UNIVERSITY OF CINCINNATI

MAY28 , 2003

I, ANTHONY DAVID MOSLEY ,
hereby submit this as part of the requirements for the degree of:

MASTER OF ARTS

in:

SPEECH-LANGUAGE PATHOLOGY

It is entitled:

INDUCED VOCAL FATIGUE IN HEALTHY OLDER MEN WITH NORMAL VOICES

Approved by:

Linda Lee

[illegible]

INDUCED VOCAL FATIGUE IN HEALTHY OLDER MEN WITH NORMAL VOICES

A thesis submitted to the

Division of Research and Advanced Studies
of the University of Cincinnati

in partial fulfillment of the
requirements for the degree of

MASTER OF ARTS

in the Department of Communication Sciences and Disorders
of the College of Allied Health Sciences

May 28, 2003

by

Anthony David Mosley

B.A., Trinity University, 1991
M.S., University if Illinois, 1995

Committee Chair: Linda Lee, Ph.D.
Professor/Director of Graduate Studies
Department of Communication Sciences and Disorders
University of Cincinnati

Reader: Lisa Kelchner, Ph.D.
Assistant Professor
Department of Communication Sciences and Disorders
University of Cincinnati

Abstract

No previous study observed the effects of a vocal endurance task on the voices of older men. For the current study, healthy men between the ages of 60 and 80 read loudly for two hours. Pre- and post-task measurements were taken of acoustics (fundamental frequency, fundamental intensity, frequency range, and maximum phonation time), laryngeal physiology (including glottic closure and supraglottic hyperfunction), perceptual quality of voice, and self-ratings of voice quality and phonatory effort.

Mean fundamental frequency was found to rise (not significantly), pre- to post-task, and to remain elevated during the 20-minute recovery period. Self-ratings of effort and voice quality worsened (significantly for voice quality). Glottic closure patterns tended to change, while laryngeal hyperfunction tended to increase. These findings appear to support the conclusion that the vocal mechanism of the healthy older male shows distinctive responses to a fatiguing vocal task, probably due to age-related degradation of muscle function.

Acknowledgments and Dedication

The author wishes to thank Dr. Lisa Kelchner for her dedication to this study, and, in particular, for her willingness to spend so many of what could otherwise have been free hours in the voice laboratory. Thanks also to Dr. Linda Lee for her excellent advice and commitment to making this thesis the best possible document. Additional recognition is due to William Mabrey, who volunteered many hours of his time in helping to make this study a reality.

This thesis is dedicated, with all my love,

to my wife, Leigh, who made it possible,

and to my daughter, Eloise,

who made it worthwhile.

Table of Contents

List of Tables and Figures

Induced Vocal Fatigue in Healthy Older Men with Normal Voices

Introduction

Jackson (1940) described "myasthenia laryngis" as an asthenia of the phonatory muscles of the larynx, particularly the thyroarytenoids. This condition (which, Jackson stated, bore no connection to myasthenia gravis) resulted from overuse of the voice. Since Jackson's time, numerous studies have analyzed the effects of stressful phonation upon the voices of persons with seemingly normal larynges. A substantial number of researchers have concluded that abuse of the voice can lead to a collection of symptoms that, taken together, characterize "vocal fatigue."

Although a precise definition of vocal fatigue is yet to be agreed upon, the disorder is generally viewed in the literature as stemming from functional misuse of the vocal mechanism. Loud, prolonged talk (Jackson, 1940; Kitch & Oates, 1994) and speaking while anxious or with laryngeal tension (Gotaas & Starr, 1993; Kitch & Oates, 1994; Koufman & Blalock, 1988; Novak, Dlouha, Capkova, & Vohradnik, 1991) have been named as causes of vocal fatigue. Laryngeal fatigue has been attributed to use of an inappropriate speaking pitch, whether high (Jackson, 1940; Sander & Ripich, 1983) or low (Koufman & Blalock, 1988). Poor breath stream control (Koufman & Blalock, 1988; Sander & Ripich, 1983) has also been cited as a source of vocal fatigue. Certain professions may predispose one towards developing vocal fatigue, including teaching (DeBodt, Wuyts, Van de Heyning, Lambrechts, & Abeele, 1998; Gotaas & Starr, 1993; Kostyk & Rochet, 1998; Rantala, Paavola, Korkko, & Vilkman, 1998), acting (Kitch &

Oates, 1994; Koufman & Blalock, 1988; Novak, et al., 1991) and singing (Jackson, 1940; Kitch & Oates, 1994; Koufman & Blalock, 1988), although some researchers have also concluded that persons with trained voices may be less vulnerable to symptoms of vocal fatigue (Gelfer, Andrews, & Schmidt, 1991, 1996; Scherer, Titze, Raphael, Wood, Ramig, & Blager, 1991).

