posterior portion of the muscle belly a single band was present (not illustrated).

PCA

For this muscle, coronal section passed from anterior to posterior through the belly, and transverse sections passed from superior to inferior.

Coronal Sections: In the superior belly of the muscle MEPs located in the medial portion and in the lateral portion of the inferior belly (not illustrated). In the mid-belly MEPs were present as a short wide band (Figure 3h).

Transverse Sections: As the belly was sectioned from superior to inferior, MEPs were first located in the medial (superior) and then lateral (inferior) portions of the muscle.

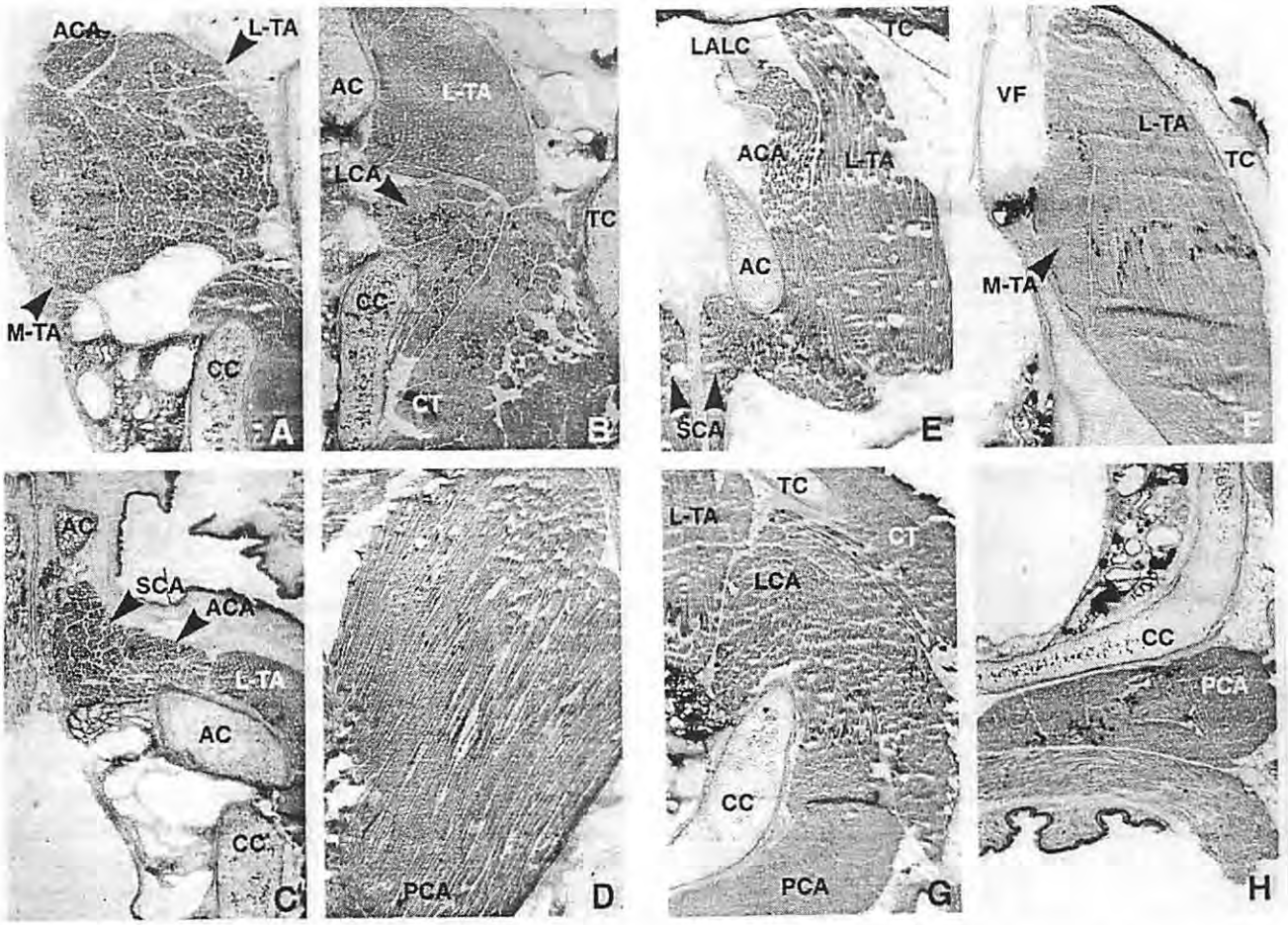


Figure 3. MEPs appear as black spots on the muscle sections. A-D: Coronal views; E-F:Transverse views. The number of the sliced section is matched to the numbers on the X grid of Figure 4 and is a distance of 0.06 mm times the number from the anterior/superior. A: MEPs are scattered throughout the lateral TA. This section number is 27, 1.62 (27 x 0.06) mm posterior to the beginning of thyroid cartilage. B: MEPs are scattered throughout the LCA on section number 49. C: MEPs of the SCA are scattered throughout the muscle in section number 58. D: MEPs run diagonally in a zigzag fashion from the medial in the superior portion of the muscle to lateral in the inferior portion (section number 79). E: MEPs of the lateral TA, ACA, and SCA run in a band at the anterior, mid-, and posterior third of each muscle belly, respectively. This section number 16 was 0.96 (16 x 0.06) mm inferior to the top of LAlC. F: MEPs of the lateral and medial TA are seen in the different bands on section number 28. G: MEPs of the CT course in a band through the muscle in a curved fashion. An MEP band is seen at the lateral part of the LCA in section number 43. h: A short, wide band of MEPs appears in the midportion of the PCA muscle, section number 58.

ACA

Both coronal (Figure 3c) and transverse (Figure 3e) sections of the ACA showed MEPs in the posterior third of the muscle belly.

SCA

Coronal Sections: In the anterior SCA, MEPs were located in the medial part of the muscle belly (not illustrated). MEPs were scattered throughout the muscle in the middle of the SCA illustrated in Figure 3c. In the posterior SCA, MEPs were located in the inferior portion of the muscle belly.

Transverse Sections: MEPs ran in a band across the lateral third of the belly in the posterior portion but, at midbelly, they were located in the medial portion (Figure 3e).

In summary, MEPs of all laryngeal muscles are located at the midbelly except for the lateral TA and ACA (Figures 4a, 4b; following page). MEPs in the lateral TA are located in the anterior third of its belly, and MEPs of the ACA are located in the posterior portion of its belly. The MEP band in the medial TA is about 0.3 mm wider than in the lateral TA and is located more anteriorly.

Myosin Fiber-Typing

Table 1 shows the percentage of MY-32-positive fibers in laryngeal muscles. The TA (Figure 5a; following page) and ACA muscles were composed exclusively of fast fibers. In contrast, the CT had the lowest percentage of fast fibers (Figure 5b). The diameter of MY-32-negative muscle fibers tended to be smaller than the diameter of positive muscle fibers.

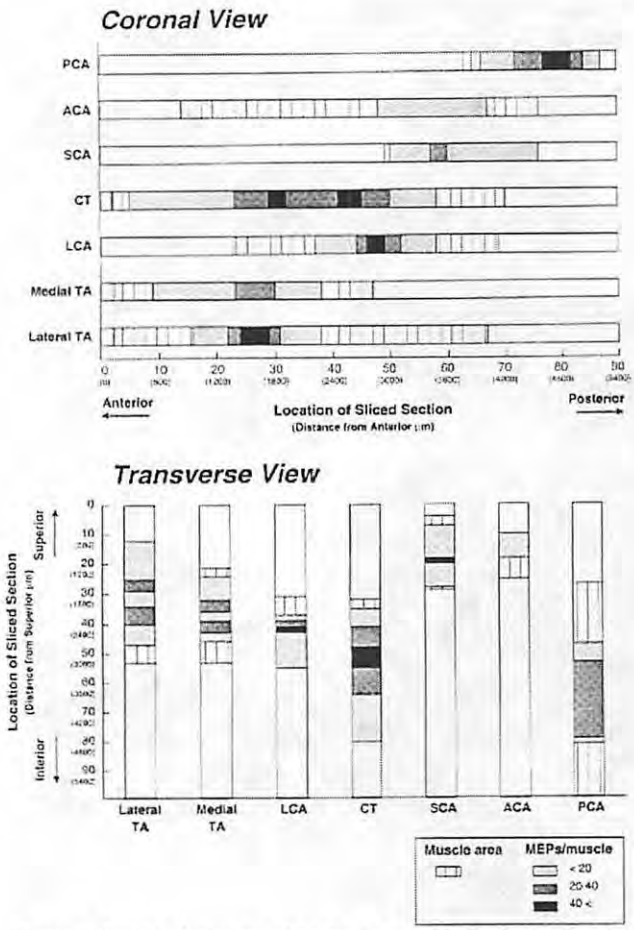


Figure 4. Graphic depictions of the density of MEP distribution based on the coronal section view (n=2, upper) and on the transverse section view (n=2, lower) of each rat laryngeal muscle. The number on the X/Y grid matches the section numbers in Figure 3; each section equals 60 im thick. MEPs of all muscles were located at the midbelly except the lateral TA and ACA. MEPs in the lateral TA are located in the anterior third of its belly, and MEPs of the ACA are located in the posterior part of the belly.

Discussion

This study was designed to investigate the basic anatomic structure of the rat larynx to further validate its use as a model to evaluate morphological and functional changes induced by botulinum toxin injection. Little previous work has been done on the myology or MEP distribution in the rat larynx

This study has elucidated four features that have not been fully described previously. We have described the location and attachments of three muscles -- the TA, ACA and SCA -- and the location and configuration of the LAlC cartilage. Similar muscles were described by Kobler et al. (8) in 1994, but our observations suggest that the ACA originates from the LAlC and inserts on both the AC and CC. In addition, we found that the LAlC is located superior to the middle portion of the vocal fold near the base of the epiglottis.

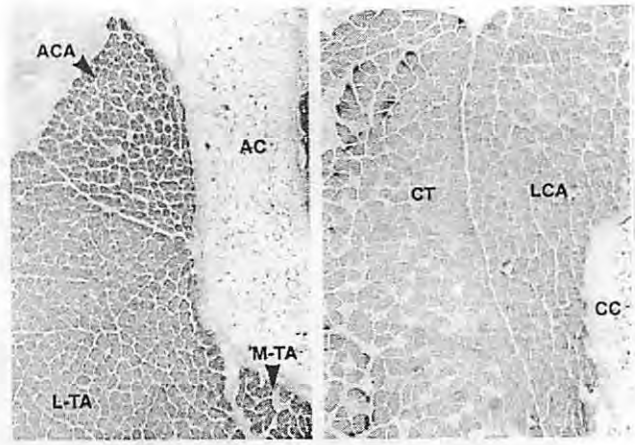


Figure 5. MY-32-positive (fast fiber) muscles are shown in ACA (left), lateral TA (left), medial TA (left), LCA (right), and CT (right).Percentages of MY-32-positive fibers in these muscles are shown in Table 1.

The functions of these additional muscles appear to compare with those of muscles in the human larynx. Based on attachments and direction of the muscle fibers in the rat, the SCA appears to be analogous to that of the interarytenoid muscle in the human larynx in that it is also in a position to adduct the AC (11,12). The reason for the arrangement of these muscles appears related to the structure of the AC. The rat AC is relatively larger than that of the human. The vocal process and body of the AC are enlarged posteriorly and this long cartilaginous extension renders the posterior commissure angle very sharp (see Fig 1). This acute angle makes it impossible to connect each AC by muscles. The SCA structure and its connection between the AC and CC provides an improved mechanical advantage, making it better able to achieve complete closure of the cartilaginous vocal folds. The ACA is connected to the LAlC and to the muscular process of both the AC and the CC. The LAlC appears tightly attached to the base of the epiglottis. Therefore, the ACA seems to function in a protective role in that it is capable of closing the laryngeal opening and depressing the epiglottis as does the aryepiglottic muscle in the human larynx. Another role for the LAlC in the rat larynx might be to maintain an open laryngeal vestibule for respiration. This is important since the rat laryngeal vestibule is otherwise wide but lax.

A good understanding of MEP distribution in each of the laryngeal muscles is helpful to obtain the maximum effect of botulinum toxin injection, for the toxin blocks presynaptic release of acetylcholine at the neuromuscular junction. Our study reveals the major patterns of MEP distribution in rat laryngeal muscles. MEPs in the PCA, LCA, CT, and SCA are concentrated mostly at the midbelly of muscle. In the lateral TA, MEPs are concentrated at the anterior third of its belly where they are scattered throughout the cross-sectional area of the muscle. The ACA and medial TA differ

from this pattern: in the ACA, MEPs are located mostly at the posterior portion of the muscle; MEPs in the medial TA are located at the midbelly. In both of these muscles the MEPs are distributed diffusely; no concentrated areas are observed.

In the human larynx, Rosen et al. (13) reported that MEPs in the TA muscle were diffusely distributed throughout the muscle. Similar findings were reported in the cat larynx. However, these findings were based on data obtained from only transverse sections of both lateral and medial portions of the TA and did not distinguish the medial and lateral portions. It was suggested that the reason for scattered MEPs in the TA is that TA muscle fibers have multiple MEPs, and/or that the TA is geometrically complex. The present study suggests that the rat TA muscle is also geometrically complex, but that the distribution area of MEPs is not spread diffusely throughout the muscle. Rather, when the muscle is viewed in at least two planes of section and the medial and lateral portions of the muscle are studied individually, the TA appears not to have as diffuse a pattern as suggested by other investigators. The diffuse appearance might be related to differences in the MEP pattern in two portions of the TA muscle. We found that the width of the MEP band distribution was not the same in the lateral and medial TAs. The MEP band in the medial TA is about 0.4 mm wider than that in the lateral TA and is located more anteriorly. Consequently, the bands in the two portions of the muscle are staggered: the MEP band of the medial TA extends further anteriorly than that of the lateral TA. When the two portions are juxtaposed, the appearance is a very wide, diffuse innervation band. Overall the distribution of the MEPs in each laryngeal muscle of the rat correlated with the distribution of MEPs in the human larynx except possibly the TA muscle. These results suggest that further studies on the human TA using multiple planes of section are warranted to show whether the pattern is the same as reported here for the rat.

The distribution of fibers with myosin subtype isoforms has been the subject of several recent publications (14) showing changes in myosin subtypes induced by age (15) and nerve injury (16). This study demonstrates that the fiber-type distribution of laryngeal muscles in rats is predominantly fast. This baseline of fiber-type distribution in each of the laryngeal muscles will be important in assessing subtle fiber-type changes induced by botulinum toxin.

The data from this study should help to establish the rat as a useful animal model for laryngeal studies and provide a baseline that can be used in future studies to assess morphological or functional changes induced by agents such as botulinum toxin. For example, the rat model could be used to assess the effect of variations in treatment technique on magnitude and duration of response of botulinum toxin injection in adductor spasmodic dysphonia patients. Other potential uses will be facilitated with further study.

Summary

Because the rat is less expensive and more accessible than dogs or primates, it is a useful experimental model. They are particularly attractive for studies requiring a great number of subjects. We studied the gross and microscopic anatomy of the rat larynx with particular attention to the myology and neuromuscular structures. Two bellies of TA, lateral and medial, were identified. An laryngeal alar cartilage, an alar cricoarytenoid muscle, and a superior cricoarytenoid muscle were identified and described.

The major patterns of MEP distribution in rat laryngeal muscle were demonstrated. PCA, LCA, CT, and SCA had MEPs concentrated mostly at the midbelly of muscle. MEPs in the medial TA were located at midbelly, while the lateral TA had MEPs concentrated at the anterior third of its belly where they were scattered throughout the cross-sectional area of muscle. MEPs in the ACA were located mostly at the posterior portion of muscle. In each muscle, MEPs were distributed diffusely. Muscle fiber-typing showed subtle differences among intrinsic laryngeal muscles. Fast fibers were predominant in the rat laryngeal muscles.

This study supports the expanded use of rats in studies on laryngeal neuromuscular function and disease with greater understanding of the neuromuscular and skeletal structure of the rat larynx.

Acknowledgments

This work was supported by the National Center for Voice and Speech through Grant P6000976 from the National Institute on Deafness and Other Communication Disorders.

The authors would like to thank Ms. Celeste Kirk for editorial assistance, Ms. Joan Kozel for illustrations and Miss. Amy Ella Mast for creating charts.

References

1) Sagantz JW, Madarasz AT, Forsell MA, Burger GT, Ayres PH, Coggins CR. Histological sectioning of the rodent larynx for inhalation toxicity testing. Toxicolog Pathol 1992;20(1):118-121.

2) Lowis DJ. Morphological assessment of pathological change within the rat larynx. Toxicolog Pathol 1991;19(4):352-357.

3) O'Halloran KD, Curran AK, Bradford A. Ventilatory and upper-airway resistance response to upper-airway cooling and CO_2 in anaesthetized rats. Pflungers Archiv-Eur J Physiol 1994;429(2):262-266.

4) Strome S, Sloman-Moll E, Samonte BR, Wu J, Strome M. Rat model for vascularized laryngeal allograft. Ann Otol Rhinol Laryngol 1992;101(11):950-953.

5) Strome S, Wu J, Strome M, Brodsky G. A comparison of preservation techniques in vascularized rat laryngeal transplant model. Laryngoscope 1994;105:666-668.

6) Lidegran M, Kjorell U, Henriksson R, Forsgren S. Bombesin-like immunoreactivity in the rat larynx: increase in response to irradiation. Regulatory Peptides 1995;55(3):321-330.

7) Inagi K, Rodriquez AA, Ford CN, Heisey D. Transoral electromyographic recordings in botulinum toxin-induced rat larynges. Ann Oto Rhino Laryngol (submitted)

8) Kobler JB, Datta S, Goyal RK, Benecchi EJ. Innervation of the larynx, pharynx, and upper esophageal sphincter of the rat. J Comp Neurol 1994;349:129-147.

9) Aikawa M, Shimozawa A. Anatomical observations on the musculi suprahyoidei and apparatus hyoideus with special reference to the musculus jugulohyoideus and os stylohyoideum of the mouse, hamster and rat. Acta Anat 1994;150(4):294-306.

10) Hebel R, Stromberg MW. Anatomy of the Laboratory Rat. The Williams and Wilkins, Baltimore, 1976:58-59.

11) Aibara R. Motor innervation of the canine arytenoid muscle. J. of Otolaryngology of Japan, 1991:805-816.

12) Kawakita S. Motor innervation of the guinea pig arytenoid muscle from the standpoint of the reinnervation process. J. of Otolaryngology of Japan, 1995: 391-401.

13) Rosen M, Malmgren LT, Gacek RR. Three-dimensional computer reconstruction of the distribution of neuromuscular junctions in the thyroarytenoid muscle. Ann Otol Rhinol Laryngol 1983;92:424-429.

14) Sawchak JA, Leung B, Shafig SA. Evidence for a new isoform of fast myosin heavy chain in rat skeletal muscle. Muscle Nerve 1992;15(12):1349-1353.

15) Larsson L, Biral D, Campione M, Schiaffino S. An age related type IIB to IIX myosin heavy chain switching in rat skeletal muscle. Acta Physiolog Scandi 1993;147(2):227-34.

16) Jakubiec-Puka A. Changes in myosin and actin filaments in fast skeletal muscle after denervation and self-reinnervation. Comp Biochem & Physiol. Comp Physiol 1992;102(1):93-98.

NCVS Status and Progress Report - 11
May 1997, 75-82

Transoral Electromyographic Recordings in Botulinum Toxin Injected Rat Larynges

Katsuhide Inagi, M.D.
Division of Otolaryngology-Head and Neck Surgery, The University of Wisconsin-Madison
Arthur A. Rodriquez, M.D.
Department of Rehabilitation Medicine and Neurology, The University of Wisconsin-Madison
Charles N. Ford. M.D.
Division of Otolaryngology-Head and Neck Surgery, The University of Wisconsin-Madison
Dennis M. Heisey, Ph.D.
Department of Surgery, The University of Wisconsin-Madison

Abstract

Objective assessment of muscle function following botulinum toxin injections in laryngeal muscles is difficult in human subjects. We developed a rat laryngeal model for the study of botulinum toxin injection. A new laryngoscopic technique has made it possible to observe the rat larynx endoscopically and to obtain electromyographic measurements during and after injection of toxin. The electromyographic interference pattern, fibrillation potentials and vocal fold movement were used for analyzing dose and volume effects of injected toxin. We conclude that the lowest dosage to produce the maximal duration of functional laryngeal impairment is 0.01U in a volume of 0.4 µl.

This model will enable us to obtain physiologic and histologic parameters that can be used to assess the selection of optimal treatment regimens with botulinum toxin for the treatment of patients with spasmodic dysphonia.

Introduction

Spasmodic dysphonia is a focal dystonia characterized by action-induced involuntary dysrhythmic contractions of vocal fold musculature. The etiology and curative treatment for this disease are unknown. Moreover since spasmodic dysphonia is not a life threatening disease, laryngeal tissue from surgery or autopsy is rarely obtained. As a result, evaluation of the treatment effect of botulinum toxin on the larynx is currently based largely on the clinical response as determined by subjective and perceptual information; such assessments are subsequently used to select optimal sites of injection, dosage and concentration of botulinum toxin.

Many studies (1-4) have utilized non-laryngeal muscles to study botulinum toxin, but the laryngeal musculature has rarely been studied in this fashion because the muscles are small and difficult to isolate. In the larynx, muscles with different functions, fiber types, and innervations are closely approximated. Botulinum toxin can be easily injected into non-targeted muscle tissue, or can diffuse into adjacent muscles. Furthermore, the small size of laryngeal muscles would suggest a dose-response relationship different from that of the larger limb muscles. For these reasons an animal model (allowing both physiologic and anatomic studies) would be useful to study the effects of botulinum toxin injection in the larynx. Cohen et al. (5) worked out an effective dose based on dose-response curves from dog larynges. George et al. (6) mapped the affected muscle fibers using glycogen depletion techniques in the canine thyroarytenoid muscle.

We developed a rat laryngeal model for assessing the effects of botulinum toxin injections. Using the rat as a model for botulinum toxin injection was facilitated through the implementation of a novel technique to observe the rat larynx endoscopically and also to obtain electromyographic measurements during and after injection of toxin. This technique has greatly improved our ability to investigate the efficacy of botulinum toxin and the development of a rational basis for decisions regarding dosage, volume, loca-

tion, and timing of botulinum toxin injections. We believe that the rat model offers certain advantages over human or canine subjects, especially in the ability to study large numbers of subjects with repeated invasive measurements or to perform histological studies. Rats are relatively inexpensive and easy to handle. Many of the anatomic issues which make the larynx unique are shared between rats and humans. Specific dosage and volume will, of course, differ from human, however we believe that much can be learned from the rat model which is applicable to humans.

Cichon et al. (4) described the compound muscle action potential (CMAP) of the gastrocnemius muscle in rats undergoing botulinum toxin injections. They noted that the CMAP amplitude is an objective and quantifiable measure of the number of functioning muscle fibers. A dose-response relationship was observed between the CMAP amplitude and the botulinum toxin dosage, but there was no apparent correlation with the concentration as a product of dose of the toxin relative to volume of injectate.

Because the CMAP cannot be easily measured in the rat larynx, it was necessary to develop other means of indexing laryngeal muscle fiber function. In this study, the anterior tibialis muscle was also selected because the CMAP could be obtained from supramaximal stimulation of the peronial nerve, previous studies of botulinum toxin injection into the anterior tibialis muscle are available, and the fiber type is similar to the thyroarytenoid. To evaluate the effect of botulinum toxin, we measured vocal fold movement, and selected laryngeal electromyographic parameters (interference pattern and fibrillations) in the thyroarytenoid. In the anterior tibialis we measured foot movement, fibrillations, and the CMAP from supramaximal stimulation of the peroneal nerve. (Table 1).

Table 1.
Vocal Fold, Foot Movement, Interference Pattern and Fibrillation

Grade	VF/Foot. Movement	Interference pattern GAP	Amplitude	Fibrillation Fib	Location
0	none	none	0	0	0
1	<49%	many	much reduced	isolated	2
2	>50%	few	little reduced	isolated	several
3	100%	no	good (equal to intact side)	profuse	most

Vocal fold and foot movement, interference pattern, and fibrillation were graded on a scale of 0 to 3, and correlated with dosage and volume of botulinum toxin.

Materials and Methods

Rats

A total of 29 adult male Sprague-Dawley rats, weighing between 360g and 550g (457.6 ± 51.5 g) were used for this study. Rats were anesthetized with ketamine 48 mg/kg and xylazine 4 mg/kg in an intraperitoneal injection at the each experimental session (three times a week; Monday, Wednesday, and Friday). For the anterior tibialis injection, the whole leg and hindquarter of the injected side were shaved.

Botulinum Toxin Injection Methods

Freshly prepared crystalline complex (30 U/ng) botulinum toxin was diluted with sterile saline to the desired dosages and volumes. Diluted toxin was precisely injected unilaterally into the region of the thyroarytenoid muscle at the mid-portion of vocal fold using a 5-µl syringe with a 25-gauge 2-inch needle using a microlaryngoscopic technique. Similar syringes and needles were used to inject the mid-belly of the anterior tibialis muscle. The side of injection was randomly chosen. Laryngeal and limb movement was graded from 0 to 3 (Table 1). Laryngeal and limb injections were performed using varying dosages as depicted in Table 2.

Laryngoscopy and Video-endoscoy

We developed an original laryngoscope and operating platform (Fig. 1A) for transoral observation and surgical manipulation of the rat larynx. The anesthetized rat was secured on an operating platform in a nearly vertical position with the mouth secured open. The laryngoscope was inserted orally and adjusted to provide the best view of the larynx using an operating microscope (Zeiss, Thornwood, New York, USA). The operating microscope was connected to the video equipment using a c-mount and micro-camera (CIRCON ZIMMER, Santa Barbara, California, USA) which allowed us to view and record the larynx simultaneously (Fig. 1B). Vocal fold movement was graded from 0 to 3 (Table 1) with 0 meaning virtually no movement and 3 meaning normal movement. In our pilot studies, no animals demonstrated marked diminution or loss of cyclic respiratory activity from the anesthesia used in this study. Our measures were not dependent on small changes in timing which might have occurred secondary to the anesthetic.

Electromyographic Methods

Electromyographic sampling of the thyroarytenoid muscle and anterior tibialis muscle was performed with a 50-mm, 26-gauge monopolar needle electrode and a monopolar subcutaneous reference electrode. Filter settings of 20 Hz to 10 kHz and a sweep speed of 10 msec/division were used. Display sensitivity of 50µV/division was used for interference pattern analysis, and 20µV/division was used for detecting fibrillations in the thyroarytenoid and the anterior tibialis muscle. A Nicollet Viking electromyograph (Mound, Minnesota, USA) was used for all electromyographic studies.

The most complete interference pattern from the thyroarytenoid muscle was obtained by manipulation of the

Table 2.
Number of rates per Injection Dosage/Volume Treatment

Thyroarytenoid muscle injection

	Volume		
Dosage	0.1µl	0.2µl	0.4µl
0.01U	2	4	2
0.001U	3	3	2
0.0001U	2	3	2
Saline	0	0	2

Anterior tibialis muscle injection

	volume					
Dosage	0.1µl	0.2µl	0.4µl	1µl	2µl	4µl
0.4U	0	0	0	0	0	1
0.1U	0	0	0	1	1	1
0.01U	0	1	0	2	4	3
0.001U	1	1	1	2	2	2
0.0001U	0	1	0	1	1	1
Saline	0	0	1	0	0	1

Injection dosage, volume, and numbers of rats are shown in this table.

electrode within the muscle to maximize high-frequency content and amplitude. Interference patterns were then semi-quantitatively graded from 0 to 3 using the uninjected side for comparison (Table 1). A larger interference grade indicates greater activity of muscle fibers.

Fibrillations were induced by sampling or probing within the muscle. In the thyroarytenoid muscle, they were observed between rhythmic bursts of respiratory activity. Waveforms in the rat thyroarytenoid were typical of these potentials with fibrillations having a tri-phasic waveform with a positive initial phase, and typical positive sharp waves both observed with the muscle at rest.Amplitudes and durations are slightly smaller in the laryngeal muscles than in the anterior tibialis with durations of approximately one msec for fibrillations and amplitudes of 20-100 µV for both positive sharp waves and fibrillation potentials. Fibrillation potentials in thyroarytenoid and tibialis anterior muscles were graded from 0 to 3 (Table 1).

The CMAP amplitude (mV) in the anterior tibialis muscle was measured with the rat limb immobilized on a metal bar. Four mm disk electrodes were placed with the G1 (active) electrode located over the mid-belly of the tibialis anterior muscle, and the G2 (reference) electrode over the tibialis anterior tendon distally. Electrode sites were marked with India ink to assure the same electrode placement over each recording section. A monopolar electrode was used to stimulate the peroneal nerve percutaneously, and adjusted for a 20% supramaximal stimulus. Stimulus duration was 0.1 msec.

Ten repetitive stimuli were then applied at either 2 Hz or 50 Hz, to evaluate for a facilitating or decremental response. The amplitude difference was determined between the first and fourth stimuli.

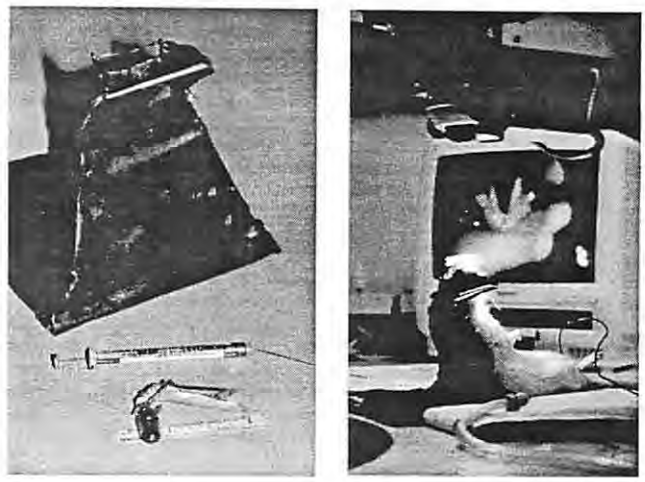


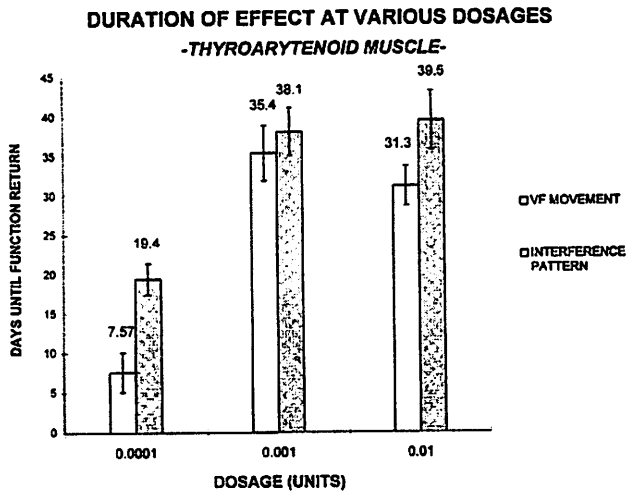
Figure 1. A (left): An original laryngoscope and operating platform for transoral observation and surgical manipulation of the rat larynx. This laryngoscope, made from aluminum, was designed to distend the oral cavity while anteriorly displacing the vallecula to visualize the larynx. The operating platform consists of a malleable lead board that can be modified as necessary for optimal visualization of the larynx. B (right): The rat is secured on the operating platform in the nearly vertical position with the mouth open. An overhead camera shows the rat larynx.

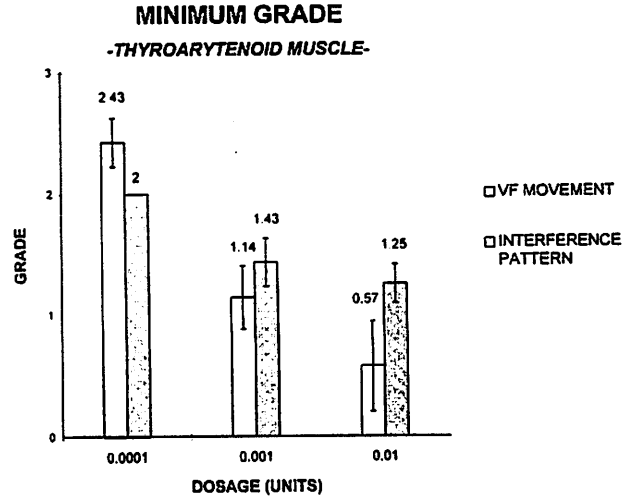
Video-laryngoscopic observation and EMG measurements were performed three times a week (Monday, Wednesday, and Friday) for 64 days. The judgments of these measurements were performed by judges blinded to the injected dosage and volume. Vocal fold movement and EMG interference pattern grading measurements were obtained in comparison to the contralateral side (Table 1).

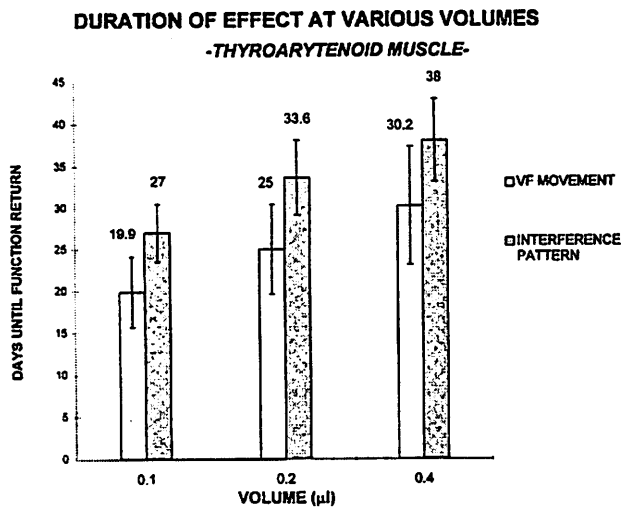
Data Analysis

Of the 29 rats, one that was injected with 0.001U toxin in a volume of 0.2 µl into the left thyroarytenoid muscle died on day 15 after injection from an anesthesia-related cause. All 28 other rats were observed for 64 days following injection.

The duration of the botulinum toxin effect was defined as the time required to recover 100% of the preinjection status. The duration of fibrillation was defined as the time interval between the day of injection to the last day they were observed. The minimum grade of each parameter (vocal fold movement, interference pattern, foot movement, and CMAP in the anterior tibialis muscle) was compared to injection dosages and volumes. Two-way analysis of variance was used to simultaneously test whether dosage, volume, or their interaction influenced the dependent variables. If the anova test for treatment effect was significant, the treatment effect was further examined by comparing all pairs of means with Fisher's Protected Least Significant Difference Procedure(7). All testing was done at the nominal 5%, the reported p-values control the comparisonwise Type I error rate within each dependent variable.







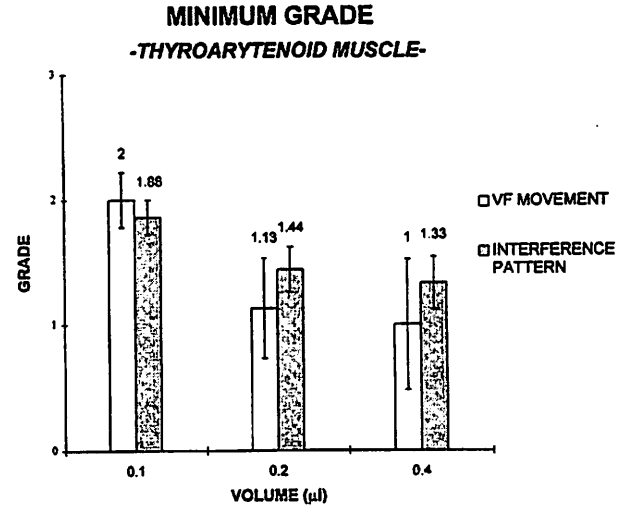


Figure 2. Time to functional return of vocal fold movement and interference pattern for various dosages (A; top) or volumes (B; bottom) of injectate. Rats injected with more than 0.001U had a significantly ($p<0.01$) longer effect duration than 0.0001U-injected rats. The 0.4 µl-injected rats had significantly ($p<0.05$) longer effect duration than 0.1 ml-injected rats.

Figure 3. Minimal grades of vocal fold movement or interference pattern for various dosages (A; top) or volumes (B; bottom) of injectate. Significant ($p<0.01$) differences in the minimum grade score of vocal fold movement and interference pattern were observed at all dosages. Rats injected with 0.0001U had significantly higher minimum grades than the other rats. Significant differences in the minimum grade score of vocal fold movement ($p<0.01$) and interference pattern ($p<0.05$) were observed as a function of volume. Rats injected with 0.1 ml had a significantly higher minimum grade than the other rats.

Results

Vocal Fold Movement

No changes in vocal fold movement were observed in control animals. The length of time needed to recover pre-injection vocal fold movement ranged from 19.4 days in 0.0001U-injected rats to 39.5 days in 0.01U-injected rats (Fig. 2A). Rats injected with the higher dosages showed significantly ($p<0.01$) longer effect durations than 0.0001U-injected rats. Rats injected with greater volumes of injectate had significantly ($p<0.05$) longer effect durations than 0.1 µl-injected rats (Fig. 2B).

Significant ($p<0.01$) differences in the minimum grade score of vocal fold movement were observed with the variation in both dosage (Fig. 3A) and volume (Fig. 3B) of injectate. The 0.0001U or 0.1 µl-injected rats had significantly higher minimum grades than the other rats.

Interference Pattern of Thyroarytenoid Muscle

No changes in the electromyography interference pattern over time were observed in control animals. The interference pattern had a positive relationship between the effect duration and the doses or volumes of the toxin. The amount of time needed to recover pre-injection status of the interference pattern was significantly ($p<0.01$) greater for rats injected with 0.001U or more than for 0.0001U-injected rats (Fig. 2A). The effect duration of botulinum toxin using

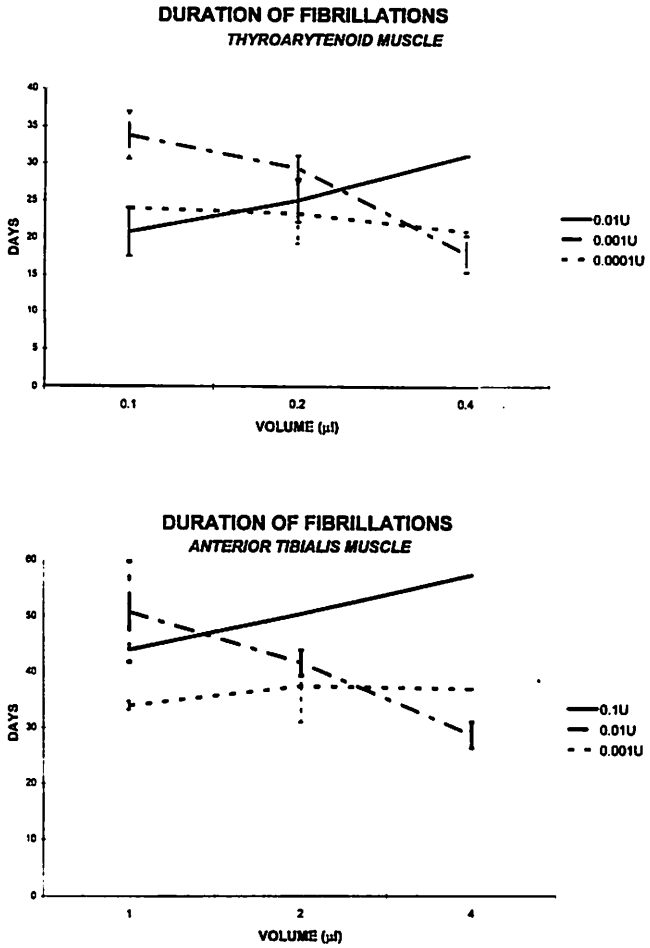


Figure 4. Duration of fibrillation of the thyroarytenoid muscle (A; top) and the anterior tibialis muscle (B; bottom) were compared with the variation in both volume and dosages. Significant ($p<0.05$) differences were obtained in fibrillation duration between thyroarytenoid and anterior tibialis muscles, but the same general pattern was observed in both.

the interference pattern was only 7.4 days in 0.0001U-injected rats. Significant differences in recovery time were also observed at different volumes (Fig. 2B). Significantly longer recovery times were recorded for the 0.4 µl-injected rats (38 days $p<0.01$) and 0.2 µl-injected rats (33.6 days $p<0.05$) than for 0.1 µl-injected rats (27 days) (Fig. 2B).

Significant differences in the minimum (most affected) scores of interference patterns were observed with variation in both dosage (Fig. 3A) and volume (Fig. 3B) of injectate. Rats injected with a dosage of 0.0001U ($p<0.01$) or a volume of 0.1 µl ($p<0.05$) had significantly higher minimum grades than the other rats.

Fibrillation Potential in Thyroarytenoid and Anterior Tibialis Muscle

A different pattern of response was observed in rats injected with the largest dosage (0.01U) in comparison with the other dosages (0.001 and 0.0001U). A positive

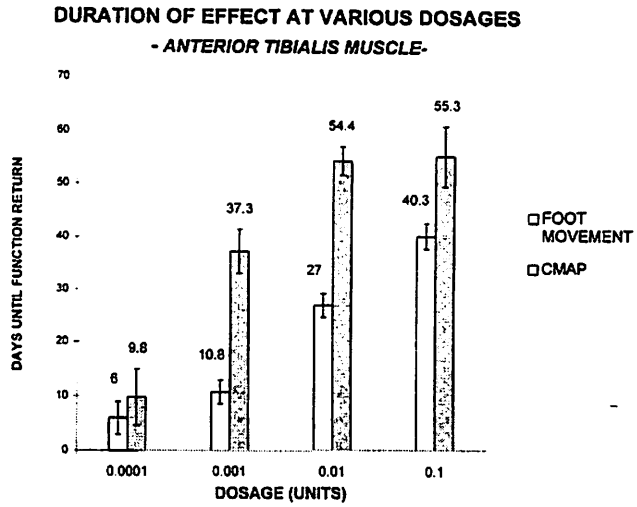


Figure 5. Duration of effect measured by recovery of foot movement and CMAP amplitude at various dosages of injectate. A significant dose-duration relationship was observed in foot movement ($p<0.05$) and CMAP($p<0.01$).

relationship was observed between volume and thyroarytenoid muscle fibrillation duration in the rats injected with 0.01U. However a negative relationship between volume and fibrillation duration was observed in those injected with 0.001U (Fig. 4A). There was no difference in thyroarytenoid muscle fibrillation duration among rats injected with various volumes using a dosage of 0.0001U. This general pattern was also observed in fibrillation of the anterior tibialis muscle (Fig. 4B); the rats injected with 0.1U dosage had a positive relationship between duration of fibrillation and volumes, but 0.001U injected rats had a negative relationship between volume and fibrillation duration. For the thyroarytenoid muscle, rats injected with 0.001U in 0.1 µl had significantly ($p<0.05$) prolonged durations of fibrillations in comparison to other rats injected with 0.001U but with larger volumes. In contrast rats injected with a larger dosage (0.01U) had significantly ($p<0.05$) longer durations of fibrillations with increasing volumes.

Foot Movement

Comparing 0.01U and 0.001U injections, a positive dose-effect relationship was observed (Fig. 5). Rats injected with 0.001U using a volume less than 1 µl recovered their foot movement significantly earlier. Comparing doses in rats injected with a volume greater than 1 µl, a significant ($p<0.05$) dose-duration relationship was observed, but there were no significant differences with the variation of volume (Table 3; following page). No significant relationship between minimum grade and injection volume was observed in foot movement (data not shown).

Table 3.
Effect at Various Volumes
- Anterior Tibialis Muscle -

Measurements \ Volume		1µl	2µl	4µl
Duration of Effect (days)	Foot Movement	19.5±6.9 (4)	26.0±3.1 (5)	19.2±4.9 (5)
	CMAP	34.5±9.2 (6)	45.5±5.7 (7)	37.6±6.8 (7)
Minimum Relative CMAP		0.54±0.13 (6)	0.49±0.08 (7)	0.57±0.12 (7)

Mean ± Standard Error (n)

Time to functional return of foot movement and CMAP amplitude, and minimum relative CMAP (the recorded minimum CMAP divided by pre-injection CMAP amplitude) was compared with the variation in the volumes of injectate.

CMAP Amplitude in Anterior Tibialis Muscle

Over the 26 measurements in control animals the relative CMAP (i.e., the recorded CMAP divided by pre-injection CMAP amplitude) varied from 0.92 to 1.22.

A positive relationship was observed between the dose and time to recovery of the CMAP amplitude (Fig. 5). Rats injected with 0.1U or 0.01U took a significantly ($p<0.01$) longer time to recover than other rats. There was no correlation between volume and duration (Table 3).

A strong relationship ($p<0.01$) was observed between the dose and minimum relative CMAP over the time series (Fig. 6). However no correlation of volume to minimum relative CMAP was observed (Table 3).

Repetitive Stimulation of the Anterior Tibialis Muscle

No consistent increment or decrement at 2 Hz stimuli was seen in the rats injected with small dosages (0.0001 to 0.01U) or with small volumes (0.1 to 0.4 µl). Using 50 Hz stimulation, seven rats (7/27; 26%) had incremental responses. The largest increment was 138%.

Inter-rater Agreement

The rater went back and rescored a representative sample of 13 rats. On the botulinum toxin injected side, exact agreement was 61% (8/13), and agreement within one unit was 100%. Kappa, a measure of exact agreement corrected for chance (8), was 0.43, which was highly significant ($p=0.0049$). On the contralateral side, exact agreement was 100%, and kappa = 1 ($p=0.0001$). On the injected side, the rank correlation coefficient was $r = 0.78$ ($p=0.005$), and on the contralateral side, $r = 1$ ($p = 0.001$). These results indicate a fairly high degree of inter-rater agreement and association.

Discussion

Efficacy of botulinum toxin for treatment of spasmodic dysphonia has been assessed largely on the basis of

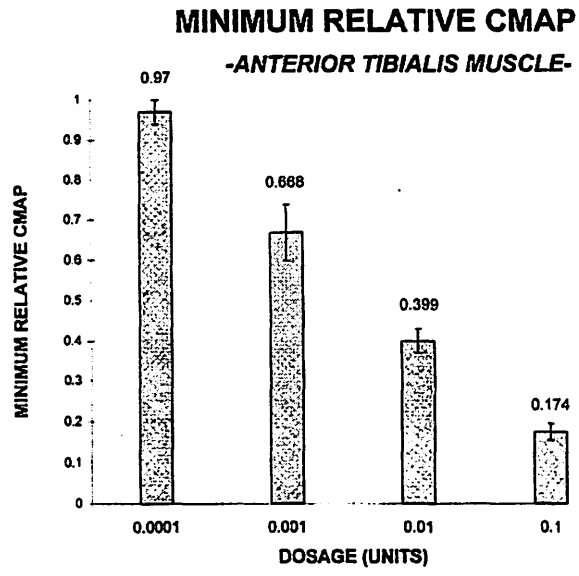


Figure 6. Minimum relative CMAP (the recorded minimum CMAP divided by pre-injection CMAP amplitude) was compared with the variation in the dosages of injectate. A strong relationship ($p<0.01$) was observed between the dose and minimum relative CMAP.

patient subjective reports. Decisions about dosage, placement of injectate, and intervals between injections have likewise been decided on an empirical basis influenced by the patient's demand and the preferences of the clinician. The scientific basis for making these decisions has been inadequate. Our clinical study (9) suggested that injection of both the thyroarytenoid and the lateral cricoarytenoid muscles simultaneously gave the best voice results. The exact reason for these results is unclear. It is also unclear where the injectate spreads following injection.

Despite technical difficulties in studying laryngeal musculature(10) we believe laryngeal muscles should be used for the study of botulinum toxin effects in spasmodic dysphonia. For this reason, studies of botulinum toxin injection in the larynx have used animal models in the hope that it would be possible to extend the findings from animal models to humans. The canine larynx is anatomically different from the human larynx, yet dogs have been used because of convenient laryngeal size, accessibility, and availability. Rats (10) (11) have a laryngeal anatomy somewhat different from dogs or humans (12) (13) but they are attractive for studies requiring great numbers of subjects, for they are relatively inexpensive and easy to handle. The problems with laryngeal studies in rats have been largely technical. Our development of a specialized laryngoscope and operating platform has solved many of these problems and appears to render studies using the rat laryngeal model possible.

This study was designed to investigate whether measurements of vocal fold movement and selected electromyographic parameters (interference pattern and fibrillation potentials) are useful indicators of functioning laryngeal muscle fibers. Our purpose was to study the relationship between laryngeal function and botulinum toxin dosage, injectate volume, and concentration in the rat larynx.

The response to botulinum toxin injection in the rat larynx was documented within 6 hours and the maximum effect was observed 1 to 7 days after injection. In this period, fibrillation potentials developed. Some recovery of vocal fold movement was observed 11 to 15 days after injection. The appearance of fibrillation potentials soon after injection and the rapid return of vocal fold function suggest that return of function was quicker than occurs in the human. Therefore, the rat laryngeal model appears well suited to study the effects of repeated botulinum toxin injection.

This study demonstrated a strong dose-response relationship between botulinum toxin dosage and duration of CMAP amplitude diminution, as well as foot movement impairment in the anterior tibialis muscle. The CMAP amplitude is the summation of simultaneously stimulated action potentials of all of the muscle fibers within the muscle and is therefore an index of the number of functioning muscle fibers (14). However, it is not technically possible to obtain the CMAP in the rat larynx, so other parameters are needed to assess the effect of botulinum toxin in the larynx. The duration of vocal fold movement impairment and interference pattern diminution showed similar dose-volume relationship patterns. Significant differences were observed for both movement impairment duration and interference pattern diminution duration between rats injected with more than 0.001U and those injected with 0.0001U. A statistically significant relationship was observed between both dosage and volume with the degree of impairment of vocal fold movement and electromyographic measures. These results suggest that the vocal fold movement and the interference pattern grade are reliable parameters for measuring the effects of botulinum toxin.

Variation of injectate volume from less than 1 µl to 4 µl did not significantly alter the duration of the botulinum toxin effect in the anterior tibialis muscle using any parameter. Apparently, the relatively small differences in injectate volume had little effect in such a large muscle. In one study, the estimated number of muscle fibers paralyzed by botulinum toxin in the rat anterior tibialis muscle was increased 10% by increasing injectate volume from 1 to 10 µl (2). However, in the larynx, significant differences were observed between injections of 0.1 µl and those of 0.2 µl and more. One can postulate that diffusion to adjacent muscles such as the lateral cricoarytenoid may have occurred, resulting in a greater degree of the vocal fold weakness. We are currently studying the histological relationship between botulinum toxin dosage and extent of denervation in the TA and adjacent muscles.

Examination of fibrillation duration results suggests that the concentration of botulinum toxin affected fibrillation duration. Ten times greater volumes and dosages were injected into the anterior tibialis muscle, because it is much larger than the thyroarytenoid. Botulinum toxin concentrations of 10U/ml or more significantly prolonged the fibrillation duration. We suggest the possibility that these concentrations of botulinum toxin represent levels sufficient to induce complete paralysis in the single muscle fiber.

Because no consistent incremental or decremental responses of the CMAP amplitude to 2 Hz or 50 Hz repetitive stimulation were observed when we calculated the difference between the first and fourth stimuli, this parameter was not helpful in determining the duration of action of botulinum toxin. Variable responses to repetitive stimulation have also been reported in human botulism (15).

The possibility of laryngeal muscle injury from the monopolar needle EMG electrode might be considered. However, no control animals demonstrated changes in vocal fold movements, interference pattern scores, or the presence of fibrillations. Thus, these parameters were unaffected by repeated electromyographic sampling. Morphological changings following botulinum toxin injection will be described in another publication.

Conclusion

These data allow us to select suitable botulinum toxin dosages and volumes for the rat larynx. We conclude the lowest dosage to produce the maximal duration of functional laryngeal impairment is 0.01U with a volume of 0.4 µl. Using this animal model we hope to refine the selection of appropriate treatment variables, and to develop ways of enhancing botulinum toxin efficacy for the treatment of patients with spasmodic dysphonia.

Acknowledgments

This work was supported by the National Center for Voice and Speech through Grant P6000976 from the National Institute on Deafness and Other Communication Disorders.

The authors would like to thank Ms. Celeste Kirk for editorial assistance.

References

1) Brown MC, Holland RL, Hopkins WG, Keynes RJ. An assessment of the signal for terminal sprouting within and between muscles. Brain Res 1980;210:145-151.

2) Shaari CM, Sanders I. Assessment of the biological activity of botulinum toxin. Therapy with botulinum toxin., 1st ed. New York. Marcel Dekker, 1994:159-170.

3) Duchen LW. Changes in motor innnervation and cholinesterase localization induced by botulinum toxin in skeletal muscle of the mouse: Difference between fast and slow muscles. J Neurol Neurosurg Psychiatry 1970;33:40-54.

4) Cichon JV, McCaffrey TV, Litchy WJ, Knops JL. The effect of botulinum toxin type A injection on compound muscle potential in an in vivo rat model. Laryngoscope 1995;105: 144-148.