Although research into the causes and characteristics of vocal fatigue has been conducted for decades, relatively little is known about the mechanisms by which vocal fatigue occurs in older persons. Age-related changes to the larynx and other physical structures can occur to members of both sexes even in persons free of detectable pathology (Mueller, 1997; Mueller, Sweeney, & Baribeau, 1984; Sinard & Hall, 1998). In men, age-related anatomical changes may contribute to altered function that can lead to discernible differences even in the voices of healthy individuals (Han & Yang, 1990; Hollien & Shipp, 1972; Mueller et al., 1984; Orlikoff, 1990; Ramig & Ringel, 1983). However, it is not currently known whether age-related changes to the vocal mechanism may cause older men to show distinct responses to conditions that fatigue the voice. In particular, no study has systematically observed the impact of a vocal endurance task upon the voices of healthy men who sound normal for their ages. The study described here investigated the effects of such a task upon the voices of healthy men in their seventh and eighth decades of life whose pre-task voices were free of perceptible dysphonia.

Review of the Literature

The following section examines research to-date on physiologic, acoustic, aerodynamic, and perceptual/self-perceptual characteristics of vocal fatigue. The current literature regarding the consequences of aging for the male voice is also reviewed.

Characteristics of Vocal Fatigue

Physiologic features. Change in glottic closure has been a frequent finding in studies of persons with healthy voices who undertook a fatiguing task (Linville, 1995; Stemple, Stanley, & Lee, 1995) or who complained of chronic laryngeal fatigue (Eustace, Stemple, & Lee, 1996). Eight of ten subjects studied by Stemple et al. showed abnormal glottic configurations (anterior chink, anterior and posterior chink, or incomplete closure) following two hours of loud reading. Although Linville's subjects were more likely to show post-task glottal areas that were changed rather than unchanged, the glottal configurations she found were more likely to be smaller or unchanged rather than larger, a result that may have been related to the shorter time period of Linville's fatiguing task (15 minutes) than that of Stemple et al. By comparison, of the 88 subjects complaining of chronic laryngeal fatigue whose files were reviewed by Eustace et al., the "vast majority" presented with anterior chinks, anterior and posterior chinks, or spindle-shaped closure (i.e. vocal fold bowing) (Eustace et al., 1996, p. 151). Stemple et al. suggested that the development of an anterior glottal chink may point to TA involvement and may be considered a diagnostic sign of vocal fatigue, while Eustace et al. suggested that other glottal configurations may also develop secondary to weakening of the laryngeal musculature, leading to increasingly severe symptoms of vocal fatigue.

A mild decrease in phase symmetry has been found in subjects complaining of vocal fatigue (Eustace et al., 1996), and an increased amplitude of glottal opening was found by Gelfer et al. (1996). Laryngeal hyperfunction, as indicated by supraglottic compression both anterior-posterior (Colton & Casper, 1996) and medial (Linville, 1995), has been found in vocally fatigued subjects. Sander and Ripich (1983) suggested that laryngeal hyperfunction, as an inappropriate means for increasing vocal intensity, may explain vocal fatigue following loud talking in persons who do not use adequate respiratory support to regulate intensity. Edema of the vocal fold mucosa (or "vascular engorgement," to use Jackson's term) has also been detected in vocally fatigued persons (Jackson, 1940; Scherer et al., 1991).

Acoustic features. Stemple et al. (1995) found a significant increase in average fundamental frequency when subjects chose a comfortable reading pitch following a two-hour-long fatiguing task. Toner (2002) also found a significant increase in mean fundamental frequency after her adolescent male subjects read loudly for two hours. Other studies (Eustace et al., 1996; Neils and Yairi, 1987; Novak et al., 1991) found no significant relationship between vocal fatigue and changes of speaking or reading fundamental frequency (SF_0). These differing results may reflect the fact that the fatiguing task was shorter (45 minutes) in the Neils and Yairi study than in those of Stemple et al. and Toner, while Novak studied changes in vocal characteristics of actors before and after performances, again suggesting vocal activities that may have been less fatiguing than the two hours of loud reading employed by both Stemple et al. and Toner. Eustace et al. researched the case histories of voice patients who had been diagnosed with chronic laryngeal fatigue, hence standardization of fatiguing tasks was not possible. They also found no deviations from the norm in the average fundamental frequencies of sustained vowels produced by their subjects, whereas both the vocally trained and the untrained subjects studied by Gelfer et al. (1991) showed significant increases in sustained fundamental frequencies for some (but not all) vowels following a one-hour fatiguing task.