5) Cohen SR, Thompson JW. Use of botulinum toxin to lateralize true vocal cords: A biochemical method to relieve bilateral abductor vocal cord paralysis. Ann Otol Rhinol Laryngol 1987;96: 534-541.

6) George EF, Zimbler M, Biller HF, Wu B, Sanders I. Quantitative mapping of the effect of botulinum toxin injections in the thyroarytenoid muscle. Ann Otol Rhinol Laryngol 1992;101:888-892.

7) SAS Institute Inc. SAS/STAT User's guide Version 6. 4th ed. Vol 2. Cary NC, 1989.

8) Cohen J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement 1960; 20:37-46.

9) Inagi K, Ford CN, Bless DM, Heisey D. An analysis of factors affecting botulinum toxin results in spasmodic dysphonia. J. Voice 1996;10(3):306-313.

10) Kobler JB, Datta S, Goyal RK, Benecchi EJ. Innervation of the larynx, pharynx, and upper esophageal sphincter of the rat. J. Comp Neurol 1994; 349:129-147.

11) Inagi K, Schultz E, Ford CN. An anatomical study of the rat larynx. Otolaryngol Head Neck Surg (in press)

12) Rosen M, Malmgren LT, Gacek RR. Three-dimensional computer reconstruction of the distribution of neuromuscular junctions in the thyroarytenoid muscle. Ann Otol Rhinol Laryngol 1983;92:424-429.

13) Claassen H, Werner JA. Fiber differentiation of the human laryngeal muscles using the inhibition reactivation myofibrillar ATPase technique. Anat Embryol 1992 ;186:341-346.

14) Valls-Sole J, Tolosa, ES, Ribera G. Neurophysiological observation on the effects of botulinum toxin treatment in patients with dystonic blepharospasm. J Neurol Neurosurg Psychiatry 1991;54:310-313.

15) Oh SJ. Botulism: Electrophysiological studies. Ann Neurol 1977;1(5): 481-485.

NCVS Status and Progress Report - 11
May 1997, 83-89

Physiologic Assessment of Botulinum Toxin Effects in the Rat Larynx

Katsuhide Inagi, M.D.
Charles N. Ford, M.D.
Division of Otolaryngology-Head and Neck Surgery, The University of Wisconsin Medical School
Edward Schultz, Ph.D.
Department of Anatomy, The University of Wisconsin Medical School
Arthur A. Rodriquez, M.D.
Department of Rehabilitation Medicine and Neurology, The University of Wisconsin Medical School
Diane M Bless, Ph.D.
Division of Otolaryngology-Head and Neck Surgery, The University of Wisconsin Medical School
Thomas Pasic, M.D.
Division of Otolaryngology-Head and Neck Surgery, The University of Wisconsin Medical School
Dennis M. Heisey, Ph.D.
Department of Surgery, The University of Wisconsin Medical School
Nadine Connor, Ph.D.
Division of Otolaryngology-Head and Neck Surgery, The University of Wisconsin Medical

Abstract

Botulinum toxin (BT) is a currently used treatment for spasmodic dysphonia (SD) and other related focal dystonias. Because objective assessment of physiological and histological effects of BT in humans is difficult, this study provides a basis for using a rat laryngeal model. A new laryngoscopic technique was developed which made it possible to observe and manipulate the rat larynx endoscopically. In this experiment, dosages and volumes of BT were varied and three physiological parameters were measured. Results indicated significant reductions in vocal fold motion and spontaneous laryngeal muscle activity as a function of increased BT dosage. In addition, the optical density of PAS stained laryngeal muscle after electrical stimulation was increased following BT injection, which is an indirect measure of denervation. Significant volume effects in optical density were observed in the lateral thyroarytenoid and lateral cricoarytenoid muscles on the contralateral side. We concluded that the rat laryngeal model is suitable for assessing BT effects. In addition, the three physiological variables provided useful and reliable measures of laryngeal function. It is our intention to use the rat laryngeal model to further examine the physiological and histological effects of BT with the goal of developing new methods for the treatment of patients with SD.

Introduction

Spasmodic dysphonia (SD) is a focal dystonia of unknown etiology which affects the intrinsic laryngeal muscles and interferes with phonation. While there is no definitive cure, periodic injection of botulinum toxin (BT) has proven effective in temporarily relieving symptoms of the voice disorder. However, optimal dosages[1-4], injectate volumes, timing and sites of repeated injections[5-6], the reinnervation process, and diffusion patterns into adjacent muscles are not well understood. Knowledge of these processes could provide clinicians with a rational basis for making treatment decisions.

Because SD is not life-threatening, human laryngeal tissue from surgery or autopsy is rarely obtained. Therefore, investigations concerning the actions of the toxin in SD treatment require an animal model. Although dog larynges have been used extensively for physiological studies[7-9], a rat preparation may be more useful since these

animals are inexpensive and easy to handle. Use of the rat larynx as an animal model is also supported by numerous basic investigations of rat skeletal muscle, including the larynx[10-13], and the development of new techniques to accommodate the small size of the preparation[13].

The aim of this study was to determine whether the rat larynx can provide a useful animal model for assessing the physiologic effects of BT. Further, it was of interest to establish measurable parameters which could serve as indicators of laryngeal functioning in rats. To this end, the dosage and volume of BT injection into the rat larynx were varied to address three issues: 1) the dosage and volume of BT appropriate for injection into the rat larynx to simulate therapeutic injections in humans, 2) the measurable physiologic changes which correlate with variable BT dosage and volume, and 3) the physiologic changes which may occur in muscles adjacent or contralateral to the site of injection.

Materials and Methods

Seventy-two adult male Spague-Dawley rats weighing between 477 and 541 g (mean of 505g ± 20g) were used in this study (Table 1). Rats were anesthetized with an intraperitoneal injection of ketamine (48 mg/kg) and xylazine (4mg/kg) prior to periods of vocal fold observation or BT injection. In our pilot studies, no animals demonstrated marked diminution or loss of cyclic respiratory activity from the anesthesia used in this study. Further, our measures were not dependent upon small changes in timing which may result secondary to anesthetic

BT Injection Methods

Freshly prepared crystalline complex (30 U/ng) BT was diluted with sterile saline into dosages of 0.0001 to 1 U and volumes of 0.1 to 2.5µl. Using a microlaryngoscopic technique, diluted BT was precisely injected unilaterally at varying dosages and volumes (Table 1) into the thyroarytenoid muscle at the mid-portion of the vocal fold using a 5µl syringe with a 25-gauge, 2-inch needle. Saline (0.2µl) was injected into three control rats using the same technique. The side of injection was randomly chosen.

Table 1.
Number of Subjects

Dosages \ Volumes	0.1µl	0.2µl	0.4µl
1 U	2	0	0
0.1 U	2	2	2
0.01 U	6	6	7
0.001 U	7	7	6
0.0001 U	7	7	7
Saline	0	3	0

Larynoscopy and Video-endoscopy

A customized laryngoscope and operating platform[13] for transoral observation and surgical manipulation of the rat larynx were used. The anesthetized rat was secured on an operating platform in a nearly vertical position with the mouth fixed open. The laryngoscope was inserted orally and adjusted to provide the best view of the larynx using an operating microscope (Zeiss, Thornwood, New York, USA). The operating microscope was connected to video equipment and a micro-camera (CIRCON ZIMMER, Santa Barbara, CA, USA) which allowed us to view and record the larynx simultaneously[13].

All rat vocal fold movements were recorded at 21X magnification. Recordings were made prior to injection and at 3 days after injection. As shown in Figure 1, objective measurements were obtained from digitized glottal images of videostroboscopic recordings. Two glottal frames were selected for each rat which represented full adduction and full abduction during respiration. Four reference points were determined for each glottal frame:

1) posterior commissure, which was designated point P,
2) absolutely vertical line at point P, which was designated line H,
3) the edge of the right vocal fold at 2 mm from point P, which was designated point R,
4) the edge of the left vocal fold at 2 mm from point P, which was designated point L.

The following measurements were then derived from the reference points :

1) lateral angles of the right and left vocal folds between line H and a line connecting point P with either point R or point L were designated as angles H-RP and H-LP, respectively, and,
2) movement angles of the right and left vocal folds were the difference between angles H-RP or H-LP in the adduction frame and the corresponding angles in the abduction frame, respectively.

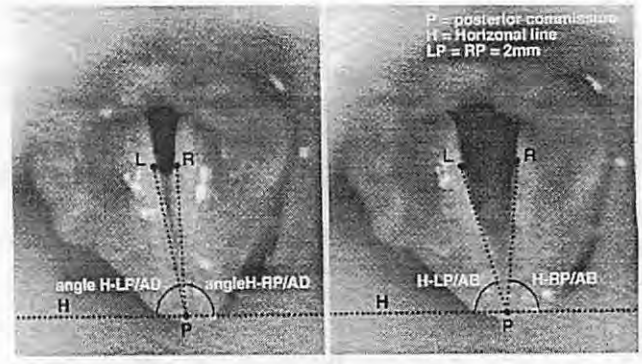


Figure 1. Digitized glottal images of videostrobscopic recordings. The angle H-RP/AD was 89.0°, the angle H-LP/AD was 73.0°, the angle H-RP/AB was 59.5°, and the angle H-LP/AB was 56.0°. Movement angles of the right vocal fold was 19.5°, and left vocal fold was 17.0°.

Table 2.
Rating Criteria of EMG Interference Patterns

Grade	Interference pattern	
	Gap	Amplitude
0	no	0
1	many	much reduced
2	few	little reduced
3	no	good (equal to intact side)

EMG interference pattern was graded on a scale of 0 to 3

All measurements were performed using Image 1.59 software (USDHSS, National Institutes of Health, Bethesda, MD, USA) and a Macintosh II fx computer (Apple, Cupertino, CA, USA).

Electromyographic Methods

Electromyographic (EMG) sampling of the thyroarytenoid muscle (TA) was performed both pre-injection and again 3 days post-injection with a 50 mm, 26 gauge monopolar needle electrode and a monopolar subcutaneous reference electrode (Nicollet Viking electromyograph, Mound, MN USA). Filter settings of 20 Hz to 10KHz and a sweep speed of 10 msec/div were used. The most complete EMG interference pattern from the TA muscle was obtained by manipulation of the electrode within the muscle to maximize high-frequency content and amplitude.

Pre and post-injection EMG signals were displayed on its monitor with a sensitivity of 50 μV/div. As shown in Table 2, EMG interference patterns were rated from 0 to 3, with larger ratings indicating greater muscle activity relative to the uninjected side. Ratings were performed by a judge (Rodriquez) blind to the injected dosages and volumes.

Periodic Acid Schiff (PAS) Staining Methods and Optical Density Measurements

Intensity of PAS staining, as measured in optical density (pixels), reflects failure of glycogen depletion and is an indirect measure of denervation. As such, a PAS staining and optical density measurement protocol was developed to assess glycogen depletion within 36 hours once these larynges were sectioned. Three days after injection, both recurrent laryngeal nerves were identified and stimulated with bipolar stimuli for 20 minutes at 10 mA and 10 Hz. Following euthanasia, larynges were excised and frozen immediately with liquid nitrogen and sectioned at a thickness of 20 μm. All reagents were freshly made just before staining to prevent denaturing.

Coronal sections at the level of the midbelly of the following intrinsic laryngeal muscles were selected for optical density measurements:

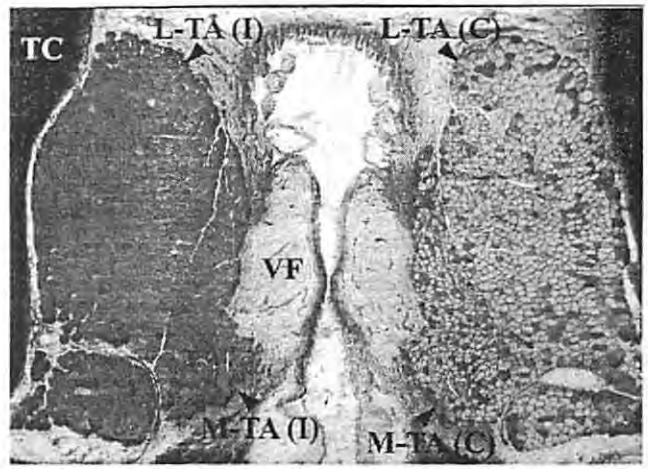


Figure 2. The mean actual optical density of the lateral thyroarytenoid (L-TA) muscle was 57.3 pixels in injected (I) and 43.3 pixels in contralateral (C) side, and the medial thyroarytenoid (M-TA) muscle was 59.1 pixels in injected (I) and 52.5 pixels in contralateral (C) side after glycogen depletion and PAS staining in 0.001 units with a volume of 0.1 μl injected larynx.

1) the whole area of the medial thyroarytenoid (MTA),
2) the whole area of the lateral thyroarytenoid (LTA) ,
3) the whole area of the ala cricoarytenoid (ACA)[10],
4) the whole area of the lateral cricoarytenoid (LCA),
5) the whole area of the superior cricoarytenoid (SCA)[10], and
6) the whole area of the posterior cricoarytenoid (PCA).

As shown in Figure 2, optical density gray scales, which ranged from 0 to 256 pixels, were digitized using a microscope (Leits Wetzlar, Germany) and CCD camera (CCD 72, Dage MTI, Michigan City, IN) connected to a computer (Macintosh II fx; Image 1.59 software). All measurements were performed three times using the same magnification (i.e., 2.5X objective) and the same light source power. To minimize error due to the light source, calibrations were performed by absolute optical density using Photographic Step Tablet No. 2 (Kodak, Rochester, NY, USA) before and after measurements.

Data Analyses

Analysis of variance (ANOVA) of the ranked transformed data was used for analyzing the associations among treatment factors and physiological parameters of vocal fold movement, EMG interference pattern rating and optical density. The ANOVA was used to simultaneously test whether dosage, volume or their interaction influenced the dependent variables. Fisher's Protected Least Significant Difference[14] was used for post-hoc tests of significant ANOVAs. Spearman ranked correlation was used for analyzing interactions among the three physiological parameters. All statistical tests were performed at the nominal 5% level; the reported p-values controlled the comparisonwise Type I error rate within each dependent variable.

Animal Care

This study was performed in accordance with the PHS Policy on Human Care and Use of Laboratory Animals, the NIH Guide for Care and Use of Laboratory Animals, and the Animal Welfare Act (7 U.S.C. et seq). The animal use protocol was approved by the Institutional Animal Care and Use Committee (IACUC) of the University of Wisconsin.

Results

Rats injected with less 0.1 U of BT survived the procedure, with the exception of one rat injected with 0.01 U of BT in a volume of 0.4 µl. In contrast, all rats injected with over 0.1 U of BT died within 48 hours due to choking and aspiration.

In the control rats, no changes in vocal fold movements or EMG interference pattern ratings were found pre-injection versus 3 days post-injection. All control rat laryngeal muscles were diluted in glycogen bilaterally, with the exception of the cricothyroid muscle.

Effect of BT Dosage and Volume on Vocal Fold Movement

As shown in Figure 3, vocal fold movement angle on the injected side was progressively reduced with increasing BT dosage at statistically significant levels ($p<0.01$). In contrast, alterations in BT volume on the injected side had no effect. On the contralateral side, also shown in Figure 3, volume effects were also absent and a significant BT dosage effect was found only between 0.001 and 0.01 U ($p<0.01$).

Effect of BT Dosage and Volume on TA Muscle Activity

On the injected side, dosage and volume effects on spontaneous TA muscle activity were similar to those reported for vocal fold movement. As observed in Figure 4, a significant, progressive reduction in muscle activity was found as BT dosage increased ($p<0.01$). However, no volume effect was observed on the injected side. On the contralateral side, only rats injected with dosage of 0.01 U in a volume of 0.4 µl resulted in a significant diminution of the EMG interference pattern rating ($p<0.01$).

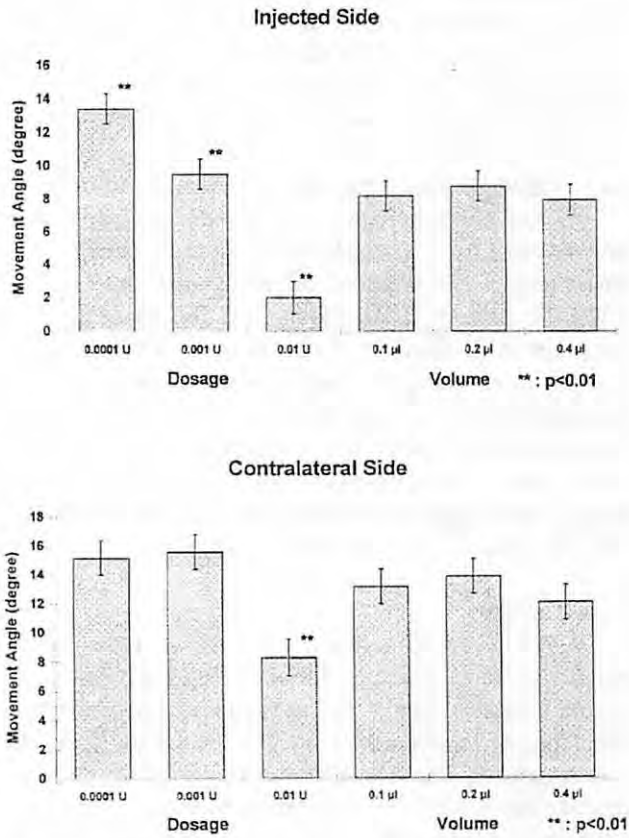


Figure 3. The comparison between digitized vocal fold movement and the BT dosages or volumes. Significant differences were obtained in dosages, but not volumes. Obtained significant difference was between 0.001U, 0.001U, and 0.01U in the injected side and betweeen 0.01U and 0.001U or 0.0001U in the contralateral side.

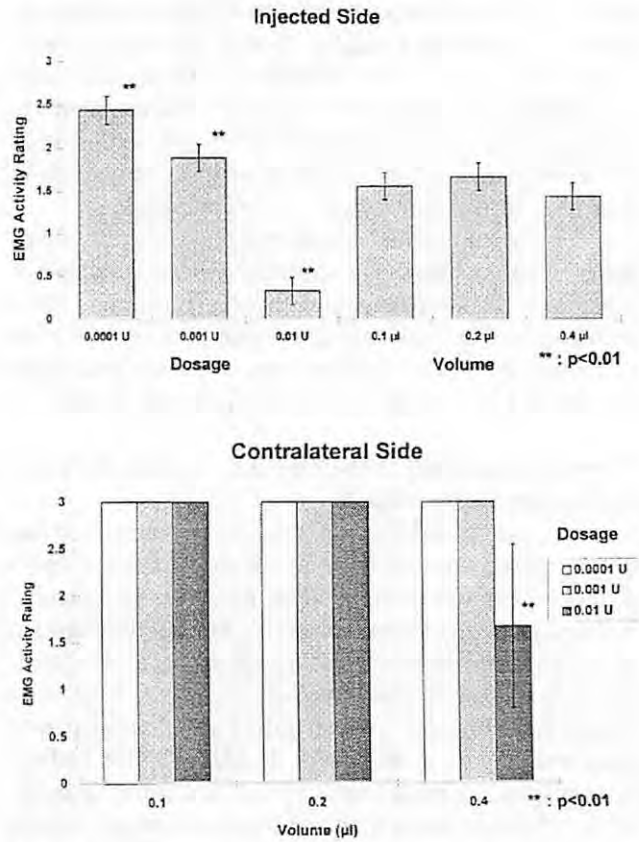


Figure 4. The upper chart presents the results of dose- and volume effect on the EMG interference pattern rating on injected side. A significant dose effect was observed. Regarding the contralateral side (lower chart), only rats injected with a dosage of 0.01 units in a volume of 0.4 µl had a significant decrease in EMG interference pattern ratings compared to rats injected with any other combination of BT dosages and volumes.

Effect of BT Dosage and Volume on Optical Density

As shown in Figure 5a, all of the laryngeal muscles on the injected side had progressively greater optical densities (pixels) with increases in BT dosage, which indicates failure of glycogen depletion (MTA, LTA, ACA, SCA and PCA at $p<0.01$; LCA at $p<0.05$). As mentioned, failure of glycogen depletion following nerve stimulation is an indirect measure of denervation. With regard to injectate volumes (Fig. 5a), all muscles injected with more than 0.4 µl had greater optical densities than muscles injected with smaller volumes. However, this volume relationship was statistically significant only for the ACA muscle ($p<0.05$). On the contralateral side, as shown in Figure 5b, significant increases in optical density were again observed with increases in BT dosage for all muscles (MTA, LTA, SCA, and PCA at $p<0.01$; ACA and LCA at $p<0.05$). Post-hoc tests revealed that significant differences were generally observed between dosages of 0.0001 and 0.01 U. When injectate volumes were compared, higher optical density measures in contralateral LTA, ACA, and LCA were indicative of significant differences in failure of glycogen depletion for volumes of 0.4µl versus 0.1 and 0.2 µl.

Correlation Among Physiologic Parameters

Measured decreases in vocal fold movement were significantly related to increases in optical density of the LTA, PCA and SCA muscles ($p<0.01$; Table 3; following page). Similarly, decreased EMG activity ratings were significantly correlated with increased optical densities in most laryngeal muscles (LTA, SCA, and PCA at $p<0.01$; ACA and MTA at $p<0.05$). A significant positive correlation was observed between digitized vocal fold movement and EMG interference pattern rating.

Discussion

Although BT is currently being used to manage the symptoms of spasmodic dysphonia and other focal dystonias, the physiological and histological effects of various dosage

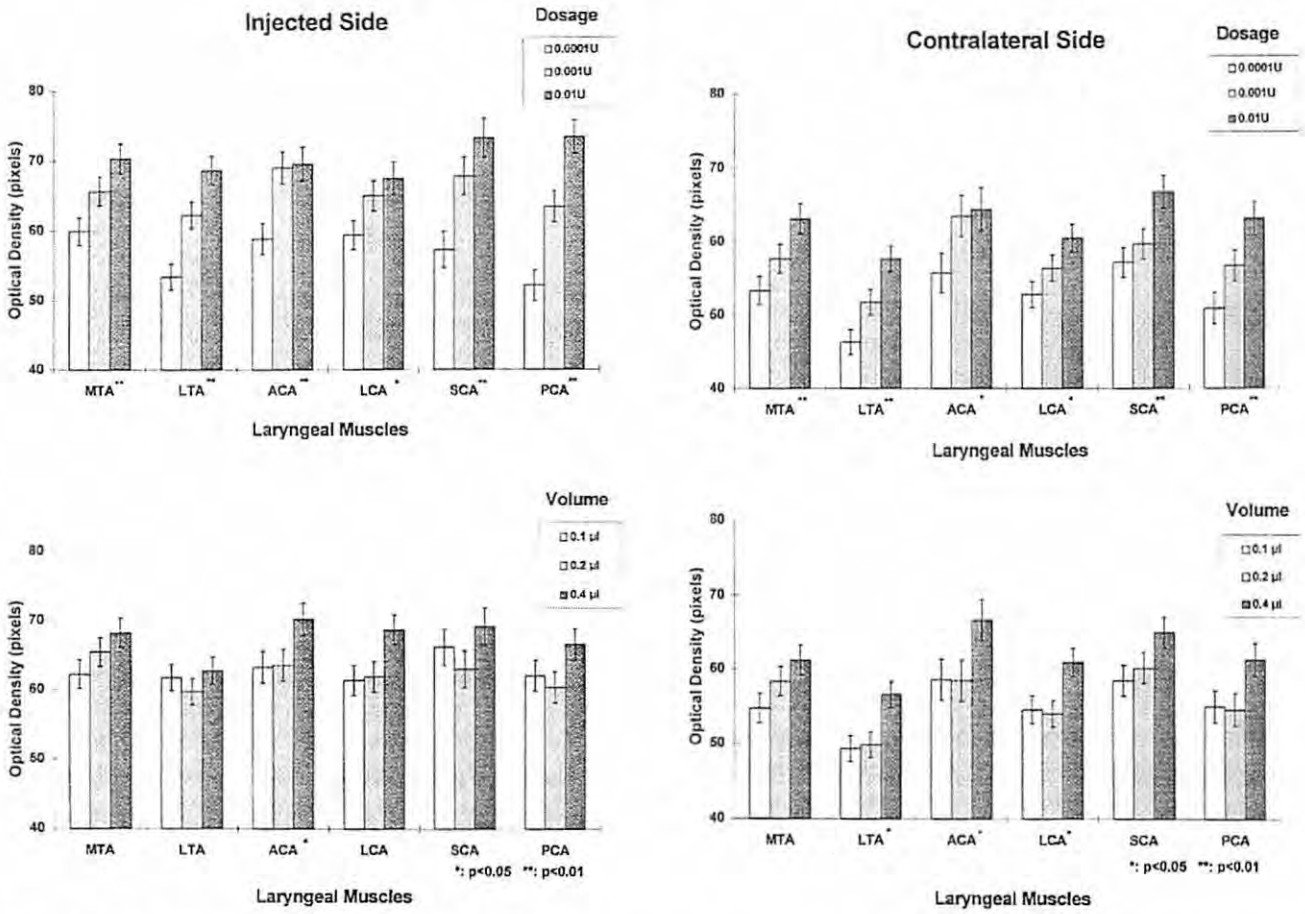


Figure 5. (A; left) The dose (upper chart) and the volume (lower chart) relationships determined by measuring the optical density in each injected laryngeal muscle after stimulation and PAS-staining are presented in these figures. A significant dose relationship is observed in all laryngeal muscles. Only the ACA shows a significant volume relationship. (B; right) A significant dose (upper chart) relationship is observed in all contralateral laryngeal muscles. The LTA, ACA, and LCA of the contralateral side show a significant volume (lower chart) relationship.

Table 3.
Correlation Among the Physiologic Parameters

	Optical Density						Interference
	ACA	MTA	LTA	LCA	SCA	PCA	Pattern
Digitized VF	n.s	n.s	p<0.01	n.s	p<0.01	p<0.01	p<0.01
Movement	-	-	(-0.35)	-	(-0.35)	(-0.50)	(0.73)
Interference	p<0.05	p<0.05	p<0.01	n.s	p<0.01	p<0.01	\
Pattern	(-0.36)	(-0.44)	(-0.62)	-	(-0.58)	(-0.72)	

( ): R

regimens are unknown. Despite the technical difficulties in studying laryngeal musculature in the rat[15], the results of this experiment clearly suggest that the rat larynx is a suitable animal model for studying physiological and histological effects of BT. It is our hope that the findings reported here in an animal model can be extended to humans for the development or enhancement of techniques for treating SD and other related focal dystonias.

Our technique allowed us frequently to observe rat laryngeal functioning under various conditions without surgery. Using our technique, BT dosages from 0.01 to 0.0001 U and volumes from 0.1 to 0.4 μl were injected unilaterally into the vocal fold and allowed observations of safety and associated side-effects. For example, it was found that rats were lost when injected with over 0.1 U of BT. Further, by varying dosage and volume parameters, simulation of the desired BT therapeutic effect of vocal fold paresis, or the undesirable side effect of vocal fold paralysis could be achieved. Vocal fold paresis versus paralysis in the rat may simulate clinical conditions of low[3,4] versus high[1,2] dosages of BT injections in humans with SD symptoms.

Three physiological measures, which included vocal fold movement angle, EMG interference pattern rating and optical density, were used to examine the effects of BT on laryngeal functioning. The results of statistical analyses indicated that vocal fold movement and muscular activity were diminished on the injected side following BT injection. In addition, failure of glycogen depletion following nerve stimulation was also observed, which is an indication of denervation. As such, these measures appear to be useful and reliable indicators of laryngeal function in a rat model.

Alterations in the dosage of BT had measurable physiologic effects on both the injected and contralateral sides. Specifically, progressive decreases in vocal fold movement, muscular activity and increased optical densities were observed. However, no volume related effects were observed on the injected side, which suggests that an injectate volume of 0.1 μl is sufficient to affect adjacent muscles on the injected side.

As shown schematically in Figure 6, optical density measurements revealed progressively larger regions of laryngeal muscle denervation as BT dosage or volume increased. Speacially, a BT dosage of 0.001 U at any

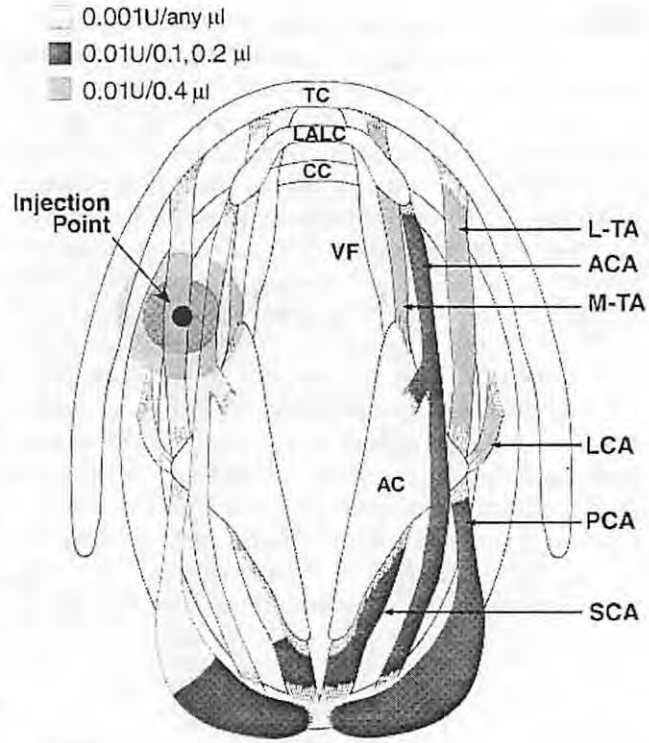


Figure 6. The schematic drawing of the BT infiltration related to the BT dosages and volumes. TC : thyroid cartilage, LALC: laryngeal ala cartilage, CC: cricoid cartilage, AC: arytenoid cartilage, L-TA: lateral thyroarytenoid muscle, ACA: ala-cricoarytenoid muscle, M-TA : medial thyroarytenoid muscle, LCA: lateral cricoarytenoid muscle, PCA: posterior cricoarytenoid muscle, SCA: superior cricoarytenoid muscle, VF: vocal fold.

volume affected only muscles ipsilateral to the site of injection. Evidence of BT diffusion was observed in ipsilateral MTA, LTA, LCA, ACA, SCA, and PCA. When BT dosage was increased to 0.01 U, in a volume of 0.1 or 0.2 μl, contralateral infiltration was observed in the PCA, SCA, and ACA; these muscles were likely affected due to a posterior distribution of motor end plates[10]. At the highest volume (i.e., 0.4 μl), a BT dosage of 0.01 U affected the contralateral LTA and LCA, which are located furthest away from the injection site.

The results reported here suggest that measures of digitized vocal fold movement, EMG interference pattern rating and optical density provide an accurate and reliable assessment of laryngeal functioning in the rat after BT injection. However, many questions remain with regard to the effects of BT. Further investigations from our laboratory will be undertaken with the goal of achieving greater understanding of the mechanisms of BT action on the larynx. By use of this animal model we hope to refine the selection of appropriate treatment variables and to develop ways of enhancing BT efficacy in the treatment of patients with SD.

Conclusions

This study provides a basis for using a rat laryngeal model for investigating the effects of BT. Significant BT dosage effects were observed relative to vocal fold motion, EMG interference pattern rating in laryngeal muscles and optical density of PAS stained muscle after electrical stimulation.

The major findings of this study may be summarized as follows: 1) BT dosages over 0.001 U significantly affected adjacent muscles, 2) a dosage of 0.01 U in a volume of 0.1 or 0.2 µl infiltrated the nearest contralateral laryngeal muscles, and 3) a dosage of 0.01 U in a volume of 0.4 µl infiltrated to the most distant contralateral laryngeal muscles, such as the LTA and LCA muscle.

The animal model presented in this paper will be used in future experiments to further delineate the mechanisms of BT and in the development of new methods for treating and managing SD and related focal dystonias.

Acknowledgment

This work was supported by the National Center for Voice and Speech through Grant P60 00976 from National Institute on Deafness and Other Communication Disorders.

The authors would like to thank Ms. Joan Kozel for illustrations.

References

1) Miller RH, Woodson GE, Jankovic J. Botulinum toxin injection of the vocal fold for spasmodic dysphonia. *Arch Otolaryngol Head Neck Surg* 1987;113: 603-605.

2) Ludlow CL, Naunton RF, Sedory SE, Schulz GM, Hallet M. Effect of botulinum toxin injections on speech in adductor spasmodic dysphonia. *Neurology* 1988; 38: 1220-1225.

3) Ford CN, Bless DM, Lowery JD. Indirect laryngoscopic approach for injection of botulinum toxin in spasmodic dysphonia. *Otolaryngology-Head and Neck Surgery* 1990;103(5):752-758.

4) Ford CN, Bless DM, Patel NY. Botulinum toxin treatment of spasmodic dysphonia: Technique, indications, efficacy. *J Voice* 1992;6(4):370-376.

5) Inagi K, Ford CN, Rodriquez AA, Schultz E, Heisey DM. Efficacy as a function of timing in repeated botulinum toxin injections. *Ann Otol Rhinol Laryngol* 1996 (submitted).

6) Inagi K, Ford CN, Bless DM, Heisey D. An analysis of factors affecting botulinum toxin results in spasmodic dysphonia. *J. Voice* 1996; 10(3):306-313.

7) Cohen SR, Thompson JW. Use of botulinum toxin to lateralize true vocal cords: A biochemical method to relieve bilateral abductor vocal cord paralysis. *Ann Otol Rhinol Laryngol* 1987; 96: 534-541.

8) Cohen SR, Thompson JW, Camilon FS. Botulinum toxin for relief of bilateral abductor paralysis of the larynx: Histologic study in as animal model. *Ann Otol Rhinol Laryngol* 1989; 98: 213-216.

9) George EF, Zimbler M, Biller HF, Wu B, Sanders I. Quantitative mapping of the effect of botulinum toxin injections in the thyroarytenoid muscle. *Ann Otol Rhinol Laryngol* 1992;101:888-892.

10) Inagi K, Schultz E, Ford CN. : An anatomical study of the rat larynx: Establishing the rat model for neuromuscular function. *Otolaryngol Head and Neck Surgery* 1997 (in press).

11) Merati A, Bodine SC, Bennett T, Jung H, Furuta H, Ryan AF. Identification of novel myosin heavy chain gene expressed in the rat larynx. *Biochimica et Biophysica Acta* 1996;1306:153-159.

12) Merati A, Turner WJD, Dang V, Bodine-fowler S. Fiber type size and composition of rat thyroarytenoid muscle in the normal and denervated state. *Otolaryngol Head and Neck Surgery* 1995;113(2);118-119.

13) Inagi K, Rodriquez AA, Ford CN, Heisey DM. Transoral electromyographic recordings in botulinum toxin injected rat larynges. *Ann Otol Rhinol Laryngol*, 1997 (in press).

14) SAS Institute Inc. SAS/STAT User's Guide Version 6. 4th ed. Vol 2. Cary NC, 1989.

15) Kobler JB, Datta S, Goyal RK, Benecchi EJ. Innervation of the larynx, pharynx, and upper esophageal sphincter of the rat. *J. Comp Neurol* 1994; 349:129-147.

NCVS Status and Progress Report - 11
May 1997, 91-94

Impact of Signal Type of Validity of Voice Perturbation Measures

Michael P. Karnell, Ph.D.
Department of Otolaryngology-Head and Neck Surgery, The University of Iowa
Department of Speech Pathology and Audiology, The University of Iowa
Ann Chang, M.A.
Department of Speech Pathology and Audiology, The University of Iowa
Alice Smith, Ph.D.
Henry Hoffman, M.D.
Department of Otolaryngology-Head and Neck Surgery, The University of Iowa

Abstract

This project examined the impact of signal stability as described by Titze (1995) on the validity of jitter, shimmer and noise-to-harmonic ratio measures obtained from 61 patients with voice disorders. The results showed that the validity of the measures *decreased* when Type II signals were excluded from analysis. This finding was contrary to the theory that perturbation measures are clinically valid only for Type I signals.

Introduction

Perturbation measures have been found to be very sensitive to both large and small instabilities in the voice waveform (Rabinov et al., 1994, Hoyt et al., 1994, Walton & Orlikoff, 1994, Debruyne et al., 1994, Sanderson et al., 1993). When the instability is large, the validity of the measure becomes questionable.

Titze (1995) suggested that perturbation measures are only appropriate for waveforms that have a continuous periodic quality (Type I signals). Such signals are easily analyzed by pitch and amplitude tracking programs commonly used for voice perturbation analyses. Type II signals are those that have periods of stability interspersed with sudden, substantial shifts or breaks in stability which he described as "bifurcations". Pitch and amplitude tracking systems can track the relatively stable portions of the waveform but errors occur when the bifurcations are encountered. Type III signals are extremely unstable cannot be analyzed with common voice analysis systems.

Analysis of a Type I signal is shown in Figure 1. Below the acoustic waveform trace are the corresponding continuous frequency and amplitude plots (MDVP, Kay Elemetrics). A Type II signal, shown in Figure 2, illustrates a bifurcation which appears as a gap in the frequency and

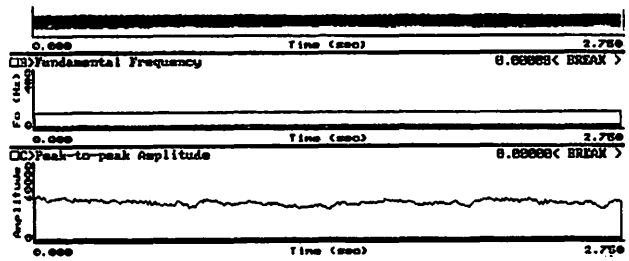


Figure 1. A graphic example of a Type I signal showing the acoustic waveform at the top and the corresponding narrow-band spectrogram at the bottom. Note the continuous harmonic structure over time in the spectrographic display.

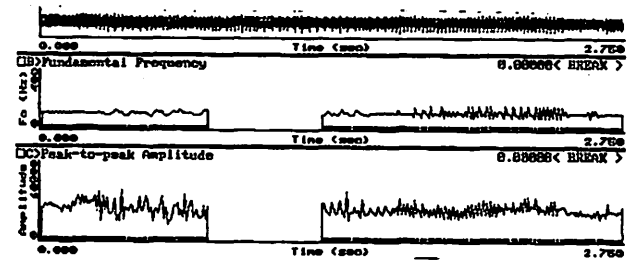


Figure 2. A graphic example of a Type II signal. Note the "bifurcation" that appears is a discontinuity in the harmonic structure of the spectrographic display.

amplitude plots. Such gaps clearly illustrate where periods of waveform instability caused brief failures of the pitch and amplitude analysis algorithms. Analysis of a Type III signal produces no plot at all but rather an error message stating the signal is not voiced.

Titze (1995) recommended that signals that have "bifurcations" (Type II signals) or that are "chaotic" (type III signals) are appropriately described by visual or perceptual means respectively. For example, he recommended using a narrow band spectrographic displays for visual documentation of Type II signals, not unlike how radiographs are used to document tumors or broken bones. For Type III signals, he suggested using perceptually based descriptors (e.g., severely dysphonic or aphonic) rather than attempting acoustic analyses that would result in invalid measurements.

Titze's suggestions directly impact the application of acoustic perturbation measures in the clinical voice evaluation process. If this approach to signal typing is correct, perturbation analysis may be appropriate only for a small subgroup of individuals with relatively mild or minor voice disorders. The purpose of this project is to examine the relative impact of signal type on the validity of several select perturbation measures.

Methods

Subjects

A series of 61 adult patients with suspected voice or laryngeal disorders were evaluated in the Department of Otolaryngology-Head and Neck Surgery. Evaluations consisted of perceptual, acoustic, and stroboscopic analysis of voice and laryngeal function. In this study, the assessment power of the perceptual and acoustic analyses were further examined with consideration of acoustic signal type.

Tasks

Patients were instructed to produce conversational speech, a reading task, and multiple sustained [a] and [i] productions during the course of the evaluation. Sustained [a] vowels were recorded for acoustic analysis. Sustained [i] vowels were recorded during videostroboscopic recordings of laryngeal vibratory function. For the purposes of this project, only the acoustic recordings for the sustained [a] vowels were considered.

Perceptual Analysis:

Clinician judgments of voice quality were recorded according to the GRBAS system described by Hirano (1984) where G = Grade, R = Roughness, B = Breathiness, A = Asthenia, and S = Strain. Each parameter was rated on a four point scale where 0 = normal, 1 = mildly involved, 2 = moderately involved, 3 = severely involved by a certified speech-language pathologist with extensive (+10 years) experience in the assessment of voice disorders. A single composite rating was recorded after the clinician listened to all conversational, read, and sustained phonations produced by the patient during the evaluation.

Acoustic Analysis

Each patient was instructed to produce three sustained [a] productions at comfortable pitch and loudness levels and to attempt to make each production as steady as possible. Recordings were made direct to disk for subsequent analysis using the MutiDimensional Voice Program (MDVP, Kay Elemetrics). Each production was subsequently analyzed for fundamental frequency, jitter (MDVP measure "Jitt" derived from Pinto & Titze, 1990), shimmer (MDVP measure "Shim" derived from Pinto & Titze, 1990), and noise-to-harmonic ratio (MDVP measure "NHR" described by Delyski, 1993).

Fundamental frequency, jitter, shimmer, and noise-to-harmonic ratio were averaged across the three sustained [a] productions and recorded for subsequent analysis. Each voice sample was approximately 2.75 seconds in duration and did not include voice onset or offset segments.

Signal Typing

Subsequent to recording, each sustained [a] vowel was visually and perceptually inspected for bifurcations and noise Titze (1995). Narrow band spectrograms, pitch contours, and amplitude contours produced via CSL (Kay Elemetrics) from the recorded signals were reviewed. Signals that contained no bifurcations were judged to be Type I. Signals that contained bifurcations were judged to be Type II. Type III signals were those that were essentially non-periodic or chaotic signals. Type III signals were excluded from this analysis.

Validity

An approach to validity analysis was employed that examined test sensitivity, specificity, efficiency, positive predictive power and negative predictive power. This approach may be described as follows:

Let

a=true positives (screening positive and patient affected)
b=false positive (screening positive; patient unaffected)
c=false negative (screening negative; patient affected)
d=true negative (screening negative; patient unaffected)
sensitivity = a/a+c
specificity = d/b+d
positive predictive power = a/a+b
negative predictive power = d/c+d
efficiency = a+d/a+b+c+d

Normative perturbation thresholds were established in order to determine that each patient was either positive (abnormal) or negative (normal) on the basis of the perturbation measures. Patients were considered to have abnormal jitter if the mean jitter measurement exceeded (1.04%).

Abnormal shimmer was considered as a measure that exceeded (3.81%). Abnormal noise to harmonic ratio was judged to be a mean N/H ratio that exceeded (.19). These values were those recommended in the MDVP acoustic analysis system (Kay Elemetrics) on the basis of analysis of 15 individuals with normal voices and 53 patients with laryngeal diseases (Deliyski, 1993). Patients were considered to be "affected" if the "Grade" perceptual rating of voice quality was greater than "0" at the time of voice evaluation.

Results

Voice Quality

Nineteen of the 61 patients were judged to have normal voice quality (G=0). The remaining 42 patients had abnormal voices. Twenty patients were judged to have mildly dysphonic voice quality (G = 1). Eighteen were judged to have moderately dysphonic voice quality. Four were judged to have severely dysphonic voice quality.

Signal Type

Of the 61 patients examined, 38 (62%) produced only type I signals. Of the remaining 23, four patients produced one type II signal, six produced two type II signals, and 13 produced three type II signals. None of the signals of the 61 patients included in this study were considered to be type III signals.

Validity

Assessment validity was first considered without consideration of signal type. The results are shown in Table 1.

Jitter: The jitter algorithm correctly identified 71.4% (30) of the 42 patients who had abnormal voice quality (sensitivity). However, only 63.2% (12) of the 19 patients considered to have normal voices also had jitter values that were within normal limits. Therefore, while positive predictive power for jitter was relatively high (.811), negative predictive power was no better than chance (.500) and overall measurement efficiency was .689.

Table 1.
Validity Analysis for All 61 Patients Regardless of Signal Type

	JITTER	SHIMMER	N/H RATIO
TRUE POSITIVES	30	34	15
FALSE POSITIVES	7	4	0
FALSE NEGATIVES	12	8	27
TRUE NEGATIVES	12	15	19
SENSITIVITY	.714	.810	.357
SPECIFICITY	.632	.789	1.000
POS. PREDICTIVE POWER	.811	.895	1.000
NEG. PREDICTIVE POWER	.500	.652	.413
EFFICIENCY	.689	.803	.557

Shimmer: The shimmer algorithm was successful at correctly identifying 81.0% (34) of the 42 patients with abnormal voice quality. It also identified 78.9% (15) of the 19 patients who had normal voice quality. Both positive and negative predictive power for shimmer was higher than for jitter. Not surprisingly, therefore, overall measurement efficiency for shimmer (.803) was also higher.

N/H Ratio: Of the three algorithms examined, noise-to-harmonic ratio performed least well at identifying abnormal patients. It successfully identified only 35.7% (15) of the 42 patients with dysphonic voices. However, it accurately identified all 19 patients who had normal voice quality. Because there were no false positives, positive predictive power for N/H ratio was a perfect 1.000. However, the high rate of false negatives resulted in a negative predictive power of only .413. Overall measurement efficiency was only .557, lower than either jitter or shimmer.

According to Titze (1995), perturbation measurement may be innappropriate when applied to signals with bifurcations (type II signals). It may be expected, therefore,

Table 2.
Validity of 38 Patients Who Produced no Type II Signals

	JITTER	SHIMMER	N/H RATIO
TRUE POSITIVES	13	14	2
FALSE POSITIVES	5	3	0
FALSE NEGATIVES	9	8	20
TRUE NEGATIVES	11	13	16
SENSITIVITY	.591	.636	.091
SPECIFICITY	.688	.812	1.000
POS. PREDICTIVE POWER	.722	.824	1.000
NEG. PREDICTIVE POWER	.550	.619	.444
EFFICIENCY	.632	.711	.474

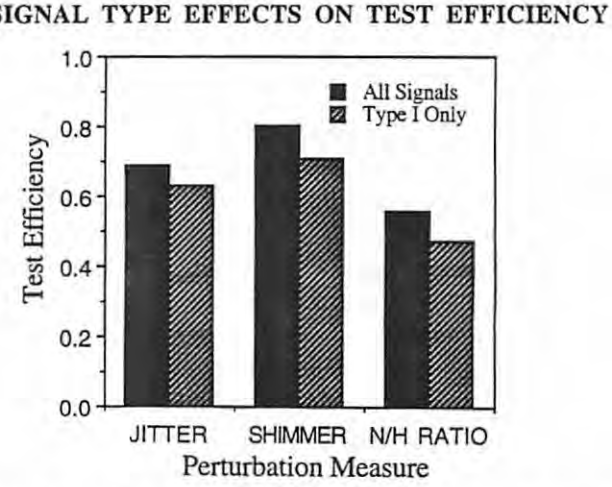


Figure 3. Test efficiency reduced when only Type I signals were included in the analysis.

that overall measurement validity of perturbation measures would improve if they were applied only to signals free of such bifurcations. However, when patients who produced Type II signals were eliminated from the analysis, overall validity of all three perturbation measures suffered (Table 2).

Effects of removing type II signals from the analysis are shown graphically in Figure 3. Overall test efficiency for each measure was reduced when only type I signals were included in the analysis.

Discussion

Perturbation measures are considered very sensitive indicators of vocal stability. Because they are very sensitive to small perturbations in the voice waveform, Titze has argued that they should not be applied to waveforms that contain sudden large changes in periodic stability (bifurcations) or waveforms that lack periodic stability altogether (chaotic). We argue, however, that the clinical appropriateness of these measures depends upon how they are applied to the clinical process and how the measurement system employed deals with bifurcated or chaotic waveforms.

One clinical application of perturbation measures involves determining whether the measured perturbation in a given voice signal is within normal limits. In this sense, the clinical question is binary in nature and depends on whether the acquired measure falls on one side or the other of an established, accepted norm. When used in this manner, the measures obtained for this study appear to show greater test efficiency when bifurcated signals are included in the analysis because of the manner in which the measurement system deals with bifurcations. The measurement system used for this project automatically attempts to measure perturbation regardless of the size of the perturbation. When the perturbation is very large it consistently reports a number higher than the normal threshold. Such waveforms would, therefore, be categorized as abnormal on the basis of perturbation measures alone. The perceptual effect of a bifurcation is a "voice break" which usually occurs as a relatively large, short term alteration in voice quality. Such voices may be categorized as abnormal on the basis of this percept alone, or they may be categorized as abnormal because the voice breaks occur in conjunction with other more longer term abnormalities in voice quality. Therefore, the liklihood of agreement between the perceptual and the perturbation findings for a given signal will be strong when type II signals are included in the analysis.

Variations in perturbation in Type I signals may be subtle compared to Type II signals. Subtle variations in these waveforms also have less dramatic perceptual effects. As the perceptual effect becomes more subtle, categorization of the signal as normal or abnormal on the basis of percept becomes more difficult and subject to error. Consequently, agreement with perturbation measures suffers.

The absolute accuracy of the actual perturbation measurement in type II signals remains in question for the reasons offered by Titze (1995). However, provided whatever error there may be included in those measures results in a substantial increase in the measurement, it is unlikely that signals with bifurcations will be considered normal on the basis of perturbation measures. Clinical application of perturbation measurements in the manner described here, therefore, seems appropriate for both type I and type II signals.

Although they were not included in this analysis, we agree that type III signals, those that have little or no periodic component, require other means of measurement and analysis. The perturbation analysis system in our laboratory rejects such signals as being "unvoiced" resulting in no measurements based on periodicity. Such signals may be best documented using signal printouts, spectrograms, or other approaches.

Acknowledgements

This research was funded in part by PHS Research Grant DC-00976, National Institute for Deafness and other Communicative Disorders.

References

Debruyne F. Delaere P. Wouters J. Uwents P. (1994) Acoustic analysis of tracheo-oesophageal versus oesophageal speech Journal of Laryngology & Otology. 108:325-8.

Hirano, M. (1981) CLINICAL EXAMINATION OF VOICE New York: Springer-Verlag Wien.

Hoyt DJ. Lettinga JW. Leopold KA. Fisher SR. (1992) The effect of head and neck radiation therapy on voice quality. Laryngoscope. 102(5):477-80.

Rabinov CR. Kreiman J. Gerratt BR. Bielamowicz S. (1995) Comparing reliability of perceptual ratings of roughness and acoustic measure of jitter. Journal of Speech & Hearing Research. 38(1):26-32.

Sanderson RJ. Anderson SJ. Denholm S. Kerr AI. (1993) The assessment of alaryngeal speech. Clinical Otolaryngology. 18(3):181-3.

Titze, I.R. Workshop on acoustic voice analysis: Summary statement. The National Center for Voice and Speech for the National Institute on Deafness and other Communication Disorders, 1995.

Walton JH. Orlikoff RF. (1994) Speaker race identification from acoustic cues in the vocal signal. Journal of Speech & Hearing Research. 37 (4:738-45).

NCVS Status and Progress Report - 11
May 1997, 95-105

An Anatomic Study of the Tensor Veli Palatini and Dilatator Tubae Muscles in Relation to Eustachian Tube and Velar Function

Raffi Barsoumian, M.S.
University of Rochester School of Medicine
David P. Kuehn, Ph.D.
Department of Speech and Hearing Sciences, University of Illinois at Urbana-Champaign
Jerald B. Moon, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa
John W. Canady, M.D., F.A.C. S.
Department of Otolaryngology, The University of Iowa Hospitals and Clinics
Department of Surgery, The University of Iowa Hospitals and Clinics
Department of Orthopedics, The University of Iowa Hospitals and Clinics

Abstract

In a gross anatomic study of 20 sides in 16 human head specimens, the tensor veli palatini, dilatator tubae, and tensor tympani muscles were studied. The tensor veli palatini was observed to insert onto the anterior one-third of the pterygoid hamulus, whereas the dilatator tubae rounded the middle one-third of the ptervgoid hamulus without an insertion. Thus, the dilatator tubae, not the tensor veli palatini, could serve to tense the anterior velum. An insertion from the superior pharyngeal constrictor muscle onto the posterior one-third of the hamulus could provide a curbing function for the dilatator tubae muscle. Adipose tissue, located at the hamulus, could provide lubrication for the tendinous fibers of the dilatator tubae as they round the hamulus. The dilatator tubae was observed to attach to the hook of the eustachian tube and is accepted as the tubal dilator. The tensor yen palatini was observed to be continuous with the tensor tympani. Full color versions of the figures are available at the following Website: http://www.shc.uiowa.edu/papers/tensor/

The literature provides numerous anatomic and histologic descriptions of the eustachian tube and velum but lacks an accurate description of the muscles that are common to both structures and how these muscles interact to influence their function. Of these, the tensor veli palatini is believed to be the paratubal muscle most directly involved in eustachian tube function.

Tensor Veli Palatini Anatomy

The tensor veli palatini is described as a thin, fan-shaped muscle originating from the membranous portion of the eustachian tube, the scaphoid fossa, and the spine of the sphenoid bone to become tendinous and form an acute angle as it rounds the pterygoid hamulus to insert onto the hard palate and to form the palatine aponeurosis (Urbantschitsch, 1884; Simkins, 1943; Ross, 1971; Proctor, 1973; Rood, 1973; Dickson, 1975; Rood & Doyle, 1978). It is innervated by the mandibular division of the trigeminal nerve (Rich, 1920b; Perlman, 1991). The tensor veli palatini is a bipenniform muscle according to Dickson & Maue-Dickson (1982). Upon inserting into the velum, it unites with its complement arising from the opposite side (Bateman & Mason, 1984). An additional attachment of the tensor veli palatini to the medial pterygoid lamina has been reported (Ross, 1971; Rood, 1973; Rood & Doyle, 1978).

The tensor tympani muscle has been observed to be a direct continuation of the tensor veli palatini muscle (Rudinger as reported by Urbantschitsch, 1884; Mayer as reported by Politzer, 1926; Wolff, 1934; Lupin, 1969; Proctor, 1973; Rood & Doyle, 1978; Seif & Dellon, 1978). However, neither Graves and Edwards (1944) nor Gray (1973) report a continuity between the tensor tympani and the tensor veli palatini. Although Mayer as reported by Politzer (1926), Proctor (1973) and Rood and Doyle (1978) have observed this continuity, they report a fibrous band

demarcating the junction of the two muscles. Rood and Doyle (1978) describe the tensor tympani as a bipennate muscle with its medial portion arising from the sphenoid bone and the top of the cartilaginous eustachian tube, whereas its lateral portion arises from the tensor veli palatini. It inserts onto the manubrium of the malleus. Akin to the tensor veli palatini, the tensor tympani is innervated by the mandibular division of the trigeminal nerve (Perlman, 1991).

Gray (1973) and Proctor (1973) report a bursa to exist between the tendon of the tensor veli palatini and the hamulus, whereas Salins and Bloxham (1989) and Kronman et al. (1991) also report a bursa but contend that the tendon of the tensor veli palatini passes through the bursa as it rounds the hamulus. Moore (1992) describes a bursa as an enclosed sac which contains synovial fluid limited by a synovial membrane. The function of a bursa is to reduce the friction generated by a tendon passing over bone. According to this definition, Ross (1971) did not observe a bursa after careful microscopic dissection of eight hamulus-tendon complexes. In one cadaver, a gelatinous substance that strongly resembled a bursa was found between the tendon of the tensor veli palatini and the hamulus. However, it was suggested that this structure was pathologic and resulted from the degeneration of the tensor tendon caused by friction at its pressure point with the hamulus. Although Kronman et al. (1991) and Salins & Bloxham (1989) observed a bursa to exist, they did so from pathologic specimens; consequently, the accuracy of these findings must be suspected.

With respect to the composition of the tensor veli palatini muscle, Proctor (1973) described the top half to be muscular, whereas the bottom half, approaching the hamulus, is tendinous. Internally, the top half of the muscle is tendinous whereas its bottom half is muscular (Proctor, 1973). The location and dimensions of the human tensor veli palatini is consistent among individuals (Kuehn & Kahane, 1990). However, controversy abounds concerning the relation of the tensor veli palatini with the pterygoid hamulus.

A number of authors have observed the tensor veli palatini to insert onto the hamulus (Weber-Leil as reported by Urbantschitsch, 1884; Rich, 1920a; Politzer, 1926; Ross, 1971), whereas others, (Dickson & Maue, 1970; Rood & Doyle, 1978) report this insertion to be absent. In a gross anatomy study, Ross (1971) observed this insertion in 14 of 20 embalmed cadavers and in 13 of 32 fresh cadavers. Barring only one fresh cadaver in which this insertion was sufficiently anchored to be of functional significance, Ross (1971) attributed these findings to be a consequence of tissue shrinkage due to embalming. Conversely, Rich (1920a) observed a firm attachment of the tensor veli palatini onto the hamulus in the dog and considered this attachment to be critical to the function of the tensor veli palatini in the dog.

Aside from Rich (1920a), previous investigators have diminished the importance of the relation between the tensor veli palatini and the hamulus. For example, Rood (1973) declared, "that the tensor veli palatini is the tubal dilator, regardless of the absence of a hamular insertion." Notwithstanding its validity, this assertion must not negate the importance of furnishing the surgeon with an accurate description of this relation, because it is essential to reconstructing a palatal cleft in the most propitious manner. If, in fact, a hamular insertion does exist, the biomechanics of the tensor veli palatini would differ from the situation in which an insertion is absent. Consequently, a corresponding biomechanical model may be determined by ascertaining this relation. This biomechanical model may elucidate the manner in which the tensor veli palatini functions. Much of the confusion that surrounds the anatomy and physiology of the tensor veli palatini has been generated, albeit unintentionally, by the investigators who have studied this muscle. Traditionally, the authors have generalized their anatomic or physiologic findings to be true for the muscle as a whole, when in fact, they presumably described only a part of the tensor veli palatini. This confusion may be due, in part, to nomenclature. As greater insight has been gained into its function, the tensor veli palatini has been identified by various names.

Nomenclature

Historically, the tensor veli palatini has been identified by the following names listed in order of their appearance in the literature: spheno-salpingo-staphylinus; circumflexus palati (Politzer, 1926); tensor palati by Jonnesco as reported by Poirier and Charpy (1912); abductor tubae by Tortual as reported by Politzer (1926); dilator tubae by von Troltsch as reported by Urbantschitsch (1884); tensor veli palatini by Politzer (Politzer, 1926). Of these, tensor palati, dilatator tubae, and tensor veli palatini are the terms presently used, with tensor veli palatini being the newest and most commonly used name.

Tensor palati was the name used for this muscle by English anatomists during the nineteenth century (Gray, 1859; Quain, 1909). German anatomists (Urbantschitsch, 1884) referred to this muscle as the tensor veli. In 1926, Politzer combined these two names and renamed the muscle, tensor veli palatini. Although tensor veli palatini is the most current and commonly used name, Ross (1971) used the dated name, tensor palati, when reporting his observations.

In addition, some authors, for example von Troltsch and Tortual (as cited by Politzer, 1926), have used the name dilator tubae or abductor tubae interchangeably with tensor palati. Rudinger, who is credited for establishing the name dilatator tubae, denotes it as a muscle band, distinct from the tensor veli palatini, which arises from the eustachian tubal cartilage and joins with the tensor palati (Quain, 1909). Rood and Doyle (1978) reported the dilatator tubae to blend with the tensor veli palatini prior to approaching the hamulus, at which point the distinction of these two muscle

bundles was obscured, contrary to descriptions by Ross (1971) and Rood (1973) who observed distinct tendons corresponding to these muscle bundles. Unaware of this distinction, Proctor (1973) used the name dilatator tubae interchangeably with tensor veli palatini.

Rood and Doyle (1978) adopted the name dilatator tubae and expanded its description based on their observations and those of Terracol et al. (1949) and Ross (1971) who described the tensor veli palatini as being divided anatomically and functionally into two distinct planes. Terracol et al. (1949) classified the layers of the tensor veli palatini as superficial and deep, with the superficial layer originating from the muscle's osseus origins whereas the deep layer originated from the membranous portion of the eustachian tube, thereby implicating it as the functional component of the tensor veli palatini responsible for tubal function. Ross (1971) reported 75-80% of the tensor veli palatini to originate from the fibrous and cartilaginous portion of the eustachian tube, whereas the remaining portion of the muscle arises from bone. However, it is unclear whether or not Ross (1971) was referring to the membranous portion of the eustachian tube when reporting about a fibrous origin. Also, Rood and Doyle (1978) observed a portion of the tensor veli palatini to originate "from the entire membranous wall of the posterior one-third to one-half of the membrano-cartilaginous tube," though an attachment to the lateral cartilaginous lamina was not observed in either the fetal or adult specimens. Rood & Doyle (1978) declared their findings and those of Terracol et al. (1949) and Ross (1971) to be consistent with the description of the dilatator tubae by Gray (1973).

In agreement with Rood and Doyle (1978), the names tensor veli palatini and dilatator tubae will be used in this report. The latter term will be used to refer to the specific portion of the tensor veli palatini responsible for tubal dilation.

Muscle Function

The controversy that surrounds the function of the tensor veli palatini, especially the means by which it influences velar and tubal function, is abundant. The tensor veli palatini has been reported to: 1) tense the velum (Quain, 1909; Gray, 1973; Proctor, 1973; Hairston & Sauerland, 1981; Doyle et al., 1990); 2) possess no velar function (Rich, 1920a; Rood & Doyle, 1978; Honjo et al., 1979); 3) have no major effect on movement or overall positioning of the velum (Kuehn, 1990); 4) close the eustachian tube (Blakeway, 1913; Simkins, 1943); 5) assist in tubal content flow (Sief & Dellon, 1978); and 6) open the eustachian tube (Rich, 1920a; McMyn, 1940; Terracol et al., 1949; Dickson & Maue, 1970; Ross, 1971; Proctor, 1973). Some authors credit the tubal dilation to the dilatator tubae (Urbantschitsch, 1884; Gray, 1973; Rood & Doyle, 1978). The authors who report the dilatator tubae to be the eustachian tube dilator are in agreement with the authors who charge the tensor veli palatini with this function. The former authors describe a specific portion of the tensor veli palatini, as previously mentioned.

Hairston and Sauerland (1981) reported that the tensor veli palatini is active during the inspiratory phase of respiration. From this observation they concluded that during the inspiratory phase of respiration, a taut velum is favored over a flaccid velum. According to the authors, a flaccid velum would promote air turbulence and a subsequent increase in upper airway resistance by narrowing the oral cavity. Therefore, the activation of the tensor veli palatini during the inspiratory phase of respiration must result in tensing the velum because this is most advantageous.

Doyle et al. (1990) bilaterally paralyzed the tensor veli palatini with botulinum toxin in three juvenile rhesus monkeys. On the second post-operative day, the animals stopped feeding and their health was noted to be failing. The animals were diagnosed with aspiration pneumonia and treated. Also, nasal reflux of liquids was observed suggesting palatal dysfunction. Two of these monkeys died on the fourth and fifth post-operative days; the third was sacrificed on the sixth day. An autopsy performed on the third monkey revealed a food bolus obstructing the larynx. They attributed the death of these monkeys to "failure of intrinsic palatal function with concomitant failure of the epiglottis to close the airway during feeding." From this, they concluded that the bilateral paralysis of the tensor veli palatini resulted in velar dysfunction, which suggests the tensor veli palatini has a velar function and that function is to tense the velum.

According to Kuehn (1990), in order for this conclusion to be valid: 1) "a sensory-motor interaction between the velum and larynx to account for the aspiration pneumonia with the tensor veli palatini muscle serving as a common catalyst," must exist and 2) the tensor veli palatini must play a significant role in the biomechanics of the velar mechanism, for its paralysis to result in the nasal reflux. Although Kuehn cited literature which satisfied the first component of this conclusion, he opposed the second component. Kuehn argued that the tensor veli palatini cannot raise or lower the velum appreciably because, for this purpose, it is at a severe mechanical disadvantage; "the applied force delivered to the velum by the tensor tendon winding around the hamulus is very close to the functional pivot of the velum." Bolstering the position of Kuehn, Honjo et al. (1979) observed no velar movement during stimulation of the tensor veli palatini in the dog.

Whether or not the tensor veli palatini may tense the human velum is not certain. It is possible that some tension could occur without appreciable movement of the velum as a whole.

Simkins (1943) reported the tensor veli palatini to passively close the eustachian tube by compressing the

laterotubal tissue. Others have argued for an opposite function. Seif and Dellon (1978) suggested that the tensor veli palatini does not actively open the eustachian tube. Rather, a subsequent increase in muscle girth exerts an inward "pumping" force against the lumen of the eustachian tube upon contraction, thereby "facilitating tubal content flow under the constant force of gravity."

Rich (1920a) demonstrated that damage to the tensor veli palatini incapacitates eustachian tube function in the dog. From this, he concluded that the tensor veli palatini is the only muscle that opens the eustachian tube. Terracol et al. (1949) attributed the dilatory function of the tensor veli palatini to its deep layer whereas Ross (1971) reported the tensor veli palatini to dilate the eustachian tube along two vectors, inferior and anterolateral. Proctor (1973) reported the tensor veli palatini to dilate the eustachian tube upon every swallow. In a more current study, Leider et al. (1993) reported the eustachian tube to dilate 74% of the time upon swallow. Aside from Rudinger (as reported by Urbantschitsch, 1884), Rood and Doyle (1978) are the only authors that use the name dilatator tubae to specify the portion of the tensor veli palatini which dilates the eustachian tube. They ascribe this function to the dilatator tubae because it is the only portion of the tensor veli palatini that attaches to the eustachian tube.

Purpose

The purpose of the present study is to describe the relation between the 1) tensor veli palatini and the pterygoid hamulus, 2) dilatator tubae and the pterygoid hamulus, 3) tensor veli palatini and the eustachian tube, 4) dilatator tubae and the eustachian tube, 5) tensor veli palatini, dilatator tubae, and their tendons with the hard and soft palates and 6) tensor veli palatini and tensor tympani muscle.

Method

Sixteen embalmed adult human cadavers were obtained from the University of Iowa's Deeded Body Program. Of these, 9 female and 7 male specimens were dissected, totaling 20 sides. Seven female specimens were dissected unilaterally and two were dissected bilaterally. Five male specimens were dissected unilaterally and two were dissected bilaterally (Table 1). All of the specimens were Caucasian. The specimens, which are listed in Table 1, satisfied the following selection criteria: (1) an intact pterygoid hamulus, (2) an intact tensor veli palatini, (3) an intact hard palate, and (4) absence of pathology in the oropharyngeal and nasopharyngeal regions. The tensor veli palatini was dissected on one side of 13 bisected heads, numbered as follows: 1-11, 13, 15. The corresponding bisected halves of specimens 8-11, 13, and 15 were discarded due to fracture of the hamulus. A bilateral dissection of the tensor veli palatini was performed on 3 heads; 12, 14, and 16.

The dissection of these specimens included examining the: (A) tensor veli palatini muscle, (B) tendon of the tensor veli palatini rounding the hamulus, (C) dilatator tubae muscle, (D) relation of the tendon of the tensor veli palatini with the hard palate, (E) origin of the tensor veli palatini from the scaphoid fossa and spine of the sphenoid bone, (F) relation of the tensor veli palatini with the medial pterygoid plate, (G) relation of the dilatator tubae with the eustachian tube, (H) relation of the tendon of the tensor veli palatini and dilatator tubae with the velum, and (I) relation of the dilatator tubae with the tensor tympani muscle. A display of the areas examined in each specimen, the side of examination, and specimen sex is included in Table 1.

Following bisections of the head in the midsagittal plane, dissections were performed from a medial, lateral, and inferior aspect in each case, isolating the tensor veli palatini and the dilatator tubae. Dissection from the medial aspect required removal of mucosa, the levator veli palatini

Table 1.
Dissected Areas in Sixteen Specimens

Specimen	A	B	C	D	E	F	G	H	I	Sex	Side
1	x	x	x				x			F	left
2	x	x	x	x	x	x	x	x		M	right
3	x	x	x	x		x	x	x		F	left
4	x	x	x	x		x	x	x		F	left
5	x	x	x	x		x		x		M	left
6	x	x	x	x		x	x	x		F	left
7	x	x	x	x	x	x	x			M	bilateral
8	x	x	x	x		x	x	x	x	F	left
9	x	x	x	x	x	x		x		M	left
10	x	x	x	x	x	x		x		M	left
11	x	x	x			x		x		F	left
12	x	x	x		x	x		x	x	F	bilateral
13	x	x	x		x	x			x	F	right
14	x	x	x			x	x	x		F	bilateral
15	x	x	x	x				x		M	right
16	x	x	x	x		x		x		M	bilateral

A = Tensor veli palatini (medially and laterally)
B = Tendon of tensor veli palatini rounding hamulus
C = Dilatator tubae (medially and laterally)
D = Relation of tendon of tensor veli palatini with hard palate
E = Tensor veli palatini origin from scaphoid fossa and spine of sphenoid bone
F = Relation of tensor veli palatini with medial pterygoid plate
G = Dilatator tubae relation with eustachian tube
H = Relation of tendon of tensor veli palatini and dilatator tubae with velum
I = Relation of dilatator tubae with tensor tympani

muscle, the cartilaginous eustachian tube, and adipose tissue. Dissection from the lateral aspect required removal of the mandible, musculature superficial and deep to the mandible, and adipose tissue. Dissection from the inferior aspect was performed from the oral surface of the palate, requiring removal of glandular and adipose tissue along the oral surface of the hard and soft palate and the molar teeth to expose the fibers of the tensor veli palatini rounding the hamulus and inserting onto the hard palate and extending into the soft palate and palatine aponeurosis.

In addition to standard dissecting instruments, an illuminated magnifying lens, a circular saw, and rongeurs were used. Photographs of the dissections were made in the Otolaryngology - Head and Neck Surgery Department at the University of Iowa Hospitals and Clinics by a professional photographer, D. Kay Klein.

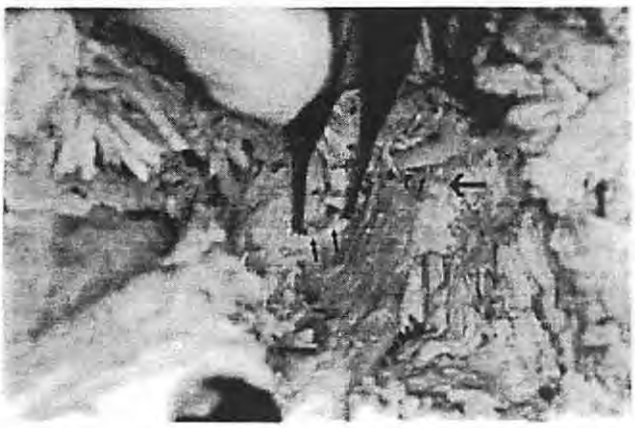


Figure 1. This photograph, a medial view of the right tensor veli palatini, depicts it taking origin from the scaphoid fossa (small arrows) and along the sphenoid bone (smallest arrows) to its spine (large arrow). The dilatator tubae muscle has been removed in order to show the origin from the sphenoid bone and sphenoid spine.

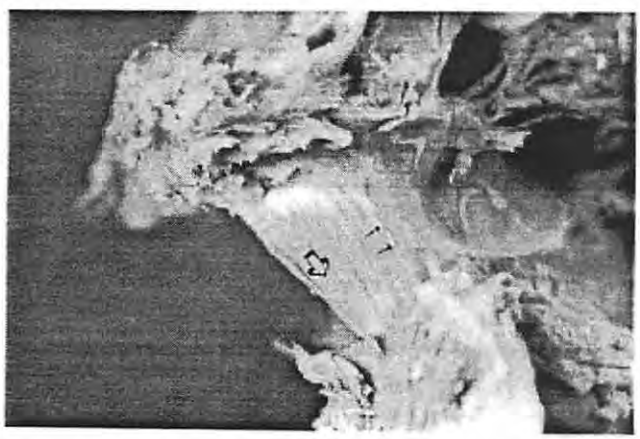


Figure 2. A medial view, depicts the relation of the left tensor veli palatini (small arrows) with the medial pterygoid plate. The tensor veli palatini is attached to the lateral surface of the medial pterygoid plate. The dilatator tubae (large empty arrow) intermingles extensively with the tensor veli palatini in this specimen, obscuring their distinction.

Results

In these descriptions the tensor veli palatini, dilatator tubae, and tensor tympani muscles are considered separately, based on their functional significance. Unless otherwise stated, the following observations were made on 20 of 20 sides, in 16 specimens.

Tensor Veli Palatini

Superomedially, the tensor veli palatini originates from the scaphoid fossa and adheres to an osseous origin as it travels laterally to the spine of the sphenoid bone, its superolateral origin as shown in Figure 1. (Full color versions of the figures are available at our Website as is identified in the Abstract.) From its superomedial origin it descends immediately lateral to the medial pterygoid plate along its periphery. On 18 of 20 sides in 14 specimens, the tensor veli palatini attached to the medial pterygoid plate (Figure 2). However, this attachment cannot be considered a true insertion because a fascial sheath envelops the tensor veli palatini preventing an insertion onto the medial pterygoid plate.

Although the size and location of the tensor veli palatini was constant across the specimens, the composition of the tensor veli palatini varied across specimens. Tendinous fibers that were interspersed throughout the muscle merged into a tendon upon approaching and rounding the hamulus. However, some muscle fibers persisted in this tendon (Figure 3) and along with it, inserted onto the hamulus, hard palate, and into the velum. The tendon of the

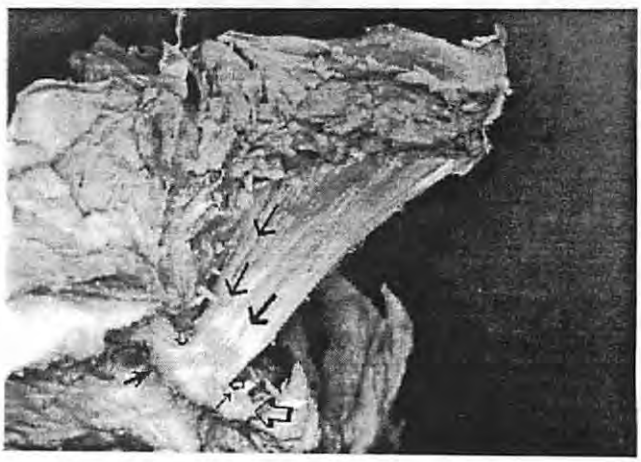


Figure 3. A lateral view, depicts the left tensor veli palatini (large thin arrows) and the left dilatator tubae (thick arrow) merging into a tendon approaching the hamulus at an acute angle (small empty arrows). After the tendinous fibers wrap around the hamulus, those corresponding to the tensor veli palatini (arrow with slanted head) continue anteriorly toward the hard palate, whereas those corresponding to the dilatator tubae (small arrow) fan out posteriorly into the velum. Also, the superior constrictor (large empty arrow) is shown originating from the tip of the hamulus and serving as a boundary limiting lateral movement of the dilatator tubae. The junction of the dilatator tubae and the superior constrictor is represented by the double-headed arrow.

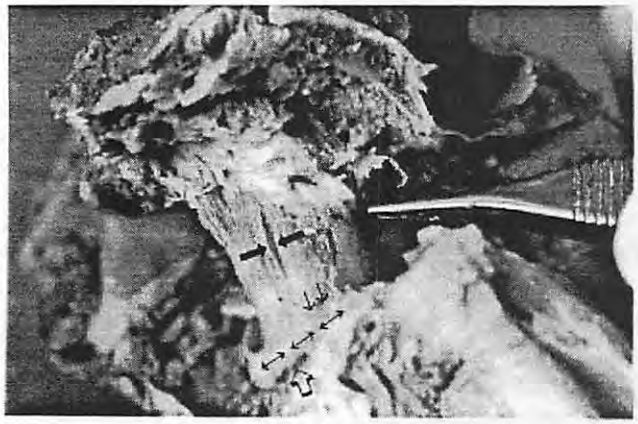


Figure 4. A medial view, depicts the left tensor veli palatini inserting onto the hamulus (thin arrows). Also, the distinction between the left tensor veli palatini and the left dilatator tubae (thick arrows) is apparent. A connective tissue junction and intermingling of some muscle fibers is shown between the thick arrows. The tendinous fibers of the left dilatator tubae are shown to round the hamulus devoid of an insertion (large empty arrow). The superior constrictor muscle has been removed in order to show the entire hamulus (double-headed arrows).

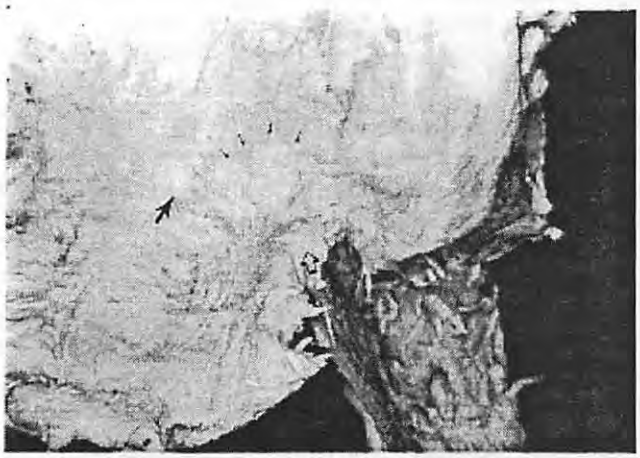


Figure 5. An inferior view of the hard palate and velum, depicts the left tensor veli palatini (empty arrow) rounding the hamulus and continuing superoanteriorly to insert onto the posterior border of the hard palate (small arrows). The posterior nasal spine is located at the tip of the arrow with the slanted head.

tensor veli palatini firmly inserted onto the anterior one-third of the hamulus as it rounded this structure (Figure 4). After rounding the hamulus (Figure 5), the tendon of the tensor veli palatini inserted into the half of the posterior border of the hard palate corresponding to its side of origin, whereas its mate inserted onto the other half (Figure 6). In the velum, the tensor veli palatini fanned out and joined, but did not interdigitate with its mate, at the midline raphe.

Tensor Tympani

From the level of its superomedial origin, the tensor veli palatini extended into the tympanic cavity, to become continuous with the tensor tympani (Figure 7). A tendinous junction delineated the two muscles (Figure 7).

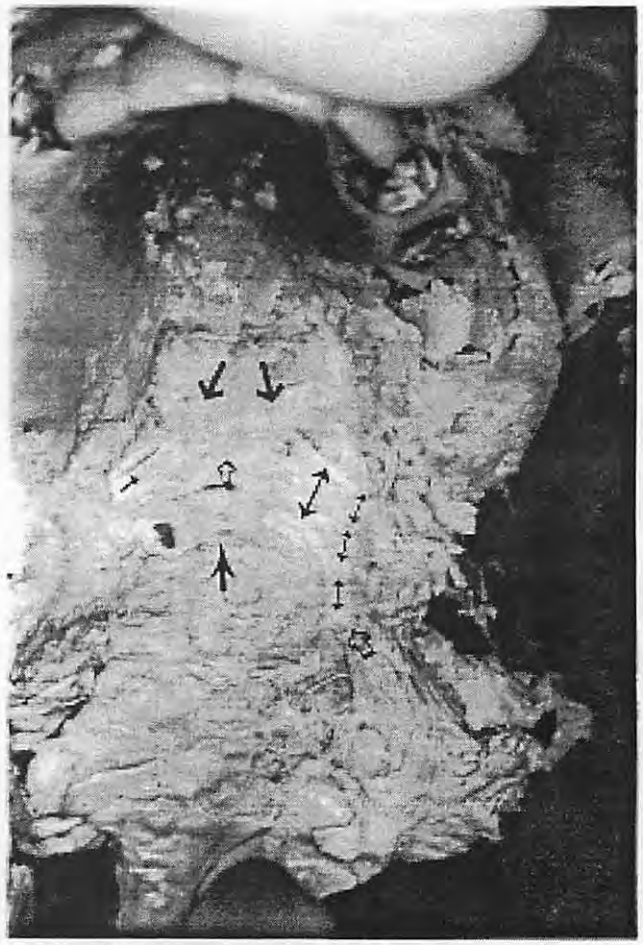


Figure 6. An inferior view of the hard palate and velum, depicts the tendon of the tensor veli palatini and some of its muscle fibers (small arrow), inserting onto the posterior border of the hard palate, bilaterally (large arrows). The palatine aponeurosis (double-headed arrow) is seen rounding the hamulus (small double-headed arrows). The superior constrictor is shown inserting onto the tip of the hamulus (small empty arrow). The velar midline, where the tendons of the tensor veli palatini from either side join, is designated by the arrow with the slanted head. The posterior nasal spine is located at the tip of the small empty arrow.

Dilatator Tubae

The dilatator tubae attached to the hook of the eustachian tube (Figure 8), medially overlapping the superolateral origin of the tensor veli palatini (Figure 9). Due to both a membranous and cartilaginous composition, the lateral wall of the eustachian tube, commonly referred to as the hook, was observed to be more pliant than the massive, exclusively cartilaginous medial portion. A representative cross-sectional view of this configuration can be found in Rood and Doyle (1978), Figure D. The size and location of the dilatator tubae was observed to be consistent among individuals. The anteromedial descent of the dilatator tubae formed an acute angle with the tensor veli palatini at the hamulus, collectively resulting in a distinctive triangular shape with an apex at the hamulus and a base spanning from the scaphoid fossa to the sphenoid spine (Figure 3). In its

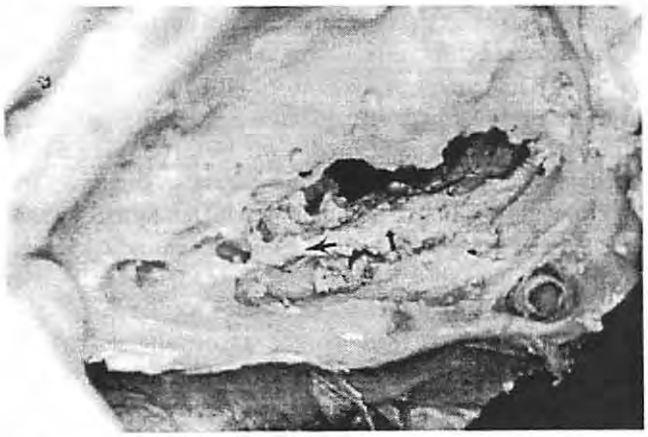


Figure 7. This photogragh, a superior view of the left tympanic cavity, depicts the tensor veli palatini (small arrows) continuing into the tensor tympani (arrow with slanted head) at the tendinous junction of the two muscles (small thick arrow). The dilatator tubae (large empty arrow) and the temporalis muscle (small empty arrow) are identified. The top right corner of the photograph represents the rostral part of the specimen, whereas the bottom left corner represents the caudal part of the specimen.

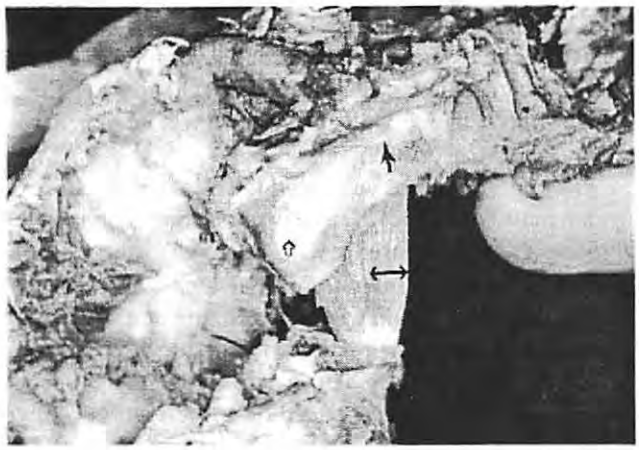


Figure 8. A medial view, depicts the width of the right dilatator tubae (double-headed arrow) and its origin from the hook of the eustachian tube (arrow with slanted head). The orifice of the eustachian tube is represented by the small empty arrow.

anteromedial descent, fibers of the dilatator tubae observed on 5 of 20 sides, in 5 specimens (1, 7-left, 10, 13, and 14-right), intermingled extensively to blend with the tensor veli palatini obscuring the separation, and therefore, the distinction between the dilatator tubae and the tensor veli palatini (Figure 10). As observed on 13 of 20 sides in 11 specimens (2, 3, 4, 5, 8, 9, 11, 12-bilateral, 14-left, 15, and 16-bilateral), the dilatator tubae, despite a connective tissue alliance and intermingling of some muscle fibers with those of the tensor veli palatini, remained distinct (Figures 4 and 11). In addition, the dilatator tubae appeared deficient on 2 of 20 sides in 2 specimens (6 and 7-right). In specimen 6, although tendinous fibers were present, the muscular bulk was virtually absent.

Figure 9. A medial view, depicts the left dilatator tubae overlapping the left tensor veli palatini. The hook of the eustachian tube is represented by the thin arrow. Fibers of the tensor veli palatini originating from the scaphoid fossa (arrow with slanted head) are shown. The width of the dilatator tubae is represented by the double-headed arrow.

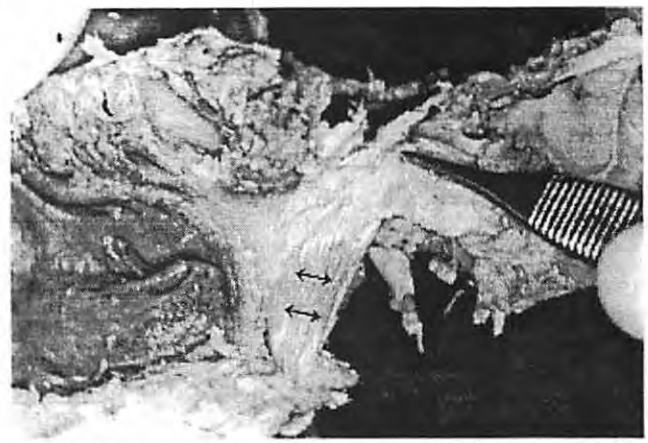


Figure 10. A medial view of the left tensor veli palatini and the left dilatator tubae, depicts extensive intermingling of their fibers (double-headed arrows), making them indistinguishable.

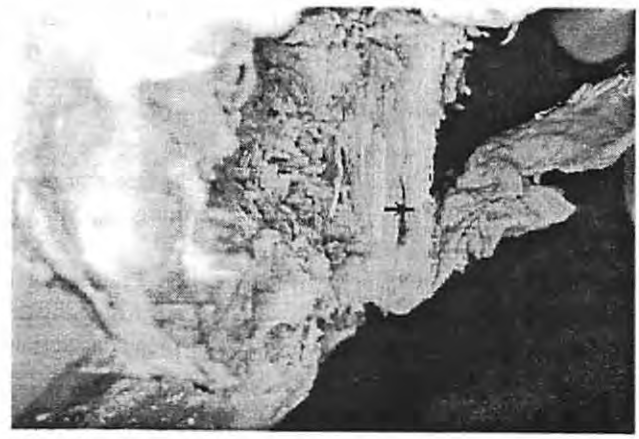


Figure 11. A lateral view of a deficient left dilatator tubae, depicts the absence of muscle bulk while tendinous fibers persist in the absence of muscle (arrow with slanted head).

Although the dilatator tubae was not observed to insert onto the hamulus (Figures 4 and 11), diffuse strands of fascia adjoined it to the hamulus, but did so without hampering its mobility. Also at this location, adipose tissue infiltrated and lubricated this diffuse mesh of fascia, which may be confused for a bursa. However, a true bursa was not identified in any specimens.

Upon rounding the middle one-third of the hamulus, the dilatator tubae became tendinous. Tendinous fibers from the tensor veli palatini fanned out posteriorly and intermingled with the tendon of the dilatator tubae which rested above the tensor veli palatini as the dilatator tubae rounded the hamulus, making the two tendons indistinguishable at this point.

Lastly, a portion of the superior pharyngeal constrictor muscle was observed to originate from the posterior one-third of the hamulus across specimens. This portion of the superior constrictor muscle may limit the potential lateral displacement of the dilatator tubae muscle.

Discussion

Tensor Veli Palatini

The tensor veli palatini has five osseus relations. The osseus origins of the tensor veli palatini observed in this study, from the scaphoid fossa and sphenoid spine, are consistent with the description of previous authors (Urbantschitsch, 1884; Simkins, 1943; Graves & Edwards, 1944; Ross, 1971; Proctor, 1973; Rood, 1973; Dickson, 1975; Rood and Doyle, 1978). Ross (1971), Rood (1973), and Rood & Doyle (1978), observed the tensor veli palatini to insert onto the medial pterygoid plate. However, such an observation could not be supported by the results of this gross anatomic study. Although an insertion of muscle fibers onto the medial pterygoid plate was absent, a firm attachment between the tensor veli palatini and the medial pterygoid lamina, mediated by a fascial sheath, was observed on 18 of 20 sides. In either case, the presence or absence of this additional insertion would neither hinder nor augment the function of the tensor veli palatini because it is anchored by its established osseus relations.

The stable osseus attachment sites for the tensor veli palatini could provide an anchor for the free margin of the muscle that is adjacent to and intermingles with the dilatator tubae muscle. In this fashion, the tensor veli palatini muscle could offer a stiffness gradient against which the dilatator tubae muscle could exert its force in opening the eustachian tube. The greater the stiffness, upon contraction, of the tensor veli palatini, the more effective its anchoring function would be in assisting the dilatator tubae muscle in opening the eustachian tube. Thus, if the tensor veli palatini muscle is slackened for a particular swallow, the eustachian tube may not open for that swallow event. This would be consistent with the results of Leider et al. (1993) who reported that the eustachian tube does not open during every swallow event.

The tendon of the tensor veli palatini, ascending from the hamulus medially and inserting onto the hard palate to form the palatine aponeurosis, was observed to be a taut, durable, and implacable structure across specimens. The inferolateral position of the hamulus with relation to the hard palate, which was observed to be consistent across specimens, lengthens the distance that the tendon of the tensor veli palatini must travel to its insertion. Based on this hamular position, the tendon of the tensor veli palatini must travel inferiorly as it rounds and inserts onto the hamulus and then must travel superiorly to insert onto the posterior border of the hard palate and medially to fan out along the velar midline. Therefore, if the hamulus were broken or if a hamulotomy were performed, the tendon of the tensor veli palatini would become slackened, resulting in a slack velum. In the past, cleft palate surgeons would break the hamulus in order to reduce the tension of palatal repair, though this procedure is not practiced at present (Millard, 1980).

Although not previously reported, a few sparse muscle fibers from the tensor veli palatini were observed to persist with the tendon of the tensor veli palatini and insert onto the hard palate. However, it is unlikely that these muscle fibers significantly influence the function of the tendon of the tensor veli palatini.

For the tensor veli palatini to tense the velum, its tendon must be displaced laterally upon contraction. In the absence of a hamular insertion, the tensor veli palatini would possess a biomechanical mechanism which would enable it to tense the velum. In this case, upon contraction, shortening of the tensor veli palatini would pull its corresponding tendon by drawing it laterally toward the hamulus, thereby tensing the velum. As described by several authors previously, this biomechanical mechanism may be likened to a pulley, the tensor veli palatini and its corresponding tendon being the rope, the hamulus being the rotor, and the hard palate and the anterior velum being the load. This biomechanical mechanism is considered a highly advantageous class I lever system with a load nearest the fulcrum or rotor, in this case, the hamulus. If, in this pulley system, the rope were anchored to the rotor, the pulley could not function and would therefore be unable to displace the load.

In agreement with Weber-Liel (as reported by Urbantschitsch, 1884), Rich (1920a), Simkins (1943), Ross (1971), the tensor veli palatini was observed to insert firmly onto the hamulus in all specimens. Though not previously reported, this insertion was limited to the anterior one-third of the hamulus. Because the tensor veli palatini and its tendon are attached firmly to the hamulus, contraction of the tensor veli palatini could not displace its tendon laterally to tense the velum. Therefore, the tensor veli palatini lacks the biomechanical capability required for tensing the velum and in full agreement with Rich (1920a), Rood & Doyle (1978),

Honjo et al. (1979) and Kuehn (1990), does not tense the velum upon contraction based on its anatomic constraints.

Although contraction of the tensor veli palatini does not result in tensing the velum, the influence that the tendon of the tensor veli palatini imparts to the velum must not be dismissed, especially in palatal reconstruction. The tendon of the tensor veli palatini, which comprises the entire anterior portion of the palatine aponeurosis, bridges the junction of the hard palate and velum. In agreement with Kuehn (1990), this bridge may buffer the stress of continuous velar elevation and depression. According to Kuehn, a direct interface of soft tissue (muscle, gland, mucosa, and adipose) with the hard palate cannot tolerate the abrasion sustained from velar activity. For this reason a taut, durable, and implacable tendon is required that provides a stiff integrity to the anterior portion of the velum. The few muscle fibers that persist may add to the stiffness of this region upon contraction even if they function isometrically.

Contraction of the tensor veli palatini may influence the function of the tensor tympani. With regard to their latency response times, Kamerer (1978) reported the tensor veli palatini and the tensor tympani to act in similar fashion when they contract, as they do in swallowing. Based on the results of this study, it appears that contraction of the tensor veli palatini may serve to anchor the tensor tympani during activities such as vocalization, chewing, and swallowing.

Dilatator Tubae

The dilatator tubae, but not the tensor veli palatini, was observed to attach to the lateral wall of the eustachian tube in all specimens of this study. The entire width of the dilatator tubae was attached directly to the hook of the eustachian tube in all specimens. This observation is consistent with Terracol et al. (1944), Ross (1971), Rood (1973), and Rood & Doyle (1978). Therefore, in agreement with Terracol et al. (1944), Ross (1971), Gray (1973) and Rood & Doyle (1978), the dilatator tubae is accepted as the tubal dilator.

Much controversy surrounds the attachments of the dilatator tubae muscle. Ross (1971), reporting on the tensor veli palatini but presumably describing the dilatator tubae, observed the dilatator tubae to attach to the fibrous and cartilaginous portion of the eustachian tube. However, Terracol et al. (1944) and Rood & Doyle (1978) contend that the dilatator tubae attaches exclusively to the membranous eustachian tube. This point is of great relevance to the function of the dilatator tubae. Because membranous tissue is generally more pliant than cartilaginous tissue, the dilatator tubae could exert less force upon contraction and still achieve tubal dilation if, in fact, it attaches exclusively to the membranous tissue.

In this gross anatomic study, the membranous portion of the eustachian tube could not be accurately distinguished from the cartilaginous portion. A thorough histologic analysis is required to determine this issue. Although this study may not provide a definitive answer to this question, the hook of the eustachian tube, to which the dilatator tubae attaches, was observed to be pliant across specimens. From this, it can be inferred that the dilatator tubae attaches primarily to the membranous portion of the eustachian tube, but may not be limited to this area. Also, the hook may be more pliant than the thickened cartilaginous region and a hinge-type arrangement might exist at the location where the hook attaches to the thicker cartilaginous region.

The intermingling of dilatator tubae muscle fibers with the tensor veli palatini was observed to be extensive on 5 of 20 sides in 5 specimens. In the other specimens, however, this intermingling was observed to be more subtle and mediated by connective tissue, as reported by Rood and Doyle (1978). Although the fibers of the dilatator tubae intermingle with the tensor veli palatini, this relation does not immobilize the dilatator tubae. Rather, it secures the anterior border of the dilatator tubae. In addition, a portion of the superior pharyngeal constrictor muscle, originating from the posterior one-third of the hamulus, prevents the dilatator tubae from being displaced laterally along the hamulus (Figure 3). The body of the dilatator tubae was observed to be less taut than the body of the tensor veli palatini, across specimens. Due to a slightly relaxed nature and in the absence of a hamular insertion, the dilatator tubae may be translated from its proper location. As a result, to ensure proper function, its potential anterior and posterior displacement must be restricted, which could be accomplished by intermingling with tensor veli palatini fibers and the curbing effect provided by the superior constrictor muscle.

From its superior attachment, the dilatator tubae travels anteriorly, dives inferiorly, and courses laterally as it rounds the middle one-third of the hamulus. Based on the specimens studied, it was determined that the dilatator tubae did not insert onto the hamulus, upon rounding this structure. Thus, the middle one-third, but not the anterior one-third of the fibers rounding the hamulus are rather freely mobile. This is an important finding because there has been considerable confusion about whether the tendon of the "tensor veli palatini" muscle (in actuality, tendons from both the tensor veli palatini and dilatator tubae) moves freely around the hamulus. Our results demonstrate that if one pulls on the anterior one-third, considerable resistance will be met because of the firm attachment. However, pulling on the middle one-third will lead to rather easy movement because the fibers in that region are attached to the hamulus only by very loose fascia.

Contrary to the statements of Ross (1971) and Rood (1973), it was found that both the dilatator tubae and the tensor veli palatini become tendinous and blend just before rounding the hamulus such that two individual tendons were not identified at the hamulus, as reported by Rood and Doyle (1978).

Hamular-Tendinous Relation

Although previous authors, (Gray, 1973; Proctor, 1973; Salins & Bloxham, 1989; Kronman et al., 1991), have reported a bursa to exist between the tensor veli palatini and the hamulus, no such structure was observed in this study. This confusion may have been created by Gray (1859) who described a synovial membrane located between the tensor veli palatini and the hamulus. Subsequent investigators may have associated a bursa with this structure.

As reported in the results, adipose tissue infiltrated the diffuse mesh of fascia located between the dilatator tubae and the hamulus. It appears that this adipose tissue lubricates the tendon of the dilatator tubae as it rounds the hamulus, thereby decreasing the friction on this tendon. Decreasing the friction on the tendon of the dilatator tubae is advantageous for two reasons: 1) it facilitates the medial displacement of the tendon of the dilatator tubae by providing a slick interface between muscle and bone, thus requiring less force to achieve tubal dilation and 2) it protects the tendon from injury due to abrasion.

Tensor Tympani

Rood and Doyle (1978) reported the tensor veli palatini to originate from the tensor tympani. Although this continuity was observed in all three specimens in which the tensor tympani was studied, it is improper to designate a muscular origin to a muscle. In general, muscles originate from bone or cartilage, not from muscle. Therefore, one may argue that the tensor veli palatini may have an additional osseus origini, the manubrium of the malleus. This contention would challenge the integrity of the tensor tympani as being separate from the tensor veli palatini. However, the tensor tympani and the tensor veli palatini have been reported as e muscles which fuse during development to become continuous. Wolff (1934) reported that in a 3-month-old fetus the tensor tympani is neither continuous with the tensor veli palatini nor is attached to the manubrium of the malleus.

In each of the three specimens of the current study, a short tendinous band joined the tensor tympani and the tensor veli palatini as reported by others (Proctor, 1973; Holborow, 1975; Rood & Doyle, 1978). Given this rather firm connection between the tensor tympani and the tensor veli palatini, it appears that the connection could serve as an anchor point, thus the origin, for effecting inward movement of the malleus. Clearly, however, the anchor point would be more stable if the tensor veli palatini muscle were contracting simultaneously.

Future Research

Although a number of investigators have studied the paratubal musculature, few have studied the dilatator tubae muscle. The biomechanical models of middle ear aeration which have been proposed are generally based on animal models which have different paratubal anatomy from that of the human as reported (Rood & Doyle, 1978). Furthermore, none of the animals that have been studied have a dilatator tubae muscle that is separate from a tensor veli palatini muscle. Establishing an accurate biomechanical model of middle ear aeration hinges on precise anatomic descriptions of the paratubal musculature, especially the human dilatator tubae muscle. In addition, the complex relation among the tensor veli palatini, the dilatator tubae, and the velar muscles, especially the palatopharyngeus and the palatoglossus require detailed investigation to sort out their mutual function in tensing the velum during swallowing and concomitantly opening the eustachian tube.

Acknowledgments

This research was supported in part by PHS Research Grants DC-00976 from the National Institute on Deafness and Other Communication Disorders and DE-10436 from the National Institute of Dental Research.

References

BATEMAN HE, MASON RM. Applied Anatomy and Physiology of the Speech and Hearing Mechanism. Springfield, IL: Charles C. Thomas, 1984.

BLAKEWAY, H. Investigations in the anatomy of the palate. Journal of Anatomy and Physiology 1913; 48:409-416.

DICKSON DR. Anatomy of the normal velopharyngeal mechanism. Clinics in Plastic Surgery 1975; 2:235-248.

DICKSON DR, MAUE WM. Human Vocal Anatomy. Springfield, Illinois: Charles C. Thomas, 1970.

DICKSON DR, MAUE-DICKSON WM. Anatomical and Physiological Bases of Speech. Boston: Little, Brown and Company, 1982.

DOYLE WJ, CASSELBRANDT ML, SWARTS JD, BLUESTONE CD. Observations on the role for the tensor veli palatini muscle in intrinsic palatal function. Cleft Palate J 1990; 27: 317-318.

GRAVES G, EDWARDS L. The eustachian tube, review of its descriptive, microscopic, topographic, and clinical anatomy. Archives of Otolaryngology. 1944;63:359-397.

GRAY H. Anatomy, Descriptive and Surgical. Philadelphia: Blanchard and Lea, 1859.

GRAY H. Anatomy of the Human Body, 29th American Ed., C.M. Goss, Ed. Philadelphia: Lea & Febiger,1973.

HAIRSTON LE, SAUERLAND EK. Electromyography of the human palate: Discharge patterns of the levator veli palatini and tensor veli palatini. Electromyography and Clinical Neurophysiology 1981; 21:287-297.

HOLBOROW C. Eustachian tubal function: Changes throughout childhood and neuro-muscular control. J Laryngology and Otology 1975; 89:47-55.

HONJO I, OKAZAKI N, NOZOE T. Role of the tensor veli palatini muscle in movement of the soft palate. Acta Otolaryngologica 1979; 88:137-141.

KAMERER DB. Electromyographic correlation of tensor tympani and tensor veli palatini muscles in man. Laryngoscope 1978; 88:651-662.

KRONMAN JH, PADAMSEE M, NORRIS LH. Bursitis of the tensor veli palatini muscle with an osteophyte on the pterygoid hamulus. Oral Surgery, Oral Medicine, and Oral Pathology 1991; 71:420-422.

KUEHN DP. Commentary. Cleft Palate J 1990; 27:318-319.

KUEHN DP, KAHANE JC. Histologic study of the normal adult human soft palate. Cleft Palate J 1990; 27:26-34.

LEIDER J, HAMLET S, SCHWAN S. The effect of swallowing bolus and head position on eustachian tube function via sonotubometry. Otolaryngology-Head and Neck Surgery 1993; 109:66-70.

LUPIN A. The relationship of the tensor tympani and the tensor veli palatini muscles. Annals of Otol Rhinol Laryngol 1969; 78:792-796.

MCMYN JK. The anatomy of the salpingo-pharyngeus muscle. J Laryngology and Otology 1940; 55:172-186.

MILLARD RD. Cleft Craft: the evolution of its surgery, 2nd Ed. Boston: Little, Brown and Co, 1980.

MOORE KL. Clinically Oriented Anatomy. Baltimore: Williams and Wilkins, 1992.

PERLMAN AL. The neurology of swallowing. Seminars in Speech and Language 1991; 12:171-183.

POLITZER A. Text-book of the diseases of the ear for students and practicioners, 6th Ed. Philadelphia: Lea and Febiger, 1926.

POIRIER P, CHARPY A. Traite d'Anatomie Humaine, 3rd Ed. Paris: Masson et Cie, 1912.

PROCTOR B. Anatomy of the eustachian tube. Archives of Otolaryngology 1973; 97: 1-8.

QUAIN J. Elements of Anatomy, 11th Ed. London: Longmans and Green, 1990.

RICH AR. A physiologic study of the eustachian tube and its related muscles. Bulletin of the Johns Hopkins Hospitals, 1920a; 31:206-214.

RICH AR. The innervation of the tensor veli palatini and levator veli palatini muscles. Bulletin of the Johns Hopkins Hospital, 1920b; 31:305-310.

ROOD SR. The morphology of m. tensor veli palatini in the five-month human fetus. American Journal of Anatomy, 1973; 138:191-196.

ROOD SR, DOYLE WJ. Morphology of the tensor veli palatini, tensor tympani, and dilatator tubae muscles. Annals of Otolaryngology 1978; 87:202-210.

ROSS MA. Functional anatomy of the tensor palati. Archives of Otolaryngology 1971; 93:1-8.

SALINS PC, BLOXHAM GP. Bursitis: A factor in the differential diagnosis of orofacial neuralgias and myofascial pain dysfunction syndrome. Oral Surgery, Oral Medicine, and Oral Pathology 1989; 68:154-157.

SEIF S, DELLON AL. Anatomic relationship between the human levator and tensor veli palatini and the eustachian tube. Cleft Palate J 1978; 15:329-336.

SIMKINS CS. Functional anatomy of the eustachian tube. Archives of Otolaryngology 1943; 33:476-484.

TERRACOL J, CORONE A, GUERRIER Y. La Trompe d'Eustache. Paris: Masson et Cie, 1949.

URBANTSCHITSCH V. Lehrbuch der Ohrenheilkunde. Wein u. Leipzig: Urban and Schwartzenberg, 1884.

WOLFF D. The microscopic anatomy of the eustachian tube. Annals of Otol Rhinol Laryngol 1934; 43:483-494.

NCVS Status and Progress Report - 11
May 1997, 107-118

Velopharyngeal Closure Force and Levator Veli Palatini Activation Levels in Relation to Varying Phonetic Contexts

David P. Kuehn, Ph.D.
Speech and Hearing Sciences, The University of Illinois at Urbana-Champaign
Jerald B. Moon, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa

Abstract

The purpose of this study was to measure velopharyngeal closure force in relation to varying phonetic contexts for normal adult male and female subjects. Levator veli palatini muscle activity also was measured. Place and manner of articulation, voicing, and the affects of consonant sequencing were studied in different vowel contexts. When the data were grouped within sexes, no differences were found in absolute values of velopharyngeal closure force for the male versus female subjects. As expected, nonnasal consonants were produced with greater velopharyngeal closure force than nasal consonants. High vowels were produced with greater closure force than low vowels. Closure force was greater for voiceless than for voiced consonants but only for the males and only within /i/ and /u/ contexts. The lingua-dorsal consonant was associated with greater closure force than the lingua-apical consonant but only for the males and only in the high-back vowel environment. Significant differences in closure force were not found between fricatives or stops. A tendency for greater closure force for the fricative consonant was observed when the fricative followed rather than preceded the nasal consonant. Vowel identity had an affect on closure force during consonant production in the males in that closure force was greater for /s/ and /n/ in high versus low vowel contexts. Males exhibited a larger number of significant differences in closure force than the females which could be related to a more constant and limited locus of velar positioning for the females. The results suggest that velopharyngeal closure force is not controlled by a single muscle, the levator veli palatini, but that other muscles and mechanical factors are likely contributors.