Regarding the related measure of pitch range, Stemple et al. (1995) reported that the majority of their subjects had difficulty reaching their pre-task low pitch levels for sustained vowels, which the authors attributed to fatiguing of the TA muscle. The authors noted that high end measures were not affected, and in fact subjects were ultimately able to maintain their pre-test frequency ranges, both high and low. Similarly, Toner (2002) found no significant differences in pre- to post-task highest and lowest attainable pitches, although she did observe raised lowest pitches for eight of her 10 experimental subjects. Eustace et al. (1996) also found their vocally fatigued subjects to be within the normal limits for frequency range. However, Koufman and Blalock (1988) found that their subjects with functional musculo-skeletal tension disorders (tension-fatigue syndrome, or “Bogart-Bacall Syndrome”) tended to have significantly reduced vocal ranges compared to normal controls.

Gelfer et al. (1991, 1996) observed significantly increased speaking intensity in both their vocally trained and untrained subjects following fatiguing tasks, while Kostyk and Rochet (1998) found no significant changes in intensity across time in either their experimental subjects (teachers reporting five or more symptoms of vocal fatigue) or their control group. Gelfer et al. (1996) suggested that an increase in post-task intensity could be related to increased vocal fold contact, a less open glottal configuration during the most closed part of the cycle, and increased amplitude of vocal fold excursion. It may be surmised that their one-hour fatiguing task may have been sufficient to create the above glottal conditions in the absence of the more open (and presumably more fatigued) glottal configuration observed by Stemple et al. (1995).

Perturbation measures have varied across studies of vocal fatigue. Some researchers have detected altered jitter ratios in fatigued subjects (Eustace et al., 1996; Gelfer et al., 1991; Scherer, et al., 1991; Stemple et al., 1995), while other have detected no changes in frequency perturbation (Verstraete et al., 1993). Eustace et al. (1996) noted abnormal levels of jitter in male subjects aged 70-79, but found no other group to be outside norms. Scherer et al. (1991) noted a significant increase in jitter measures in their vocally trained subject (who spoke loudly for two and one-half hours) but not for their untrained subject (who spoke for one hour). The authors attributed the jitter observed in the first subject to vocal fold edema and/or neuromuscular fatigue of the cricothyroid muscle, and possibly of the vocalis. By comparison, Gelfer et al. (1991) saw an increase in jitter ratio (singing) for their untrained subjects and a decrease (sustained vowel) for their trained subjects. Stemple et al. (1995) also found reduced jitter measures (high pitch level) on post test, although none of their subjects showed abnormal jitter pre- or post-task. Verstraete et al. (1993) acknowledged that their 25-minute speaking task may not have been sufficiently stressful to alter the level of jitter in their subjects' voices.

Regarding intensity perturbation (shimmer), Scherer et al. (1991) again found no significant changes for the untrained subject, whereas the trained subject showed increased shimmer pre- to post-task. Shimmer values recovered more quickly than jitter values during the post-task rest period, which the authors attributed to a faster recovery of the respiratory system than of the laryngeal system. Verstraete et al. (1993) found no significant increase in shimmer following phonation at any pitch level, again perhaps owing to the relatively short period of the fatiguing task. Gelfer et al. (1991) found no significant changes (sustained vowels) in shimmer, pre- to post-task, for their vocally trained subjects, while untrained singers showed some post-task improvement in shimmer during singing.

Aerodynamic measures. Eustace et al. (1996) found abnormally reduced maximum phonation times for all but their youngest subjects, and increased flow rates for the males at comfortable and high pitch levels. The authors reported that, in addition to the presence of abnormal glottal configurations, the most consistent measures found to be unusual in their subjects were decreased maximum phonation times and abnormally high airflow rates. The authors stated that this combination of symptoms supported the hypothesis that symptoms of vocal fatigue were related to muscle weakness caused by voice abuse, misuse, or overuse.