The velopharyngeal mechanism functions as a valve during speech to direct the airstream through the oral cavity, the nasal passage, or both. Linguistically, because phonemes are generally characterized as being either [+nasal] or [-nasal], it is tempting to ascribe a binary function for the control of the velopharyngeal mechanism (Moll & Daniloff, 1971; Moll & Shriner, 1967). Unfortunately, for the sake of simplicity, it appears that function of the velopharyngeal mechanism may be more complicated than that implied in relation to binary control. For example, it is likely that velopharyngeal control for vowels and probably also for sonorants, such as /w/ and /l/, involves a specified position or positions different from that of either nasal consonants or nonnasal obstruent consonants (Bell-Berti, Baer, Harris, & Nimi, 1979; Bell-Berti & Krakow, 1991; Benguerel, Hirose, Sawashima, & Ushijima, 1977a; Karnell, Linville, & Edwards, 1988; Kent, Carney, & Severeid, 1974; Kuehn, 1976; Vaissiere, 1988). Velar positional variation also appears to be language-specific (Clumeck, 1976).

Consistent with velar positional variation, it has been demonstrated that activity for the major muscle of velar elevation, the levator veli palatini, not only varies for nasal versus nonnasal sounds, as expected, but also varies within the class of nonnasal speech sounds in normal individuals (Bell-Berti, 1976; Benguerel, Hirose, Sawashima, & Ushijima, 1977b; Fritzell, 1969; Kuehn, Folkins, & Cutting, 1982; Kuehn & Moon, 1994; Lubker, Fritzell, & Lindquist, 1970; Moon, Smith, Folkins, Lemke, & Gartlan, 1994; Seaver & Kuehn, 1980) and in subjects with cleft palate as well (Kuehn & Moon, 1995). Recent comprehensive reviews have dealt in greater detail with influences in velopharyngeal activity at the intrasegmental (Bell-Berti, 1993) as well cross-segmental (Krakow, 1993) levels, the

latter including such factors as linguistic stress variation and speaking rate changes.

The focus of many studies on velar positioning has been on defining relations between velar height and the sound segment being produced. Lubker et al. (1970) suggested that velar position control should not be modeled as a one-to-one relation between phoneme and velopharyngeal muscle activity. They proposed a model in which levator veli palatini muscle activity during speech might be controlled in relation to certain phoneme categories. Voiceless versus voiced consonants and high versus low vowels were offered as examples. Lubker et al. suggested further, however, that variability in muscle activity within a category might be seen as a function of phonetic context. For example, they found that levator muscle activity was always much greater for nonnasal consonants following nasal rather than other nonnasal consonants. Based on such results, Lubker et al. posited that velopharyngeal muscle activation levels depend on momentary and variable starting and ending states that are determined in relation to phonetic context and that are adjusted accordingly to prevent excessive nasal coupling.

The conclusion offered by Lubker et al. (1970) implies that there may exist a velopharyngeal target specification for different phoneme categories, subject to contextual variation. However, it is not clear from their data, or any other available data, what the nature of those targets might be. Several candidates are possible, such as velar position, air pressure or airflow, acoustic or perceptual goals, or velopharyngeal closure force. From a neural control perspective, closure force is especially appealing in that activity in the primary motor cortex has been shown to be closely related to certain parameters of force output (Evarts, 1968). Stein (1982) has argued that force is an important control parameter in motor behavior.

Barlow and Bradford (1992) discussed force control, in relation to speech, for the lips, tongue, and jaw. They defined "resultant" force, distinct from isometric force, as the "geometric sum of constituent forces resulting from the activation of separate force generators" (p.8). The force generators are individual muscles working synergistically in a confined system. With regard to the velopharyngeal mechanism, the resultant force would be that of contact between the velum and pharyngeal wall. The concept of resultant force, in relation to the velopharynx, is important because it has been shown that although the levator veli palatini muscle is the primary muscle of velopharyngeal closure, it is not the sole determiner. Muscles in addition to the levator participate in a cooperative fashion to position the velum (Kuehn et al., 1982; Moon et al., 1994; Seaver & Kuehn, 1980).

Although investigators in previous studies have alluded to the necessity for tight velopharyngeal closure, the force of velar to pharyngeal wall contact has yet to be studied systematically. In fact, very few investigators have measured velopharyngeal closure force during speech (Cohn, Fisher, McWilliams, Seth, Ferketic, Beery, & Vallino, 1986; Goto, 1977; Moon, Kuehn, & Huisman, 1994). Airtight closure during nonnasal speech is often assumed and airflow reportedly rarely occurs in nonnasal contexts (Hoit, Watson, Hixon, McMahon, & Johnson, 1994; Lubker & Moll, 1965; Thompson & Hixon, 1979). However, significant nasal airflow can occur even during normally produced obstruent consonants within low vowel or nasal consonant environments (Andreassen, Smith, & Guyette, 1992). Therefore, it is likely that velopharyngeal closure force would vary in those situations as well.

It is important to gain a greater understanding of velopharyngeal closure force for at least two reasons. First, even if velopharyngeal closure is achieved, the magnitude of velopharyngeal closure force may need to vary to offset the opposing mechanical and aerodynamic forces generated in the oral cavity. In a similar vein to that suggested by Lubker et al. (1970), closure force might be adjusted to meet the momentary demands of the system, which may be determined in part by the category of the phoneme being produced, to avoid excessive nasal coupling. It has been shown that the velum may continue its upward movement past the point at which velum-to-pharyngeal wall contact is made (Kent et al., 1974; Kuehn, 1976). This may be a reflection of the need to exert greater force of closure in certain contexts. A second reason for a better understanding of velopharyngeal closure force is that force levels might be very different for normal mechanisms compared to anatomically or physiologically impaired mechanisms involving, for example, cleft palate or a neuromuscular disorder. Kuehn and Moon (1995) found that subjects with cleft palate exhibit higher levels of levator muscle activity for speech in relation to the muscle's total operating range, compared to subjects with a normal mechanism. However, Kuehn and Moon did not measure velopharyngeal closure force in those subjects. If individuals with cleft palate use levels of closure force that are near or exceed the threshold of fatigue for their muscles, this may lead to excessive nasal coupling and hypernasality. Therefore, such higher levels of force are likely avoided in those individuals.

The purpose of this investigation was to obtain measures of velopharyngeal closure force and concomitant levator veli palatini muscle activity in relation to varying phonetic contexts for a group of subjects with normal velopharyngeal mechanisms. It is expected that the information derived will be useful in further delineating motor control mechanisms for speech, especially as it pertains to velopharyngeal functioning, and will be useful as a basis against which to compare force measures obtained in subjects with impaired velopharyngeal mechanisms.

Method

Subjects

Seven male and seven female college-age persons with no reported speech, language, or hearing deficits served as subjects. Normal speech production, including normal oral-nasal resonance balance, was verified at the time of testing by the investigators.

Speech Sample

A number of sustained vowels (V), vowel-nonnasal consonant-vowel (VCV), vowel-nasal consonant-vowel (VNV), vowel-nonnasal consonant-nasal consonant-vowel (VCNV), and vowel-nasal consonant-nonnasal consonant-vowel (VNCV) utterances were recorded. The speech sample used is shown in Table 1. The speech sample was designed to allow comparison among high, low, front, and back vowels and within the consonant categories of voicing (for fricatives), place (lingua-alveolar versus lingua-palatal plosives), and manner (plosives, fricatives, nasals). Ten productions of each utterance were recorded. Order of production of the speech samples was randomized across subjects.

Force Bulb

Velopharyngeal closure force was transduced using a silastic bulb device described in detail by Moon et al. (1994). The bulb is teardrop-shaped, flattened anteriorly and posteriorly, and measures 5 mm front-to-back and 10 mm side-to-side. Although the bulb is larger than a standard endoscope, it is rather easy to insert transnasally because it is soft, pliable, and collapsible upon insertion. The advantages of the size and shape of the bulb have been discussed previously (Moon et al., 1994; Moon, Kuehn, & Huisman, 1995). Most subjects experience little discomfort with insertion or placement of the bulb. The bulb is attached to a silastic tube with a 3 mm outside diameter and the tube is connected to a Honeywell Microswitch (model 162PC01D) transducer. The transducer output was amplified using a Biocommunications Electronics (model 205) amplifier.

Table 1.
Speech Sample. V=vowel, C-nonnasal consonant, N=nasal consonant

V	VCV	VNV	VCNV	VNCV
i	ɑtɑ	ɑnɑ	ɑsnɑ	ɑnsɑ
æ	iti	ini	isni	insi
ɑ	utu	unu		
u	ɑsɑ			
	isi			
	usu			
	ɑzɑ			
	izi			
	uzu			
	ɑkɑ			
	iki			
	uku			

The bulb was calibrated externally from the subjects by using applications of gram weights and, therefore, all force measures will be reported in grams. Although the general calibration procedure has been described previously (Moon et al., 1994), additional calibration measures were obtained for this study. These additional measures, described below, were motivated by 1) a desire to compare force measures between groups, particularly males versus females, and 2) to determine the possible effects of intraoral air pressure on velopharyngeal closure values.

To simulate amount of velar contact with the force bulb in an inferior-superior dimension, several thin wooden beams were constructed to lay across the bulb at its most sensitive location. Within subjects, the bulb is positioned at the location producing maximum output as described below. Moon et al. (1994) showed that subjects exhibited velopharyngeal force values that hovered around 50 grams for vowels, and this was chosen as a convenient weight for calibration. Thus, a 50-gram weight was hung from each wooden beam which, in turn, compressed the bulb. The wooden beams measured 3, 6, 9, 12, and 15 mm in their "vertical" dimension and all measured 13 mm in their "left-to-right" dimension. Output voltages for the "vertical" dimensions were as follows: 1.44, 1.42, 1.29, 1.24, and 1.14 V for the 3-15 mm beam measures respectively. Thus, as the surface area of contact increased, the output voltage decreased. We attribute this to the fact that the perimeter of the bulb is somewhat stiffer than the face of the bulb and as the surface area of contact increases, the perimeter offers relatively more resistance to deformation and output voltage decreases. In situ, then, as the area of velar contact with the bulb increases, the output voltage (force measure calibrated in grams) would be expected to decrease.

In their groups of normal young adult subjects, McKerns and Bzoch (1970) found a mean vertical contact between the velum and posterior pharyngeal wall of 5.7 mm for males and 9.5 mm for females in lateral-view x-ray studies. These values are close to the 6-mm and 9-mm wooden beams which produced output voltages of 1.42 and 1.29, differing from each other by only about 10%. Therefore, we decided to compare force values for subject groups statistically knowing that the margin of absolute error might be on the order of 10% for such comparisons. It may be possible in future studies to control for absolute difference between subjects more precisely.

Given that the velum is incompressible but compliant, it is possible that any external force applied to the velum, such as heightened intraoral air pressure, might be transmitted through the velar tissue and sensed by the adjacent force bulb. To test for this possibility, we embedded the force bulb in a portion of store-bought pork liver to simulate the physical characteristics of the velum. The liver portion was trimmed to be consistent with the size of the adult velum. The liver with embedded force bulb was then placed in an air-tight cannister with a plunger and piston at the opposite end to vary the pressure within the cannister. At pressures that simulated intraoral air pressures for conversational (below 10 cm H_2O) or loud (10-20 cm H_2O) speech, the force bulb did follow pressure variations that occurred in those ranges. Thus, it can be concluded that if the oral cavity is closed, or nearly so, as in the case of obstruent consonants, the force bulb is capable of sensing not only direct pressure applied to it by virtue of velar contact provided by its own musculature, but also intraoral pressure variation that it senses as a force transmitted through the substance of the velum.

Electromyography

Levator veli palatini muscle activity was recorded using stainless steel electrodes, 110 micrometers in diameter. The EMG signals were amplified using Biocommunications Electronics preamplifiers (model 301) and amplifiers (model 205). The audio signal from a dynamic microphone was amplified using a Nakamichi preamplifier and Tascam tape recorder (model 22-4) amplifier.

Placement Procedure

A light spray of 4% lidocaine topical anesthetic was applied to the more patent nostril for force bulb insertion and to the oral cavity for EMG electrode insertion. A negative pressure was exerted on the bulb and tube assembly to decrease the size of the bulb initially and aid in bulb insertion. The bulb and attached tube were then slid through the nasal cavity. The tube, exiting the nostril, was attached to the transducer and the output of the transducer was monitored on an oscilloscope (Tektronix model 5111A). Force varied depending on the vertical position of the bulb in the velopharynx. Gross placement could be gauged initially by observing the lower portion of the bulb through the oral cavity and then elevating the bulb so that most of it was out of view. Final adjustments then were made by maneuvering the tube in and out of the nasal passage (to move the bulb up and down in the velopharynx) while the subject produced a series of /s/ sounds exerting oscillating and incrementally changing force on the bulb. The peak force was noted on the oscilloscope and, at that level, the tube was taped to the nose to secure the bulb in place.

The hooked-wire electrodes were inserted perorally, using disposable hypodermic needles, within the dimple of the velum at an angle following the course of the levator muscle, that is, in a superior, posterior, and lateral direction. The two wires were placed approximately 4 mm apart and 10 mm deep into the levator muscle on the subject's right side. Placement criteria included electromyographic (EMG) activity that was observed in association with sustained /s/ production.

Data Analysis

Bulb force, EMG activity, and the audio signals were monitored on an oscilloscope (Tektronix model 5111A) and recorded on a Sony digital instrumentation recorder (model PC108M). EMG signals were full-wave rectified and low-pass filtered with a 40 ms time constant. Voice and force bulb signals were digitized at 1000 Hz sampling rate using a laboratory computer and commercially available analog-to-digital conversion software, CSpeech (Read et al., 1990). Closure force signals then were digitally low-pass filtered at 30 Hz. Data then were displayed and analyzed using custom graphics and analysis routines.

Figure 1 shows an example of rectified and smoothed levator veli palatini EMG activity and a force bulb trace for a sustained vowel. For each sustained vowel, average EMG activity and closure force were measured over a 500 ms segment of the steady state portion of the vowel (between time points "a" and "b" in Figure 1). For all other EMG and force measures, either peak (nonnasal consonants) or valley (nasal consonant) values were used as described below.

Figure 2 shows an example of measures made from a VNCV utterance. The peak EMG value (point "a" in Figure 2) was used as the measure for nonnasal consonant activity immediately following the nasal consonant. Peak force was measured (point "b" in Figure 2) as the force representing velopharyngeal closure for the nonnasal con-

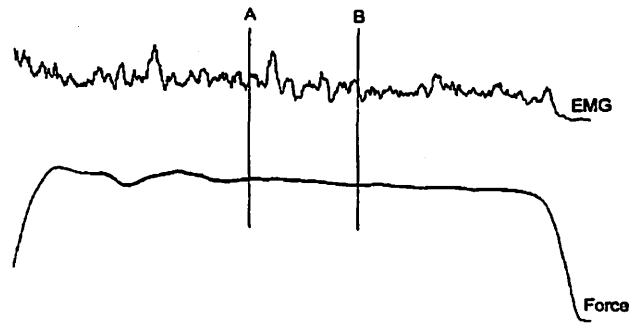


Figure 1. Example of rectified and smoothed levator veli palatini EMG activity (top) and velopharyngeal force bulb trace (bottom) for sustained /i/ vowel. Measures were averaged over a 500 ms segment; a=onset, b=offset.

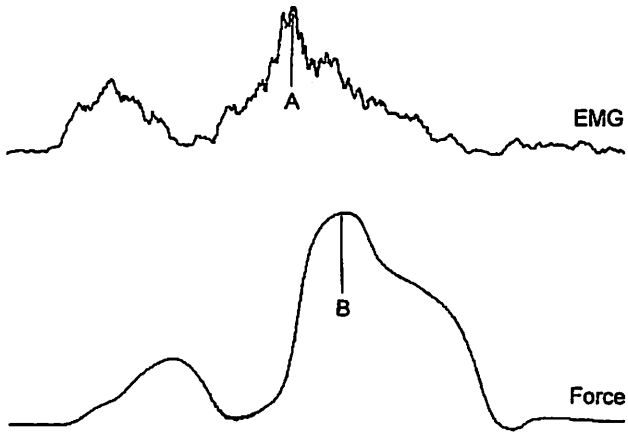


Figure 2. Example of rectified and smoothed levator veli palatini EMG activity (top) and velopharyngeal force bulb trace (bottom) for a VNCV (/insi/) utterance. Peak EMG activity and peak force for the nonnasal consonant measured at points a and b respectively.

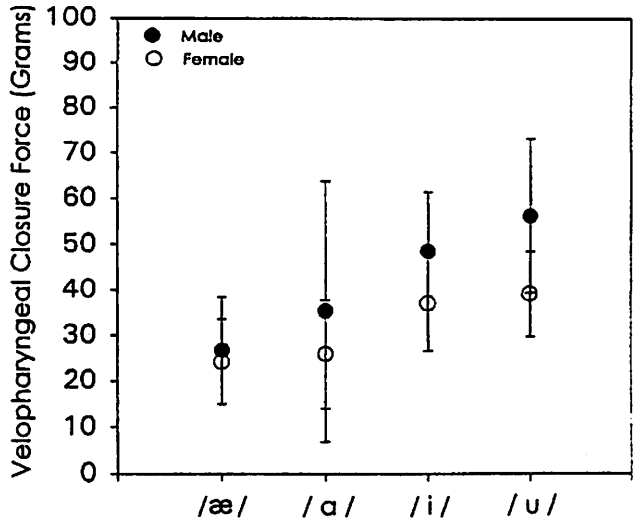


Figure 3. Velopharyngeal closure force for the four sustained vowels.

sonant immediately following the nasal consonant. In a similar fashion, for VCNV utterances, peak values for EMG and force were measured for the nonnasal consonant immediately preceding the nasal consonant. Peak EMG and force measures also were determined for the nonnasal consonant between the vowels in VCV utterances. Finally, for the VNV utterances, measures were obtained at the lowest point in the traces for EMG and force corresponding to the nasal consonant between the two vowels. These points are similar to the valleys for EMG and force preceding the peak values ("a" and "b" respectively) in Figure 2.

Levator EMG activation levels were normalized within each subject. This was accomplished by identifying the maximum peak value across the entire data set for a given subject and setting that value at 100%. All other EMG values recorded for that subject were referenced to the maximum value.

Statistical Analyses

The data were analyzed using multivariate analysis of variance (Morrison, 1976, pp. 170-193) in the framework of the profile analysis model for the two independent groups (Morrison, 1976, pp. 153-160). All hypotheses were tested using the Wilks' lambda statistic (Morrison, 1976, 222-223). The analyses were carried out using the GLM procedure of SAS (SAS, 1987). As explained above, it was deemed appropriate to compare velopharyngeal closure force measures between as well as within subjects. However, it is rarely appropriate to compare EMG measures between subjects, at least in absolute terms. Therefore, only within-subject EMG measures are presented below.

Results

Vowels

Means and standard deviations of velopharyngeal closure force for the four sustained vowels are shown in Figure 3. Although the mean force values were greater for the males compared to the females for all four vowels, these differences did not reach statistical significance beyond the 0.05 level, presumable owing to the variability for each of these productions. Within sex groups, it was found that the means differ significantly across the four vowels for both the male ($p<0.0003$) and female ($p<0.03$) subjects. Post hoc comparisons indicated the following significant differences at the pre-established level ($p<0.007$) for seven post hoc tests (individual vowel comparisons plus high /i,u/ versus low /æ, ɑ/ vowel comparisons): 1) high vowels greater than low vowels (both sex groups), 2) /u/ greater than /æ/ (both sex groups), 3) /u/ greaterthan /ɑ/ (both sex groups), and 4) /i/ greater than /æ/ (males).

Figure 4 (following page) shows the results for levator EMG across the four sustained vowels. There was not a significant difference across the vowels for either males or females. Post hoc tests further indicated lack of significant EMG differences for individual vowel comparisons and high versus low vowel comparisons.

Voicing

Means and standard deviations for closure force measures in each of the vowel contexts for voiced (/z/) versus voiceless (/s/) consonants are shown in Figure 5. When the data were grouped across vowels, no significant difference in closure force was observed for males versus females. Moreover, no significant difference in closure force was found for /s/ versus /z/. This was true for the males and females and suggests that across vowel categories there

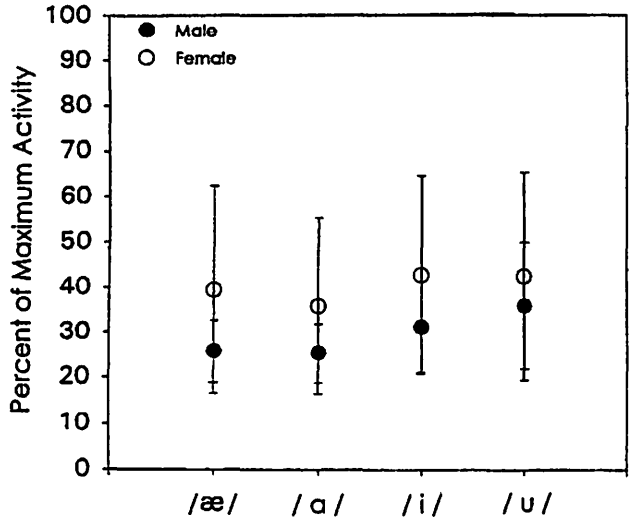


Figure 4. Levator veli palatini muscle activity for the four sustained vowels.

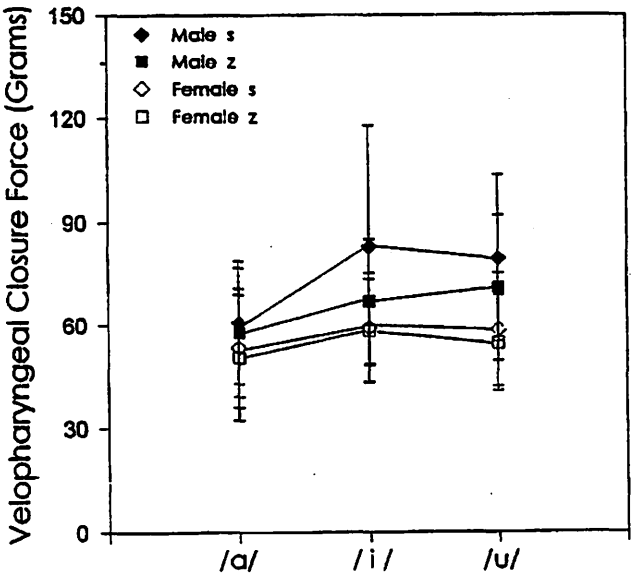


Figure 5. Velopharyngeal closure force in relation to voicing (/s/ versus /z/) across the three vowel contexts.

is no difference in velopharyngeal closure force between voiceless and voiced consonants for either sex group.

When the vowel data were separated for the males, the following differences were found: 1) closure force was significantly greater ($p<0.05$) during /s/ within /isi/ and /usu/ contexts than within the /ɑsɑ/ context, 2) closure force for /s/ in /isi/ was greater ($p<0.05$) than that for /z/ in /izi/, and 3) closure force for /s/ in /usu/ was greater ($p<0.05$) than that for /z/ in /uzu/. No closure force comparisons with vowels separated reached statistical significance at the 0.05 level for the females.

Few significant differences were found with regard to the normalized EMG comparisons (Figure 6). Specifically, for the females it was found that: 1) EMG activity for

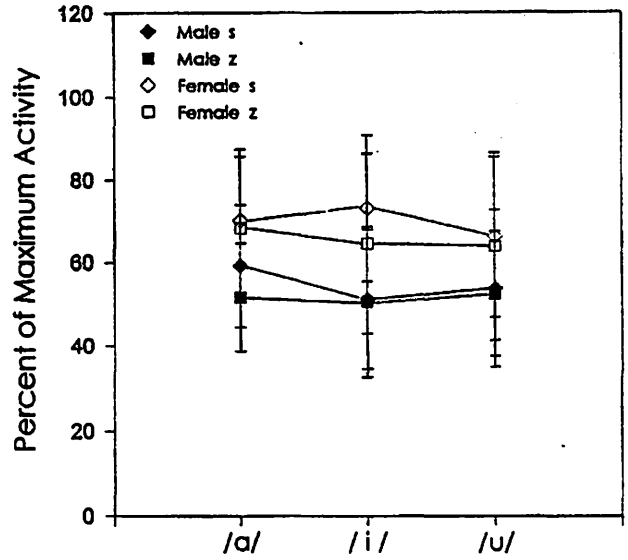


Figure 6. Levator veli palatini muscle activity in relation to voicing (/s/ versus /z/) across the three vowel contexts.

the /s/ in /isi/ is greater ($p<0.05$) than that for the /s/ in /usu/ and 2) EMG activity for the /s/ in /isi/ is greater ($p<0.05$) than that for the /z/ in /izi/. No significant differences were found for the males in EMG activity for any of the comparisons depicted in Figure 6.

Comparison of Figures 5 and 6 indicates that levator muscle activity cannot be the *sole* determiner of changes in velopharyngeal closure force because there is not an exact parallelism between the corresponding functions of these two graphs. For example, for the males, it can be seen in Figure 5 that closure force is greater for /s/ in the /isi/ and /usu/ contexts than in the /ɑsɑ/ context. However, no significant differences across these same contexts were found for EMG activity in the males (Figure 6).

Place

Means and standard deviations for closure force measures in each of the vowel contexts for dorsal (/k/) versus apical (/t/) consonants are shown in Figure 7. When the data were grouped across vowels, it was found that closure force during /k/ is significantly greater ($p<0.05$) than that during closure for /t/ for the male subjects but not the female subjects. However, /k/ was associated with greater force for the males only in comparing /uku/ to /utu/ and not in comparing /iki/ to /iti/ or /ɑkɑ/ to /ɑtɑ/. This suggests that velopharyngeal closure may be tighter for dorsal consonants than apical consonants only for adult males in a high-back vowel environment. No other significant differences in force were observed involving the place feature. Although the mean values of closure force were consistently greater for the males compared to the females (Figure 7), no significant differences were observed owing to the large variances involved.

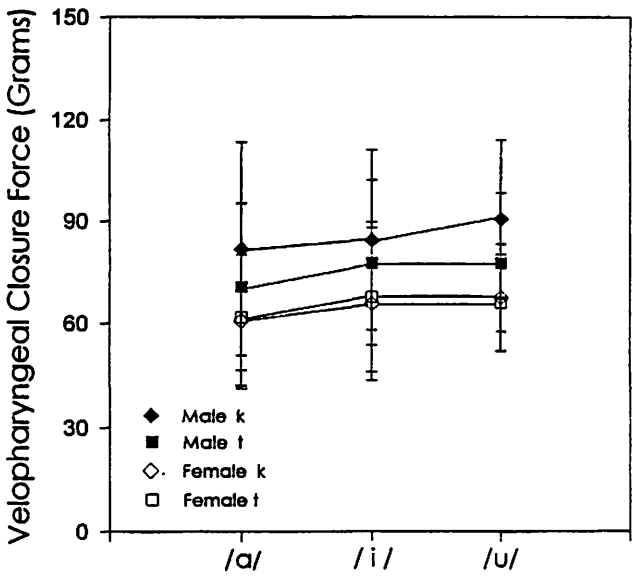


Figure 7. Velopharyngeal closure force in relation to place of articulation (/k/ versus /t/) across the three vowel contexts.

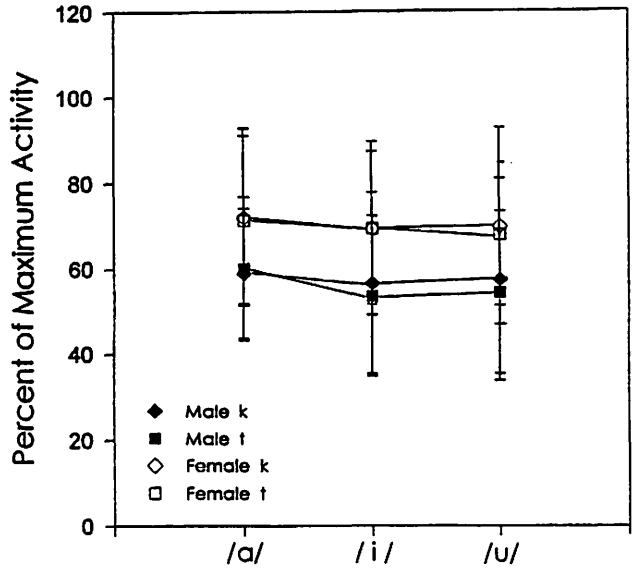


Figure 8. Levator veli palatini muscle activity in relation to place of articulation (/k/ versus /t/) across the three vowel contexts.

The EMG data for dorsal versus apical consonants are shown in Figure 8. Significant differences were not found for either the male or female subjects when the data were grouped across vowels, nor were any significant differences found when the vowel data were separated for either sex group.

Manner

Means and standard deviations for closure force measures in each of the vowel contexts for nasal (/n/), plosive (/t/), and fricative (/s/) consonants are shown in Figure 9. No significant differences were found in closure force for the male versus female subjects. When the data

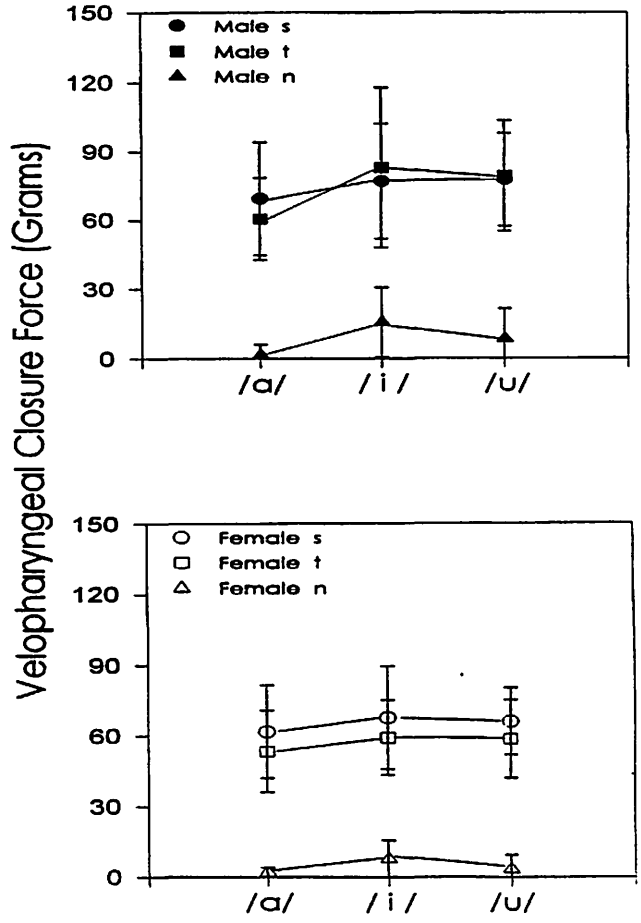


Figure 9. Velopharyngeal closure force in relation to manner of production (/s/, /t/, /n/) across the three vowel contexts.

were grouped across vowels, it was found that closure force during /n/ is significantly less ($p<0.001$) than that during closure for /t/ and /s/ for the male and female subjects. This indicates that velopharyngeal closure is tighter for plosive and fricative consonants than nasal consonants. With vowels grouped for the females, although the mean level of force of closure for /s/ was consistently greater than that for /t/ (Figure 9), these differences did not reach significance at the 0.05 level owing to the variability involved.

When the vowel data were separated, the following significant differences were found for the males: 1) closure force for /s/ was greater ($p<0.05$) in the contexts /isi/ and /usu/ than in /ɑsɑ/ and 2) closure force for /n/ was greater ($p<0.05$) in the contexts /ini/ and /unu/ than in /ɑnɑ/. Thus, although the force for /n/ in the low vowel /ɑ/ context is near zero, the force for /n/ is significantly greater than zero in the high vowel (/i,u/) contexts.

No significant differences were found for any comparison when the vowel data were separated for the females. Thus, closure force during /s,t,n/ did not vary for these consonants as the vowel context changed for the females.

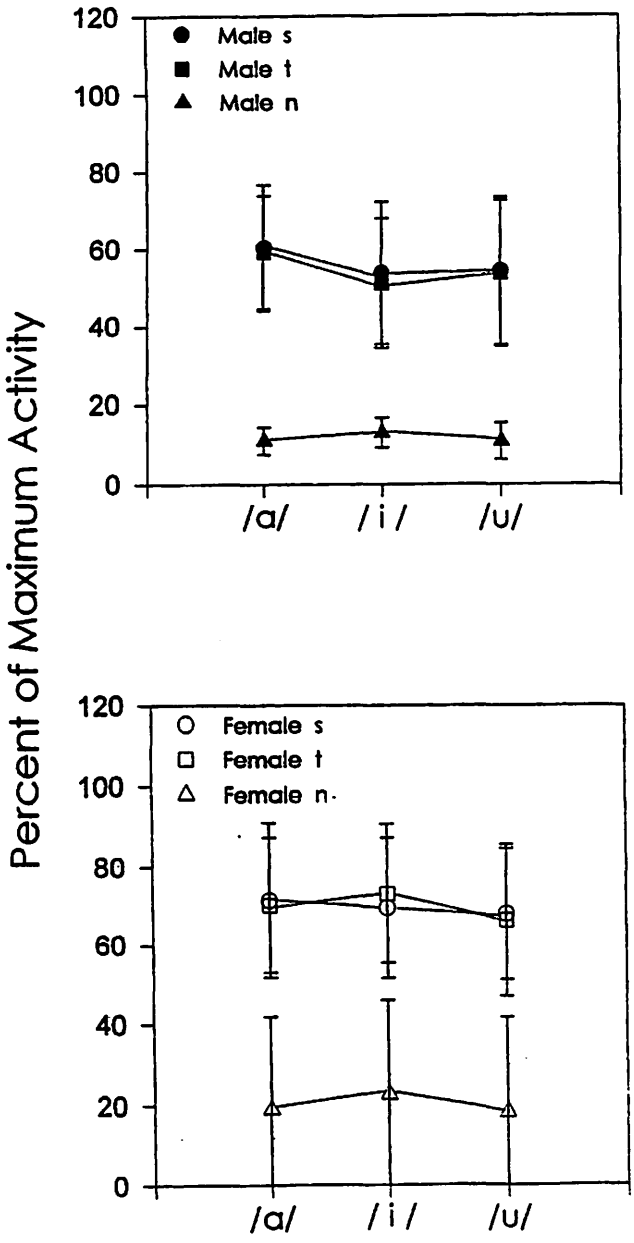


Figure 10. Levator veli palatini muscle activity in relation to manner of production (/s/, /t/, /n/) across the three vowel contexts.

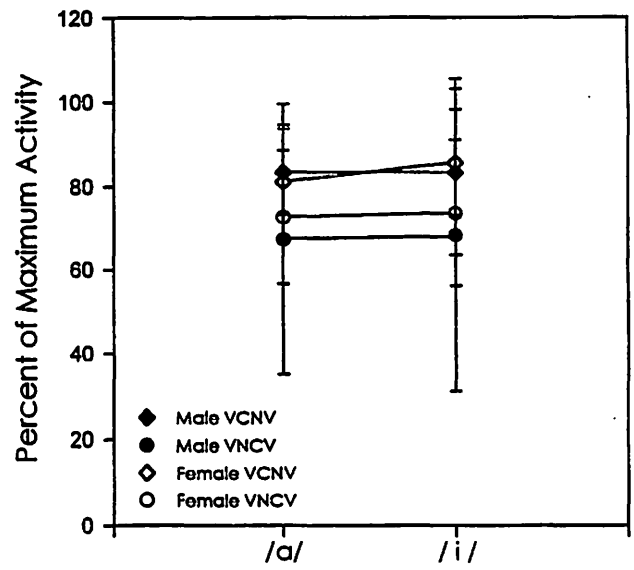


Figure 11. Velopharyngeal closure force in relation to nasal-nonnasal sequencing across the two vowel contexts.

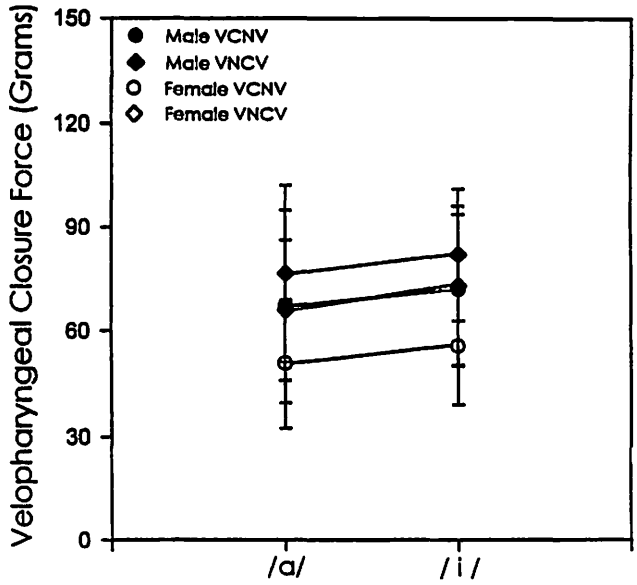


Figure 12. Levator veli palatini muscle activity in relation to nasal-nonnasal sequencing across two vowel contexts.

The EMG data for manner of production (/s,t,n/) are shown in Figure 10. When the data were grouped across vowels, it was found that levator EMG activity for /n/ is significantly less ($p<0.0001$) than that for /t/ and /s/ for the male and female subjects. Thus, levator EMG activity, as well as velopharyngeal closure force, is significantly less for nasal consonants than for the plosive or fricative consonants.

Nasal-Nonnasal Sequencing

Means and standard deviations for closure force measures in the vowel contexts /ɑ/ and /i/ for /sn/ and /ns/ sequencing are shown in Figure 11. As in previous comparisons, no significant differences in closure force were observed between the male and female subjects. When the data were grouped across the two vowels, it was found that closure force for /s/ in the context /VnsV/ is significantly greater ($p<0.01$) than that during closure for /s/ in the /VsnV/ sequence for the female subjects. This was true for the male subjects also but only for the /ɑnsɑ/ versus /ɑsnɑ/ context.

The EMG data for nasal-nonnasal sequencing are shown in Figure 12. Unlike that for the force data, there were no significant differences for the males or the females for EMG activity. Even though consistent differences in the

means can be seen for /VnsV/ sequencing versus /VsnV/ sequencing (Figure 12), no significant differences emerged owing to the variability involved.

Discussion

A major purpose of this study was to provide information about velopharyngeal closure forces in different phonetic contexts, which is largely lacking in the literature. We were motivated, in large measure, by the earlier work of Lubker and his colleagues (Lubker, 1975; Lubker et al., 1970). Many of the findings of the current study are in agreement with and support the work of Lubker and other investigators who have published electromyographic results and information regarding positioning of the velum. Thus, the force data of the current study appear to be predictable, to some degree, based on EMG and positional data. In general, the results of this study demonstrate the versatility of the normal velopharyngeal mechanism in meeting the demands imposed by varying phonetic contexts.

Vowels

The force data shown in Figure 3 are consistent with what might be expected based on many studies investigating positioning of the velum during vowel production. Aside from differences in velar position between nasal versus nonnasal consonants, differences between velar position for high versus low vowels appear to be the most robust. In general, high vowels are produced with higher velar positions than low vowels (e.g., Moll, 1962) and the results of this study indicate that velopharyngeal closure force follows that trend. This suggests that there is a relation between velar height and the degree of velopharyngeal closure force for vowel production. It is likely that the influence of tongue positioning for vowels on velopharyngeal closure extends to adjacent speech sounds as well. Thus for the male subjects, closure force was significantly less for /s/ in the /ɑsɑ/ context than in the /isi/ and /usu/ contexts (Figure 5) probably owing to the downward pull on the velum by the tongue within the /ɑ/ environment. These effects are due mainly to tongue height. Significant differences were not found in closure force between front versus back vowels suggesting an affinity within these vowel classes with regard to force control.

Unlike that for force, no significant differences were found for the levator veli palatini EMG measures across vowels (Figure 4). These results support the conclusion of Kuehn et al. (1982) and Moon et al. (1994) that muscle (or other) forces in addition to that of levator are important in velopharyngeal closure. Thus, from a motor control perspective, it is not a simple matter of specifying activation levels for a single muscle. Instead, it is likely that activation levels for several muscles must be specified for some common goal, such as force of closure. In the case of vowel production the goal may be greater force of closure for high vowels than low to satisfy acoustic and perceptual requirements involving appropriate oral-nasal impedance balances (Lubker, 1975).

Voicing

When the vowel data were separated, three significant differences were observed in the male force data (Figure 5) and two significant differences were noted in the female EMG data (Figure 6). The males exhibited some tendency to exert less force of closure during voiced than for voiceless consonants. This might allow some leakage of air during the voiced obstruents (Yanagihara & Hyde, 1966) which is compatible with the maintenance of a sufficient transglottal air pressure differential during the production of voiced obstruents (Bell-Berti, 1975). However, less closure force does not necessarily imply velopharyngeal opening. The relation between degrees of closure force and potential air leakage through the velopharyngeal port requires additional investigation.

For the females, levator EMG activity was significantly greater for /s/ in the /isi/ context than for /z/ in the /izi/ context. It is possible that the levator muscle requires a greater activity level with the tongue in a high front position to achieve equal force of velopharyngeal closure for voiceless versus voiced consonants in females. It is known that intraoral air pressure is greater for voiceless than for voiced fricatives (Arkebauer, Hixon, & Hardy, 1967; Subtelny, Worth, & Sakuda, 1966) and that the palatoglossus muscle exerts a downward pull on the velum during high front vowel production (Moon et al., 1994). Thus, the levator muscle may have to contract more forcefully during voiceless fricative production to offset the opposing force of the palatoglossus muscle for /i/ that might otherwise result in release of the higher air pressure through the velopharyngeal port. Given the propensity of the velum to elevate to a higher level in males compared to females (Kuehn, 1976; McKerns & Bzoch, 1970), there would be less need for greater levator activity to offset possible release of air pressure for the voiceless consonants in males.

Place

Previous velar position (Kuehn, 1976) and EMG (Bell-Berti, 1993) studies have indicated that variations in place of production do not have a large affect on velar activity and such affects may be indirectly related to tongue positioning (Bell-Berti, 1993). The results of the current study suggest that the place feature might have some influence on force of velopharyngeal closure but the affect is not large. In comparing the two place categories involving the tongue dorsum and tongue apex, it was found that force of closure was significantly different only for the males; greater closure force for /k/ than /t/ and only in a high-back vowel

environment. Neither closure force for the females nor EMG for either sex group were significantly different. It is possible that greater closure force for /k/ than /t/ in the males is primarily a mechanical effect through the anterior faucial pillars, perhaps, owing to the typically higher tongue position for /k/ than for /t/. Given that levator activity was not significantly different for /k/ versus /t/ and that palatoglossus, if active, would most likely function to assist in tongue elevation and thus velar lowering, it is probable that greater closure force for /k/ was not related to active muscle adjustment, at least from levator and palatoglossus. Additional information about this matter is needed that would combine velopharyngeal force measures with EMG measures from the palatoglossus muscle in addition to the levator muscle.

Manner

As expected, both velopharyngeal force and levator EMG activity were significantly less for nasal versus fricative or stop production for both males and females. A relatively open velopharyngeal port is necessary for the production of nasal consonants. However, for the males, force was significantly greater for nasals in an /i/ and /u/ context and suggests that some degree of velopharyngeal closure occurred for these nasal productions. It has been demonstrated previously that the entire locus of velar position changes for nasal production (i.e., velum lowered) as well as that for nonnasal production depending on the vowel context (Kuehn, 1976; Moll & Shriner, 1967). Given that levator activity was not significantly different across these same contexts for the males suggests that factors in addition to levator suppression determine velopharyngeal positioning during nasal consonant production. Mechanical influence attributable to tongue positioning is a probable candidate.

Although there was some suggestion in the female subjects that velopharyngeal closure force may be greater for stops than fricatives (Figure 9), these differences did not reach statistical significance. Perhaps with a larger subject sample, such differences might emerge.

Nasal-Nonnasal Sequencing

The data shown in Figure 11 indicate that, in general, closure force for a nonnasal consonant is greater following a nasal consonant than preceding it. This suggests that velopharyngeal closure for a nonnasal consonant is tighter when it immediately follows, rather than precedes, a nasal consonant. The mean values in Figure 12 suggest that levator EMG activity is greater when it works to raise the velum following a nasal consonant than when the velum is already in a raised position by virtue of its preceding state. However, these differences did not reach statistical significance.

Lubker et al. (1970, p.16) stated that "... palatal musculature would be predicted to contract as forcefully as necessary to move the soft palate from wherever it is to wherever it must be to prevent excessive nasal coupling." In the examples shown in the current study in Figures 11 and 12, it is possible that the levator (and presumably other muscles as well), in closing the velopharyngeal port, produces an inertial effect such that when contact is made for the nonnasal consonant, that contact (i.e., velopharyngeal closure force) continues to increase.

Sex Differences

When the data were grouped across conditions for comparisons involving the four vowels (Figure 3), six voicing conditions (Figure 5), six place conditions (Figure 7), nine manner conditions (Figure 9), and four sequencing conditions (Figure 11), no significant differences were found between levels of velopharyngeal closure force between the male and female subjects. Although the amount of surface contact between the velum and the force bulb was not measured in the current study, it is possible that, as a group, the females had greater contact area than the males (McKerns & Bzoch, 1970). As stated in the Method section, with the force measuring device used in this study, the greater the area of contact, the *less* the measured force. Thus, the consistent differences in the means (though not reaching significance) shown in Figures 3,5,7,9,11 in which the force values for males are greater than that for females might be even smaller if the force bulb measures did not have a tendency to vary inversely with amount of surface contact. As a rough estimate, if the mean values for the females are increased by 10% (see explanation in the Method section concerning the 10% amount) it can be seen that the values are very close to the mean values for the males.

When the data for males and females were separated, it was found that the male subjects generally exhibited greater differences across the various consonant categories for velopharyngeal force than the females. It is possible that females, as a group, may exhibit less upward velar displacement along the posterior pharyngeal wall, that is, less velar "height" compared to males (Kuehn, 1976; McKerns & Bzoch, 1970) and thus, a more limited range for changes in velopharyngeal closure force. In this regard, McKerns and Bzoch (1970) reported that male subjects elevated their velums an average of 10 mm above the hard palatal plane compared to an average of 5.9 mm for the females.

Force Categories

It is plausible that velopharyngeal closure force is conditioned in relation to phonemic or phonetic rules. Within languages, such rules might be consistent with the original notion of Lubker and his colleagues (Lubker, 1975; Lubker et al., 1970). They suggested that there may be different categories of closure such as that for high versus low vowels, voiceless versus voiced consonants, etc.

With regard to possible differences between languages, Solé (1995) presented convincing evidence that velopharyngeal control for nasalization in /CVVN/ sequences is fundamentally different for native speakers of Spanish versus English. Based on her data, she argues that velar movements in Spanish for such sequences are "automatic" and under low level physiologic control whereas for English such movements are controlled by higher level "preprogrammed" activity. She describes the latter as "...phonologization of a mechanical-articulatory effect" (Solé, 1995, p.19). Thus, low level physiologic constraints might be modified and conditioned by speakers of a language until those modifications become incorporated into the language's normal phonologic rule system.

General Discussion

In general, the results of this study indicate that normal velopharyngeal closure force for speech involves a versatile process that is under the control of the individual speaker to a large degree but is also constrained by mechanical factors and probably also is subject to the phonologic rules of one's language system. There were a number of significant differences between phonemes and phoneme categories involving force measures in which the corresponding levator EMG measures were not significantly different. We attribute this to the likely contribution of muscles in addition to the levator muscle in positioning the velum (Kuehn et al., 1982; Moon et al., 1994; Seaver and Kuehn, 1980). Also, mechanical factors are probably important, particularly tongue positioning with accompanying faucial pillar adjustments as they affect velopharyngeal force, apart from or in addition to the forces provided by the velopharyngeal muscles acting directly on the velum. In addition, intraoral air pressures might influence closure force independently of muscle activity or mechanical factors.

The fineness of velopharyngeal force control may be conditioned by two major factors: 1) the speaker's use of the mechanism for connected speech versus nonspeech activities and 2) physical limitations that may be imposed on the mechanism. With regard to the first, the degree of volitional control appears limited to, perhaps, two or three different positional states (Moon & Jones, 1991; Moon et al., 1994; Shelton, Harris, Scholes, & Dooley, 1970). That is, upon command, it is difficult for an individual to control the precise position (and probably velopharyngeal closure force) of the velum although some individuals appear to be better at such tasks than other individuals. In connected speech, more degrees of positional freedom may be developed and incorporated in the motor program to accommodate the demands of the speaker's particular linguistic system. With regard to the nonspeech activity of swallowing, although not formally included in the protocol of this study, it is our impression in observing many swallowing episodes that velopharyngeal closure force for swallowing is very different from that for speech. For example, maximum force tends to be lower in the pharynx for swallowing than for speech probably due to the increased involvement of the constrictor muscles for swallowing.

With regard to the second conditioning factor related to fineness of velopharyngeal force control, physical limitations as those imposed by surgical scarring as a sequella to cleft palate, such individuals may exhibit reduced versatility. That is, velopharyngeal closure force may be more constant than that for normals because of a reduced range of movement. As a result, individuals with velar impairment might tend to elevate the velum to a position that is somewhat of a compromise between a fully elevated and a fully lowered position. Interestingly, in this study, females tended to exhibit a more constant closure force than the males. However, the locus of that more constant positioning involved a velar positioning, presumably, that was sufficiently elevated to avoid excessive nasal coupling. The velar positional locus for individuals with impaired velums, as related to scarring for example, may be too low, thus leading to excessive nasal coupling. Yet, this compromise involving a locus that is too low (and accompanying force that is not sufficient) may be necessary so as to avoid a state of physiologic fatigue that could occur if the locus were habitually higher. We are currently examining this possibility.

Acknowledgments

This study was supported in part by PHS Research Grants DC-00976 from the National Institute on Deafness and Other Communication Disorders and DE-10436 from the National Institute of Dental Research. We would like to acknowledge the statistical support of the Biostatistics Core of the National Center for Voice and Speech and Wendy Edwards for her assistance in data analysis and graphics.

References

Arkebauer, H.J., Hixon, T.J., & Hardy, J.C. (1967).Peak intraoral air pressure during speech. *Journal of Speech and Hearing Research,10,* 196-208.

Andreassen, M.L., Smith, B.E., & Guyette, T.W. (1992). Pressure-flow measurements for selected oral and nasal sound segments produced by normal adults. *Cleft Palate-Craniofacial Journal,29,* 1-9.

Barlow, S.M. & Bradford, P.T. (1992). Measurements and implications of orofacial muscle performance is speech disorders. *Journal of Human Muscle Performance.1,* 1-32.

Bell-Berti, F. (1975). Control of pharyngeal cavity size for English voiced and voiceless stops. *Journal of the Acoustical Society of America,57,* 456-461.

Bell-Berti, F. (1976). An electromyographic study of velopharyngeal function in speech. *Journal of Speech and Hearing Research,19,* 225-240.

Bell-Berti, F. (1993). Understanding velic motor control: studies of segmental context. In M.K. Huffman & R.K. Krakow (Eds.), *Phonetics and Phonology; Nasals, Nasalization, and the Velum, Volume 5 (pp. 63-85).* San Diego: Academic Press.

Bell-Berti, F., Baer, T., Harris, K., & Nimi, S. (1979). Coarticulatory effects of vowel quality on velar function. *Phonetica,36,* 187-193.

Bell-Berti, F. & Krakow, R.A. (1991). Anticipatory velar lowering: a coproduction account. *Journal of the Acoustical Society of America,90,* 112-123.

Benguerel, A.-P., Hirose, H., Sawashima, M., & Ushijima, T. (1977a). Velar coarticulation in French: a fiberoptic study. *Journal of Phonetics,5,* 149-158.

Benguerel, A.-P., Hirose, H., Sawashima, M., & Ushijima, T. (1977b). Velar coarticulation in French: an electromyographic study. *Journal of Phonetics,5,* 159-167.

Cohn, E., Fisher, S., McWilliams, B., Seth, A., Ferketic, M., Beery, Q., & Vallino, L. (1986). An investigation of velopharyngeal sphincteric pressures during speech and non-speech tasks. Presented at the annual meeting of the American Cleft Palate Association.

Clumeck, H. (1976). Patterns of soft palate movements. *Journal of Phonetics,4,* 337-351.

Evarts, E.V. (1968). Relation of pyramidal tract activity to force exerted during voluntary movement. *Journal of Neurophysiology,31,* 14-27.

Fritzell, B. (1969). The velopharyngeal muscles in speech: an electromyographic-cineradiographic study. *Acta Otolaryngologica,* Supplement 250.

Goto, T. (1977). Tightness in velopharyngeal closure and its regulatory mechanism. *Journal of the Osaka University Dental Society,22,* 1-19.

Hoit, J.D., Watson, P.J., Hixon, K.E., McMahon, P., & Johnson, C.L. (1994). Age and velopharyngeal function during speech production. *Journal of Speech and Hearing Research,37,* 295-302.

Karnell, M.P., Linville, R.N., & Edwards, B.A. (1988). Variations in velar position over time: a nasal videoendoscopic study. *Journal of Speech and Hearing Research,31,* 417-424.

Kent, R.D., Carney, P.J., & Severeid, L.R. (1974). Velar movement and timing: evaluation of a model of binary control. *Journal of Speech and Hearing Research,17,* 470-488.

Krakow, R.A. (1993). Nonsegmental influences on velum movement patterns: syllables, sentences, stress, and speaking rate. In M.K. Huffman & R.K. Krakow (Eds.), *Phonetics and Phonology; Nasals, Nasalization, and the Velum, Volume 5 (pp. 87-116).* San Diego: Academic Press.

Kuehn, D.P. (1976). A cineradiographic investigation of velar movement variables in two normals. *Cleft Palate Journal,13,* 88-103.

Kuehn, D.P., Folkins, J.W., & Cutting, C.B. (1982). Relationships between muscle activity and velar position. *Cleft Palate Journal,19,* 25-35.

Kuehn, D.P. & Moon, J.B. (1994). Levator veli palatini muscle activity in relation to intraoral air pressure variation. *Journal of Speech and Hearing Research,37,* 1260-1270.

Kuehn, D.P. & Moon, J.B. (1995). Levator veli palatini muscle activity in relation to intraoral air pressure variation in cleft palate subjects. *Cleft Palate-Craniofacial Journal,32,* 376-381.

Lubker, J.F. (1975). Normal velopharyngeal function in speech. *Clinics in Plastic Surgery,2,* 249-259.

Lubker, J.F., Fritzell, B., & Lindquist, J. (1970). Velopharyngeal function: an electromyographic study. *Speech Transmission Laboratory Quarterly Progress and Status Report, Royal Institute of Technology, Stockholm,4,* 9-20.

Lubker, J.F & Moll, K.L. (1965). Simultaneous oral-nasal air flow measurements and cinefluorographic observations during speech production. *Cleft Palate Journal,2,* 257-272.

McKerns, D. & Bzoch, K.R. (1970). Variations in velopharyngeal valving: the factor of sex. *Cleft Palate Journal,7,* 652-662.

Moll, K.L. (1962). Velopharyngeal closure on vowels. *Journal of Speech and Hearing Research,5,* 30-37.

Moll, K.L. & Daniloff, R.G. (1971). Investigation of the timing of velar movements during speech. *Journal of the Acoustical Society of America,50,* 678-684.

Moll, K.L. & Shriner, T.H. (1967). Preliminary investigation of a new concept of velar activity during speech. *Cleft Palate Journal,4,* 56-69.

Moon, J.B. & Jones, D.L. (1991). Motor control of velopharyngeal structure during vowel production. *Cleft Palate-Craniofacial Journal,28,* 267-273.

Moon, J.B., Kuehn, D.P., & Huisman, J.J. (1994). Measurement of velopharyngeal closure force during vowel production. *Cleft Palate-Craniofacial Journal,31,* 356-363.

Moon, J.B., Kuehn, D.P., & Huisman, J.J. (1995). Author response: Letter to the editor. *Cleft Palate-Craniofacial Journal,32,* 263.

Moon, J.B., Smith, A.E., Folkins, J.W., Lemke, J.H., & Gartlan, M. (1994). Coordination of velopharyngeal muscle activity during positioning of the soft palate. *Cleft Palate-Craniofacial Journal,31,* 45-55.

Morrison, D.F. (1976). *Multivariate statistical methods.* New York: McGraw-Hill.

SAS Institute. (1987). *SAS/STAT guide for personal computers, Version 6.* Cary, NC: SAS Institute, Inc.

Seaver, E.J. & Kuehn, D.P. (1980). A cineradiographic and electromyographic investigation of velar positioning in non-nasal speech. *Cleft Palate Journal,17,* 216-226.

Shelton, R.L., Harris, K.S., Scholes, G.M., & Dooley, P.M. (1970). Study of nonspeech palate movements by scaling and electromyographic techniques. In J.F. Bosma (Ed.) *Second symposium on oral sensation and perception* (pp. 432-441). Springfield, IL: Charles C. Thomas.

Solé, M. (1995). Spatio-temporal patterns of velopharyngeal action in phonetic and phonological nasalization. *Language and Speech,38,* 1-23.

Stein, R.B. (1982). What muscle variables does the central nervous system control? *Behavior and Brain Sciences,4,* 535-578.

Subtelny, J.D., Worth, J.H., & Sakuda, M. (1966). Intraoral pressure and rate of flow during speech. *Journal of Speech and Hearing Research,9,* 498-518.

Thompson, A.E. & Hixon, T.J. (1979). Nasal air flow during normal speech production. *Cleft Palate Journal,16,* 412-420.

Vaissiere, J. (1988). Prediction of velum movement from phonological specifications. *Phonetica,45,* 122-139.

Yanagihara, N. & Hyde, C. (1966). An aerodynamic study of the articulatory mechanism in the production of bilabial stop consonants. *Studia Phonologica,4,* 70-80.

NCVS Status and Progress Report - 11
May 1997, 119-126

Viscosities of Injectable Biomaterials in Vocal Fold Augmentation Surgery

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

Abstract

Vocal fold vibration depends critically on the viscoelasticity of vocal fold tissues. For instance, phonation threshold pressure, a measure of the "ease" of phonation, has been shown to be directly related to the viscosity of the vibrating mucosa. Various injectable biomaterials have been used in vocal fold augmentation surgery, with injection sites sometimes close to or inside the mucosa. Yet their viscosities or other mechanical properties are seldom known. This study attempts to provide some data on viscosities of some commonly used phonosurgical biomaterials.

Using a parallel-plate rotational rheometer, oscillatory shear experiments were performed on injectable collagen (Zyderm, Collagen Corporation), glutaraldehyde cross-linked (GAX) collagen (Phonagel or Zyplast), absorbable gelatin (Gelfoam, Upjohn Company), and human abdominal subcutaneous fat. Samples of human vocal fold mucosal tissues were also tested.

Under sinusoidal oscillatory shear at 10 Hz and at 37°C, the dynamic viscosity was 21 Pascal-second (Pa-s) for gelatin, 8-13 Pa-s for the two types of collagen, 3 Pa-s for fat, and 1-3 Pa-s for vocal fold mucosa. Results extrapolated to 100 Hz also show similar differences among the biomaterials, but all values are an order of magnitude lower because of the typical inverse frequency relation for biological materials.

Our data suggest that the use of fat for vocal fold augmentation is more conducive to the "ease" of phonation because of its relatively low viscosity, which is closest to physiological levels.

Introduction

Injection laryngoplasty is a common practice in phonosurgery. The usual objective is to medialize a paralyzed vocal fold, to soften a scar, or to augment an atrophic vocal fold. In the management of scarring or atrophy, the injection is close to or directly into the vibrating vocal fold mucosa, so as to repair focal defects, to smooth the vocal fold margin and to soften the stiffened scar tissue[1,2]. A number of synthetic injectable materials have been used historically in vocal fold augmentation surgery, including paraffin, Teflon, and silicone (in Japan). Today, the use of biological materials is becoming more popular, primarily because of the promising results, but also because of persisting problems associated with synthetic materials like Teflon[3,4].

Injectable collagen and fat have been the most widely used biological materials, while gelatin is a common biomaterial used for temporary vocal fold medialization. Bovine dermal collagen suspension (Zyderm Collagen Implant, Collagen Corporation) was first used by Ford *et al.*[5] in phonosurgery and was shown to be a viable alternative to Teflon[2,6]. Glutaraldehyde cross-linked (GAX) collagen (Phonagel or Zyplast) was better tolerated immunologically and was more stable against gradual post-surgery resorption[6,7]. More recently, autologous collagen has been harvested from patients' own abdominal skin, which has virtually eliminated any immunological risk[8]. Autologous fat from abdominal liposuction or excision has also been used clinically for augmenting paralyzed vocal folds, with encouraging results[9,10,11]. Favorable results have also been reported in fat-graft reconstruction of injured canine vocal fold mucosa[12,13] and scarred human vocal fold mucosa[11,14]. Gelatin (Gelfoam, Upjohn Company) paste was first used by Schramm *et al.*[15] as a biomaterial for the temporary restoration of glottal competence, e.g., prior to radiation therapy[16].

Many factors need to be considered when choosing an appropriate biomaterial for injection into the vocal

folds. Some of the more important considerations include the ease of injection, risk of immunologic responses, potential malposition and overinjection, and resorption and stability of the implant. The overriding consideration, however, is the ability of the repaired vocal fold to return to normal or near-normal vibration, which depends critically on the viscoelasticity of vocal fold tissues. Viscous and elastic properties are both affected by the introduction of the foreign biomaterial, especially in the management of vocal fold scarring. It has been shown that the *phonation threshold pressure* for vocal fold oscillation is directly related to the viscosity of the vocal fold mucosa[17]. This pressure has been defined as the minimum subglottal pressure required to produce vocal fold oscillation[17,18]. It is an objective indicator of the "ease" of phonation, and can be useful as a diagnostic parameter of vocal function. Titze[17] showed analytically that phonation threshold pressure increases with viscosity of the vibrating vocal fold mucosa. The validity of this relationship has been verified by different kinds of experimental results[19-22].

Unfortunately, few data on viscosity or other mechanical properties of the commonly used biomaterials are available in the voice and speech literature. This study attempts to remedy this situation by presenting measurements of the *dynamic viscosity* of injectable collagen, gelatin, fat, and human vocal fold mucosa. Dynamic viscosity characterizes the internal friction of a material during deformation and can be defined in the context of sinusoidal oscillatory shear[23]. When a sinusoidal oscillatory shear stress is applied to a sample of viscoelastic material, a steady-state condition can be reached (for linear small-amplitude oscillations) where the sinusoidal shear stress will have a constant relation to the sinusoidal shear strain (deformation) at a given frequency of oscillation. The strain amplitude will be proportional to the stress amplitude, and a constant phase lag will exist. Dynamic viscosity can be derived from this phase lag and the ratio of stress to strain amplitude. Mathematically, the viscosity is defined by the analytical relationship between shear stress, strain, and strain rate[23,24]:

$$\tau = G'\gamma + \eta'\dot{\gamma} \qquad (1)$$

where τ is shear stress (in Pa), γ is shear strain (in radians), $\dot{\gamma}$ is strain rate (in s^{-1}), G' is elastic shear modulus (in Pa) and η' is dynamic viscosity (in Pascal-second, Pa-s).

Materials and Methods

Injectable collagen samples included both 0.0075% glutaraldehyde cross-linked (GAX) bovine dermal collagen suspension (Phonagel or Zyplast, Collagen Corporation, Palo Alto, CA) and the noncross-linked Zyderm (Collagen Corp.) identical to those used for phonosurgery[5-7]. They were both solubilized, purified and reconstituted preparations of Type I collagen suspended in a phosphate buffered saline solution (35 mg/ml) and were contained in syringes ready for injection. For each type of collagen, three samples (each of about 0.5 cc) were taken from three different syringes and tested in a rheometer. Autologous collagen was not included in the present study, but will be tested in future studies. Three injectable gelatin samples were prepared by mixing 0.375 g of Gelfoam (Upjohn Company, Kalamazoo, MI) absorbable gelatin powder with 1.5 cc (0.9%) saline solution, resulting in a gelatin sponge suspension of 250 mg/ml, a concentration also used in phonosurgery[15,16].

Fat samples were harvested from a 70-year-old female cadaver from autopsy, about 24-hour post-mortem. An approximately 4-cm^3 block of subcutaneous fat tissue was excised from the periumbilical area of the cadaver's anterior abdomen. CT observations have shown two layers of abdominal subcutaneous fat, superficial and deep[25], but it was not known from which layer the sample was harvested. Possible differences between the two may be studied in future experiments. The block of fat was minced into smaller pieces (about 0.1 cm^3), and any fascia and fibrous tissue was removed. It was then rinsed thoroughly with saline solution to remove any remaining debris. This procedure of fat-graft preparation was similar to that described for phonosurgery[11]. Three 0.5 cc samples were taken from the pieces and tested on the rheometer.

Samples of vocal fold mucosa were also tested. They were obtained from two excised larynges (a 62-year-old male and a 72-year-old male), about 24-hour post-mortem from autopsy. Mechanical properties of the mucosa might have been affected by post-mortem changes but fresher human tissues were not available at the time of testing. The mucosa consisted of the superficial layer of lamina propria and the epithelium, which remained attached to the superficial layer during testing and hence served as a natural boundary of attachment to the rheometer. Two samples of mucosa were obtained from each larynx, one from the left and the other from the right vocal fold. The volume of each sample was about 0.1 cc.

Rheometer and Oscillatory Shear Experiments

Rheological or viscometric measurements were made with a Bohlin Controlled Stress (CS-50) Rheometer, controlled by a computer with commercial CS Oscillation data acquisition and analysis software. A plate-on-plate (parallel-plate) testing geometry was used, with a stationary lower plate and a rotating upper plate (diameter = 20 mm). The parallel-plate geometry was used so that the sample volume required to fill the gap could be adjusted by the gap size. To match a sample volume of 0.1 cc (typical of mucosal

tissue samples), the gap size was 0.3 mm. Complete sample filling of the gap was ensured by first mounting samples with slightly larger volume than that of the gap (0.1 cc). Excessive material was then wiped off the rim of the gap while care was taken to prevent removing too much.

The sample was subject to a precisely controlled sinusoidal torque from the upper plate, driven by a motor (torque range 0.001 to 10 mNm, resolution < 0.0002 mNm). The resulting angular displacement and angular velocity of the upper plate were monitored by a sensitive transducer (linear angular deflection range, resolution $\approx 2 \times 10^{-5}$ rad) as a function of time. Sinusoidally varying shear stress, shear strain and strain rate were calculated from the preset torque and the measured angular velocity by the computer. From the shear stress and strain rate functions, the CS Oscillation software calculated dynamic viscosity using the linear viscoelasticity theory mentioned above. Measurements of dynamic viscosity were made at a frequency range of 0.01 to 15 Hz, covering 32 frequencies over three decades. Testing at higher frequencies was not performed because linearity of the underlying viscoelasticity theory could not be ensured when rotor (upper plate) and sample inertial effects become significant at high frequency. Linearity was also maintained by using small-amplitude oscillation with the shear strain amplitude set at 0.01 rad throughout the experiments. This target shear strain was located by strain (or stress) sweep tests done prior to frequency sweep experiments, where steady dynamic viscosity was obtained at low strain up to at least 0.01 rad.

Calibration of the rheometer was done by the manufacturer and was double checked by measuring the steady-shear viscosities of standard polydimethylsiloxane solutions (Dow Corning 360 Medical Fluids) with known viscosities of 0.1, 0.35, 1, and 12.5 Pa-s at 20°C. The measured viscosities showed a less than 6% deviation from the stated viscosities at all strain rates in all cases. Throughout the experiments, the sample under test was maintained at a temperature of 37°C ± 0.1°C by a Bohlin Temperature Control Unit which circulated water into the lower plate. The vocal fold mucosa samples were also kept viable during testing by bathing in physiological saline solution.

Results

Figure 1 shows the dynamic viscosity of three samples of injectable gelatin (Gelfoam) as a function of oscillation frequency, on a log-log scale. Because the three gelatin samples came from a single preparation, discrepancies in results between them reflect an estimation of the experimental error or repeatability of the oscillatory shear experimental procedure. Possible sources of errors include differences in the process of sample mounting (e.g., during the removal of excessive sample material) and the rheometer

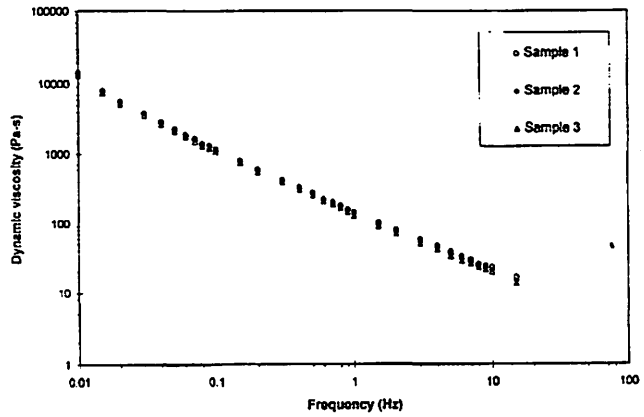


Figure 1. Dynamic viscosity of injectable gelatin (Gelfoam) as a function of frequency.

(e.g., errors in torque output and measurement of angular displacement). According to the largest differences between the results in Figure 1 (between samples 1 and 3), experimental error or repeatability of the experimental procedure is estimated to be within 10% of the measured dynamic viscosity values across all frequencies.

Figures 2 to 5 (following page) show the dynamic viscosities of noncross-linked collagen (Zyderm), glutaraldehyde cross-linked (GAX) collagen (Phonagel, or Zyplast), human abdominal subcutaneous fat, and human vocal fold mucosa, respectively. As in Fig.1, they are plotted against frequency on a log-log scale. For injectable collagen and fat, results of three samples of each material are shown. For vocal fold mucosa, results of two samples from each larynx are shown, one from the left and one from the right vocal fold. It can be seen that discrepancies between samples of the same material in Figs. 2 to 5 are larger than those in Fig. 1. Since the discrepancies in Fig. 1 reflect the experimental error, it indicates that in addition to the experimental error there are slight variations in dynamic viscosity of different samples of the same material.

As is the case of most biomacromolecules and polymeric materials[23], the dynamic viscosities of injectable biomaterials and vocal fold mucosal tissues are a monotonically decreasing function of frequency. Moreover, the decrease with frequency is about linear on the log-log scale as shown in Figs. 1 to 5, which indicates that the relationship between dynamic viscosity and frequency can be modelled by a power law:

$$\eta' = k f \tag{2}$$

or

$$\log \eta' = \log k + n \log f \tag{3}$$

where k and n are constants. Data for each material shown in Figs. 1 to 5 were averaged across samples and then fitted to

Table 1.
Results of Linear Least-Squares Regressopm Curve-Fitting ([Eq. (3)]

Material sample	k	n	R^2
Gelatin (Gelfoam)	149.39	-0.9030	0.997
GAX collagen (Phonagel, or Zyplast)	99.851	-0.9145	0.998
Noncross-linked collagen (Zyderm)	66.395	-0.9154	0.998
Human abdominal subcutaneous fat (70-year-old female)	23.576	-0.9508	0.994
Vocal fold mucosa (72-year-old male)	15.758	-0.8741	0.995
Vocal fold mucosa (62-year-old male)	6.5601	-0.9109	0.998

Values of k, n, and R^2 for injectable biomaterials and human vocal fold mucosal tissues were computed on averages of results shown in Figs. 1 to 5.

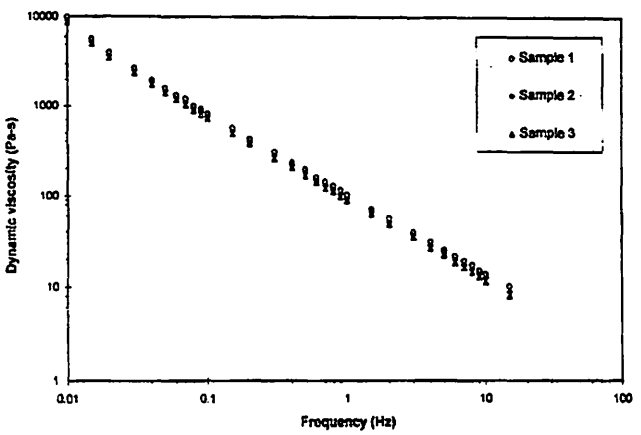


Figure 2. Dynamic viscosity of injectable GAX collagen (Phonagel, or Zyplast) as a function of frequency.Figure 2. Dynamic viscosity of injectable GAX collagen (Phonagel, or Zyplast) as a function of frequency.

Eq. (3) by simple linear least-squares regression. Table I summarizes the results of curve-fitting by giving the values of k, n, and the square of the correlation coefficient R for different materials. Data are very well matched by the regression equation, yielding correlation coefficients above 0.99. Dynamic viscosity decreased with frequency at about the same rate for different materials, with very similar slopes (n). This is also clearly shown in Figure 6, which plots the composite average values of dynamic viscosity versus frequency for all materials.

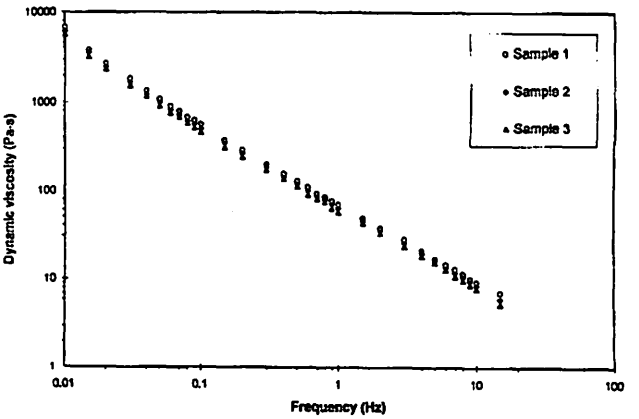


Figure 3. Dynamic viscosity of injectable noncross-linked collagen (Zyderm) as a function of frequency.

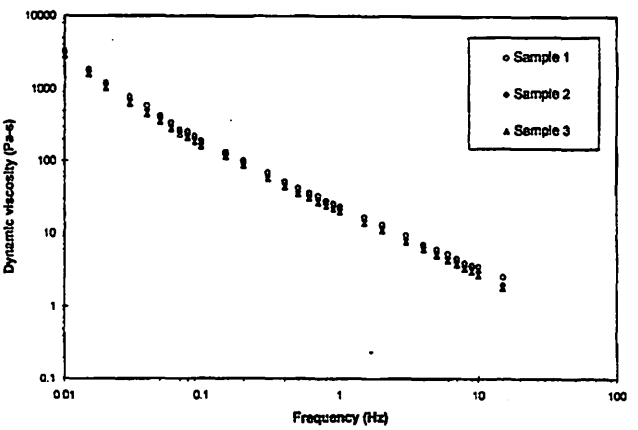


Figure 4. Dynamic viscosity of human abdominal subcutaneous fat (from a 70-year-old female) as a function of frequency.

The main difference between the materials can be described by the vertical separation of the curves in Fig. 6, or the difference in the values of k in Table I. Injectable gelatin is the most viscous, followed by GAX collagen, noncross-linked collagen, fat, and vocal fold mucosa. Fig. 6 also shows the results of data extrapolation to 100 Hz, a typical male vocal fold vibration frequency, based on the regression equation [Eq. (3)]. Table II summarizes the measured dynamic viscosity values at 10 Hz and the extrapolated values at 100 Hz. Results show that the order and magnitude of differences among the materials in dynamic viscosity extrapolated to 100 Hz are about the same as for the measured frequencies, because of similarity in slopes of the curves. However, it should be noted that a deviation from this similarity in slope is possible at frequencies higher than 15 Hz, especially for vocal fold mucosal tissues that show some degree of "nonlinearity" (on the log-log scale) when higher frequencies are approached. Unfortunately, the highest frequency of oscillation allowed in the rheometer is

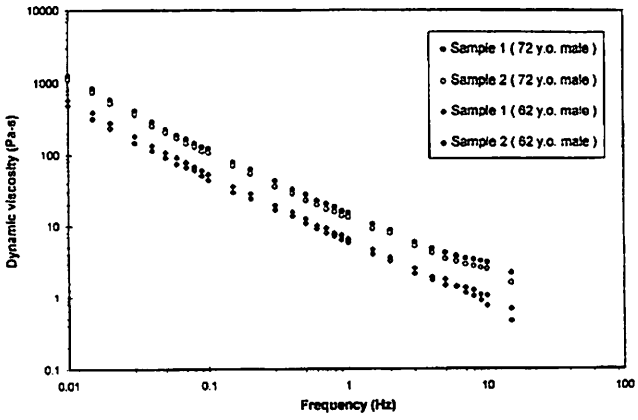


Figure 5. Dynamic viscosity of human vocal fold mucosa (superficial layer of lamina propria) as a function of frequency.

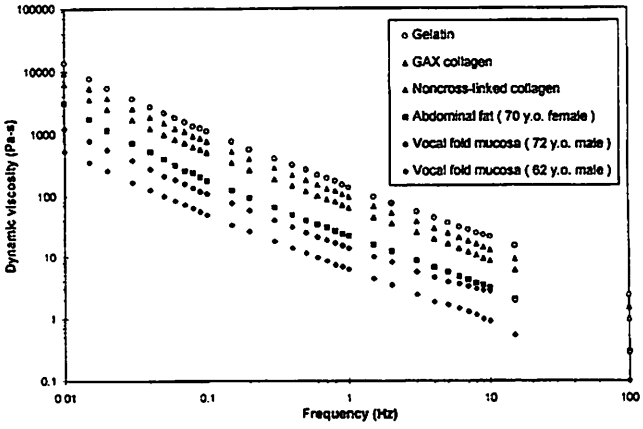


Figure 6. Dynamic viscosity of injectable biomaterials and human vocal fold mucosa as a function of frequency. Data points are averages of those in Figs. 1 to 5. The values at 100 Hz are extrapolated by linear least-squares regression [Eq. (3)].

limited by inertia of the rotating plate and sample inertia, as mentioned above. In future studies, measurements made at higher frequencies are needed to validate the extrapolation.

Discussion

The viscosity of a material is a measure of its resistance to shear flow, and is always associated with dissipation of internal energy, typically as heat. A material with a low viscosity flows easily, with little internal energy being dissipated in the shearing process. A highly viscous material flows more slowly and dissipates more energy in the process. The resistance to flow for a viscoelastic material is quantified by dynamic viscosity. At the microscopic level, this resistance depends on the ease of relative shear or slippage between molecules of the material, which is determined by different kinds of intramolecular and intermolecular interactions. Some of the most significant molecular

Table 2.
Dynamic Viscosity of Injectable Biomaterials and Human Vocal Fold Mucosal Tissues Measured at 10 Hz and Extrapolated to 100 Hz

Material sample	Dynamic viscosity (Pa-s)	
	measured at 10 Hz	extrapolated to 100 Hz
Gelatin (Gelfoam)	21.297	2.335
GAX collagen (Phonagel, or Zyplast)	12.844	1.480
Noncross-linked collagen (Zyderm)	8.563	0.980
Human abdominal subcutaneous fat (70-year-old female)	3.026	0.296
Vocal fold mucosa (72-year-old male)	2.702	0.281
Vocal fold mucosa (62-year-old male)	0.897	0.099

Extrapolation was based on simple linear least-squares regression (see Eq. 3).

interactions include physical entanglement, electrostatic forces (e.g., hydrogen bond), hydrophilic and hydrophobic interactions, and chemical interactions (e.g., covalent cross-link formation)[26]. The stronger the molecular interactions, the more resistant the molecules are to slippage, and the higher the viscosity. These interactions are especially important in polymeric materials (the biomaterials tested in this study), where long chains of molecules occupy a large space relative to their atomic dimensions.

Gelatin is a denatured or partially hydrolyzed collagen. Our results show that injectable gelatin has a higher dynamic viscosity than collagen (see Fig. 6 and Table II). It can be explained by the fact that gelatin exists as a heterogeneous aggregate of partially hydrolyzed fragments of collagen fibrils and tropocollagen molecules. These fragments of gelatin molecules are basically randomly coiled together with intramolecular and intermolecular interactions, like physical entanglements, hydrogen bonds, and covalent cross-links[27]. In addition, the gelatin samples used were prepared with a higher concentration (250 mg/ml) than that of collagen (37 mg/ml). These two factors contributed to the higher viscosity of gelatin. Owing to its relatively high viscosity, injectable gelatin should only be used for medialization of paralyzed vocal folds in phonosurgery, with injection sites out of the vibrating vocal fold mucosa.

The chemically cross-linked GAX collagen showed a higher dynamic viscosity than noncross-linked collagen, as expected from the formation of many covalent cross-links in tropocollagen molecules after glutaraldehyde treatment. Molecular interactions between and within the collagen fibrils are therefore stronger for GAX collagen; hence its higher viscosity.

Fat is stored in droplets in cellular adipocytes of subcutaneous adipose tissue. The typical lipid molecules are the highly nonpolar and hydrophobic triacylglycerols. Our results show that fat has a lower dynamic viscosity than collagen, probably due to its very different molecular structure. Unlike collagen molecules, fat molecules do not have significant electrostatic forces between them, particularly for fatty acids with relatively saturated hydrocarbon chains. Collagen, including "noncross-linked" collagen, also has numerous covalent cross-links among individual molecules, while the relatively inert fat molecules do not have reactive functional groups to form cross-links. Hence fat has relatively weak molecular interactions and a lower viscosity.

Since the vocal fold mucosal tissues tested in this study were obtained 24-hour post-mortem, there may well be considerable post-mortem changes in their mechanical properties after 24 hours. Nevertheless, the dynamic viscosity of vocal fold mucosa measured at 10 Hz was around 1 Pa-s, with a considerable difference between the two subjects. The vocal fold mucosa of the 72-year-old male was about three times more viscous than that of the 62-year-old male. This finding alone is interesting and consistent with the observations of significant inter-subject differences in molecular constituents of the lamina propria[28,29]. The dynamic viscosity of canine vocal fold mucosa reported by Kakita *et al.*[30] at 0.1 Hz and 19°C was about 400 Pa-s, compared to 50 Pa-s and 100 Pa-s for our two subjects. Their much higher viscosity was probably due to the fact that they tested the deep layer of lamina propria instead of the superficial layer. Besides, the use of human versus canine tissues is certainly a major difference between the two studies.

Vocal fold molecular composition has been investigated to some depth previously[28,29,31-33]. Cellular structures found in the lamina propria include mainly fibroblasts and macrophages, while the extracellular matrix of lamina propria consists of fibrous proteins (collagen and elastin), glycosaminoglycans (e.g., hyaluronate, chondroitin sulfate, keratan sulfate, heparan sulfate), proteoglycans (glycosaminoglycans covalently attached to a protein core, e.g., versican, fibromodulin), and structural glycoproteins (e.g., fibronectin) [28,29,31-33]. As cellular structures are sparse in the superficial layer of lamina propria, the extracellular matrix probably plays an important role in metabolism as well as mucosal vibration[33]. Among the many constitutents found in the extracellular matrix, hyaluronate (conjugate base of hyaluronic acid at pH 7.0-7.4) and proteoglycans are important mechanically because of their huge size and potential for molecular interactions. However, their concentrations and molecular weights in the superficial layer are not known at present, making the correlation between molecular structure and mechanical properties difficult.

Balazs and Gibbs[34] measured the dynamic shear modulus (an elastic modulus) of hyaluronate solution (5.59 mg/ml) from the human umbilical cord. From their results, dynamic viscosity of hyaluronate is estimated to be about 0.4 Pa-s at 10 Hz. The viscosity of a hyaluronate solution depends strongly on its molecular weight and concentration[35], as expected from the differences in molecular interactions due to different polymeric chain lengths and closeness of packing. Dynamic viscosity of link-stable proteoglycan aggregate solution (40 mg/ml) is estimated to be also about 0.4 Pa-s at 10 Hz, from the dynamic shear modulus data reported by Zhu and Mow[36]. Considerable difference was found between proteoglycan aggregates and monomers, and between aggregates with and without stabilizing link-proteins. This again shows that differences in molecular interactions correspond to differences in mechanical properties.

Our data on dynamic viscosity of vocal fold mucosal tissues are quite different from those of the molecular constituents, namely collagen, hyaluronate and proteoglycans. Collagen suspension has the highest dynamic viscosity, followed by vocal fold mucosa, then hyaluronate and proteoglycan solutions (see Table II). It is likely to be a result of the complex molecular interactions among collagen, hyaluronate and proteoglycans that have yet to be fully understood. As in many other extracellular matrices, collagen and elastin likely form a fibrous scaffolding mesh in the superficial layer of lamina propria, showing various molecular interactions (e.g., electrostatic) with hyaluronate and proteoglycans[37]. Hyaluronate with its large size and extended shape and proteoglycans with their branching might also physically entangle with the mesh and become "space fillers". These molecular interactions contribute to the shear elasticity and viscosity, and explain why viscosity of vocal fold mucosa is higher than that of hyaluronate and proteoglycans. On the other hand, interactions among collagen molecules (e.g., covalent cross-links) might be somewhat weakened by the infiltration of hyaluronate and proteoglycan molecules, rendering the fibrous mesh more susceptible to shear than the cross-linked molecules in pure collagen suspension, thus producing a lower viscosity.

Conclusion

The dynamic viscosity of four commonly used injectable biomaterials for vocal fold augmentation surgery was measured by sinusoidal oscillatory shear experiments using a parallel-plate rotational rheometer. The biomaterials include noncross-linked Type I bovine dermal collagen suspension (Zyderm Collagen Implant, Collagen Corpora-

tion), glutaraldehyde cross-linked (GAX) collagen (Phonagel or Zyplast, Collagen Corp.), absorbable gelatin suspension (Gelfoam, Upjohn Company), and human abdominal subcutaneous fat. Results show that injectable gelatin is the most viscous, followed by GAX collagen, noncross-linked collagen, and fat. The dynamic viscosity of fat is close to that of the vocal fold mucosa (epithelium and superficial layer of lamina propria), while the viscosities of two key constituents of the superficial layer of lamina propria, hyaluronic acid and proteoglycans, are lower than that of the vocal fold mucosa according to the literature.

Our results suggest that the use of fat, or other materials with relatively low viscosity, would be more conducive to the "ease" of phonation, especially in phonosurgery involving the vocal fold mucosa (e.g., repair of vocal fold scarring or atrophy). Use of materials with relatively high viscosity would make vocal fold vibration more difficult. The mechanical properties of a material should be considered together with medical and other considerations in the choice of materials for vocal fold augmentation surgery.

Acknowledgment

This study was supported by a grant (No. P60 DC00976) from the National Institutes on Deafness and Other Communication Disorders, National Institutes of Health. The authors thank Dr. Brad H. Story of the Wilbur Games Gould Voice Research Center, Denver Center for the Performing Arts for assistance in viscometric measurements.

References

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

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

3. Horn KL, Dedo HH. Surgical correction of the convex vocal cord after Teflon injection. *Laryngoscope* 1980; 90:281-286.

4. Rubin HJ. Misadventures with injectable Polytef (Teflon). *Arch Otolaryngol* 1975; 101:114-116.

5. Ford CN, Martin DW, Warner TF. Injectable collagen in laryngeal rehabilitation. *Laryngoscope* 1984; 94:513-518.

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

7. Remacle M, Dujardin JM, Lawson G. Treatment of vocal fold immobility by glutaraldehyde-cross-linked collagen injection: Long-term results. *Ann Otol Rhinol Laryngol* 1995; 104:437-441.

8. Ford CN, Staskowski PA, Bless DM. Autologous collagen vocal fold injection: A preliminary clinical study. *Laryngoscope* 1995; 105:944-948.

9. Mikaelian DO, Lowry LD, Sataloff RT. Lipoinjection for unilateral vocal cord paralysis. *Laryngoscope* 1991; 101:465-468.

10. Brandenburg JH, Kirkham W, Koschkee D. Vocal cord augmentation with autogenous fat. *Laryngoscope* 1992; 102:495-500.

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

12. Wexler DB, Jiang J, Gray SD, Titze IR. Phonosurgical studies: Fat-graft reconstruction of injured canine vocal cords. *Ann Otol Rhinol Laryngol* 1989; 98:668-673.

13. Jiang JJ, Titze IR, Wexler DB, Gray SD. Fundamental frequency and amplitude perturbation in reconstructed canine vocal folds. *Ann Otol Rhinol Laryngol* 1994; 103:145-148.

14. Sataloff RT, Spiegel JR, Hawkshaw MJ, Rosen DC, Heuer RJ. Autologous fat implantation for vocal fold scar: A preliminary report. *J Voice* (in press).

15. Schramm VL, May M, Lavorato AS. Gelfoam paste injection for vocal fold paralysis: Temporary rehabilitation of glottic incompetence. *Laryngoscope* 1978; 88:1268-1273.

16. Hoffman HT, Sullivan MJ, Winter P. Gelfoam injection for vocal cord paralysis prior to radiation therapy. *Ear, Nose, and Throat J* 1991; 70:385-386.

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

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

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

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

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

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

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

24. Whorlow RW. *Rheological techniques.* Chichester, West Sussex: Ellis Horwood, 1980: 243-307.

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

26. Barnes HA, Hutton JF, Walters K. *An introduction to rheology.* Amsterdam: Elsevier, 1989: 97-114.

27. Ward AG, Courts A. *The science and technology of gelatin.* London: Academic Press, 1977: 137-177.

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

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

30. Kakita Y, Hirano M, Ohmaru K. Physical properties of the vocal fold tissue: Measurements on excised larynges. In Stevens KN, Hirano M, eds. *Vocal fold physiology.* Tokyo: University of Tokyo Press, 1981: 377-397.

31. Gray SD, Hirano M, Sato K. Molecular and cellular structure of vocal fold tissue. In Titze IR, ed. *Vocal fold physiology: Frontiers in basic science.* San Diego: Singular Publishing Group, 1993: 1-35.

32. Hirano M. Structure of the vocal fold in normal and disease states anatomical and physical studies. In Ludlow CL, Hard MO, eds. *Proceedings of the conference on the assessment of vocal pathology.* Rockville: American-Speech-Language-Hearing Association, 1981: 11-30.

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

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

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

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

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

NCVS Status and Progress Report - 11
May 1997, 127-134

Correspondence of Electroglottographic Closed Quotient to Vocal Fold Impact Stress in Excised Canine Larynges

Katherine Verdolini, Ph.D.
Division of Otology and Laryngology, Harvard Medical School
Voice/Speech/Swallowing Clinic, Beth Israel Deaconess Medical Center and Brigham and Women's Hospital
Department of Communication Sciences and Disorders, Massachusetts General Hospital Institute of Health Professions
Roger Chan, B.Sc.
Department of Speech Pathology and Audiology, The University of Iowa
Ingo R. Titze, Ph.D.
Department of Speech Pathology and Audiology, The University of Iowa
Wilbur James Gould Voice Research Center, The Denver Center For The Performing Arts
Markus Hess, M.D.
Department of Audiology and Phoniatrics, Freie Universitaet Berlin
Wolfgang Bierhals, Dipl. Ing.
Acoustic Research Lab, Freie Universitaet Berlin

Abstract

The purpose of this study was to explore the possible use of the electroglottographic closed quotient (EGG CQ) as a non-invasive estimate of vocal fold impact stress. Two excised canine larynges were used. Each larynx was mounted and vocal fold oscillation was induced using a humidified air source. Twenty-seven experimental trials were conducted for each larynx. Trials involved variations in vocal process gap, vocal fold elongation, and subglottic pressure. Simultaneous measures were made of vocal fold impact stress (SI) at the midpoint of the membranous vocal folds, and EGG CQ (dimensionless ratio). The results indicated that when threshold and saturation effects were excluded, the SI and the CQ were strongly related (linear correlation $\underline{r} = .83$ and .96 for the two individual larynges, and .81 for the combined data). Within the region of linear relation, an increase of .15 in the CQ corresponded to about 1 kPa increase in SI, for the combined data. Discussion focuses on possible clinical implications and also likely reasons for threshold and saturation phenomena.

Vocal fold contact stress (force/area) has been considered a primary variable in the pathogenesis of common, traumatic lesions within the larynx. The vocal folds undergo various types of stress during oscillation, including tensile stress, contractile stress, arytenoid contact stress, inertial stress, aerodynamic stress, and impact stress (Titze, 1994). Of these, impact stress (force per unit area) was judged to be the most likely source of trauma to the membranous vocal folds because of perpendicular delivery of force to the contralateral fold (Titze, 1994). A variety of tissue changes may result, including microvilli obliteration, surface desquamation, reduction in basal membrane-lamina propria adherence, and appearance of collagen Type IV and fibronectin-generating fibroblasts in the lamina propria which contribute to the histological basis of nodules (Gray, 1991; Gray, Hirano, & Sato, 1993; Gray, Pignatari, & Harding, 1994; Gray, Titze, & Lusk, 1987). Interarytenoid impact stress is likely relevant in the pathogenesis of some granulomas, with tissue changes including arytenoid inflammation, the appearance of grainy, vascularized, and possibly epithelialized tissue. (Miko, 1989).

Because of its importance in pathogenesis, attempts have been made over the past 100 years to assess impact stress using canine (Jiang & Titze, 1994; Kakeshita,

1927; Rethi, 1897;), bovine (Scherer, Cooper, Alipour-Haghighi, & Titze, 1984; Scherer, Cooper, Alipour-Haghighi, & Titze, 1985; Scherer & Titze, 1982; Titze & Scherer, 1981), feline, and human (Hess, Verdolini, Bierhals, Mansmann, & Gross, in review; Shipp, Reed, Doherty, & Morrisey, 1993) subjects, and computer and mathematical models (Titze, 1994; Verdolini, Berry, Titze, & Chan, unpublished data) (see also Baken & Isshiki, 1977; Reed, Doherty, & Shipp, 1990). Thus far, limited information has become available about the range of stresses that can be expected under different phonatory and non-phonatory conditions. In humans, measured interarytenoid stresses that ranged between about 1 - 200 kPa, with the greatest values occurring during non-speech functions such as coughing, throat-clearing, and thoracic fixation (Hess et al, in review). In human and canine subjects, impact stresses between the membranous vocal folds during phonation appear to range between about 1-4 kPa depending on the phonatory task (Jiang & Titze, 1994; Hess et al, in review).

These initial estimates are important for ultimately understanding the contribution of impact stress to pathogenesis. Further work in this direction is important and is ongoing. Clinically, what is needed is a non-invasive estimate of vocal fold impact stress. The primary purpose of the present study was to address this issue by exploring the possible use of the electroglottographic closed quotient (EGG CQ) as a non-invasive indicator of impact stress across the membranous vocal folds.

The EGG signal is obtained non-invasively by affixing surface electrodes to the skin overlying the thyroid cartilage alae, bilaterally. The signal reflects electrical conductance across and around the vocal folds. When the vocal folds are in contact, conductance is greater than when there is an air gap (Scherer, Druker, & Titze, 1988; simulated data indicated in Titze, 1984). Thus, across the vibratory cycle, the time-course of vocal fold approximation can be evaluated. The CQ is a dimensionless ratio (closed time divided by total cycle time) indicating the proportion of time that the glottis is closed.

The CQ might be a non-invasive indicator of impact stress for several reasons. Previous empirical work has indicated that the CQ increases with increasing adduction (for example, Peterson, Verdolini-Marston, Barkmeier, & Hoffman, 1994; Scherer, Vail, & Rockwell, 1995). Furthermore, increased adduction tends to lead to increased impact stress, within constant subglottic pressure and frequency conditions, based on results from excised canine larynx studies (Jiang & Titze, 1994; Verdolini et al, unpublished data). Direct evidence of a relation between the CQ and intramuscular impact stress is provided by Scherer and Titze (1982), whose data showed a strong correspondence between EGG and impact stress measures, at least at large subglottal pressures.

Based on these considerations, the specific research question asked in this protocol was: What are relations between the EGG CQ and impact stress (SI) in excised canine larynges, for different phonation conditions? The phonation conditions involved varying inter-vocal process width, vocal fold elongation, and subglottic pressure.

Methods

Subjects

Two excised canine larynges were obtained from the cardiac research laboratory at The University of Iowa. The animals had been used in previous cardiology experiments; neither was sacrificed for this experiment. The animals were one male and one female, each weighing approximately 20 kg.

Experimental set-up

The basic experimental set-up has been described in detail elsewhere (Durham, Scherer, Titze, & Druker, 1987). Briefly, each larynx was mounted using a three-pronged device affixed to a micrometer, for the manipulation of inter-vocal process width. A separate micrometer was used to vary vocal fold length. During phonation trials, pre-established inter-vocal process widths were maintained using wooden shims, with shim thickness verified using a digital caliper (Mitutoyo Digimatic). When width and elongation had been established for a given trial, vocal fold oscillation and thus phonation were induced by delivering humidified, warmed air to the larynx using a laboratory air compressor and a ConchaThermIII heater-humidifier (Respiratory Care, Inc.). Subglottic pressures were established and monitored using an open-ended manometer (Dwyer No. 1211). For all experimental trials, vocal fold impact stress SI was measured in kPa at the anteroposterior midpoint of the membranous vocal folds. The midpoint was identified with an Ethilon 9-0 surgical microsuture positioned prior to data collection, using the digital caliper to guide placement. The suture was placed on the superior-medial edge of the vocal fold, but slightly on top so that the pressure sensor would not be in contact with the suture. Impact stress measures were made using a round, Precision Measurement Type 060 piezoresistive transducer (Ann Arbor, Michigan) with frequency response from DC to 50 kHz, and a dynamic range of 0-50 psi (approximately 0-345 kPa). Transducer diameter was less than 2 mm and thickness was about 1 mm. The transducer line was encased in a thin L-shaped metal cannula with plastic insulation for steady placement, with the transducer itself extruding from the encasement. The transducer was stabilized between metal prongs prior to data collection, after visually ensuring that the superior one-third of the transducer would extend above the vocal folds during oscillation. This positioning was chosen because deeper

embedding would have precluded visual verification of positioning and even risked extending the sensor subglottally. More superficial positioning would have limited the area of the sensor in contact with the oscillating folds. Impact stress signals were amplified by a Wheatstone bridge housed in the cannula. Calibration procedures conducted with a sphygmomanometer (Nissei D-267038) and Fluke 87 digital multimeter indicated response linearity between 0-40 mm mercury (0 - 5.26 kPa).

A three-electrode (SynchroVoice Inc.) electroglottograph was used to collect EGG data. After mounting the larynx, the two main electrodes were pinned to the thyroid cartilage alae at about the level of the cricothyroid muscle. The third, ground electrode was pinned to the canine trachea, which remained attached. Both impact stress and EGG signals were high-pass filtered at 10 Hz during data collection, using a Wavetek Rockland Dual Hi/Lo Filter, Model 432. The high-pass filtering did not distort relevant frequencies of the signal, which were greater than 10 Hz. A BectroVoice (ElectroVoice, Inc.) microphone was used to collect audio signals, including the experimenters' voices. Audio signals would be used as a chatter channel during data analysis, but were not otherwise pertinent to the experiment.

EGG and impact stress signals were monitored online during the experiment using a Tektronik 2212 60 MHz Digital Storage Oscilloscope. All signals, including EGG, impact stress, subglottic pressure, and acoustic signals were digitized at 10 kHz using anti-aliasing filtering and stored using a Sony PC-108M 8-channel Instrumentation Cassette (Digital Audio Tape, DAT) Recorder. All trials were video taped for supplementary visualization and for monitoring the larynx during data collection, using a Panasonic AG-1960 SVHS Hi-Fi MTS Multiplex Video Cassette Recorder with Panasonic WJ-810 Time Date Generator, Sennheiser microphone Model MD 441 U3, Pioneer Digital Stroboscope by Kernco, Model DS 303, and Sony Trinitron Model No. CVM-1900 Color Receiver/Monitor.

Experimental design and procedures

The experiment involved a 3 (vocal process gap) x 3 (vocal fold elongation) x 3 (subglottic pressure) within-subjects design. Data collection was thus attempted in each of 27 experimental conditions, for both larynges. Specifically, impact stress and EGG data were collected for each larynx at three glottal widths measured at the vocal processes: -1 mm, 0 mm, and 3 mm. The negative width or vocal process gap refers to a compression of the vocal processes relative to a barely touching configuration rather than tissue overlap; the vocal processes were pressed together to yield a micrometer reading that was 1 mm less than the reading when the tissue was barely touching (0 reading). Within each vocal process gap condition, data were further collected for each of three vocal fold elongations (resting length, 80% resting length, and 120% resting length). Finally, within each of the width/elongation conditions, data collection was attempted at each of three subglottic pressures (0.8-0.9 kPa, 1.4 kPa, and 2.0 kPa). In several cases periodic vocal fold oscillation was not obtained at the specified subglottic pressure. Thus, there were missing cells in the data matrix in the final results. Trial order was varied across larynges.

To provide some check on the extent to which the vocal vibratory pattern, and thus the CQ, might be affected by sensor placement, after all the primary experimental trials were completed, three trials for one larynx (Larynx A) were repeated without the introduction of the pressure sensor. Those trials were conducted with a vocal process gap of 1.0 mm and an elongation corresponding to resting length. All three subglottic pressure conditions were used (0.8-0.9 kPa, 1.4 kPa, 2.0 kPa). Further verification trials were not conducted in either larynx because of tissue deterioration and difficulty inducing oscillations after a relatively lengthy procedure.

Data analysis

For data analysis, impact stress signals were first high-pass filtered at 1 Hz, and EGG signals were high-pass filtered at 10 Hz using a Kemo Dual Variable Filter Type VBF/4 (DC) to control for signal drift. Both EGG and impact stress data were then digitized at 10 kHz using the CODAS/WINDAQ A/D conversion board (12-bit) and CODAS software. WINDAQ functions were used to display the signals. One segment of 20 stable EGG cycles was then identified for analysis. For the analysis, the CQ was determined using the horizontal-intersection method, and the 35% criterion described by Rothenberg and Mahshie (1988) to indicate vocal fold closure, with the signal oriented so that closing was upward. The average CQ (closed time in cycle divided by total cycle time) was calculated for the entire 20-cycle segment, using locally developed software. Peak-to-peak values in the AC impact stress signal (differences between maximum and minimum stress) were also computed by the software, and averages were obtained, for the same segment. Peak-to-peak measures were used to limit the influence of any possible DC drift. The paired values (average CQ and average impact stress SI) for the segment were then used to compute simple correlations between the parameters, within and across larynges. Correlations were tested for significance using two-tailed t-tests with a null hypothesis of $r = 0.0$. Regression equations were also computed.

Results

Potential effect of sensor placement on vocal fold vibratory pattern and CQ

Of the three reliability trials conducted both with and without placement of the pressure sensor between the vocal folds, two yielded complete, analyzable data sets: (1) glottal width = 1.0 mm, vocal fold elongation = 100% resting length, and subglottic pressure (Ps) = 0.8-0.9 kPa, and (2) glottal width = 1.0 mm, vocal fold elongation = 100% resting length, and Ps = 1.4 kPa. The third trial, performed with a Ps of 2.0 kPa yielded analyzable data only when the sensor was not present. For the two successful trial pairs at Ps = 0.8-0.9 kPa, the CQ was .78 both with and without the pressure sensor. At Ps = 1.4 kPa, the CQ was .77 without sensor, and .81 with sensor. Based on these limited observations, together with the relatively large number of missing data points in the data set in general (below), it appeared that the sensor might interfere with systematic vocal fold oscillations in some cases. However, when systematic oscillations did occur, the CQ did not appear strongly affected by the sensor's presence, based on two trials.

Relation between EGG CQ and SI

Raw data for CQ and impact stress (SI) measures are shown in Tables 1 and 2. The tables indicate that systematic vocal fold oscillations yielding interpretable data were not obtained in a large number of trials attempted. As already discussed, it is possible that the placement of the sensor transducer interfered with oscillations for those trials. However, also as discussed, when oscillations were ob-

Table 1.
Raw Data for Larynx A. (* Indicates Data Not Obtained)

Vocal Fold Elongation = 80% Resting Length				
Glottal Width	**-1.0 mm**	**1.0 mm**	**3.0 mm**	**Average**
Subglottic Pressure				
0.8-0.9 kPa	SI=1.05 kPa (CQ=.730)	SI=1.149 kPa (CQ=.655)	*	*
1.4 kPa	SI=0.955 kPa (CQ=.634)	SI=1.153 kPa (CQ=.697)	*	*
2.0 kPa	SI=0.573 kPa (CQ=.568)	SI=0.962 kPa (CQ=.651)	*	*
Average	**SI=0.859 kPa (CQ=.644)**	**SI=1.088 kPa (CQ=.668)**		
Vocal Fold Elongation = 100% Resting Length				
Glottal Width	**-1.0 mm**	**1.0 mm**	**3.0 mm**	**Average**
Subglottic Pressure				
0.8-0.9 kPa	SI=0.333 kPa (CQ=.560)	*	*	*
1.4 kPa	SI=0.671 kPa (CQ=.533)	SI=1.145 kPa (CQ=.807)	SI=0.346 kPa (CQ=.482)	$_{SI}$=0.721 (CQ=.607)
2.0 kPa	SI=0.699 kPa (CQ=.597)	SI=2.224 kPa (CQ=.776)	*	*
Average	**SI=.568 kPa (CQ=.563)**	*	*	
Vocal Fold Elongation = 120% Resting Length				
Glottal Width	**-1.0 mm**	**1.0 mm**	**3.0 mm**	**Average**
Subglottic Pressure				
0.8-0.9 kPa	*	*	*	*
1.4 kPa	SI=0.574 kPa (CQ=.509)	SI=1.084 kPa (CQ=.699)	SI=0.427 kPa (CQ=.457)	**SI=0.695 kPa (CQ=.555)**
2.0 kPa	SI=0.978 kPa (CQ=.559)	SI=1.603 kPa (CQ=.617)	SI=0.544 kPa (CQ=.390)	**SI=1.042 kPa (CQ=.522)**
Average	*	*	*	

tained, CQs appeared reliable, based on the limited number of reliability trials that could be conducted. Graphic relations between CQ and SI are shown in Figures 1-3. These figures display the data in the way that is most interesting clinically, that is, with SI as the dependent variable which might be predicted from the non-invasively measurable CQ.

By our own choice of interpretation, we identified three regions in the relationship between SI and CQ. The first region involved a threshold phenomenon ("t" in the figures). Within this region (below about 0.3 kPa for SI and .40 for CQ), neither SIs nor CQs were registered. A second region in the figures involved a debatable linear relation between the SI and the CQ. The boundaries of this second region were identified by computing a series of linear correlations, each one progressively eliminating outlying data points based on visual inspection, until the maximum correlation was shown. The corresponding data points were then considered as belonging to the region of linear relation. Recursive correlations obtained in this way were .83 for Larynx A, .96 for Larynx B, and .81 for the combined data). All of these correlations were significantly different from zero ($t(14) = 5.57, p < .001$ for Larynx A; $t(5) = 7.67, p < .001$ for Larynx B; $t(21) = 6.33, p < .001$ for the combined data). Finally, we identified a third region in the figures in which SI appeared unrelated to the CQ. In this area, SIs were sometimes large. This area of the function is indicated by "s" (saturation) in Figures 1-3, and occurred for CQs between about .70-.90. The Tables indicate that the largest SIs overall, which occurred in this region, were obtained for shortened vocal folds and large subglottic pressures (LarynxB).

Table 2.
Raw Data for Larynx B. (* Indicates Data Not Obtained)

Vocal Fold Elongation = 80% Resting Length				
Glottal Width	**-1.0 mm**	**1.0 mm**	**3.0 mm**	**Average**
Subglottic Pressure				
0.8-0.9 kPa		*	*	*
1.4 kPa		SI=4.715 kPa (CQ=.669)	SI=1.30 kPa (CQ=.566)	*
2.0 kPa		SI=5.340 kPa (CQ=.767)	*	*
Average	*	*	*	

Vocal Fold Elongation = 100% Resting Length				
Glottal Width	**-1.0 mm**	**1.0 mm**	**3.0 mm**	**Average**
Subglottic Pressure				
0.8-0.9 kPa	*	SI=2.173 kPa (CQ=.788)	SI=0.514 kPa (CQ=.517)	*
1.4 kPa	*	*	SI=0.929 kPa (CQ=.621)	* *
2.0 kPa	*	SI=3.623 kPa (CQ=.821)	*	*
Average	*	*	*	

Vocal Fold Elongation = 120% Resting Length				
Glottal Width	**-1.0 mm**	**1.0 mm**	**3.0 mm**	**Average**
Subglottic Pressure				
0.8-0.9 kPa	*	*	*	*
1.4 kPa	*	SI=2.755 kPa (CQ=.860)	*	*
2.0 kPa	*	SI=3.960 kPa (CQ=.847)	*	*
	*	*	*	

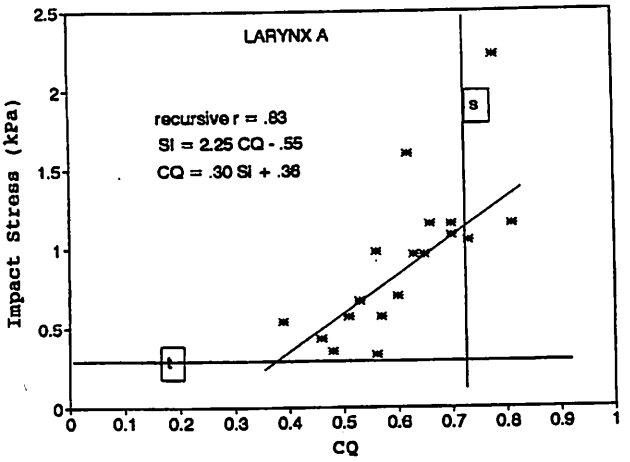


Figure 1. Data for Larynx A. X-axis indicates Closed Quotient (CQ, dimensionless unit). Y-axis indicates Impact Stress SI (kPa). Symbol "t" indicates threshold phenomenon for both variables; "s" indicates saturation in the CQ.

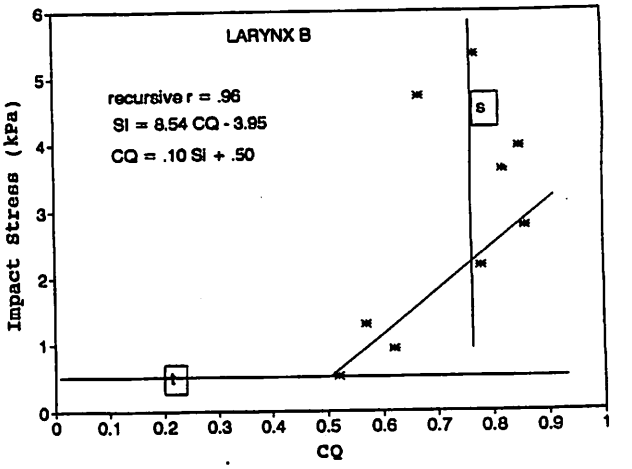


Figure 2. Data for Larynx B. X-axis indicates Closed Quotient (CQ, dimensionless unit). Y-axis indicates Impact Stress SI (kPa). Symbol "t" indicates threshold phenomenon for both variables; "s" indicates saturation in the CQ.

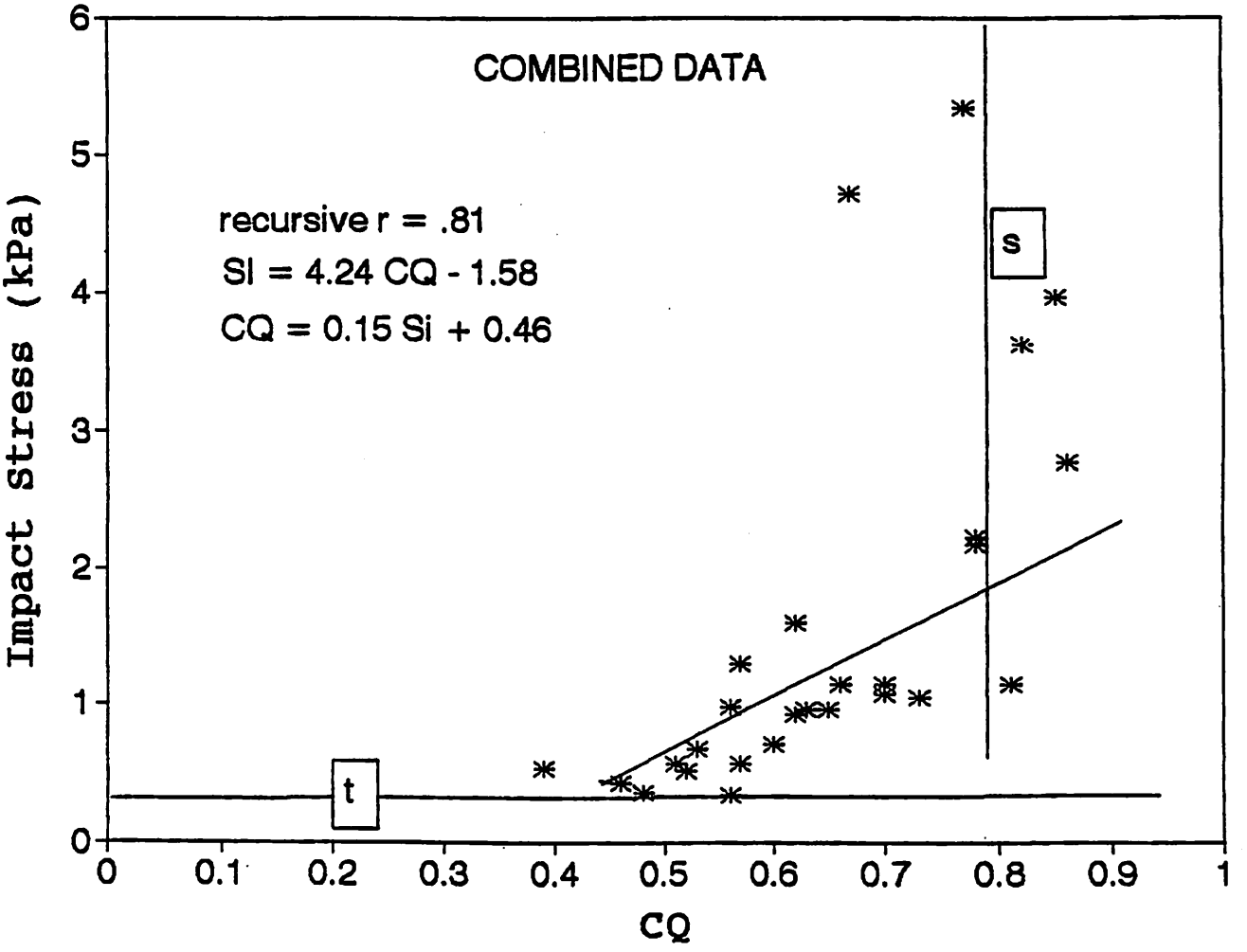


Figure 3. Data for combined larynges. X-axis indicates Closed Quotient (CQ, dimensionless unit). Y-axis indicates Impact Stress SI (kPa). Symbol "t" indicates threshold phenomenon for both variables; "s" indicates saturation in the CQ.

For both of the individual larynges, the largest SIs--again all occurring in this saturated region--were obtained with large subglottic pressures. Thus, when CQs approached saturation, apparently subglottic pressure was an important factor in influencing further growth in SI.

Regression equations for the linear portion of the SI to CQ function are also given in Figures 1-3. These equations show that within this linear region, an increase of about .15 in the CQ reflected an approximately 1 kPa increase in SI increase, combining all data. For the individual larynges, the specific values of CQ corresponding to a 1 kPa increase in the SI ranged from .10 (Larynx B) to .30 (Larynx A).

Relation of present findings to previous studies

Tables 1 and 2 show that across the experimental conditions assessed, SIs ranged from 0.333 - 5.34 kPa. These values are in the range of those reported in previously published studies on canines (for example, Jiang & Titze, 1994). An analysis of how SI varied with the experimental manipulations is not warranted for the present data because of the many missing data cells. However, inspection of the data in Tables 2 and 3 indicates that for the averages that could be computed for the different experimental cells, the results were not necessarily consistent with previous reports. For example, whereas prior studies have indicated increases in vocal fold impact stresses with increasing adduction, or vocal fold "pressing" (Jiang & Titze, 1994), in the present data this trend was not seen: An increase in adduction from 1.0 mm inter-vocal process width to -1.0 mm actually appeared to result in a decrease in impact stresses, from 1.088 kPa to 0.859 kPa (Table 1). However, the available average data for subglottic pressure manipulations did show an increase in impact stress with increased subglottic pressure, from 0.695 kPa SI at 1.4 kPa Ps, to 1.042 kPa SI at 2.0 kPa Ps (see Table 2), consistent with previous reports (Jiang & Titze, 1994). In general, the data are too scant to make any further comments about the variation in vocal fold impact stresses with the experimental manipulations and the experimental questions were not focused on this issue.

Discussion

In this experiment, three qualitatively different regions were identified in the SI/CQ function. Within the first region, below about .30-.40 for CQ and about 0.3 kPa for SI, neither CQs nor SIs were recorded. Apparently, the experimental conditions were insufficient to produce extremely low values for either variable, or perhaps there is little likelihood of registering the values under any experimental conditions because vocal fold contact is so minimal that electrical conductance is too poor to compute a CQ, and there is little if any impact stress at the sensor location.

In a second region of the function, clear recursive linear correlations were evident between the SI and the CQ. In this region, which involved CQs up to about .70-.90, a .15 increase in the CQ corresponded to an increase of about 1 kPa of SI, based on the combined data. The specific relation did appear to depend somewhat on the individual larynx, ranging from a .10-.30 increase in the CQ for a 1 kPa increase in the SI. Apparently, within this range of CQs, as vocal fold contact area increased between the vocal folds and also across the pressure transducer, vocal fold amplitudes and thus impact force increased even more and resulted in linear SI (force/area) growth.

Finally, in a third region of the function, for CQs ranging from about .70-.90, vocal fold contact area appeared to approach saturation. However, impact stress still demonstrated the capacity for growth, apparently regulated by subglottic pressure.

There was concern that placement of the pressure sensor between the vocal folds might disrupt or otherwise alter vocal fold oscillation, and thus CQs. In that case, the oscillatory patterns would have been unclear and the data's validity would have been questionable. However, only trials for which systematic oscillation occurred were evaluated in the data analysis. Furthermore, the small number of data points that assessed reliability indicated roughly similar CQs--and presumably similar oscillatory patterns--with and without sensor placement. For these reasons, it seems likely that for the trials included in data analysis, vocal fold behavior was at least generally similar to normal patterns.

The present results indicate some promise for further investigations of the CQ as a potential indicator of intra-vocal fold impact stress. If subsequent trials yield similar results, it is not excluded that the CQ might be used to non-invasively estimaterelative changes in SI for clinical purposes, with certain constraints. One constraint is that is seems unlikely that absolute CQs can predict absolute SIs for individual subjects, without first obtaining a subject-specific regression equation which would be invasive and defeat the goal of non-invasive measures; in the present experiments, regression equations were different across the larynges assessed. In clinical subjects, it is likely that such differences might be even greater. For example, in addition to factors such as neck tissue, thyroid angle, and vibrato affecting the CQ, the presence of pathology appears to reduce the CQ as compared with CQs obtained with similar laryngeal configurations in normal subjects (Verdolini, Druker, Palmer, & Samawi, in press).

Beyond these constraints and cautions, the present data suggest that changes in supra-threshold, non-saturated CQ may indicate relative changes in SI. Within this region of CQs (about .40-.70), an increase of .10-.30 in the CQ may reflect an increase of about 1 kPa in SI, at least for canine data. These data should be confirmed for animal as well as human subjects. In the interim it is interesting that this range

of CQs closely matches the range that CQs may vary across phonatory extremes in humans. In one previous human study of healthy adults, CQs differed by about 0.10-0.25 from extremely pressed to extremely breathy voice (Peterson et al, 1994). In another study of healthy adults and adults with nodules, CQs varied by about 0.10-0.15 across the same voice types, for both subject groups (Verdolini et al, in press). The point is that small changes in CQs--corresponding to perceptible differences in voice quality--may indicate clinically significant changes in SI.

In summary, data from the present study suggest that it may be reasonable to pursue the EGG CQ as a gross, non-invasive indicator of relative intraglottal impact stress within individual subjects, within a given experimental (or clinical) session. Other research should be conducted to assess the reliability of the present findings for canine and human subjects, and the potential applicability to clinical situations.

Acknowledgements

This project was supported by Grant No. P60 DC00976 and Grant No. 1 K08 DC00139-01 from the National Institute on Deafness and Other Communication Disorders and by the Joint Center for Otolaryngology, Harvard Medical School. The authors acknowledge the kind support of Ms. Gudrun Wegerich in data analysis, and Dr. Erich Luschei, Dr. Marvin Fried, and Dr. Manfred Gross in their facilitation of the project.

References

Baken, R.J., & Isshiki, N. (1977). Arytenoid displacement by simulated intrinsic muscle contraction. Folia Phoniatrica, 24, 321-336.

Durham, P., Titze, I.R., & Scherer, R. (1987). Development of excised larynx procedures for studying mechanisms of phonation. Technical Report VABL-1, Voice Acoustics and Biomechanics Laboratory, Department of Speech Pathology and Audiology, University of Iowa, Iowa City, Iowa.

Gray, S.D. (1991). Basement membrane zone injury in vocal nodules. In J. Gauffin & B. Hammarberg (Eds.), Vocal fold physiology conference (pp. 22-28). San Diego: Singular Publishing Company.

Gray, S.D., Hirano, M., & Sato, K. (1993). Molecular and cellular structure of vocal fold tissue. In I.R. Titze (Ed.), Vocal fold physiology: Frontiers in basic science (pp. 1-33). San Diego: singular Publishing Company.

Gray, S.D., Pignatari, S.N., & Harding, P. (1994). Morphologic ultrastructure of anchoring fibers in normal vocal fold basement membrane zone. Journal of Voice, 8, 48-52.

Gray, S.D., Titze, I.R., & Lusk, R.P. (1987). Electron microscopy of hyperphonated canine vocal cords. Journal of Voice, 1, 109-115.

Hess, M.M., Verdolini, K., Bierhals, W., Mansmann, U., & Gross, M. (in review). Endolaryngeal contact pressures. Journal of Voice.

Jiang, J.J., & Titze, I.R. (1994). Measurement of vocal fold intraglottal pressure and impact stress. Journal of Voice, 8, 132-145.

Kakeshita, T. (1927). Uber eine neue Methode zur Messung der beim Stimmbandverschluss wirksamen Krafte. [On a new method of measuring effective vocal fold closure forces.] Pflugers Arch. f.d. ges. Physiol d. Menschen u.d. Tiere, 215, 19-21.

Miko, T.L. (1989). Peptic (contact ulcer) granuloma of the larynx. Journal of Clinical Pathology, 42, 800-804.

Peterson, K.L., Verdolini-Marston, K., Barkmeier, J.M., & Hoffman, H.T. (1994). Comparison of aerodynamic and electroglottographic parameters in evaluating clinically relevant voicing patterns. Annals of Otology, Rhinology, and Laryngology, 103, 335-346.

Reed, C., Doherty, T., & Shipp, T. (1990). Piezoelectric transducers for the direct measurement of medial forces in the glottis. Journal of the Acoustical Society of America (Supplement 1), 88, S151-152.

Rethi, A. (1897). Die Stimmbandspannung, experimentell gepruft. [Vocal fold tension, experimentally assessed.] Sitzungsbericht d. Wien. Akad. d. Wiss., Mathem.-naturw. Kl. Abt. III, 106, 244. (Cited in Kakeshita, T.)

Scherer, R.C., Cooper, D., Alipour-Haghighi, F., & Titze, I.R. (1984). Vocal process contact pressure. In V. Lawrence (Ed.), Transcripts of the Twelfth Symposium: Care of the Professional Voice (pp. 61-68). NY: The Voice Foundation.

Scherer, R.C., Cooper, D., Alipour-Haghighi, F. & Titze, I.R. (1985). Contact pressure between the vocal processes of an excised bovine larynx. In I.R. Titze & R.C. Scherer (Eds.), Vocal fold physiology: Biomechanics, acoustics and phonatory control (pp. 292-303). Denver: The Denver Center for the Performing Arts.

Scherer, R.C., & Titze, I.R. (1982). Vocal fold contact stress during phonation. Journal of the Acoustical Society of America, 71,(S1), S55(A).

Scherer, R.C., Vail, V.J., & Rockwell, B. (1995). Examination of the laryngeal adduction measure EGGW. In K.S. Harris, F. Bell-Berti, & L.J. Raphael (Eds.) Producing speech: Contemporary issues (pp. 269-290). Woodbury, NY: American Institute of Physics. (1988).

Shipp, T., Reed, C., Doherty, E.T., & Morrisey, P. (1993). Current progress on transducing intervocal fold pressure during vegetative and phonatory behaviors. Paper presented at the Twenty-Second Symposium: Care of the Professional Voice. Philadelphia, Pennsylvania, June 7-12. Titze, I.R. (1994). Mechanical stress in phonation. Journal of Voice, 8, 99-105.

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

Titze, I.R., & Scherer, R.C. (1981). Measurement and simulation of vocal fold contact stress: Preliminary results. In V. Lawrence (Ed.), Transcripts of the Ninth Symposium: Care of the Professional Voice (pp. 21-25). NY: The Voice Foundation.

Verdolini, K., Berry, D.A., Titze, I.R., & Chan, R. (unpublished data). Vocal economy from a therapeutic perspective.

Verdolini, K., Druker, D.G., Palmer, P.M., & Samawi, H. (in press). Laryngeal adduction in resonant voice. Journal of Voice.

Part II

Tutorial reports

NCVS Status and Progress Report - 11
May 1997, 135-138

Benign Pathologic Responses of the Larynx

Steven D. Gray, M.D.
School of Medicine, The University of Utah

Traditionally, pathologic examination of laryngeal and other tissues have used staining techniques designed for cellular tissue. The area of pathology for most benign laryngeal lesions occurs in the lamina propria. This tissue is largely extracellular and thus the traditional staining techniques for pathology are not very useful. Laryngologists interested in the pathology of benign lesions have had to become familiar with tissue composition and tissue stains not learned in their conventional training. This is not altogether bad, since this knowledge has allowed us some insight into not only etiology of lesions, but also a better understanding of surgical planes and appropriate extent of surgical dissections. This chapter will first discuss the normal composition of the membranous vocal fold and then will review papers relating to pathology.

Normal Histology of the Membranous Vocal Fold

Epithelium

The surface of the vocal fold is lined with stratified squamous epithelium. The surface cells of the epithelium have a microridge appearance[1]. These laryngeal microridges seem to be extensions of cell processes, do not have specialized cellular functions, probably serve to help distribute laryngeal secretions, and may help in membranous fold contact. Their function is speculative. The remaining squamous cells are bound together with desmosomes (strong specialized intercellular attachment sites) until the basal cells are reached. The basal cells are attached to the basement membrane zone. Normally the squamous cells show no keratinization and do not show hyperkeratosis or parakeratosis.

Basement Membrane Zone (BMZ)

The BMZ is a complex arrangement of proteins designed to secure the epithelium to the lamina propria. (See Figure 1) The epithelium, as mentioned previously, is predominantly cellular and is securely tightly with desmosomes. The lamina propria is a loose, almost gelatinous tissue. Because of the diference in tissue composition, there is nothing for the cells of the epithelium to firmly attach to in the lamina propria. Also, the two largely different types of tissue creates biomechanical problems during vibration. The junction of the two tissues will be areas of great stress and potential injury.

To overcome the problems of binding two different tissue types together, the BMZ is arranged with horizontal layers of proteins. Figure 2 describes the BMZ as seen with electron microscopy. In laryngology only a few of the

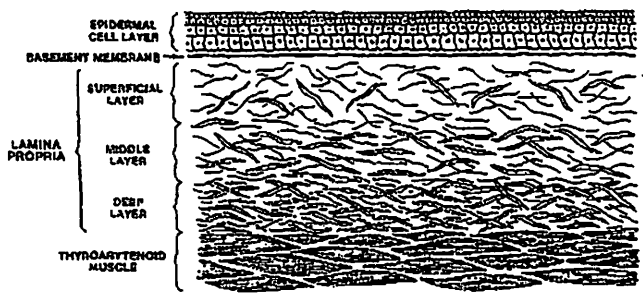


Figure 1. Arrangement of structure visible with the electron microscope from the basement cell to the fibers of the superficial layr of the lamina propria. AP - attachment plagues; DP - subbasal dense plate; AFL - anchoring filaments; AF - anchoring fibers.

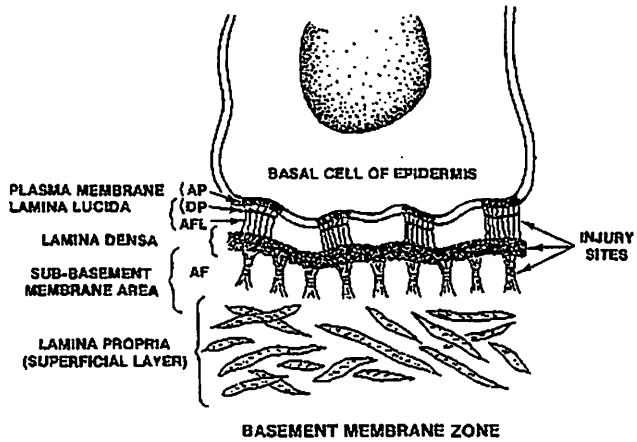


Figure 2. Basement membrane zone between epidermis and lamina propria.

proteins have been examined in normal and diseased states[2,3]. Currently the most studied proteins of the BMZ have been collagen type 4 and fibronectin. Collagen type 4 (a component of the lamina densa) has been useful to study because it occurs only in the BMZ and therefore is a very good marker of BMZ abnormalities. A wide band of collagen type 4 would indicate injury and thickening of the BMZ. Fibronectin occurs in the BMZ, but also is present in the lamina propria and therefore is less specific about the site of injury.

A comment about the significance of the exquisitely organized BMZ is appropriate. Each particular location of a protein in the BMZ is significant. Its orderly arrangement is designed to secure the epithelium to the lamina propria in a strong, stable, narrow band. If that arrangement is injured or interfered with, then integrity of the vocal fold to withstand stress maybe lost. Fortunately, the BMZ, when allowed to heal, is a prompt, thorough healer. Repair can occur within days. More permanent injury probably occurs when repair is not allowed due to repetitive occurrence. One more item should be mentioned about the BMZ but probably applies to the lamina propria as well. At least one of the protein components which secure the epithelium to the lamina propria has an individual genetic relationship. Anchoring fibers (collagen type 7) extend from the lamina densa into the superficial layer of the lamina propria. The population density (this means the number of fibers present within a certain area) of anchoring fibers is determined genetically[4,5]. Some of us have many,, some of us have few, those unlucky enough to have too few or none have a disease known as epidermolysis bullosa. This may or may not have import in vocal diseases, but it seems that continued research is leading to the conclusion that not all vocal folds are created equal and that this vocal fold composition likely has a genetic basis.

Lamina Propria

The lamina propria is mostly an extracellular matrix of proteins and carbohydrates with mild to moderate vasculature. Cellular constituents consist of fibroblasts which maintain the matrix of proteins, myofibroblasts which repair the matrix, and macrophages (some mast cells have been found) which are involved in inflammatory responses. Thus far most of the pathological responses have been defined by extracellular matrix changes and therefore we will focus on the matrix.

Extracellular matrix (ECM) is divided into two areas: the fibrous proteins and the filler or interstitial proteins. The fibrous proteins, which are the collagens and elastins, provide the scaffolding or framework of the ECM. Through study of the location and concentration of the fibrous proteins, Hirano separated the layered structure of the lamina propria into the superficial, intermediate, and the deep layer of the lamina propria[6]. The superficial layer contains few elastin and collagen fibers and in general is sparse in any mature fibrous proteins, the intermediate layer has an abundance of elastin fibers and the deep layer is characterized by the predominance of collagenous fibers.

The interstitial proteins provide the properties of viscosity, tissue and fluid flow and adhesion. The interstitial proteins are divided into the proteoglycans and the glycoproteins. Although many proteins are being investigated the more important ones seem to be fibronectin (glyoprotein) and hyaluronic acid (proteoglycan) [7]. Fibronectin is a protein laid down normally in all vocal folds, but is deposited more abundantly in response to injury. It has a very binding quality and is a precursor to scar formation. Hyaluronic acid is strongly implicated in being a key molecule in determining the hydration state of the vocal fold. Hyaluronic acid is a voluminous molecule which binds water and therefore helps establish tissue viscosity. Hyaluronic acid deposition is either partially genetically or hormonally influenced since there in a significant difference in amount of deposition based on gender, with males having a 3 times more per unit area[8].

Pathological Responses

Epithelium

Pathologic descriptions of the epithelium of benign laryngeal lesions has not been shown to be diagnostically useful. Various benign lesions have contained hyperkeratotic, hyperplastic and atrophic epithelium, irrespective of the clinical appearance[9]. Not all benign lesions have hyperproliferative responses of the epithelium. A interesting report by Thompson and Griffin looked at Langerhans cells, which are antigen presenting cells located in the epithelium, and function as part of the immune system in normal and pathologic laryngeal mucosa[10]. They found that a high number of Langerhans cells were present in vocal fold polyps. The significance of that finding is not known, but may suggest some immunological response as part of the pathogenesis of polyps.

Basement Membrane Zone

As mentioned earlier, because of different tissue composition, there will be areas of stress points when the tissue is put into vibration. One of these areas, the junction of the epithelium to the lamina propria or BMZ is superficial and therefore would be expected to receive considerable trauma during excessive phonation. Chronic BMZ injury was first described with electron microscopy showing BMZ reduplication and some disorganization[11]. The complete extent of BMZ injury has been better elucidated through studies using collagen type 4 and fibronectin. Basement membrane zone injury of considerable thickness, associated with abnormal protein deposition was described as a com-

mon type of patholoIgic pattern for some laryngeal lesions [3]. This pattern type was described to repetetive vocal abuse lesions such as nodules by Courey et al [2]. They looked at many laryngeal lesions and found that nodules had a thickened BMZ and dense fibronectin arrangement when compared to normals or polypoid lesions. The consequence of this injury is that the disorganized BMZ may leave the vocal fold in a predisposed state and the fibronectin deposition may lead to increased stiffening of that part of the membranous fold.

Courey et al also studied polypoid lesions and reported that these lesions generally did not show the increased BMZ thickening and showed less fibronection deposition than that found in nodules [2]. Others have reported increased BMZ thickening with polypoid lesions[12,13].

Lamina Propria

Most benign laryngeal lesions occur predominantly in the superficial layer of the lamina propria. Much of the pathological response of these lesions remains to be elucidated. It appears that nodules may have increased fibronectin deposition and a normal amount of vasculature structures [2,3]. This is opposed to the findings present in polyps or Reinke edema. Both of these latter lesions may have signs of recent hemorrhage, fibrin and edematous lakes. Dikkers and Nikkels indicate that "the combination of signs of recent bleeding, depositions of iron and fibrin, and thrombosis confirms the clinical diagnosis of polyp. The combination of epithelial BM thickening, edematous lakes, extravascular erythrocytes, and increased thickness of submucosal vessel walls confirms the clinical diagnosis of Reinke edema [12]." (Table 1)

Discussion

It is still speculative to determine initiating events leading to the formation of benign laryngeal lesions. However, it appears that the recent pathophysiologic papers regarding this topic are in agreement that nodule lesions show predominantly BMZ injury and increased fibronectin deposition. Courey states "these findings are consistent with current theories that nodules represent a response to **repeated** trauma that leads to BMZ thickening and a disordered fibronectin pattern secondary to attempted wound healing [2]". Dikkers and Nikkels suggest that "the trauma responsible for formation of nodules might be a tearing force...directed medially ...where amplitude is maximal." These small "high frequency tearing forces result in subepithelial" injury [12]. In nodules it seems that the response to injury is confined to the BMZ and the extracellular matrix of the superficial and perhaps intermediate layer of the lamina propria.

Polyps have less BMZ injury, less fibronectin deposition and more vascular injury. These findings suggest that polyp formation may be from a single, acute, event as well as repetetive injury. It also suggests that injury may occur primarily to the vascular structures of the lamina

Table 1.
Main Histopathological Features in Clinically Most Important Diagnoses

	Thick Basement Membrane		Edematous Lakes		Fibrin		Signs of Recent Bleeding		Iron		Inflammation		Thrombosis		Increased Vessel Wall Thickness	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Polyp (n=30)	18	60	21	70	18	60	16	53	11	37	15	50	7/26	27	13/26	50
Reinke Edema	13	81	14	88	5	31	12	75	1	6	8	50	1/15	7	9/15	60
Vocal Fold Nodules (n=11)	11	100	4	36	0	0	1	9	0	0	5	45	1	9	3/8	38
Unilateral Broad-Based Swelling (n=10)	9	90	7	70	0	0	5	50	0	0	4	40	0	0	1/6	17
Unilateral Broad-Based Swelling (n=9)	9	100	3	33	0	0	3	33	0	0	4	44	1	11	1	11

Table reprint courtesy:
Annals Publishing Company
4507 Laclede Avenue
St. Louis, MO 63108
(314) 367-4987

Dikkers FG, Nikkels PG: Benign lesions of the vocal folds: Histopathology and Phonotrauma. *Ann Otol Rhinol Laryngol* 1995, 104:698-703.

propria. Dikkers and Nikkels speculate that "the initiating trauma may cause capillary damage leading to edema, bleeding, and leakage of fibrin. Repair is then probably hampered by phonotrauma, inducing recurrent capillary trauma... [12]"

The pathogenesis and biology of Reinke edema, chronic erythematous conditions, chronic edematous conditions and disorders causing vocal fold stiffening (including scar) have not been well described. It appears that disorders causing chronic laryngeal edema are variable [13]. More focused studies on the interstitial proteins and on cellular activity in this conditions may be insightful. The current advances in sorting out the pathophysiologic mechanisms for nodules and polyps should be applauded by all interested in voice. This information allows better management of the vocal lesions, but also will help us determine risk factors for such lesions and subsequent prevention programs.

References of Special Interest

**Courey M, Shohet J, Scott M, Ossoff R: Immunohistochemical characterization of benign laryngeal lesions. This paper looks at Collagen Type IV and Fibronectin with immunohistochemical stains of clinically known benign vocal fold lesions. Precise measurements of the basement membrane zone were performed. The paper includes an excellent description of various lesions including nodules, polyps and cysts.

**Dikkers F, Nikkels P: Benign lesions of the vocal folds: Histopathology and phonotrauma. This paper looks at a variety of benign pathologic lesions with both electromicroscopy and histologic stains. This paper is strong because of the number of lesions examined and the excellent descriptions of the pathology. This paper is particularly good at looking at some of the vascular aspects of benign laryngeal lesions.

*Gray S, Hammond E, Hanson D: Benign pathologic responses of the larynx. This paper again uses Collagen IV and Fibronectin stains to assess laryngeal injury and response in benign laryngeal lesions. It is similar to Dr. Courey's paper with the exception that this paper's weakness is not knowing what type of clinical lesions they were examining, other than information provided by the surgeon and pathologist.

*Pawlak A, Hammond T, Hammond E, Gray S: Immunocytochemical study of proteoglycans in vocal folds. This paper is of interest because it describes more of the interstitial proteins present in the vocal folds. The paper does not address pathologic responses, but does address what types of proteins may be present in normal and provides a basis to search better pathologic states.

Figures reprinted with permission by Singular Publishing Group, Inc., 401 West "A" Street, #325, San Diego, CA 92101-7904 (Phone: 800-521-8545).

References

[1] Gray S, Hirano M, Sato K: **Mollecular and cellular structure of vocal fold tissue.** *Vocal Fold Physiology: Frontiers of Basic Science*, edition 1. Edited by Ingo R. Titze, M.D. San Diego, CA: Singular Publishing; 1993:1-34.

[2] **Courey M, Shohet J, Scott M, Ossoff R: **Immunohistochemical characterization of benign laryngeal lesions.** *Ann Otol Rhinol Laryngol* 1996, **105**:525-531.

[3] *Gray S, Hammond E, Hanson D: **Benign pathologic responses of the larynx.** *Ann Otol Rhinol Laryngol* 1995, **104**: 8-13.

[4] Gray S, Pignatari S, Harding P: **Morphologic ultrastructure of anchoring fibers in normal vocal fold basement membrane zone.** *J of Voice* 1994, **8**(1): 48-52.

[5] Tidman M, Eady R: **Evaluation of anchoring fibrils and other components of the dermal-epidermal junction in dystrophic epidermolysis bullosa by a quantitative ultrastructural technique.** *J of Investigative Derm* 1985, **84**: 374-377.

[6] Hirano M: **Structure of the vocal fold in normal and disease states. Anatomical and Physical Study.** *ASHA Report* 1981, **11**: 11-30.

[7] *Pawlak A, Hammond T, Hammond E, Gray S: **Immunocytochemical study of proteoglycans in vocal folds.** *Ann Otol Rhinol Laryngol* 1996, **105**(1): 6-11.

[8] Hammond T, Zhou R, Hammond E, Pawlak A, Gray S: **The intermediate layer: A morphologic study of the elastin and hyaluronic acid constituents of normal human vocal folds.** *J of Voice* 1997, **11**(1): in press.

[9] Van Der Velden L, Schaafsma H, Manni J, Link M, Ruiter D, Ramaekers F, Kuijpers W: **Cytokeratin and Vimentin expression in normal epithelium and benign lesions of the vocal cords.** *Acta Otolaryngol (Stockh)* 1996, 116: 325-331.

[10] Thompson A, Griffin N: **Langerhans cells in normal and pathologic vocal fold mucosa.** *Acta Otolaryngol (Stockh)* 1994, **115**: 830-832.

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

[12] ** Dikkers F, Nikkels P: **Benign lesions of the vocal folds: Histopathology and phonotrauma.** *Ann Otol Rhinol Laryngol* 1995, **104**: 698-703.

[13] Volic S, Kirincic N, Markov D: **Some morphological, histological, cytological and histochemical aspects of Reinke's Oedema.** *Acta Otolaryngol (Stockh)* 1996, **116**: 322-324.

NCVS Status and Progress Report - 11
May 1997, 139-152

Velopharyngeal Incompetence

Steven D. Gray, M.D.
Division of Otolaryngology/Head Neck Surgery, University of Utah Medical Center
Judy Pinborough-Zimmerman, Ph.D.
Salt Lake City, UT

Velopharyngeal Anatomy and Function

Normal velopharyngeal function varies depending on the type of activity or speech produced. Distinct patterns have been found in velopharyngeal closure between speech, blowing, and whistling versus dry swallowing and gagging[1]. Greater velopharyngeal movement appears to accompany swallowing than blowing and phonation[2]. The general consensus among researchers is that, physiologically, velopharyngeal movements in swallowing are different from those in blowing and phonation. Physiological differences in movement between speech and non-speech activities are consistent with clinical findings in patients who are readily able to obtain good velopharyngeal closure during swallowing but are unable to obtain adequate or obtain variable closure during speech.

Normal velopharyngeal function varies according to the characteristics of the speech produced. Factors such as vowel height, consonant type, proximity of nasal sounds to oral sounds, utterance length, speaking rate, and tongue height can affect velopharyngeal patterns.

High vowels tend to be associated with greater velar height than low vowels. For example, velar height generally is higher for the high vowels /I/ and /u/ than for the low vowel /ah/[3]. No consistent differences have been found, however, for the front/back or tense/lax features of vowel sounds[4]. Magnitude of velar elevation has been found to be generally greater during the production of the /s/ sound than of low vowel sounds[5,3].

During production of oral consonants and vowels, the velopharyngeal mechanism is generally closed. There may be incomplete closure during vowel production, particularly if the vowel production is in proximity to a nasal consonant. The English language has three nasal sounds: /n, /m/, and /ng/. Low- level activity is seen during the production of these nasal consonants, usually somewhere between a relaxed and complete closure position. The velopharyngeal port will therefore change its relatively open and closed states depending on the balance of oral versus nasal consonants occurring in the speech stimuli. (Fig. 1-1)

Normal velar function may vary widely in velocity and displacement with the demand of the speaking situation. Velar displacement generally decreases with increased speaking rate[6]. The degree of velar elevation, however, is not significantly influenced by speech loudness[7]. Individuals achieve competent closure of the velopharyngeal port through a variety of sphincteric patterns utilizing muscles of the soft palate and pharynx. Muscles contributing to the function of the velopharyngeal sphincter include the five muscles composing the substance of the soft palate: the tensor veli palatine, levator veli palatine, musculus uvulae, palatatoglossus, and palatopharyngeus. The relative contribution of a sixth muscle, the superior constrictor of the

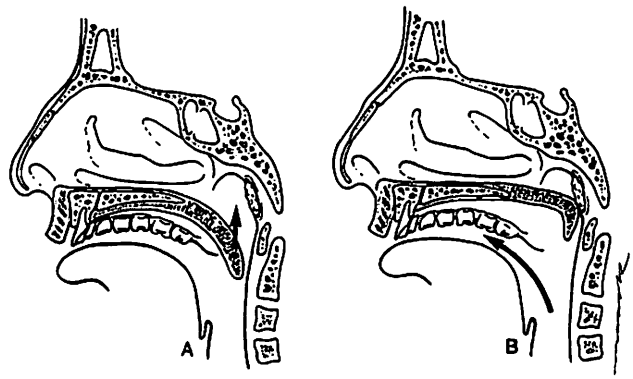


Figure 1-1. For pressurized speech sounds the flow of air needs to be channeled to the oral structures. This is done by raising the palate and sealing off the nose from the oral cavity. Velopharyngeal incompetence occurs when the velopharyngeal port is not sealed and leakage of air occurs into the nasal cavity, as shown in figure A. Figure B shows competency.

pharynx, has been debated. The most commonly held view is that this muscle is also influential in velopharyngeal closure[8].

During phonation, the velopharyngeal port closes when the velum moves in a posterior superior direction toward the posterior pharyngeal wall and the lateral pharyngeal walls move medially. In some individuals the posterior pharyngeal wall may move anteriorly. Normal velopharyngeal closure involves variability in movement. Table 1-1 presents a summary of the muscles involved in velopharyngeal function - their origination, attachment, function, and innervation.

The back and upward movement of the velum have been attributed to the action of the levator veli palatine muscle, the main muscle mass of the soft palate and the major elevator of the velum[9]. Contraction of the palatoglossus and palatopharyngeus muscles probably serves to pull downward on the velum, thereby opposing the upward pull of the levator[3]. Subtle adjustments in velar height have been attributed to the palatopharyngeus when the velum is in the elevated position[4]. Bulk to the dorsal side of the velum has been attributed to the musculae uvulae.

The contribution of lateral pharyngeal wall movement to closure varies among individuals, but generally has been found to occur during speech and is related to the speech context. The greatest pharyngeal movement has been reported to be at the level of the full length of the velum and hard palate well below the levator eminence[9]. It has been suggested that lateral movement is a result of selective contraction of the uppermost fibers of the superior constrictor muscle.

Passavant's ridge has been associated with active lateral pharyngeal wall motion. In some individuals it has been found to be the primary pharyngeal structure located on the posterior pharyngeal wall at the level of the velum. However, variations have been in the way Passavant's ridge is positioned relative to the velum. Findings suggest that in approximately one third of observed patients, Passavant's ridge is one of the primary pharyngeal structures at the level of the velopharyngeal closure[10]. Velopharyngeal closure may or may not be assisted by the presence of Passavant's ridge in some individuals.

In summary, six muscles of the soft palate and pharynx are involved in velopharyngeal closure. Normal closure varies among individuals and is expressed by the varying contributions of the velum and the lateral and posterior pharyngeal walls. Velopharyngeal function varies from individual to individual. Open and closed states of the velopharyngeal port are relative to the demands of the speech.

Table 1-1
Muscles of the Velopharyngeal Mechanism

MUSCLE	ORIGINATION	ATTACHMENT	FUNCTION	INNERVATION
Tensor Veli Palatine	Vertical portion arises from the scaphoid fossa at the base of the internal pterygoid plate; from the spine of the sphenoid and outer side of the cartilaginous portion of the eustachian tube	Terminates in a tendon which winds around the hamular process	Tenses the soft palate; opens the auditory tube during swallowing	Mandibular branch of the trigeminal nerve
Levator Veli Palatine	Arises from the under surface of the apex of the petrous portion of the temporal bone and from the inner surface of the cartilaginous portion of the eustachian tube; found to occupy the intermediate 40% of the length of the soft palate[11]	Fibers spread out in the soft palate where they blend with those of the opposite side	Acts as a sling when contracted to pull the velum in a postero-superior direction [5]; major elevator of the velum [6]; positions the velum [11]	Pharyngeal plexus derived from the glossopharyngeal and vagus nerve and the facial nerve [11]; course of the facial nerve is through the greater petrosal nerve [11]
Musculus Uvulae	Palatal aponeurosis in a circumscribed area posterior to the hard palate [12]	Inserts into the uvula	Adds bulk to the dorsal surface of the soft palate	Pharyngeal plexus; pharyngeal plexus derived from the glossopharyngeal and vagus nerve and the facial nerve[11]
Palatoglossus	Has a fan-shaped attachment from the anterior surface of the soft palate [11]	Courses through loose connective tissue within the anterior faucial pillar, and has a tapering termination in the side of tongue[11]	Elevates the tongue upward and backward to constrict the pillars, and probably lowers the velum [5]; positions the velum[11]	Pharyngeal plexus composed of branches from the glossopharyngeous and vagus cranial nerves and from the sympathetic trunk
Palatopharyngeus	Arises from the soft palate	Inserts with the stylopharyngeous into the posterior border of the thyroid cartilage	Adduction of the posterior pillars, constriction of the pharyngeal isthmus, narrowing of the velopharyngeal orifice, raising the larynx, and lowering the pharynx [4]; positions the velum[11]	Pharyngeal plexus
Superior Constrictor	Arises from the lower third of the posterior margin of the internal pterygoid plate and its hamular process	Into the median raphe	Medial movement of the lateral aspect of the pharyngeal walls [7]; high levels of activity are related to laughter [3]; may function to draw the velum posteriorly [3]; pulls the posterior wall and posterolateral angle anteriorly[11]	Pharyngeal plexus derived from the glossopharyngeal and vagus nerve and the facial nerve[11]

Velopharyngeal Closure Patterns

The motion of the velopharyngeal walls during closure have been classified into various patterns. Skolnick, McCall, and Barnes described three shapes of the persistent gap following closure: coronal, saggittal, and circular[11]. Croft, Shprintzen, and Rakoff studied closure patterns in both normal subjects and those with VPI[12]. They did not find a significant difference in the prevalence of particular closure patterns in normal subjects versus those with VPI. Using both multi-videofluroscopic views and nasoendoscopy, Croft et al. further refined these closure patterns. Four closure patterns are described based on the relative contribution of the pharyngeal walls or palate to velopharyngeal closure. (Fig. 1-2)

Coronal

A coronal pattern of closure, the most common, is present in about 55% of people. It occurs when the soft palate moving posteriorly is responsible for the majority of velopharyngeal closure. In this pattern the soft palate contacts a broad area of the posterior pharyngeal wall. Not very much medial motion of the lateral pharyngeal walls is required.

Sagittal

A sagittal closure pattern is identified when the majority of velopharyngeal closure results from medial motion of the lateral pharyngeal walls. The posterior motion of the soft palate is generally not a major component. This is the least common closure pattern, being used in only 10-15% of the population.

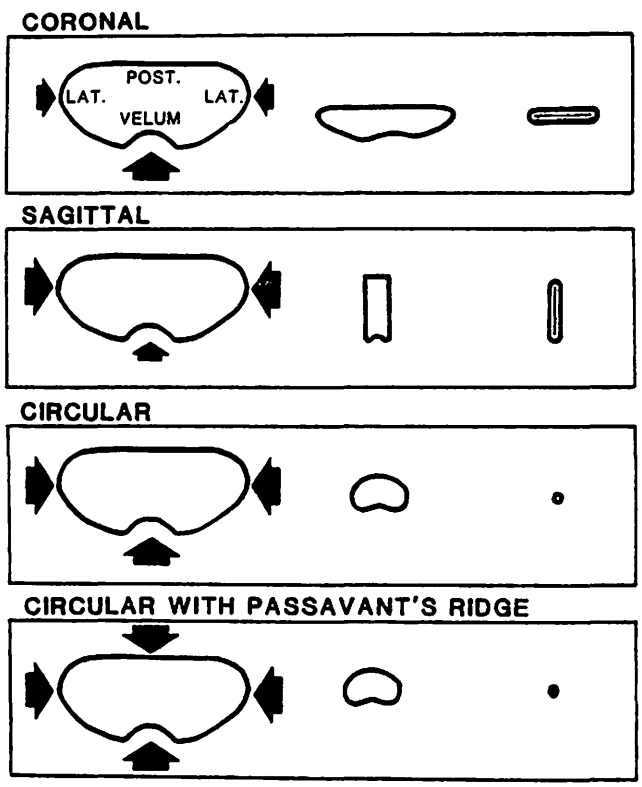


Figure 1-2. "Velopharyngeal Closure Patterns" Note that these descriptive patterns are based upon the shape of the gap as it closes.

Circular

A circular closure pattern occurs when both the lateral pharyngeal walls and the velum (soft palate) provide significant contributions to closure. In this closure pattern, not only is the soft palate making contact with the posterior surface of the pharyngeal wall, but also there is significant medial motion of the lateral pharyngeal walls so that the result of closure looks like a circle getting smaller. This pattern is present in about 10-20% of people.

Circular with Passavant's Ridge

Circular closure with Passavant's ridge is recognized when a circular closure pattern is present and the posterior pharyngeal wall moves anteriorly. This anterior movement or sphincteric action across the posterior pharyngeal wall is described as a Passavant's ridge. Therefore, this pattern of closure is descriptive of what occurs when all four pharyngeal walls are involved in closure of the velopharyngeal port. It is found in about 15-20% of the population.

Describing velopharyngeal closure patterns is clinically useful and in some cases may help decide the type of therapy required. When the velopharyngeal mechanism is dysfunctional, then an additional descriptive method is employed. Besides describing the velopharyngeal closure pattern, the relative contributions of the pharyngeal walls are rated. Although absolute measurements of the velopharyngeal gap or pharyngeal wall motion have been performed, these are difficult in a clinical setting. In 1988 an international working group proposed rating pharyngeal wall contribution to closure as a ratio of motion from one wall to the other[13]. Therefore, pharyngeal wall motion is graded by the degree of movement of that wall to the opposite wall. If complete closure occurs to the opposite wall, then a ratio of 1 is given. If motion toward the opposite wall is only halfway then a ratio of 0.5 is given. In this way pharyngeal wall dysfunction and resultant gaps in closure can be described. (Fig. 1-3, following page)

The vertical level of closure is an equally important anatomic location. In performing reconstructive velopharyngeal surgery, one would like to take advantage of existing velopharyngeal motion. Videofluroscopy or nasopharyngoscopy can be used to identify the point of maximal pharyngeal wall motion. This level of closure should be ascertained for each individual. This has generally been done using multi-videofluroscopy. Using this method, Finkelstein et al. evaluated the Passavant's ridge. They found it existed in about 40% of individuals and is formed by fibers of the superior constrictor muscle[14]. Using

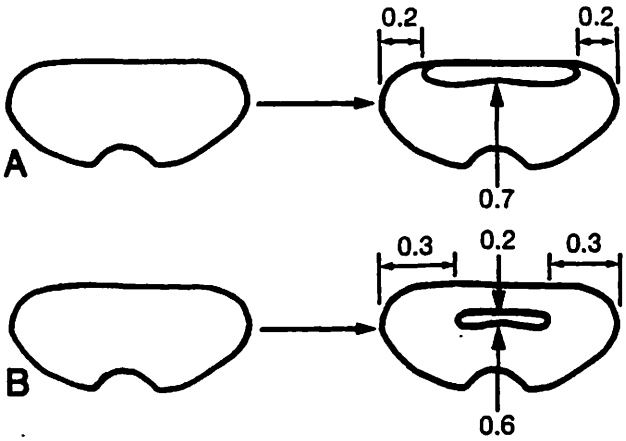
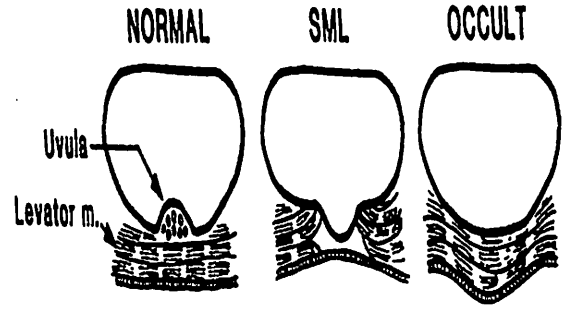


Figure 1-3. Describing velopharyngeal closure patterns is performed by first, naming the type of closure pattern involved and then showing with arrows the amount of motion present in each particular wall. A ratio of movement for the wall being described to the opposite wall is performed. In figure A this would be described as a coronal closure pattern with the velum showing 0.7 motion and the two lateral pharyngeal walls showing motion rated at 0.2. In figure B this would be rated as circular with posterior wall motion. One could also argue this as a coronal closure. This would be a very unusual closure pattern but is used here for descriptive purposes. In this case, the anterior wall shows 0.6 motion, both lateral pharyngeal walls show 0.3 and the posterior wall shows 0.2 with a resultant small coronal shaped gap.



Endoscopic view of Submucous Cleft Palate and 'Occult' Submucous Cleft.

Figure 1-4. This is self explanatory.

the torus tuberous as a landmark, they found that the upper limit of lateral pharyngeal wall motion was approximately 1 centimeter below the torus. The vertical length of closure can be variable. It can be fairly short, making it necessary to be precise about the vertical level of any reconstructive sphincter, or it can be longer, in which case precise placement may be less critical[18].

Causes of Velopharyngeal Insufficiency

Velopharyngeal insufficiency may be acquired ranging from paralytic conditions such as that encountered following strokes or base-of-skull surgery to anatomic congenital abnormalities of the palate, or congenital. Any interference with the velopharyngeal sphincter, may result in dysfunction.

Congenital anatomic conditions affecting velopharyngeal function are most commonly due to clefting of the palate and associated anomalies. Cleft palates occur in approximately 1 out of every 750 live births[15]. The incidence of VPI following repair of the cleft palate varies and depends partly on how a particular research group is defining VPI. However, it appears that about 50% of children with cleft palates have some VPI. Many of these patients can be successfully treated with speech therapy. Others require surgical intervention. Palatal clefting can be considered on a continuum ranging from frank, overt clefting of the palate to submucous clefting; an entity known as occult submucous cleft also exists. Weatherley-White et al. examined 10,836 school-aged children and determined the incidence of submucous clefting to be 1 in 1,200[16]. Most of these children were asymptomatic for VPI.

A submucous cleft is defined by a bifid uvula, muscular diastasis of the soft palate, and notching of the posterior border of the hard palate (loss of the posterior nasal spine). The bifid uvula and muscular diastasis are generally identifiable upon visual inspection. The notch in the hard palate requires palpation for identification. The presence of a bifid uvula alone does not constitute a submucous cleft palate, but only serves as a harbinger, and thus further inspection is warranted. Bifid uvulas have an incidence of 1 in 80 Caucasians[21]. An occult submucous cleft suggests that the usual triad of submucous cleft signs are not present. In the occult submucous cleft palate, there is absence of the musculus uvulae and diastasis of the levator palatini muscle. On the nasoendoscopic view, this finding is recognized as an absence of the bulge from the contraction of the musculus uvulae that is typically present on the nasal surface of the soft palate during speech. (Fig. 1-4) Many children with VPI, but no obvious cleft, have been identified as having occult submucous clefts.

Occult submucous clefts or submucous clefts predispose children toward VPI because of the abnormal musculature of their velopharyngeal sphincter. Consequently these children may acquire VPI when changes occur to the anatomy of the velopharynx, such as an adenoidectomy or even involution of the adenoid pad[17,18,19]. For this reason, a preoperative evaluation for adenoidectomy should include a careful inspection of the palate, both visually and with palpation[20].

VPI may occur following adenoidectomy even in children without a submucous cleft palate, but its occurrence in relatively healthy children can be hard to predict[21]. Huge adenoids may assist in easy velopharyngeal closure. When these adenoids are removed, considerable palatal motion may be required to obtain closure. Shorter palates, a relatively deep nasopharynx, or palatal hypotonia may all be

factors in the development of VPI following adenoidectomy. Ren, Isberg, and Henningsson showed that post-adenoidectomy VPI is often associated with enlarged tonsils or remaining prominent adenoid tissue[24]. Incomplete removal of adenoid tissue can be a cause of VPI. If concern exists that the child's palate may not be able to make the complete excursion back to the new position of the posterior pharyngeal wall, then a partial adenoidectomy can be performed. This has been described in various ways, but generally a superior half adenoidectomy is performed either with suction electrocautery or using a Thompson-Sinclair adenoid forceps or adenoid curette. If post-adenoidectomy VPI exists, then our current recommendations are for a 3 to 6 month trial of speech therapy. Useful adjuncts to this therapy are biofeedback using nasometry and occasionally C-PAP therapy for those with touch-closure VPI findings[22,23,24]. Surgically we have found both posterior wall augmentation and sphincter pharyngoplasty useful.

Many conditions or syndromes can result in or predispose children toward VPI. Furthermore, any condition that can affect the timing of the palate with speech, such as developmental delays or mental retardation, will also result in VPI. Velocardiofacial Syndrome and Down Syndrome warrant special comment.

Velocardiofacial Syndrome

Velocardiofacial Syndrome (VCF) is diagnosed with increasing frequency. This is a result of better awareness about the syndrome, as well as the availability of chromosome testing for diagnosis. This condition has rather typical facies: prominent nose, retruded mandible, cardiovascular anomalies, possibly cleft palate, and often learning disabilities. Inheritance is autosomal dominant with variable expressivity. Failure to thrive in infancy is present in about 25%, and about 35% have short stature. Obstructive sleep apnea may also occur. Mental disability is experienced by nearly all.

Facial appearance is fairly characteristic because of a long face with vertical maxillary excess, malar flatness, and mandibular retrusion. Robin's sequence may be present. The nose is usually prominent with a square nasal root and narrow nasal passages.

Cardiac anomalies are present in over 80%. Ventricular septal defect is the most common, but other cardiac anomalies such as a right-sided aortic arch, tetralogy of Fallot, and aortic valve disease are also found. Of significant consequence for the otolaryngologist is that the internal carotid arteries can be medially displaced along the posterior pharyngeal wall[25]. When considering velopharyngeal surgery in VCF patients, it is critical to consider internal carotid location. Often this is visible with nasoendoscopy. Rarely we have used radiographic techniques to precisely map the location related to the proposed surgery.

Diagnosis of VCF can be difficult. Finklestein et al., reporting on 21 patients with VCF, found that only 11 (52%) had typical manifestations. Fortunately, genetic testing is now possible for these patients. Micro-deletions of the q-11 region of chromosome 22 have been found in VCF patients. This is the same area in which micro-deletions found in Di George Syndrome occur. The micro-deletions in some patients with VCF Syndrome have encompassed the Di George chromosome locus[26]. Thus, patients with VCF have been found to have Di George Syndrome and vice versa. The possibility of both of these syndromes occurring together should be recognized by the otolaryngologist providing care for such patients.

Down Syndrome

Down Syndrome (Trisomy 21) occurs in 1.5/1,000 births and is associated with mental handicaps, hearing loss, and muscular hypotonia[27,28]. All of these can be risk factors for development of VPI. A review of 74 tonsillectomies and adenoidectomies in children with Down Syndrome identified two children with postoperative VPI[29]. These authors point out the risk of possible VPI development in children with known hypotonia and known articulation problems. Considering their predisposing factors it is somewhat surprising that post-adenoidectomy VPI does not occur more often. The predisposing factors may be offset by the narrow oropharynx and velopharynx present in Down Syndrome and also the shallow skull base at the adenoid site.

Nasopharyngoscopy

Flexible nasopharyngoscopy is an effective way of viewing the velopharyngeal orifice. The contribution of the pharyngeal walls can be visualized and nasopharyngoscopy allows fairly accurate analysis of the closure patterns and the deficiencies of velopharyngeal motion. Drawbacks of nasopharyngoscopy are that endoscopic images are prone to distortion and vertical level assessments may be more difficult[30]. Nasopharyngoscopy is usually performed so that both the video and audio can be recorded. Most children, when properly prepared, can tolerate nasopharyngoscopy. Topical anesthesia is used for the nose; however systemic sedation is generally not advisable. A speech pathologist is generally in attendance and helps perform the speech examination. Both nasal chambers can be examined, and then the nasopharyngeal area is examined during speech. Despite the shortcomings of nasoendoscopy, this method has been reliable in assessing VPI and is the standard method for viewing the nasopharyngeal area[31].

Non-Surgical Treatment

Successful treatment of velopharyngeal dysfunction depends on careful evaluation of both velopharyngeal function and speech. It is also influenced by the type and

severity of defect, client age, cognitive ability, motivation, availability of services, and expertise of service providers[32,33]. Integration of services provided by an interdisciplinary team of specialists is frequently required to develop and implement effective treatment plans. Treatment decisions must be based not only upon subjective practitioner impressions, but also should include data provided by objective procedures such as nasopharyngoscopy, videofluoroscopy, aerodynamic and/or nasometric studies, and in-depth articulation analysis[34].

For individuals who do not have the basic anatomy or function necessary to obtain adequate velopharyngeal closure for speech, surgical management is generally recommended[35]. These patients often present with consistent lack of velopharyngeal closure with varied speech demands and have hypernasality measurements in the moderate to severe range.

Non-surgical treatment options may be utilized prior to, in conjunction with, or in lieu of surgery. These options include, but are not limited to, prosthetic dental management, speech therapy, biofeedback treatment utilizing instrumentation, and resistance exercises performed while positive air pressure is introduced into the nasal cavities. To date, however, research data about the efficacy of these treatment options and factors affecting outcome are limited.

For the last 30 years, prosthetic devices have been used successfully to help assist with inadequate velopharyngeal closure[36,37]. These devices are particularly helpful when there are contraindications to surgery for significant velopharyngeal dysfunction, when dysfunction is secondary to neuromuscular involvement, or when velopharyngeal movements are poorly timed[38,39]. There are two basic types of prosthetic devices, obturators and palatal lifts.

An obturator is a device designed and constructed by a prosthetic dentist to substitute for missing or incomplete tissue; it is attached by dental wires or bands to the teeth[40]. Usually, the obturator used to improve velopharyngeal function is a thin acrylic plate which extends back into the pharyngeal area.

Customized fabrication and fitting of the obturator utilize data provided by nasopharyngoscopy, aerodynamic measurements, videofluoroscopy, and/or nasometric studies[41,42,43,44]. A gradual reduction in obturator size has been advocated for some patients; usually, this involves adjusting the size of the original obturator. It is postulated that as the size of the obturator is gradually reduced, velopharyngeal function may improve and, in some cases, eliminate the need for the obturator or for surgical intervention[42,44].

Unlike obturators, palatal lifts are designed to elevate the soft palate when the length of the palate is adequate. Lifts are particularly useful in patients with limited or no dynamic palatal movement. It is unclear at this time whether palatal lifts improve the neuromuscular ability of the velopharynx[45]. Use of obturators or palatal lifts can be difficult in young children, particularly if cooperation or cognitive ability is compromised.

Hypernasality, nasal emissions, and articulatory/phonological deficits can be the result of mislearning rather than VPI. These problems respond well to speech treatment, particularly phoneme-specific nasal emissions occurring with a select group of sounds. **Surgical intervention is contraindicated in these cases.** For children, diagnostic speech therapy is often recommended prior to physical management, particularly in cases of borderline, mild, or inconsistent hypernasality[45,46].

Speech and/or language treatment is strongly recommended for children with severe articulatory/phonological deficiencies and/or limited expressive language. Compensatory articulation error patterns, such as glottal stops, restrict airflow below the velopharyngeal valve[47,48]. Radiographic or nasopharyngoscopic observation of the velopharyngeal mechanism is appropriate for children who are able to demonstrate correct articulatory placement for some oral consonant sounds and who have enough language to provide a varied speech sample.

Diagnostic speech therapy frequently utilizes a microcomputer-based instrument called a Nasometer, which measures sound energy and has proven to be an effective biofeedback device used to reduce nasality levels in some patients[41,38]. Nasometer graphically calculates and displays the ratio of oral versus nasal sound energy or nasalance. Elevated nasalance scores on passages containing only oral sounds are significantly correlated to listener judgment ratings of hypernasality[49]. Reduced nasalance scores on passages containing primarily nasal sounds are correlated to listener judgment ratings of hyponasality[50]. In children with hypernasality nasometric scores are also significantly correlated to aerodynamic estimates of velopharyngeal area and to nasoendoscopic findings[55,51].

Treatment utilizing the Nasometer encourages such strategies as ear training, promoting a lower posterior tongue placement, increasing loudness, speaking at a lower pitch, using a more open mouth posture, controlling the type of consonants and vowels used in speech training, and slowing speech rate[41,49]. The duration of any speech therapy to reduce hypernasality should be carefully monitored. If continued improvement in hypernasality is not demonstrated within 3 to 6 months, speech therapy to reduce hypernasality should be discontinued and other treatment options explored[51].

Video-recorded nasopharyngoscopy is used successfully in biofeedback therapy for adults with VPI[29,52]. A case study using an older child demonstrated correction of phoneme-specific nasal emissions using nasopharyngoscopy[53]. Nasopharyngoscopy biofeedback therapy may prove to be a useful tool in training some children to achieve better velopharyngeal function, but further research is necessary.

A new therapy for hypernasality uses continuous positive air pressure(CPAP), which is introduced into the nasal cavities during speech production[54]. It is theorized that, in cases where hypernasality is related to physiologic limitations, resistance training might strengthen muscles involved in velopharyngeal closure. It has been demonstrated that as positive air pressure is delivered to the nasal cavities, levator veli palatini muscle activity increases[55]. Therapy can occur in the patient's home, a major treatment advantage. Preliminary case reports suggest that CPAP therapy is useful for patients with mild to moderate hypernasality, regardless of the underlying etiology.

Surgical Treatment

Pharyngoplasty

To improve velopharyngeal competence surgically, one may use operations which obturate the middle of the velopharyngeal area or the lateral portion of the velopharynx. Hynes was the first person to conceptually produce an operation which obturated the lateral part of the velopharynx[56]. His goal in designing this operation was to construct a sphincter which would be dynamic. He felt that raising lateral flaps of superior origin would lead to a neuromuscular flap which would contract. Nearly two decades later, Orticochea reported using lateral based flaps for the treatment of VPI[57]. In his description of the Orticochea pharyngoplasty, he used the posterior tonsillar pillars as donor flaps and also created a recipient site by raising a small inferior pharyngeal flap. Again, nearly one decade later, Jackson and Silverton modified the Orticochea technique[58]. It is essentially this later description of the pharyngoplasty that has gained so much popularity. The posterior tonsillar pillars are still used as the donor flaps to be sutured across the nasopharynx. However, an inferior pharyngeal flap is not created. Instead the transverse incision is made higher in the nasopharynx, and a small superior pharyngeal flap can be created if needed. Further modifications of this technique have occurred. Most of these have centered on getting an adequate superior placement and covering raw tissue areas. Several studies have assessed the success of a sphincter pharyngoplasty to improve VPI. In 1984 Riski et al. reported correction of hypernasality in 78% of patients[59]. In 1992 a more complete evaluation of the sphincter pharyngoplasty had similar success rates, although during the study the surgical technique changed[60]. An objective study was performed by Witt et al.; who found that nasal resonance improved in 79% of the patients[61]. In their study, the sphincter pharyngoplasty was employed to correct hypernasality in patients exhibiting all velopharyngeal closure patterns, instead of selecting out certain patients for other techniques. This study also appropriately documented potential problems with all types of velopharyngeal obturating surgery. They found that 30% of their patients were judged to be hyponasal postoperatively. Sloan et al. compared a modified Hynes pharyngoplasty with inferiorly and superiorly based pharyngeal flaps[62]. They concluded that the sphincter pharyngoplasty was a better technique for management of VPI.

Indications

A sphincter pharyngoplasty is a good choice when one of the following criteria are found: closure of the persistent velopharyngeal gap in which there is poor lateral-wall motion and good velar motion or significant augmentation of the posterior pharyngeal wall. A sphincter pharyngoplasty may also be indicated for those with poor motion of any of the pharyngeal walls, including the velum. These people have a relatively adynamic sphincter, and typically their velopharyngeal port is described as a black hole because there is no light reflection of the sphincter during nasoendoscopy exam[63]. Since the sphincter pharyngoplasty pulls the lateral pharyngeal walls medially and since the flaps are sewn across the back wall, both the lateral and posterior portions of the velopharyngeal port are obturated. The degree of obturation can be varied by the thickness of the flaps, the snugness or tightness with which the flaps are sutured together, and the extension of the flaps toward the uvula.

Technique

Proper patient preparation includes having recently reviewed the nasoendoscopy tape and videofluroscopy if performed. If possible, a recording of the nasoendoscopy should be present in the surgical suite so this may be reviewed at the time the operation is being performed. The need for a possible tonsillectomy or adenoidectomy should also have been previously determined. If proper positiong and placement of the sphincter will be through the adenoid pad, an adenoidectomy should have been been performed. Placing a sphincter pharyngoplasty through the adenoid can increase the difficulty of the operation. We prefer to do an adenoidectomy approximately eight weeks prior to placement of the sphincter pharyngoplasty. The same is true for a tonsillectomy. If the tonsils are so large that they preclude obtaining a healthy palatopharyngeus muscular flap, then a careful tonsillectomy is performed at approximately eight weeks prior to surgery. In doing this tonsillectomy, care is taken not to injure the posterior tonsillar pillars and to induce minimal scarring of the palatopharyngeus muscle.

Following intubation with an oral Ray tube, a Dingman or Davis mouth gag may be used to obtain adequate exposure of the posterior tonsillar pillars. Exposure to the nasopharynx is obtained by using a uvula retractor. With the soft palate retracted superiorly, the proper vertical level for the sphincter pharyngoplasty is determined by comparing anatomic landmarks found either in the video

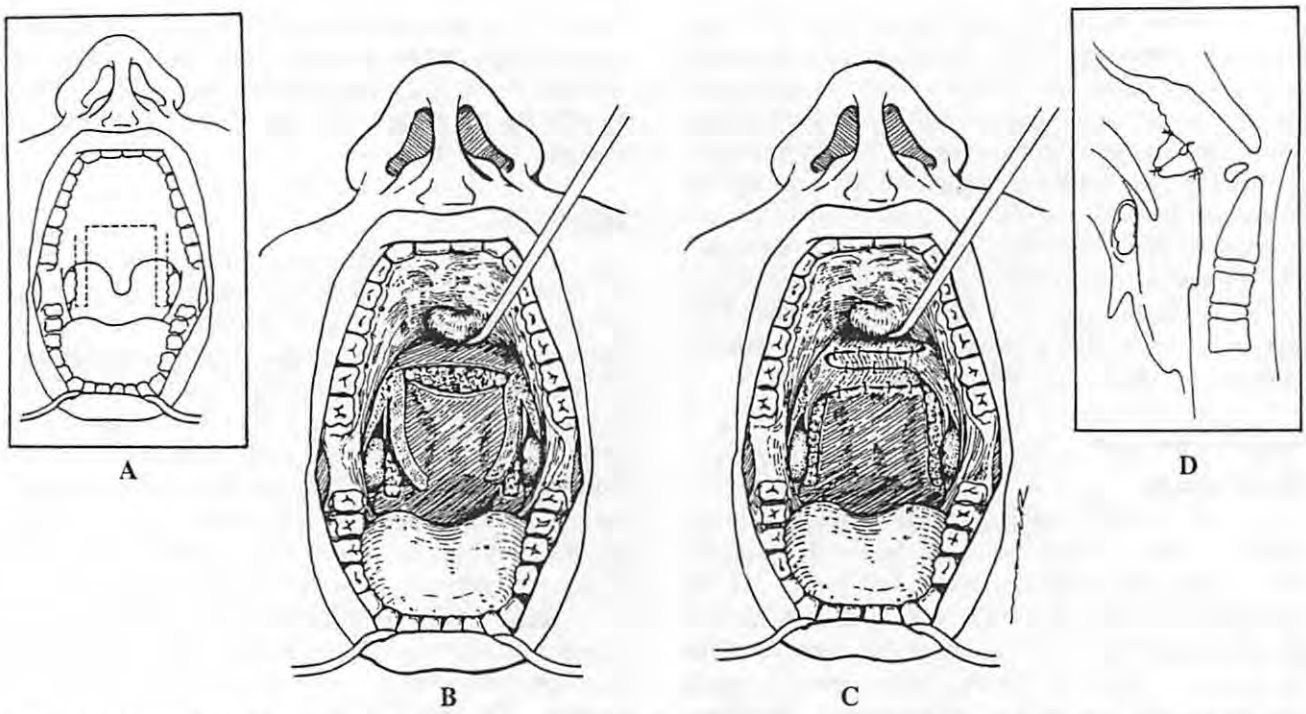


Figure 1-5. The inset figure A is an attempt to show the height or vertical level the sphincter usually needs to be placed. The soft palate needs retraction. The broken dash lines are to represent an approximation of the surgical incisions, however this is difficult to represent two dimensionally. Figure B shows the soft palate retracted superiorly, the posterior tonsillar pillars (palatal pharyngeus muscle) have been elevated. The illustration is attempting to demonstrate that the anterior incision stops at the superior pole, but may extend further vertically and medially along the nasal surface of the soft palate toward the uvula if the surgeon desires to create a narrow, tight port. Note also that the medial or posterior incision for the flaps extends up to the horizontal incision so that the flap can be easily rotated. In Figure C the flaps have been sutured into the sphincter position overlapped. As mentioned in the text, the flaps may also be sutured end to end. This illustration better demonstrates that the sphincter could be adjusted closed by pulling the flaps inmore tightly, by creating a larger posterior wall ridge (harvesting larger and thicker flaps), and by extending your anterior incision medially along the nasal side of the soft palate (this step is usually not needed and can lead to severe narrowing of the port). Figure D, the inset, shows the lateral view following the pharyngoplasty. Note that the sphincter is in the nasopharynx, above the soft palate.

fluoroscopy or nasoendoscopy with those present by viewing or palpating the nasopharynx directly in surgery. After this site is determined it is injected with a vasoconstrictive agent such as 1% Xylocaine 1:100,000 Epinephrine. The posterior tonsillar pillars are likewise mildly injected to improve hemostasis. A #15 blade is used to make an incision along the anterior surface of the posterior tonsillar pillar. The posterior tonsillar pillar is detached inferiorly just above the inferior pole of the tonsil. A combination of a #15 blade and lower lateral crural right-angle scissors is used to elevate the palatopharyngeus muscle superiorly. The posterior incision, which is close to the junction of the lateral and posterior pharyngeal wall, is carried superiorly to the level of the desired vertical placement across the nasopharynx. Then the palatopharyngeus flaps are elevated up to nearly this level. It should be mentioned that the anterior incision does not extend as superiorly as the posterior incision. In fact, the anterior incision of the palatopharyngeus muscle usually does not extend even as high as the superior pole of the tonsil, since in so doing the reconstructed velopharyngeal port can become quite narrow. (Fig. 1-5)

The further superiorly and towards the uvula the anterior incision extends, the more narrow the reconstructive port will become. If the intent of the surgeon is to keep the velopharyngeal port fairly open, but at the same time create significant obturation of the lateral and posterior portion of the nasopharynx, then the anterior incision can be extended across the back of the tonsillar pillars and curved up along the posterior and superior (nasal) surface of the soft palate/pharynx. This type of an extension allows the palatopharyngeus flap to be elevated more superiorly without necessarily constricting the velopharyngeal port.

The two palatopharyngeus muscles can also be elevated superiorly with some blunt dissection (using a Herd dissector) or finger dissection. With the two palatopharyngeus muscles elevated, an incision is made across the nasopharynx at the desired level of the sphincter placement. This usually requires palatal retraction with the uvula retractor. The palatopharyngeus flaps are sutured transversely across the nasopharynx. These flaps can either be sutured end to end or in an overlapping fashion. Vicryl 4-0 is used to suture these flaps together and to the nasopharynx. If a very snug sphincter is created, then an obstructive breathing pattern may occur in the immediate postoperative state. In this case we place a soft nasal trumpet through the sphincter until the following morning.

Pharyngeal Flaps

An operation designed to obturate the midline of the velopharyngeal port is termed a pharyngeal flap. In this procedure, a flap of tissue is pedicled superiorly or inferiorly to the posterior pharyngeal wall and then sutured to the palate. Through the years, success has been reported using both the inferiorly and superiorly based pharyngeal flap. Placement in the midline divides the velopharyngeal port into two smaller lateral ports. Mobile lateral pharyngeal walls are needed for ports to be open during breathing and closed during speech. Posterior-wall and velar motion are relatively unimportant when choosing this procedure.

In the last decade preoperative nasopharyngoscopy has facilitated better preoperative planning. More attention has been focused on the proper vertical placement of flaps. One of the criticisms of inferiorly based pharyngeal flaps is that the pharyngeal flap is tethered inferiorly and that with additional scar contracture the flap will be pulled out of optimal position[64]. For this reason, in the United States, the trend has been towards superiorly based pharyngeal flaps.

Pharyngeal flaps have been effective in reducing hypernasality. Morris et al. studied 65 subjects, more than half of which were more than ten years post surgery[65]. Eighty-three percent showed normal velopharyngeal function, and 66% showed normal or near normal speech production. Eight of the sixty-five subjects had surgical revision of the flap. One required reduction in the size of lateral ports; the other seven had either flap takedowns or lateral port enlargements performed. Likewise, Hirschberg reported a series of 500 pharyngeal flaps where 90% of the patients demonstrated improvement in hypernasality and 74% showed improved speech[66]. It was also emphasized that the best speech results were obtained when the width of the flap was appropriately matched up with relative motion of the lateral pharyngeal walls. Matching the width of the flap to obturate the remaining gap area, which is determined by maximal medial motion of the lateral pharyngeal walls, is termed tailoring of the pharyngeal flap[67]. In theory, tailor-made flaps are a good concept. In practice, achieving precise millimeter dimensions in the lateral ports can be difficult due to the many forces of scar contracture. Inconsistency of scar contracture between individuals leads to some of the complications which are reported regarding pharyngeal flaps.

Many good long-term studies have thoroughly documented the drawbacks and benefits of this procedure[68]. Sleep apnea is a reported complication from pharyngeal flaps[69,70]. Sleep apnea is more likely to occur in the early postoperative period, and within three months obstructed sleep problems should be much resolved[71,72]. In a long-term follow-up, Morris et al. did ECG's on 33 of the 65 subjects in the pharyngeal flap study[71]. Those chosen for ECG monitoring had complained of sleep difficulty or had a history of snoring. No subjects had clear findings of right ventricular hypertrophy, which can be a pathological response to obstructive sleep apnea. Their conclusion was that although symptoms of nasal airway obstruction may be present, they are rarely serious enough to cause pathology.

Indications

Indications for a pharyngeal flap are patients with relatively good lateral wall motion but a central persisting gap. Patients with a sagittal closure pattern would also be candidates for this procedure.

Technique

The pharyngeal flap is performed by placing a Dingman mouth gag and injecting the posterior pharyngeal wall and the palate with 1% Xylocaine 1:100,000 Epinephrine. Two different procedures can be used for performing the pharyngeal flap. One, termed a fish mouth technique, is essentially by-valving the soft palate (separating nasal mucosa from oral tissue) to create a pocket into which the pharyngeal flap is pulled and sutured[73]. The pocket is usually placed just superior and anterior to the edge of the soft palate on the nasal side. The other technique involves splitting the palate and has been well described by Hogan [74,75]. Both of these papers describe adjustments to the lateral ports.

The width of the flap is decided by previous nasoendoscopy or videofluroscopy. The superiorly based flap is elevated a little higher than the level ascertained to be the appropriate level for maximum medial motion of the lateral pharyngeal walls. With scar contracture, the flap usually descends a couple millimeters. The fish mouth technique or the palatal splitting technique is used to sew the flap into position. Prior to suturing the flap, small endotracheal tubes are placed in both ports to assist in postoperative breathing and to provide some guidance on lateral port size.

Postoperatively, after the small endotracheal tubes have been removed, the patient should be monitored for obstructive apnea. The tubes are usually left in place the first night, and the patient should not be discharged from the hospital until indications are clear that serious obstructive apnea is not present. Parents need to be cautioned that if obstruction seems to be getting worse, the patient should be brought back to the hospital for monitoring and reevaluation.

Augmentation of the Posterior Pharyngeal Wall

The concept of augmenting the posterior pharyngeal wall in an effort to improve velopharyngeal competence has received and will continue to receive attention for a number of reasons: (1) the risk of inducing hyponasality and nasal obstruction is theoretically lower, (2) relatively easy to perform most of the procedures, (3) the majority of persistent velopharyngeal gaps are continuous with the

posterior pharyngeal wall, and (4) in many VPI patients the velum or soft palate has relatively good motion. The concept of simply bringing the posterior pharyngeal wall forward to correct the gap in many VPI patients is appealing. In practice, many posterior pharyngeal wall augmentation techniques have failed to meet expectations. Currently one of the best techniques to augment the posterior pharyngeal wall is a sphincter pharyngoplasty since a large shelf of tissue is placed across the posterior pharyngeal wall. However, this section focuses on those techniques which augment the posterior wall.

Various materials have been used to augment the posterior wall, either by injection or implantation[76,77,78]. Tissue implants, such as cartilage and fascia have been used.[79,80,81,82]. The most common injectable material over the last few decades has been Teflon®[83]. To treat marginal VPI, Smith and McCabe injected Teflon® into the posterior pharyngeal wall creating a transverse ridge[84]. They also discuss the possible necessity of repeated Teflon® injections. A disadvantage to Teflon® is that migration may occur and that the material is not approved by the Federal Drug Administration for such use. One inadvertent carotid artery injection with Teflon® has occurred. Obviously it is important to palpate the posterior wall for great vessels prior to any sort of posterior pharyngeal wall procedure. Conversely an anecdotal advantage of Teflon® injection is that as the Teflon® migrates and the posterior wall bulges less, the velopharyngeal sphincter may slowly make adjustments to maintain competency, much like sphincter improvement in obturator reduction.

Because of the continuing theoretical advantages of posterior wall augmentation, renewed interest has occurred in using posterior wall tissue to create a shelf or ledge along the posterior wall. In this procedure a superiorly based pharyngeal flap is lifted up and buckled, so as to create a ridge across the posterior wall.

Indications

Posterior wall augmentations are selected when there is a small central midline gap and velopharyngeal closure could be complete by simply bringing the posterior wall forward. Occasionally, because patients may achieve velopharyngeal closure, but because the closure is not tight enough, pressure causes leakage of the air to occur through the velopharyngeal port. This condition is referred to as a "touch closure" problem because the pharyngeal walls touch together but do not achieve competent closure. Under these conditions a slight augmentation of the posterior pharyngeal wall is enough to provide a competent seal. Another indication for posterior wall augmentation can be VPI problems following adenoidectomy. If significant augmentation of the posterior pharyngeal wall is required, then a sphincter pharyngoplasty is usually a better technique. Posterior wall augmentation works best for gaps in the area of 1-3 millimeters or for "touch closure" problems.

Technique

Posterior wall augmentation using a buckled or folded pharyngeal flap is relatively simple and easy to

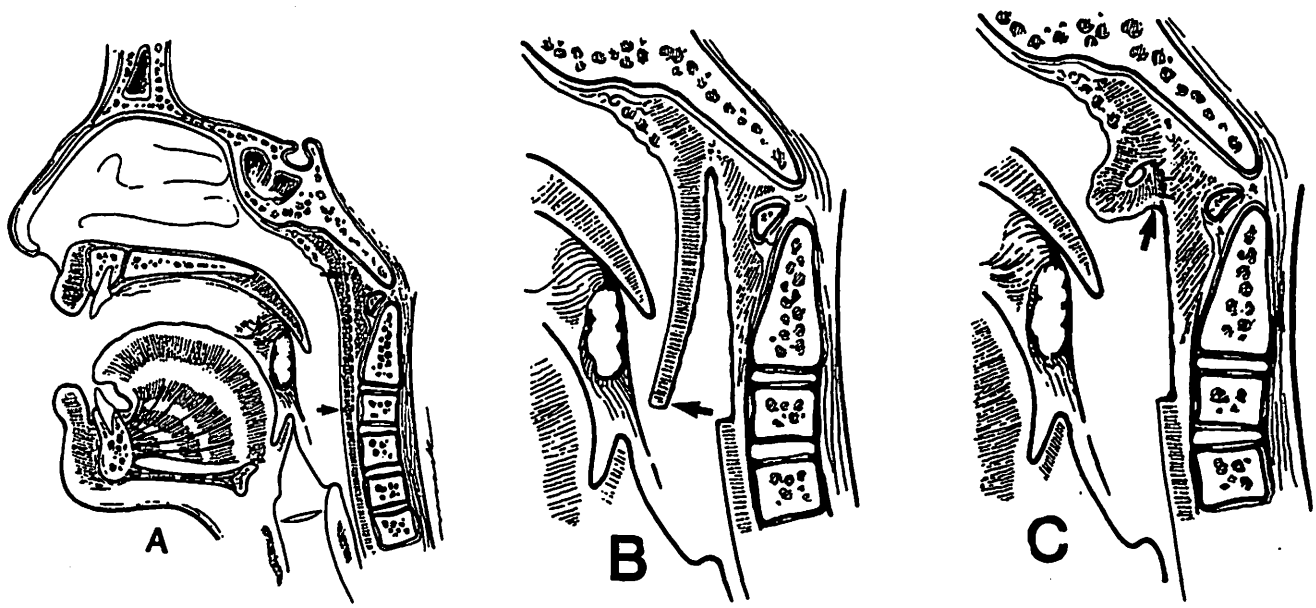


Figure 1-6. This figure demonstrates the relatively simple but effective procedure to augment the posterior pharyngeal wall. The section of the posterior pharyngeal wall to be elevated is shown in A, and shows the inferior level to be at about the inferior pole of the tonsil. Note that in figure B the constrictor muscle is raised with the flap and is elevated quite superiorly. In Figure C the flap has be folded and sutured into place. Even though the flap was elevated higher, the fold and maximum area of augmentation occurs a little lower. Although not shown in figure C, sutures are often placed on the folded flap laterally to close some of the dead space and sometimes it seems to create a little more stability.

perform. A superiorly based pharyngeal flap is raised and then folded on itself at the level of maximum pharyngeal wall motion in the area of the velopharyngeal port. This procedure is done by placing a Davis mouth gag and injecting the posterior pharyngeal wall with a vasoconstrictive solution. Then a superiorly based pharyngeal flap is elevated such that the constrictor muscle is present within the flap. The width of the flap is decided by ascertaining the width of the gap to be obturated and making the flap slightly larger. The flap does shrink once it is elevated and therefore a wider flap than anticipated is usually required. (Fig. 1-6)

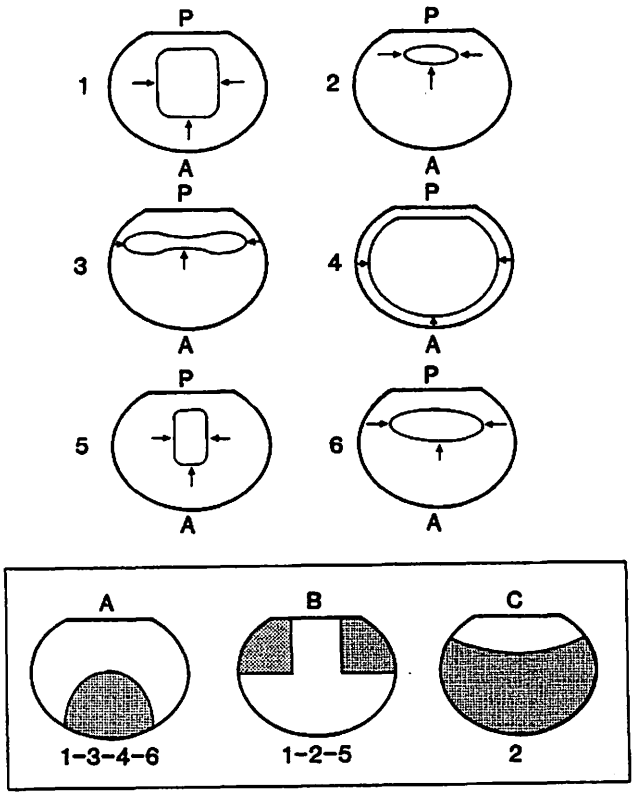


Figure 1-7. A-C represent the type of obturation one may expect with the three surgical methods described in the chapter. The shaded areas represent parts of the velopharyngeal area left open following the surgery; the white areas show what is obturated with tissue during the surgery. Remember that in all of these examples there are variations and with surgery you can obturate more or less than is demonstrated. Figure A shows the type of obturation obtained with a sphincter pharyngoplasty. Significant augmentation and closure occurs of the posterior and lateral pharyngeal walls. The area left open is midline and located just behind the uvula. This helps demonstrate why someone with good soft palatal motion will generally do well with a sphincter pharyngoplasty. Looking at figures 1-6, a sphincter pharyngoplasty would work well in treating figure 1 (this sphincter would need to be a little bit snug), figure 3, figure 4 (black hole patterns are usually best treated with a very snug or tight sphincter or with an prosthetic obturator), and figure 6. Figure B shows the obturation following a pharyngeal flap attachment to the soft palate. Since the lateral areas (ports) are left open, it is apparent why this surgery works best in people with good lateral wall motion. Matching the obturated pattern over the figures 1-6, it can be seen that a good match could be figure1 (wider pharyngeal flap would be needed), figure 2, and figure 5. Figure C shows that the folded or rolled pharyngeal flap mildly augments the posterior wall. In matching this pattern to figures 1-6, the best match is figure 2. Although figure 2 could be treated by any of the surgical methods, augmentation of the posterior wall is the least aggressive and leaves the most unobturated area. Therefore it is less likely to cause airway obstruction.

The lateral incisions are made first and carried through the constrictor muscle until the white fascia is encounteredposteriorly. At this point, a right-angle elevator can be used to elevate the remaining muscle off the fascia. The flap is detached inferiorly at about the level of the inferior pole of the tonsil. The flap is elevated superiorly to slightly above the level of the desired augmentation. When the flap is elevated above the level of augmentation, the buckle of the flap will be correctly positioned. The flap can then be sutured on itself or to the posterior wall so that the fold is in the desired location vertically. Usually 3-4 sutures are placed across the inferior edge of the flap to hold it in proper position, and then 1-2 sutures are placed on the lateral aspects of the folded flap for stability and to close any dead space between the two layers of the flap.

One of the drawbacks to this procedure is some variability in flap atrophy. Another drawback is that even though large buckles can be created, they are functionally not very useful since a large buckle has a tendency to flap (move up and down) depending on the amount of pressurized air. For this reason, small gaps of 1-3 millimeters can be closed successfully, but large gaps are inconsistently closed since a large folded flap is not vertically stable in its position.

Summary

In conclusion, treatment of VPI requires careful evaluation by a team of professionals, consisting of a speech and language pathologist and a surgeon, who have expertise in speech disorders. Care needs to be individualized. Surgical care requires preoperative assessment of the velopharynx either by nasoendoscopy, videofluroscopy or both. Selection of the surgical procedure depends upon the pattern of velopharyngeal closure, the degree of closure and associated medical conditions. A sample of common nasopharyngoscopic findings and selected surgical procedures are given for review. (Fig. 1-7)

Acknowledgments

This work was supported by grant P60 DC00976 to the National Center for Voice and Speech from that National Institute of Deafness and Other Communicative Disorders (NIDCD). Figures 1-5, 1-6, 1-7: Permission granted by the copyright owner. Facial Plastic Clinics of North America. Surgery of Cleft Lip and Palate Deformities, #4(3), August 1996.

References

1.Shprintzen RJ, et al.: A three dimensional cinefluoroscopic analysis of velopharyngeal closure during speech and nonspeech activities in normals, *Cleft Palate J* 11:412428,1980.

2.Matsuya T, Miyazaki T, Yamaoka M: Fiberscopic examination of the velopharyngeal closure in normal individuals, *Cleft Palate J* 11:286-291,1974.

3.Kuehn DP, Folkins JW, Cutting CB: Relationships between muscle activity and velar position, *Cleft Palate J* 19:25-35, 1982.

4.Seaver EJ, Kuehn, DP: A cineradiographic and electromyographic investigation of velar positioning in non-nasal speech, *Cleft Palate J* 17:216-226, 1980.

5.Fletcher *SG: Diagnosing speech disorders from cleft palate*, New York, 1978, Grune and Stratton.

6.Hawkins CF, Swisher WE: Evaluation of a real-time ultrasound scanner in assessing lateral pharyngeal wall motion during speech, Cleft *Palate J* 15:161-166, 1978.

7.Engelke W, et al.: Midsagittal velar kinematics during production of VCV sequences. Cleft *Palate Craniofacial* J 33:236-244, 1996.

8.McWilliams B, Morris HL, Shelton RL: *Cleft Palate Speech* Philadelphia, 1990, BC Decker Inc.

9.Shprintzen RJ, et al.: Selective movement of the lateral aspects of the pharyngeal walls during velopharyngeal closure for speech blowing and whistling in normals, Cleft *Palate* J 12:51-58, 1975.

10.Glaser ER, et al.: The dynamics of Passavant's ridge in subjects with and without velopharyngeal insufficiency-a multi-view videofluoroscopic study, *Cleft Palate* LT 16:24-33, 1979.

11.Boorman JG, Sommerlad BC:Levator palati and palatal dimples: their anatomy, relationship and clinical significance, *Br Jl Plastic Surgery* 38:326-32, 1985.

12.Finklstein Y, et al.:Axial configuration of the Velopharyngeal valve and its valving mechanism, *Cleft Palate-Craniofacial Journal* 32: 299-305,1995.

13.Nishio J, et al.:The motor nerve supply of the velopharyngeal muscles, *Cleft Palate J* 13:20-30, 1976.

14.Ibuki K,et al.: The course of facial nerve innervation for the levator veli palatini muscle, Cleft *Palate* J 15:209-214, 1978.

15.Azzam NA, Kuehn DP: The morphology of musculus uvlae, *Cleft Palate J* 14:78-87, 1977.

16.Kuehn DP, Azzam NA: Anatomical characteristics of palatoglossus and the anterior faucil pillar, *Cleft Palate J* 15:349-359, 1978.

17.Skolnick ML, McCall GN, Barnes M: The sphincteric mechanism of velopharyngeal closure, Cleft Palate LT 10:286, 1973.

18.Croft CB, Shprintzen RJ, Rakoff SJ: Patterns of velopharyngeal valving in normal and cleft palate subjects: A multiview videofluroscopic and nasoendoscopic study, *Laryngoscope* 91:265, 1981.

19.Golding-Kushner KJ,et al: Standardization of the reporting of nasopharyngoscopy and multiview videofluroscopy: A report from an international working group, Cleft *Palate* J 27:337347, 1990.

20.Finkelstein Y,et al: Nasopharyngeal profile and velopharyngeal valve mechanism, *Plastic Reconstruct Surgery* 92:603-613, 1993.

21.Gorlin RJ, Cerrenk J, Pruzansky S: Facial clefting and its syndromes, Birth *Defects*, 1:3, 1971.

22.Weatherley-White RCA, et al: Submucous cleft palate: incidence, natural history and implications for treatment, *Plast Reconstr Surg* 49:279, 1972.

23.Mason RM, Warren DW: Adenoid involution and developing hypernasality in cleft palate, *J Speech Hear Dis* 45:469, 1980.

24.Ren YF, Isberg A, Henningsson G: Velopharyngeal incompetence and persistent hypernasality after adenoidectomy in children without palatal defect, Cleft *Palate Craniofacial* 32(6):476-482, 1995.

25.Gibb AG: Hypernasality (rhinolalia aperta) following tonsil and adenoid removal, *J Laryngol Otol* 72:433-451, 1958.

26.Morris HL, Jrueger LJ, Bumsted RM: Indications of congenital palatal incompetence before diagnosis, *Ann Otol* 91:115-118, 1982.

27.Witzel MA,et al: Velopharyngeal insufficiency after adenoidectomy: An 8-year review, *Irit J Pediatr Otorhinolaryngol* 11:15-20, 1986.

28.Pinborough-Zimmerman J,et al: Nasometric and appliance treatment for hypernasality in one CVA patient. Presented at the American Speech-Language-Hearing Association Annual Meeting, Orlando, Florida, December 1995.

29.Witzel MA, Tobe J, Salyer K: The use of nasopharyngoscopy biofeedback therapy in the correction of inconsistent velopharyngeal closure, *Int LT Pediatr Otorhinolaryngol* 15:137-42, 1988.

30.Kuehn, DP: New therapy for treating hypernasal speech using continuous positive airway pressure (CPAP), *Plastic Reconstruct Surg* 88:959-965, 1991.

31.D'Antonio LL, Marsh JL: Abnormal carotid arteries in velocardiofacial syndrome, *Plast Reconstruct Surg* 80:471- 472, 1987.

32.Scrambler PJ, et al: Velocardiofacial syndrome associated with Chromosome 22 deletions encompassing the DiGeorge locus, Lancet 339 (8802):1138-1139, 1992.

33.Strom M: Down's Syndrome: A modern otorhinolaryngological perspective, *Laryngoscope* 91:1581-1594, 1981.

34.Evans D: Language development in mongols, *Forward Trends* 1:2325, 1974.

35.Kavanagh KT, Kahane JC: Risks and benefits of adenotonsillectomy for children with Down Syndrome, *Amer Journ of Ment Defic* 91(1):22-29, 1986.

36.Golding-Kushner K, et al: Standardization for the reporting of nasopharyngoscopy and multiview videofluroscopy: A report from an international working group, *Cleft Palate Craniofacial J* 27:337-347, 1990.

37.D'Antonio L, et al: Reliability of flexible fiberoptic nasopharyngoscopy for evaluation of velopharyngeal function in a clinical population, Cleft *Palate* J 26:217-225, 1989.

38.Fletcher, SG: Diagnosing speech disorders from cleft palate, New York, 1978, Grune and Stratton.

39.Miyazaki T: Commentary on the use of videonasopharyngoscopy for biofeedback therapy in adults after pharyngeal flap surgery, Cleft *Palate Craniofacial* LT 26:136, 1989.

40.Parameters for evaluation and treatment of patients with cleft lip and palate or other craniofacial anomalies, Cleft *Palate Craniofacial J;* 30 (Supplement 1) 1993.

41.Boone DR, McFarlane SC: The voice and voice therapy, Englewood Cliffs, New Jersey, 1994, Prentice Hall.

42.Blakely RW: Temporary speech prosthesis as an aid in speech training, Cleft Palate Bulletin 10:63-65, 1960.

43.Wolfaardt JF, et al.: An appliance based approach to the management of palatopharyngeal incompetency: a clinical pilot project, *J Prosthet Dent* 69:186-95, 1993.

44.McGrath CO, Anderson MW: Prosthetic treatment of velopharyngeal incompetence. In Bardach J, Morris HL,(eds): *Multidisciplinary management of cleft lip* and palate, Philadelphia, 1990, WB Saunders Company, pp 809-815.

45.Witt PD, et al.: Do palatal lift prostheses stimulate velopharyngeal neuromuscular activity, Cleft Palate *Craniofacial J* 32:469-475, 1995.

46.Adisman IK Cleft palate prosthetics. In Grabb WC, Rosenstein SW, Bzoch KR (eds): *Cleft Lip and Palate,* Boston, 1971, Little,Brown and Company, p 617.

47.Karnell MP, Rosenstein H, Fine L: Nasal videoendoscopy in prosthetic management of palatopharyngeal dysfunction, *J Prosthet Dent* 58:479-84, 1987.

48.Pinborough-Zimmerman J, et al.: Nasometric and appliance treatment for hypernasality in one CVA patient. Presented at the American Speech-Language-Hearing Association Annual Meeting, Orlando, Florida, December 1995.

49.Reisberg DJ, Smith BE: Aerodynamic assessment of prosthetic speech aids, J *Prosthet Dent* 54: 686-90, 1985.

50.Turner GE, William, WN: Fluoroscopy and nasoendoscopy in designing palatal lift prostheses, *J Prosthet Dent* 66:6371, 1991.

51.Gray S, Smith M: Voice disorders in children, Pediatric Clinics of North America, 1997, in press.

52.Gray S, Pinborough-Zimmerman, J: Diagnosis and treatment of velopharyngeal incompetence. In Sykes JM, Senders, CW (eds) : *Facial Plast:ic Surgery Clinics of I\Torth A-merica,* vol 4, WB Saunders Company, Philadelphia, 1996, pp 405-12.

53.Broen PA, Doyle SS, Bacon CK: The velopharyngeally inadequate child: phonologic change with intervention, *Cleft Palate Craniofacial J* 30:500-507, 1993.

54.Henningsson GE, Isberg AM: Velopharyngeal movement patterns in patients alternating between oral and glottal articulation: A clinical and cineradiographical study, *Cleft Palate Journal* 23: 1-9, 1986.

55.Dalston RM, Warren DW, Dalston ET: Use of nasometry as a diagnostic tool identifying patients with velopharyngeal impairment, *Cleft Palate Craniofacial* LT 28: 184-189, 1991.

56.Dalston RM, Warren DW, Dalston ET: A preliminary investigation concerning the use of nasometry in identifying patients with hyponasality and/or nasal airway impairment, *Jour Speech and Hearing Research,* 34:11-18, 1991.

57.Pinborough-Zimmerman J, Howe M, Romsdahl D, et al: Relationship between nasalance and nasoendoscopy in children with hypernasality. Presented at the American Speech - Language-Hearing Association Annual Meeting, Orlando, Florida, December, 1995.

58.Yamaoka M, et al: Visual training for velopharyngeal closure in cleft palate patients; a fiberscopic procedure, LT *Maxillofac Surg* 11:191-3, 1983.

59.Witzel MA, Tobe J, Salyer K: The use of nasopharyngoscopy biofeedback therapy in the correction of inconsistent velopharyngeal closure, *Int J Pediatr Otorhinolaryngol* 15:137-42, 1988.

60.Kuehn, DP: New therapy for treating hypernasal speech using continuous positive airway pressure (CPAP), *Plastic Reconstructive Surgery* 88:959-965, 1991.

61.Kuehn DP, Moon JB, Folkins JW: Levator veli palatini muscle activity in relation to intranasal air pressure, Cleft Palate Craniofacial J 30:361-368, 1993.

62.Hynes W: Pharyngoplasty by muscle transplantation, *British Journ Plast Surg.* 3:128, 1950.

63.Orticochea M: Construction of a dynamic muscle sphincter in cleft palates, *Plast and Reconstruct Surg* 41:323, 1968.

64.Jackson I: Discussion: A review of 236 cleft palate patients treated with dynamic muscle sphincter, *Plast and Reconstruct Surg* 71:187, 1983.

65.Riski JE, et al.: A rationale for modifying the site of insertion of the Orticochea pharyngoplasty, *Plast and Reconstruct Surg* 73:882, 1984.

66.Riski JE, et al.: Evaluation of the sphincter pharyngoplasty, *Cleft Palate Craniofacial J* 29(3):254-61, 1992.

67.witt PD, et al.: Sphincter pharyngoplasty: A preoperative and postoperative analysis of perceptual speech characteristics and endoscopic studies of velopharyngeal function, *Plast and Reconst Surg* 93:1155-1168, 1994.

68.Sloan GM, et al.: Surgical treatment of velopharyngeal insufficiency: Pharyngoplasty vs. pharyngeal flap, *Plast Surg Forum* 13:128, 1990.

69.Witt PD, et al.: Management of the hypodynamic velopharynx, *Cleft Palate Craniofacial J* 32(3):179-87, 1995.

70.Whitaker LA, et al.: A prospective and randomized series comparing superiorly and inferiorly based posterior pharyngeal flaps, *Cleft Palate J* 9:304, 1972.

71.Morris HL, et al.: Clinical results of pharyngeal flap surgery:The Iowa experience, *Plast and Reconstruct Surg,* 652-662, 1995.

72.Hirschberg J: Pediatric otolaryngological relations of velopharyngeal insufficiency, *Int Journ Pediatr Otorhinolaryng* 5:199, 1983.

73.Shprintzen RJ, et al: A comprehensive study of pharyngeal flap surgery: Tailor-made flaps, *Cleft Palate Journ* 16:46, 1979.

74.Valnicek SM, et al.: Perioperative complications of superior pharyngeal flap surgery in children, *Plast and Reconstr Surg* 93(5) 954-958, 1994.

75.Kravath RE, et al: Obstructive sleep apnea and death associated with surgical correction of velopharyngeal incompetence, *Journ Pediat* 96:645, 1980.

76.Ysunza A, et al.: Obstructive sleep apnea secondary to surgery for velopharyngeal insufficiency, Cleft *Palate Craniofac J* 30(4):387-90, 1993.

77.Sirois M, et al.: Sleep apnea following a pharyngeal flap: A feared complication, *Plast and Reconstr Surg* 93(5):943-947, 1994.

78.Orr WC, Levine NS, Buchanan RT: Effect of cleft palate repair and pharyngeal flap surgery on upper airway obstruction during sleep, *Plast and Reconst Surg* 80:226, 1987.

79.Rosenthal W: Zur frage der gaumenplastik, Zentralbl Chir 51:1621, 1924.

80.Hogan VM: A clarification of the surgical goals in cleft palate speech and the introduction of the lateral port control (L.P.C.) pharyngeal flap, *Cleft Palate J* 10:331, 1973.

81.Crockett DM, Bumstead RM, Van Demark DR: Experience with surgical management of velopharyngeal incompetence, *Otolaryngol Head Neck Surg* 99:1-9, 1988.

82.Gersuny R: Ueber eine subcutane prostheses, *Ztsch Heilk* 21:199, 1900.

83.Eckstein H: Paraffin for facial and palatal defects, *Dematol Ztschr* (Basel) 11:772, 1904.

84.Blocksma R: Correction of velopharyngeal insufficiency by silastic pharyngeal implant, *Plast and Reconst Surg* 31:268, 1963.

85.Perthes H: Reported by Hollweg E: Beitra zur behandlung von gaumenspalten, *Dissertation Tubingen* 1912.

86.Hill MJ, Hagerty RF: Efficacy of pharyngoplasty for speech improvement in postoperative cleft palates, Cleft Palate Bull, 10:66, 1960.

87.Halle M: Gaumennaht and gaumenplastic, Arch *Ohr Nas Kehlkopfheilk*, 12:377, 1925.

88.von Gaza W: Transplanting of free fatty tissue in the retropharyngeal area in cases of cleft palate, Lecture German Surgical Society, April 9th, 1926.

89.Ward PH: Uses of injectable Teflon in otolaryngology, *Arch Otolaryngol.* 97:637, 1968.

90.Smith JK, McCabe BF: Teflon injection in the nasopharynx to improve velopharyngeal closure, *Ann Otolaryng Rhinol and Laryng.* 86:559, 1977.

NCVS Status and Progress Report - 11
May 1997, 153-161

Volumetric Image-Based Comparison of Male and Female Vocal Tract Shapes

Brad H. Story, Ph.D.
Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts
Eric A. Hoffman, Ph.D.
Department of Radiology, The University of Iowa College of Medicine
Ingo R. Titze, Ph.D.
Wilbur James Gould Voice Research Center, The Denver Center for the Performing Arts
Department of Speech Pathology and Audiology, The University of Iowa

Abstract

A collection of 3-D vocal tract shapes corresponding to vowels and consonants of American English have been acquired for a 27 year old adult female subject using magnetic resonance imaging (MRI). Each 3-D shape was condensed into a set of cross-sectional areas of oblique sections perpendicular to the centerline of the vocal tract's long axis. Such a collection of areas is typically called an "area function. This set of images and subsequent area functions for the female subject compliments a previous similar study concerning an adult male subject [Story, Titze, and Hoffman, J. Acoust Soc. Am, 100(1), 1996]. It is the purpose of this paper to explore the morphological differences between the male and female subjects for three "cardinal" vowels /i/,/a/, and /u/. Comparisons have been made of the 3-D vocal tract shapes, area functions, and acoustic characteristics of the three vowels. The primary difference between genders is that the female pharynx is approximately 37 percent shorter than the male. Limited acoustic modeling has suggested that this shortened pharynx may play a significant role in defining male versus female voice quality.

Introduction

In the perception of human speech, most listeners typically have little difficulty discerning the gender of an adult speaker. However, computers tend to have great difficulty both in recognizing gender[1] as well as synthesizing a natural sounding female voice[2]. Thus, it is of interest to investigate the properties of the female speech production system and compare them to the characteristics of male speech. There are essentially two components of the speech production system that can contain the information that is unique to males and females; the voice source and the vocal tract. The voice source is comprised of the vibrating vocal folds, in which gender unique features are likely due to the lengths, masses, and elastic properties of the male or female vocal fold tissue. For example, a female typically has shorter, less massive vocal folds than an average male[3]. This leads to, among other things, a generally higher fundamental frequency of vibration than would have a male.

The second component, the vocal tract, filters the "raw" sound generated at the vocal folds, producing accentuated harmonics of the voice source in patterns such that the characteristic sounds of vowels and consonants are produced. In general, the female vocal tract has been observed to be about 15-20 percent shorter than that of a male[4,5,6] which in itself will raise all vocal tract resonances to higher frequencies. However, there has been some debate as to whether a particular male vocal tract shape is a uniformly or nonuniformly "stretched" version of the corresponding female vocal tract; "nonuniformly" implying that different regions of the vocal tract are stretched by differing amounts. Yang and Kasuya[5], who have performed an MRI study of a male, female, and child vocal tract, have concluded that for purposes of extracting the phonetic content of speech, a uniform scaling of the female vocal tract with respect to male is adequate. Their conclusions don't, however, address the issue of gender unique voice quality or those characteristics of the voice that provide cues to the gender of a speaker. Recently, Wong et al.[7] attempted to transform the recorded speech of a male speaker to that of a female

using the uniform vocal tract scaling proposed by Yang and Kasuya[5]. While the intelligibility remained high, the gender quality was confusing for listeners. Coleman[8] found that an artificially raised voice fundamental frequency, when used to excite a male vocal tract, still provided the perceptual cues to a male quality. Agren and Sundberg[9] determined that for alto (low range female) and tenor (high range male) singing voices, the main difference, with regard to the vocal tract, was that the third and fourth format (F and F) were more widely spaced for the female than for the male. They hypothesized that this characteristic could be a defining quality of male versus female since a narrow frequency spacing of F and F could accentuate two consecutive voice harmonics, thus placing them within the same critical band of the auditory system. Two, high-amplitude voice harmonics within the same critical band tends to produce a perception or roughness or hoarseness in the voice, more characteristic of a male[10]. Thus, a wide spacing of F and F may ensure that two consecutive harmonics will not be accentuated within a critical band, therefore maintaining a female quality. It has also been shown by Ingemann[11] and Schwartz[12] that listeners can detect gender from isolated fricative consonants in which there is no voice source, leaving the vocal tract as the primary element containing some gender characteristics. These studies suggest that the female voice quality is not entirely dependent on the voice source but also includes characteristics of the vocal tract. It seems essential, then, to acquire morphological data regarding the female vocal tract shape during production of speech sounds.

The use of MRI has, in recent years, become a viable option for obtaining 3-D information about the human vocal tract shape for static vowels and consonants. Baer et al.[13] demonstrated the use of MRI to directly measure the vocal tract shape for four vowels (2 male subjects). This study was the first demonstration of 3-D reconstructions of the vocal tract shape using imaging techniques. Dan et al.[14] used MRI to produce 3-D reconstructions of the nasal tract passages and sinus cavities. The nasal tract morphology was subsequently used to model the acoustic characteristics of the nasal system. MR volume imaging of the fricative consonants for male and female subjects has been recently reported by Narayanan et al.[15] and Narayanan[16]. These studies provide the most accurate information to date of the constrictions and air channels that produce the turbulence generated sound characteristic of fricative consonants. Yang and Kasuya[5], which was mentioned above, have reported area functions of five Japanese vowels acquired from a man, woman, and child. Story , Titze, and Hoffman[17] reported MRI based 3-D vocal tract shapes and area functions for 18 vowels and consonants of one male subject. These area functions were subsequently used to model static vowel sounds which were compared the natural vowel sounds of the imaged subject.

The aim of this study was to use MRI to acquire a set of 3-D vocal tract shapes for an adult female that correspond to vowels and consonants of American English. Such a set compliments a previous study concerning an adult male subject[17]. Some basic morphological and consequent acoustical differences between the male and female subjects are explored for the three vowels /i/, /a/, and /u/. The interest is to find those vocal tract differences that bear upon a male-like versus female-like quality of speech. The issue of normalizing female speech to an equivalent male for recognition of linguistic content is not addressed. It is recognized that these comparisons strictly apply to only one male and one female selected from the general population. Thus, no statistical power can obtained from the set, however, general characteristics from each subject are thought to be useful. The process of acquiring the imaging sets from one subject for many vowels and consonants (22) is a physically demanding and time consuming process. The acquisition and analysis of enough subjects to provide a statistically adequate sample would be prohibitively expensive and would take many years to complete. However, as more researchers begin to image the vocal tract the pool of data will eventually grow so that a reasonable statistical sample may be achieved.

Image Acquisition and Analysis

Scanning Parameters and Protocol

Volumetric imaging of the vocal tract was performed using MRI for 22 different phoneme configurations of both a male and female subject. The vocal tract shapes for the male subject have been reported in Story et al.17 and will be included in this paper for purposes of comparison with the female vocal tract. In addition, only the vowels, /i/, /a/, and /u/ will be considered. The female subject (DJ) was a 27 year-old male with no history of speech or voice disorders and is native to the state of Texas in the southern United States. For comparison, the male subject (BS) was 29 years old at the time of scanning and is native to the midwestern United States. The male subject is also the first author of this study. Additionally, the vowels /i/ and /a/ were imaged for both subjects using EBCT. However, only the MRI-based images are discussed in this paper.

The MR images were acquired using a General Electric Signa 1.5 Tesla scanner. A 24 slice series of 5 mm thick contiguous, parallel, axial sections was gathered in an interleaved acquisition. This image set extended from just cephalad of the hard palate down to the first tracheal ring. The protocol for collecting the images consisted of having the subject positioned in a comfortable supine position on the patient table and was required to phonate at a comfortable pitch and loudness. The requirement of phonation was necessary in order for the subject to maintain the desired tract shape. More details concerning the scanning protocols can be found in Story[18] and Story et al.[17, 19, 20]

Image Analysis

All image analysis operations were performed with a general image display and quantitation package called VIDA (Volumetric Image Display and Analysis) which has been developed by researchers in the Division of Physiologic Imaging at the University of Iowa[21].

The image analysis process included three main steps: 1) segmentation of the airway from the surrounding tissue, 2) three-dimensional reconstruction of the airway by shape-based interpolation, and 3) determination of an airway centerline and subsequent extraction of cross-sectional areas assessed from oblique sections calculated to be locally perpendicular to the airway centerline to create an "area function". This image analysis process is identical to previous work which can be found in Story[18] and Story et al.[17, 19, 20] and also at the Internet web site <http://everest.radiology.uiowa.edu/tutor/app/vocal/vocal.html.> Detailed explanations of the process will not be repeated here.

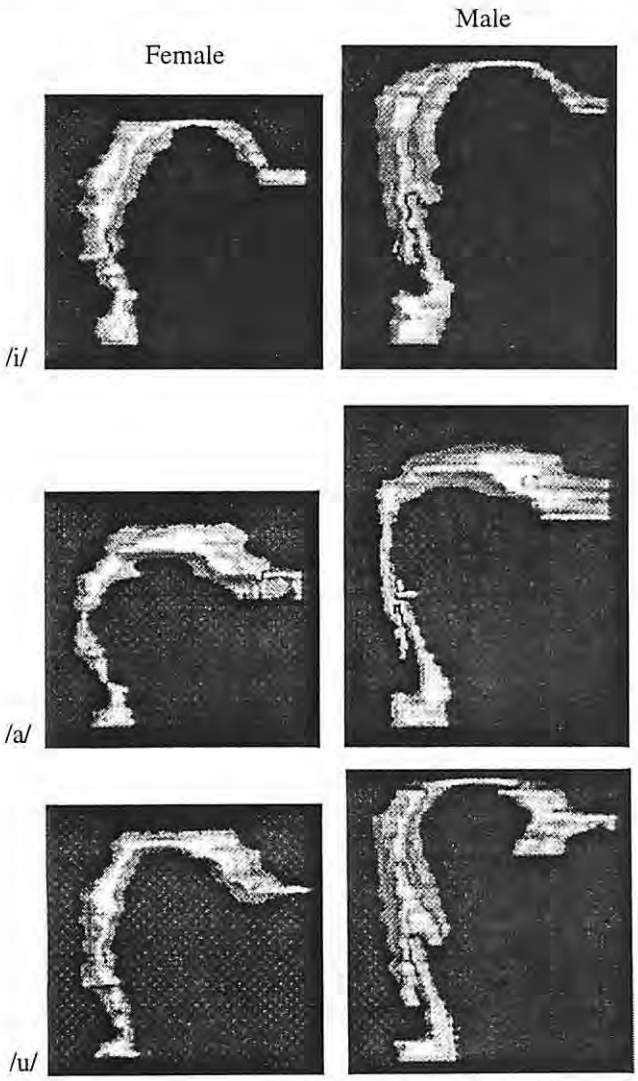


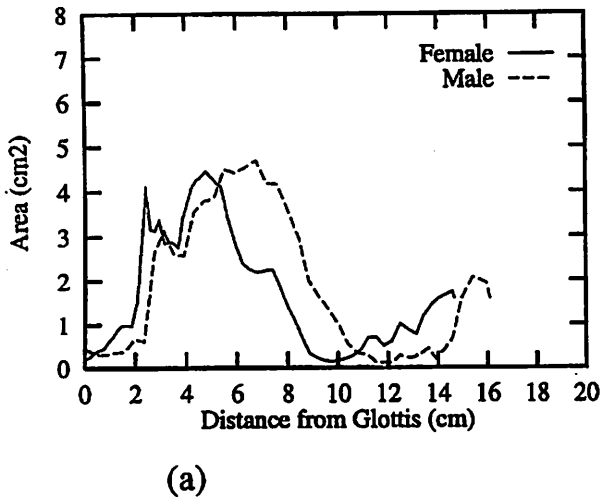
Figure 1. Shaded surface displays of three vowels (/i/, /a/., and /u/) for both a female (left column) and a male (right column).

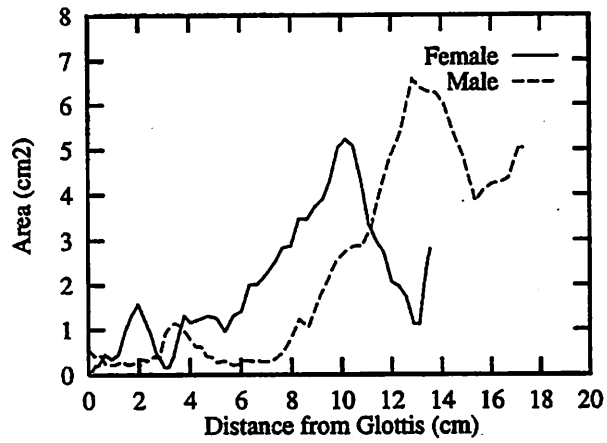
Imaging Results

The results of this experiment have been combined in Figures 1-3 with those of the male subject[17]. In Figure 1, the surface rendered representation of the 3-D vocal tract shape for the female subject is shown on the left side of the figure and the male version on the right. For each surface rendered airway, the most inferior point of the 3-D shape begins with the uppermost section of the trachea. Above the trachea, the airway becomes small in the region of the glottis and then widens, more or less depending on the vowel, into the lower pharyngeal section. In the case of the male vocal tract shapes, the fingerlike extensions that hang down below the pharynx are the piriform sinuses. In the MR data sets for the female the piriform sinuses were poorly defined and consequently were not segmented. The area functions derived from the 3-D vowel shapes are given in Figure 2 (following page) for both male and female subjects. The 0 cm location represents a point just above the glottis and the termination at the lips is at right side of the graph. Figure 3 (following page) shows the vocal tract centerline profiles (a sagittal projection of the 3-D centerline) of each vowel for both male and female. In each graph, the point located at coordinate (0,0) represents a location just above the glottis; note that this point does not correspond to the lowest part of each surface rendered airway since those included the upper portion of the trachea. The profile then moves upward through the pharynx and then curves into the oral cavity and finally terminates at the lips. The solid dot located on each tract profile is the estimated location of the uvula tip; this is used as a landmark to divide the vocal tract into pharyngeal and oral sections.

The vocal tract can be further subdivided into an epilarynx which is defined, in this study, as the region of the vocal tract that extends from just above the glottis to the point where the piriform sinuses join the main vocal tract. This is an acoustic definition more than an anatomical one in that the widening of the vocal tract at the piriform junction creates an acoustic discontinuity that is important for lower vocal tract resonance and consequent voice quality characteristics[22]. It should be noted that Yang and Kasuya[5] used the tip of the epiglottis as an anatomical landmark to define this region. The solid dot on each area function in Figure 2 indicates the location of this point. Henceforth, the pharynx is defined as the section of the vocal tract between the end of the epilarynx and the tip of the uvula. The oral cavity extends from the uvula tip to the lips.

A general observation with regard to all three vowel shapes is that the overall size of the female vocal tract is, not surprisingly, smaller than that of the male. From both the surface rendered airways and the vocal tract profiles (Figs. 1 and 3) it is also observed that the female pharynx is typically shorter than that of the male but the oral cavities are of similar length. The solid dot indicating the uvula tip is assumed to divide the tract into pharyngeal and oral sections.





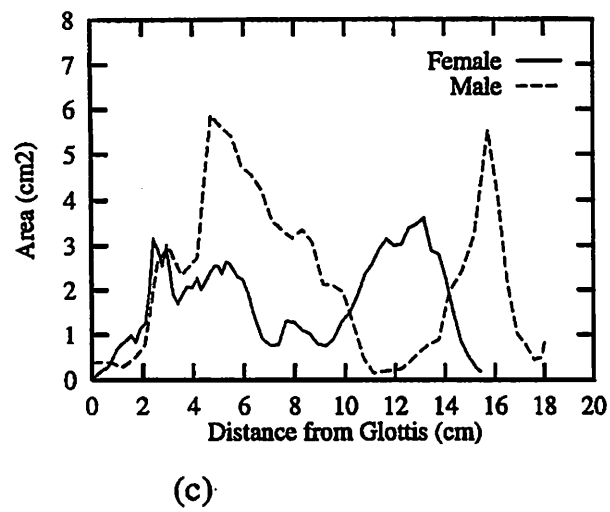
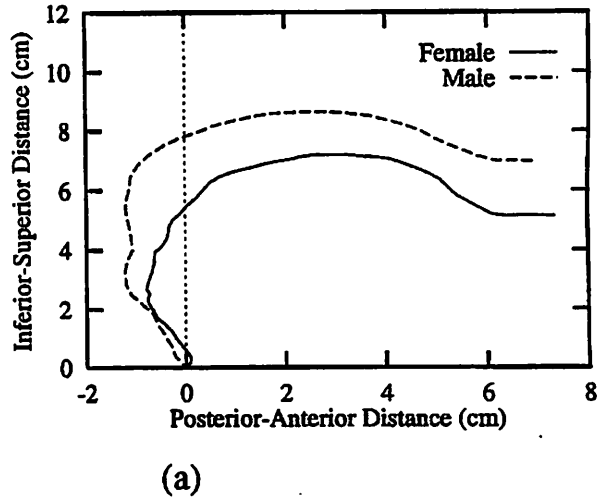
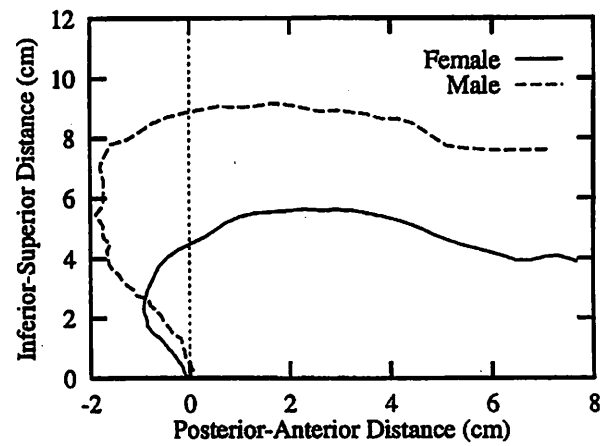


Figure 2. Area functions for female (solid) and male (dashed) vowels. The solid dot on each area function indicates the point at which the piriform sinuses join the mail vocal tract. a) /i/, b) /a/, and c) /u/.





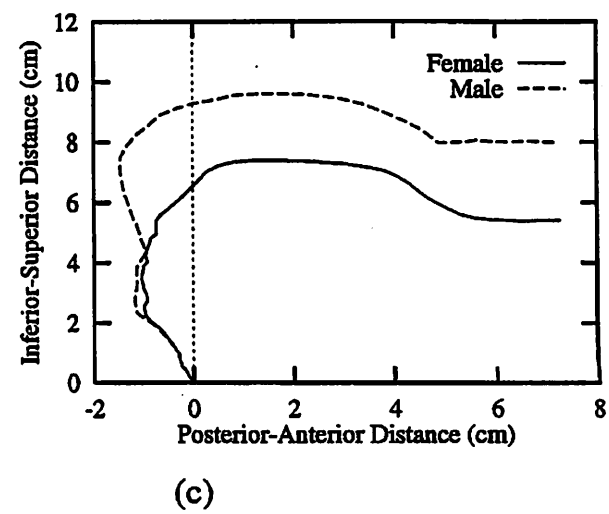


Figure 3. Sagittal projections of the vocal tract centerline for female (solid) and male (dashed) vowels. The solid dot on each projection indicates the location of the uvula tip. a) /i/, b) /a/, and c) /u/.

The maximum cross-sectional areas shown in the area functions (Fig. 2) are generally smaller for the female than for the male.

The specific case of the area functions for vowel /i/ (Figure 2a) indicates that, for the female, the region just above the glottis linearly increases from an area of 0.2 cm^2 to 1.0 cm^2 at a point approximately 2 cm above the glottis. In comparison, the male cross-sectional areas in this same region remain nearly constant at about 0.3 cm^2. Beyond this point there is an abrupt increase in area for both male and female. The female tract reaches a maximum area of 4.5 cm^2 at a distance of 5 cm from the glottis after which the area begins to decrease until a minimum of 0.14 cm^2 is reached at 9.5 cm from the glottis. Following this minimum constriction, the area again rises to a value of 1.7 cm^2 at the lip termination which is 14.2 cm from the glottis. For the male vocal tract, a maximum area of 4.7 cm^2 occurs 7.1 cm from the glottis. As was the case for the female tract, the cross-sectional area following the maximum area begins to decrease and does so until an area of 0.1 cm^2 is reached at a point 11.8 cm from the glottis. The tract remains effectively constricted with areas in the range of 0.1 to 0.4 cm^2 until about 14 cm above the glottis at which point the area abruptly rises to 2 cm^2 close to the lip termination which is 16.2 cm from the glottis. The overall length of the vocal tract is 14.7 cm for the female and 16.1 cm for the male. Based on the definitions described above, the female epilarynx and pharynx length were measured to be 2.0 cm and 4.2 cm, respectively while the oral cavity was 8.5 cm long. For the male, the epilarynx, pharynx and oral cavity lengths were measured to be 2.4 cm, 6.0 cm, and 7.7 cm, respectively.

The area function for the female vowel /a/ (Figure 2b) shows that the area gradually increases from 0.05 cm^2 just above the glottis to about 0.4 cm^2 at 1 cm from the glottis. The area then abruptly rises to a peak of 1.7 cm^2 at 2 cm above the glottis and then, just as abruptly as it rose, the area falls to 0.15 cm^2 at a point 3 cm above the glottis. Following this constriction, the area sharply increases to 1.3 cm^2 and remains nearly constant for a 1.5 cm long section of the tract. At a point about 5.3 cm above the glottis, the area begins to increase into the widened oral cavity characteristic of the /a/ vowel. A maximum area of 5.1 cm^2 occurs at 10 cm above the glottis after which the area decreases to 1.5 cm^2 before rising to 2.8 cm^2 at the lip termination. The lower pharyngeal section (or the epilarynx) of the male /a/ is quite different from the female in that a nearly constant area of approximately 0.25 cm^2 is observed from just above the glottis up to a location 2.5 cm above the glottis. Then a sharp rise in area occurs with a peak of 1.1 cm^2 located 3.4 cm above the glottis. The area gradually drops to 0.25 cm^2 before increasing from about the 7 cm point up to the 13.5 cm point reaching a peak value of 6.6 cm^2. From the point of peak area out to the mouth termination the area function decreases before rising to lip termination area of 5.0 cm^2. The vocal tract length is 13.6 cm for the female and 17.3 cm for the male. Based on the locations of the uvula tip and piriform sinus junction, the female epilarynx, pharynx and oral cavity lengths were found to be 1.5 cm and 3.0 cm while measurements for the male gave an epilarynx length of 2.8 cm, a pharynx length of 5.9 cm and oral cavity length of 8.6 cm.

The vowel /u/ for the female has an area function that begins much like the /a/ vowel with an small area of 0.05 cm^2 just above the glottis. The area then increases gradually to a value of 1.1 cm^2 before increasing abruptly to 3.1 cm^2 at a distance of 2.2 cm from the glottis. Over the course of the next 4.5 cm the area falls into the range of 2.5 cm^2 but retaining the characteristic pharyngeal expansion of an /u/. The mid-tract constriction occurs from 7 cm to 9 cm with a minimum area in this region of 0.9 cm^2. The area then rises to a peak of 3.5 cm^2 before falling to 0.2 cm^2 at the lip termination. Much like the previous two vowels, the epilaryngeal region for the male /u/ has a nearly constant area, in this case about 0.38 cm^2. This is in contrast to the female vowel where the epilarynx has a more horn-like shape. Following the epilarynx, the male /u/ has an rapid increase in area which climbs to a maximum of 6.0 cm^2 the 4.7 cm point. The area gradually decreases over the next 6.0 cm until a mid-tract constriction area of 0.15 is reached at the 11 cm point. This is considerably smaller constriction area than for the female where 0.9 cm^2 was the minimum. The constriction is approximately 2 cm long after which the area increases into the front cavity expansion where a maximum area of 5.5 cm^2 occurs. The area then drops to 0.7 cm^2 at the lip termination. The overall length of the vocal tract is 15.5 cm for the female and 18.0 cm for the male. The length of the female epilarynx was measured to be 1.9 cm, the pharynx 2.0 cm, and the oral cavity 8.7 cm. The male epilarynx was 2.0 cm long while the pharynx and oral cavity had a length of 7.1 cm and 8.9 cm, respectively.

The measurements of the total vocal tract length, lengths of the epilaryngeal, pharyngeal and oral sections and maximum cross-sectional areas are summarized in Table 1. The pharynx lengths for the female range from 1.8 cm to 3.0 cm shorter than the male while the oral cavity lengths are similar across gender. For the epilarynx lengths, vowels /i/ and /u/ are similar for the both male and female whereas the female epilarynx in the vowel /a/ is nearly half that of the corresponding male vowel. The maximum areas are, in general, slightly larger for the male than the female.

Table 1

Lengths of the epilarynx, pharynx, and oral cavities for the female and male vowels /i/, /a/, and /u/. The maximum area measured in each area function is also given. All measurements are in centimeters unless otherwise indicated.

	Female			*Male*		
Cavity	/i/	/a/	/u/	/i/	/a/	/u/
epilarynx	2.0	1.5	1.9	2.4	2.8	2.0
pharynx	4.2	3.0	4.8	6.0	5.9	7.1
oral	8.5	9.2	8.7	7.7	8.6	8.9
max area(cm^2)	4.45	5.23	3.60	4.69	6.57	5.87

Table 2.
Female to male ratios of the vocal tract region lengths and maximum areas.

Female/Male	/i/	/a/	/u/	mean
epilarynx	0.83	0.54	0.94	0.77
pharynx	0.70	0.51	0.68	0.63
oral	1.10	1.07	0.98	1.05
total length	0.91	0.79	0.86	0.85
max area	0.95	0.79	0.61	0.78

In Table 2, ratios of the female measurements to those of the male are presented for the three vowels and a mean is computed across vowels. The ratios show that all regions of the female vocal tract are shorter than the male except the oral cavity. In fact, the female oral cavity is slightly larger than the male for /i/ and /u/ and nearly the same for the /a/. This suggests that a female vocal tract is not simply a compressed version of the male vocal tract but is nonuniformly adjusted across various regions with respect to a male. The set of ratios for a given vowel provides a kind of "recipe" for transforming a male vocal tract shape into a female vocal tract or vice versa; the set of mean ratios gives a generalized "recipe". In the average sense, an vocal tract area function for the male subject could be approximately transformed into a female-like area function by compressing the epilarynx to 77 percent of its original value, the pharynx to 63 percent of its original, and expanding the oral cavity by 5 percent. The words "compress" and "expand" are deliberately used instead of "shorten" and "lengthen" to stress that the shape of the area function must be retained throughout the vocal tract; e.g. the various regions are not simply truncated to realize a shorter section length. If only the total vocal tract length were considered as the main contributor to male versus female quality, the female vocal tract would be realized by compressing the entire male vocal tract by 15 percent, on the average. The ratios in this table are quite similar to those given in Goldstein[4] , Yang and Kasuya[5] and Hogberg[6].

Acoustic Modeling of Male and Female Vocal Tract Shapes

Acoustic resonance characteristics of each male and female area function were computed with a frequency domain transmission line[23]. The results are shown in Figure 4. In each plot, the thick line indicates the format spectrum for the female vowel shape while the thinner line represents the male; each peak in the spectrum is called a format and it is the pattern of formants that defines a specific vowel or consonant.

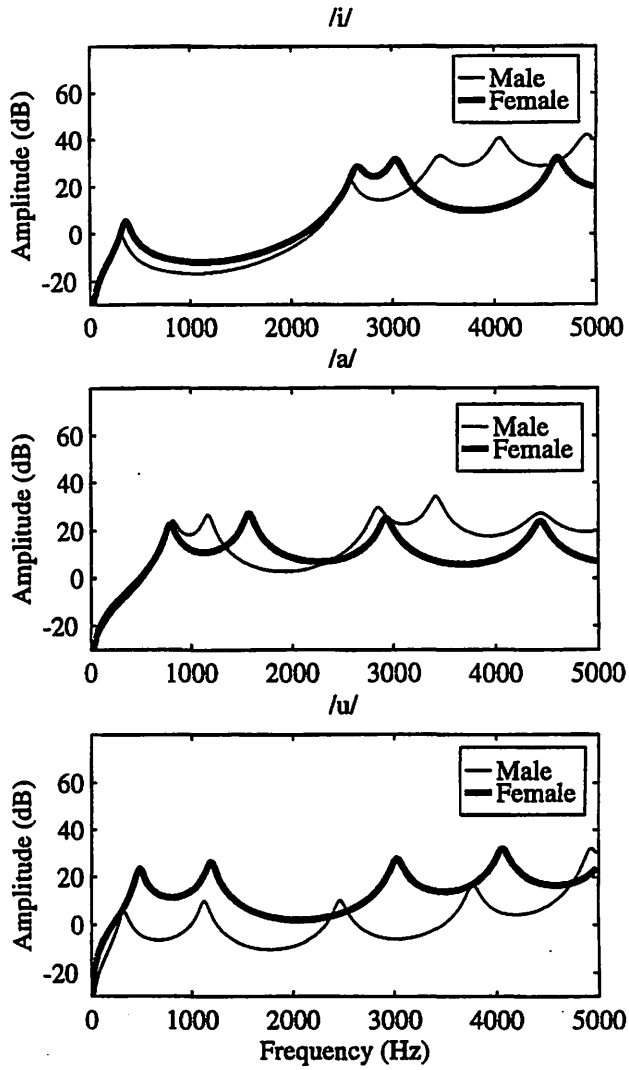


Figure 4. Formant spectra of vowels /i/, /a/, and /u/ for female (thick lines) and male (thin lines) subjects.

For the vowel /i/, the female F1 is slightly higher that the F1 of the male. F2 and F3 of the female are clustered together in the 2800 Hz to 3100 Hz range such that F2 for the female is higher than the male and F3 is lower than for the male. The female F4 is located at about 4700 Hz which gives an F4-F3 difference of approximately 1600 Hz, while the difference of F4-F3 for the male is only 700 Hz. In fact, the frequency range of 3000 Hz to 4500 Hz is highly accentuated for the male but is of relatively low amplitude for the female. This observation is similar to that described by Agren and Sundberg[9] for alto and tenor voices. For the /a/ the female F1 is slightly lower than the male but F2 is about 300 Hz higher. F3 is similar for both male and female but the female F4 appears at 4200 Hz, about 900 Hz higher

than F4 for the male. The format spectra for the /u/ shows higher format locations for the female than the male across the entire frequency spectrum.

An attempt was made to nonuniformly transform the male /i/ vowel into a female /i/ by applying the ratios calculated in Table 2. Thus, the epilarynx length, was compressed by 17 percent, the pharynx length compressed by 30 percent, and the length of the oral cavity expanded by 10 percent. The original female area function for /i/ is shown in Figure 5a as the thin solid line while the transformed male-to-female area function is shown with a thick solid line. Note that the areas in the male /i/ have not been changed, only the lengths of the three vocal tract regions have been altered. The original female and transformed male area functions are quite similar in that the points of expansion and constriction occur in nearly the same locations. The cross-sectional area throughout the length of the entire area function are also very similar. The format spectra of each of the area functions in Figure 5a is given in Figure 5b with the same line styles. The transformed male area function is observed to have format characteristics much like the female; F1 is nearly the same as the female, F2 and F3 are similarly clustered although situated about 400 Hz higher in frequency, and the F4's are located at nearly the same frequency. Thus, the male vocal tract appears to have been transformed into a vocal tract with female characteristics.

Another attempt to transform the male area function for /i/ into one with female-like qualities was performed by uniformly compressing the length of the entire area function by 9 percent. (see Table 2). Shown in Figure 6a is the original female area function for /i/ and the male area function uniformly scaled to female length. Unlike, the previous nonuniform transformation, the expansions and constrictions of the transformed male to female area function do not match closely with the original female /i/ area function. The transformed area function is shifted to the right relative to the original female area function. The format spectra for the original female /i/, the previous (nonuniform) transformation, and this simple length scaling are shown in Figure 6 (following page). The format spectrum from the uniform length scaling appears to have the same general shape as the original male /i/ vowel spectrum, except shifted upward in frequency by a small amount. This would imply that possibly the male characteristics of the vocal tract have been retained even though the format locations have been moved up in frequency. However, it should also be noted that the locations of F1 and F2 are closer to those of the female /i/ spectrum than are those from the nonuniform scaling of the male /i/. Thus, if the absolute locations of F1 and F2 are important for female voice quality then the uniform scaling may the appropriate choice. Conversely, if the overall pattern of format clustering characteristics is crucial to realizing a female quality, then the nonuniform scaling needs to be used. This would be especially true if Agren and Sundberg's[9] hypothesized significance of the distance between F3 and F4 is correct. A formal perceptual study would need to be performed to determine whether uniform or nonuniform scaling should be used to evoke a female-like quality. However, an informal listening test has suggested that the nonuniformly scaled /i/ vowel produces a slightly more female-like quality than the uniformly scaled version.

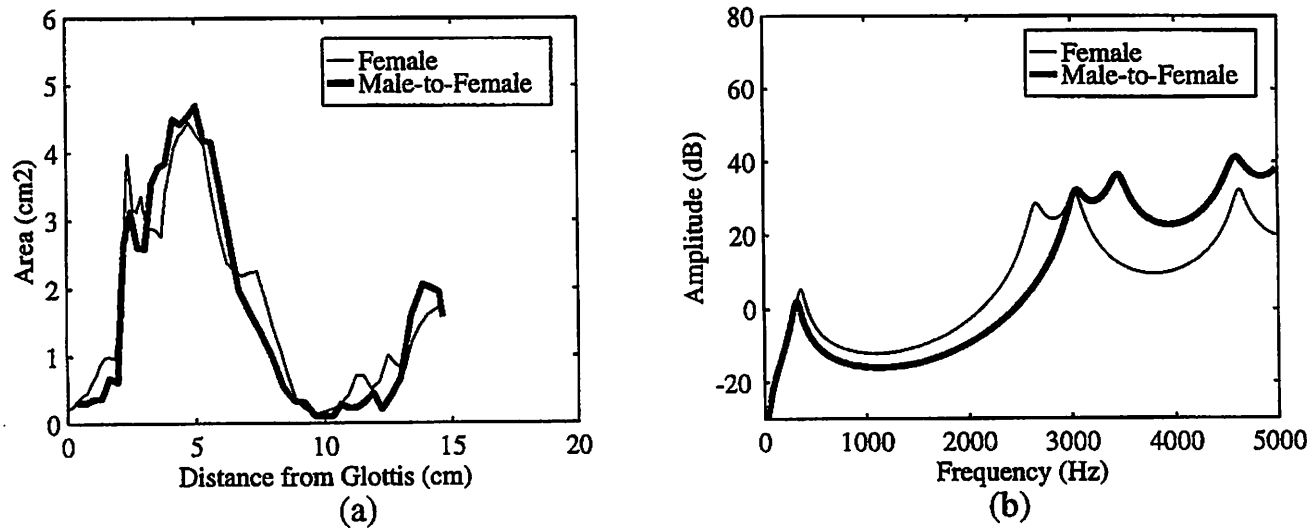


Figure 5. Original female/i/ and nonuniform transformation from male /i/ to female /i/; a) area functions, b) vocal tract transfer functions.

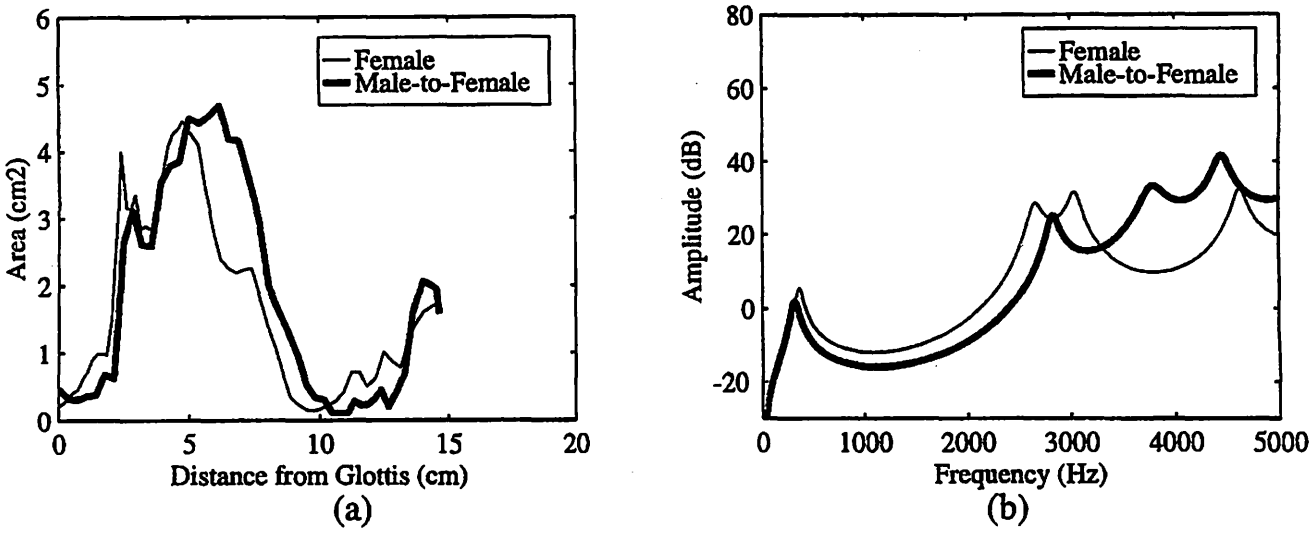


Figure 6. Original female /i/ and uniform transformation of male /i/ to female /i/; a) area functions, b) vocal tract transfer functions.

Conclusions

3-D vocal tract shapes representing three vowels (/i/, /a/, and /u/) have been acquired from one male and one female subject using MRI. The 3-D vocal tract shapes were condensed into an area function which was divided into three distinct regions representing the epilarynx, pharynx, and oral cavity. The epilarynx and pharynx were shorter than those of the male subject while the female oral cavity was nearly the same length or even slightly longer than for the male. Ratios of the female to male lengths of these three regions produced a simple recipe for transforming a male vocal tract shape to one with female qualities or vice versa. It was observed that, for the /i/ vowel, the nonuniform transformation was necessary in order to obtain the spectral qualities representative of a female. A simple uniform compression of the area functions shifted the formants higher in frequency, it did not reorganize the relationship of the formants to one another. This suggests that a female voice quality is largely dependent upon a shortened pharynx. This study has only begun to explore the male versus female vocal tract differences for vowels and consonants but based on the limited information presented it appears that a female vocal tract shape contains unique qualities that cannot be explained by a simple uniform compression of the male vocal tract.

Acknowledgments

This study was supported by grant No. R01 DC02532-02 from the National Institutes on Deafness and Other Communication Disorders.

References

1. Wu, K., and Childers, D.G., "Gender recognition from speech. Part I: Coarse analysis," J. Acoust. Soc. Am., 90(4), 1828-1840, 1991.

2. Klatt, D.H., and Klatt, L.C., "Analysis, synthesis, and perception of voice quality variations among female and male talkers," J. Acoust. Soc. Am., 87(2), 820-857, 1990.

3. Titze, I.R., "Physiologic and acoustic differences between male and female voices," J. Acoust. Soc. Am., 85(4), 1699-1707, 1989.

4. Goldstein, U.G., "An articulatory model for the vocal tracts of growing children," Ph.D. Dissertation, MIT, 1980.

5. Yang, C.S., and Kasuya, H., "Uniform and nonuniform normalization of vocal tracts measured by MRI across male, female, and child subjects," IEICE Trans. Inf. and Sys., E78-D (6), 732-737, 1995.

6. Hogberg, J., "From sagittal distance to area function and male to female scaling of the vocal tract," STL-QPSR, 4, 11-66, 1995.

7. Wong, D., Lange, R.C., Long, R.K., Story, B.H., and Titze, I.R., "Age and gender related speech transformations using linear predictive coding," J. Acoust. Soc. Am., (in review), 1997.

8. Coleman, R.O., "A comparison of the contributions of two vocal characteristics to the perception of maleness and femaleness in the voice," STL-QPSR, 2-3, 13-22, 1973.

9. Agren, K., and Sundberg, J., "An acoustic comparison of alto and tenor voices," STL-QPSR, 1, 12-16, 1976.

10. Terhardt, E., "On the perception of periodic sound fluctuations (roughness)", Acustica, 30, 201-213, 1974.

11. Ingemann, F., "Identification of the speaker's sex from voiceless fricatives," J. Acoust. Soc. Am., 44, 1142, 1968.

12. Schwartz, M., "Identification of speaker sex from isolated voiceless fricatives," J. Acoust. Soc. Am., 43, 1178, 1968.

13. Baer, T., Gore, J. C., Gracco, L. C., and Nye, P. W., "Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels," J. Acoust. Soc. Am., 90, 799-828, 1991.

14. Dan, J., Honda, K., and Suzuki, H., "Morphological and acoustical analysis of the nasal and the paranasal cavities," J. Acoust. Soc. Am., 96(4), 2088-2100, 1994.

15. Narayanan, S. S., Alwan, A. A., and Haker, K., "An articulatory study of fricative consonants using magnetic resonance imaging," J. Acoust. Soc. Am., 98(3), 1325-1347, 1995.

16. Narayanan, S. S., "Fricative consonants: An articulatory, acoustic, and systems study," Ph. D. thesis, UCLA, Dept. of Electrical Engineering, Los Angeles, CA.

17. Story, B.H., Titze, I.R., and Hoffman, E.A., "Vocal tract area functions from magnetic resonance imaging," J. Acoust. Soc. Am., 100(1), 537-554, 1996.

18. Story, B. H., "Physiologically-based speech simulation using an enhanced wave-reflection model of the vocal tract," Ph. D. Dissertation, University of Iowa, Dept. of Speech Pathology and Audiology, 1995.

19. Story, B. H., Hoffman, E. A., and Titze, I. R., "Speech simulation based on MR images of the vocal tract," SPIE Proc. Physiology and Function from Multidimensional Images, 2433, San Diego, CA, 26 Feb. - 2 Mar., 1995.

20. Story, B. H., Hoffman, E. A., and Titze, I. R., "Vocal tract imaging: A comparison of MRI and EBCT," SPIE Proc. Physiology and Function from Multidimensional Images, 2709, 209-222, Newport Beach, CA, 11-13 February, 1996.

21. Hoffman, E. A., Gnanaprakasam, D., Gupta, K. B., Hoford, J. D., Kugelmass, S. D., and Kulawiec, R. S., "VIDA: An environment for multidimensional image display and analysis," SPIE Proc. Biomed. Image Proc. and 3-D Microscopy, 1660, San Jose, CA, 10-13 Feb., 1992.

22. Titze, I.R., and Story, B.H., "Acoustic interactions of the voice source with the lower vocal tract," J. Acoust. Soc. Am., (in press), 1997.

23. Sondhi, M.M., and Schroeter, J., "A hybrid time-frequency domain articulatory speech synthesizer," IEEE Trans. ASSP, ASSP-35(7), 1987.

NCVS Status and Progress Report - 11
May 1997, 163-165

Speech Simulation and Transformation Using Empirical Articulatory Eigenmodes

Darrell Wong, Ph.D.
Brad H. Story, Ph.D.
Russel K. Long, M.S.
Wilbur James Gould Voice Research, The DenverCenter for the Performing Arts
Ingo R. Titze, Ph.D.
Wilbur James Gould Voice Research, The DenverCenter for the Performing Arts
Department of Speech Pathology and Audiology, The University of Iowa

Abstract

An articulatory model in which the vocal tract is reduced to three parameters is described. The triplet [F1 F2 F3] is mapped by lookup table to three parameters [c1 c2 c3], which represent scaling weights associated with basis eigenmodes. These modes are obtained from a Principal Component Analysis (PCA) of 2D area estimates, extracted from MRI images of the vocal tract. For sentence level speech reconstruction, area functions are recalculated from the identified weights. A time-domain vocal tract simulator then interpolates the area functions and creates a new speech signal, using as voice input a four parameter glottal flow model. Transformations to specific voice qualities, such as twang and sob can be accomplished prior to simulation by manipulating the area functions. We present plots of the formant/PCA-weight database, and audio recordings of the normal and transformed speech.

Introduction

Articulatory modeling of speech typically uses snapshots of 2d cross-sectional area vectors defined as a function of the distance from the glottis. The area functions are then used to drive a time domain simulator. Recently, area function models have been developed using statistical techniques [1],[2]. Meyer, Wilhelm and Strube used data from [1] to conduct a principal component analysis (PCA), using ten vowel shapes, each a ten element area function. Their analysis suggested that three eigenmodes could account for 93% of the variance in the data.

Story, Titze and Hoffman [3] reported a set of magnetic resonance imaging (MRI)-derived area functions for a single speaker (data in [1] came from a number of speakers). The data consisted of 12 vowels, 3 plosives, and 3 nasals, recorded for an adult male. The area functions were all normalized to 17.5 cm, with a resolution of 44 area elements per shape, rather than the 10 elements used in [2]. In a subsequent paper, Story and Titze (1996) [4] described a PCA decomposition of ten vowels from this data set, four eigenmodes representing 97% of the variance (two modes accounted for 88%). Appropriate scaling weights applied to these two modes could then reasonably approximate any area function in the set. Story et al. state that their eigenmode shapes are very close to those derived by Meyer et al, suggesting that these basis modes may be similar for other speakers.

A speech resynthesis study using the first two eigenmodes was performed by Story, Titze and Wong [5]. The 2-mode model was used to construct a database. The frequency response of each derived area function was computed with a frequency domain transmission line method, and the first two formants (F1 and F2) were extracted. Thus the database related coefficient weights to formant pairs. A glottal flow model and the database were used as inputs to a time-domain wave-reflection speech simulator to simulate speech.

Three Eigenmode Model

The above study demonstrated how a highly compact model could regenerate speech of reasonable fidelity. The two coefficient model did have some deficiencies,

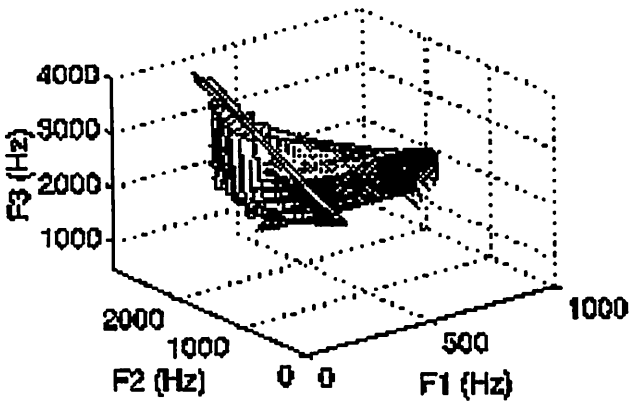


Figure 1. Formant plot of 12 vowel database.

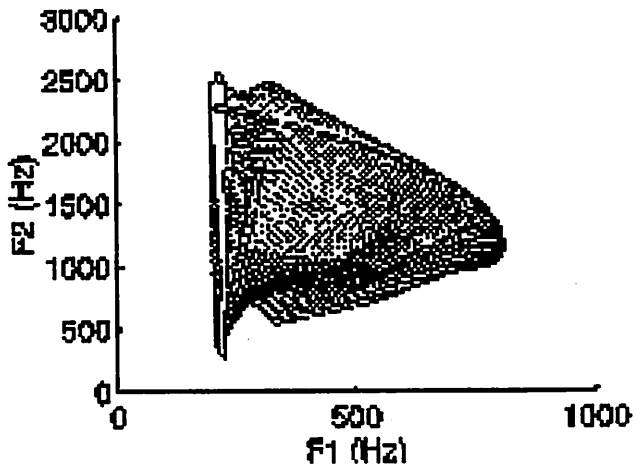


Figure 2. Formant database in F1-F2 plane.

however. The data set was limited to ten vowels. It would be of interest to extend the model to include the ability to map /er/ (as in *earth*) and /l/ sounds. Since the lack of F3 prevents accurate modeling of /er/ and /l/, it is necessary to extract F3 from the spectral responses and add another eigenmode weight to the model when generating the database.

In this paper we only briefly discuss the construction of the database and simulation system, more detail being available in [5]. Much of the synthesis system is the same. We focus instead on the enhancements made to the 2-mode system. The model is extended to a third formant and PCA weight in an attempt to accurately model F3 information. The data set is also enlarged to twelve shapes to include /er/ and /l/ vowel shapes. For a description of the PCA technique and the original 10 area functions, the reader is referred to [4].

Principal component analysis of the 12 shape set produced 88.46 % variance for three eigenmodes. An examination of the mode weights from the 10 and 12 vowel sets indicated that eigenmode 3 specifically enhanced the representation of /l/ and /er/. The maximal and minimal weight values applicable to the three eigenmodes were extracted, and the ranges subdivided into 30 intervals for c1 and c2 and 12 intervals for c3. The intervals were less dense for c3 because this eigenmode only increases the amount of variance captured by 6.4%. A database of 30x30x12 weight and formant triplets was thus created.

The formant state space appears in Figs 1 and 2. Some of the area functions produce discontinuities in the formant space, relative to similarly shaped area functions. In particular, several points at the bottom left of Fig.2, with an expected F2 of 230 Hz, are displaced to 2500 Hz. This is due to F2 converging on F1 to form a single formant. The formant picker then mis-identifies the next formant as F2 instead of F3. Similarly, in Fig.1, a shelf-like region appears at F1 near 200 Hz and F2 from 1500-2400 Hz. This is due to

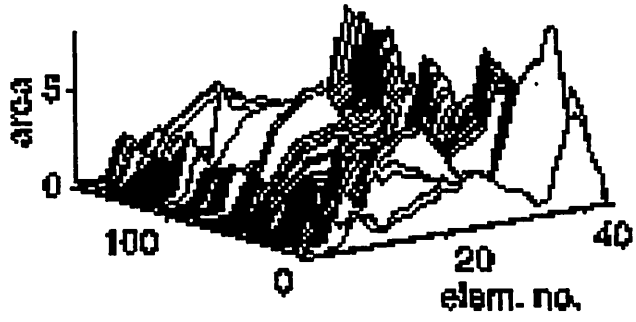


Figure 3. Selected area functions for sentence.

F3 converging toward F2. Both of these discontinuities occur when the vocal tract is very close to closure. The points were left in the database, as it was deemed likely that they could occur in reality. Note that the 12 sheets of constant c3 intersect each other, indicating a non-uniqueness problem. Each sheet attempts to span the entire F1-F2 vowel space (although some sheets 'focus' on a particular region).

The sentence "We were away a year ago" was recorded from subject BS, who was the original MRI subject. The speech was sampled at 20 kHz, and analyzed using fixed frame-size, autocorrelation-based LPC with 30 coefficients. The first three formant peaks were extracted, up to 4500 Hz. The database was searched using table lookup, the distance defined by both the formants and the previous frame's PCA coefficients. Since the database was not generated randomly, a local search was used (+/- 6 neighbours in all three dimensions) to speed up the search. Figure 3 shows the area functions selected from the database for the sentence. Speech reconstruction was accomplished by feeding the area functions into the vocal tract simulator, which inertially interpolated the area functions over time (this

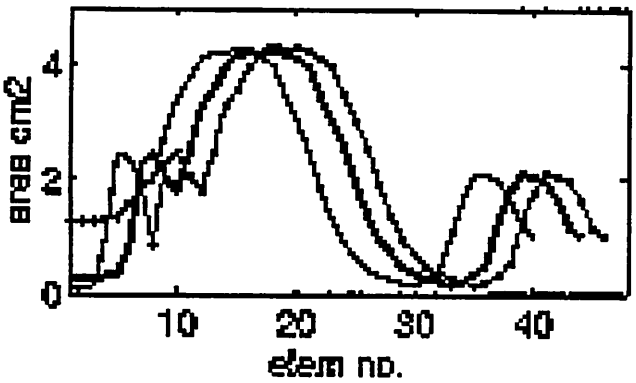


Figure 4. Normal (thick), twang (thin), sob (dotted) area functions.

reduced discontinuities between area functions). Four parameters were used to define a glottal pulse train- pitch period, LPC residual spike magnitude, open quotient and skewing quotient. Here, open and skewing quotient were held constant.

Given the articulatory nature of the database, it is easy to transform the area functions to effect different voice qualities. Two examples we have attempted are 'sob' and 'twang'. A male subject adept at mimicry was videotaped with a laryngoscope while producing these qualities. Observations indicated that sob could be accomplished by lengthening and enlarging the epilaryngeal region, while twang could be accomplished by shortening and constricting this region, along with opening the velar port. These changes were accomplished by manipulating the areas prior to reconstruction, as shown in Fig.4 for the vowel /i/. Informal listening tests of the simulated normal and affected qualities indicated that the manipulations were convincing.

Conclusions

A compact model of the vocal tract using 3 parameters has been created by using the first three eigenmodes derived from a PCA analysis of 12 vocal tract shapes. A speech-to-speech synthesis system has been implemented which easily allows vocal quality changes. An efficient model may thus consist of PCA weights transmitted at LPC frame rates, while the glottal parameters are transmitted at pitch cycle rates. Further study will focus on modeling consonants, introducing other voice source models, and speeding up the database mapping.

Acknowledgments

This work was supported by NIH Grant R01 DC02532.

References

[1] R. Harshman, P. Ladefoged, L. Goldstein, "Factor analysis of tongue shapes", JASA, 62(3), pp. 693-707, 1977.

[2] P. Meyer, R. Wilhelms, H.W. Strube, "A quasiarticulatory speech synthesiser for German language running in real-time", JASA, 86(2) pp. 523-539, 1989.

[3] B. Story, I.R. Titze, E.A. Hoffman, "Vocal tract area functions from magnetic resonance imaging", JASA, 100(1), pp. 537-554, 1996.

[4] B. Story, I.R. Titze, "Parameterization of vocal tract area functions by empirical orthogonal modes", JASA, in review

[5] B. Story, I.R. Titze, D. Wong, "A simplified model for simulation and transformation of speech". IEEE Int. Joint Symposia on Intelligence and Systems. Rockville, Maryland, Nov. 4-5, 1996.