Neils and Yairi (1987) found increased post-task air flow values, following 45 minutes of reading aloud, for two of three reading conditions (i.e. with background noise levels of 50 dB and 70dB, but not of 90 dB). Kostyk and Rochet (1998) observed different trends in the relationship of flow rate to laryngeal airway pressure in comparing the fatigued and non-fatigued teachers they studied. For the fatigued subjects, airflow significantly decreased while pressure remained relatively constant. By comparison, in the normal controls, airflow remained relatively constant while pressure was noted to increase. The latter observation, the authors speculated, may have signified that the control group more efficiently coordinated increased respiratory effort with laryngeal adjustments. On the other hand, the fatigued group may have relied on laryngeal valving alone to sustain laryngeal airway resistance.

Perceptual/self-perceptual characteristics. Neils and Yairi (1987) found no significant changes in judges' ratings of voice quality across the fatiguing task, at any background noise level. By comparison, Gotaas and Starr (1993) found that ratings of both their fatigued and non-fatigued teachers' voices tended to improve from morning to afternoon on days when the experimental subjects experienced no symptoms of vocal fatigue. However, on days when the subjects experienced fatigue, their voice ratings deteriorated as the day progressed, showing an increase in breathiness and a more strained quality. The authors suggested that teachers who fatigued had larynges and vocalization styles that were adequate unless they were subjected to unusual or excessive speaking activities, physical conditions, or psychological states.

Stemple et al. (1995) stated that eight of their ten subjects rated their speech as requiring maximum effort at the end of the vocally fatiguing task. Similarly, Toner (2002) noted significantly increased self-ratings of physical effort to speak, and significantly diminished self-ratings of voice quality, following a loud reading task; she cautioned, however, that some of her adolescent male subjects may have allowed their own expectations of how their voices should change to affect their ratings. Stemple et al.

reported that, following two hours of reading, common complaints among their subjects were of a feeling of dryness and tickle in the throat, as well as of dizziness, headache, chest tightness, upper backache, nausea, and oral numbness. Eustace et al. (1996) related subjective sensations of increased speaking effort and feelings of tension, thickness, or a lump in the throat with a failure to achieve complete glottal closure, and concomitant weakening or straining of the laryngeal muscles during speech.

Gelfer et al. (1996) reported that their subjects described subjective feelings of fatigue even in the absence of observable laryngeal symptoms. The authors theorized that perceptions of vocal fatigue may substantially precede any observable laryngeal damage. By comparison, Kostyk and Rochet (1998) reported that their subjects' perceptions of severity of vocal fatigue symptoms seemed to correspond well to changes in airflow measurements. Moreover, Scherer et al. (1991) found self-ratings of voice condition to be consistently and positively related to shimmer measures. In addition, Rantala et al. (1998) found statistically significant relationships between spectral characteristics of their subjects' voices and subjective complaints of symptoms of vocal fatigue.

Age-Related Characteristics of the Older Male Voice

Anatomic and physiologic changes due to aging. Mueller (1997) observed that "[t]he integrity of laryngeal control appears to be associated more with the health status of an individual than his or her chronological age" (Mueller, 1997, p. 162). Indeed, research with male subjects has suggested that physiologic condition (as characterized by measures such as cardiovascular health, percentage of body fat, and forced vital capacity) can be comparable in importance to chronological age in determining voice characteristics

(Orlikoff, 1990; Ramig & Ringel, 1983; Woo, Casper, Colton, & Brewer, 1992). Vocal training has also been credited with helping to limit voice changes in older professional performers (Boone, 1997; Sataloff, Rosen, Hawkshaw & Spiegel, 1997). An additional observation made by researchers, however, is that with increasing age groups of men show increasing *variability* in their voice characteristics, even when the men are considered healthy (Orlikoff, 1990; Ramig & Ringel, 1983). Orlikoff noted that, while his less healthy (in terms of cardiovascular condition) older male subjects showed approximately 105% more variability in overall acoustic measures than did his healthy younger male subjects, even his healthy older male subjects showed an increased variability of 80%. This led Orlikoff to warn of an "elderly individualism," beginning approximately at the end of the sixth decade of life, when "it becomes quite dangerous" to compare individuals with group means, in that "data merely represent general tendencies and trends" (Orlikoff, 1990, p. 454).

This leads to the point that older males considered healthy by researchers may still show differences of voice (and of vocal structures) compared to healthy younger males -- differences that we might characterize as signifying "normal aging." Numerous studies (Hirano, Kurita & Sakaguchi, 1989; Mueller, Sweeney & Baribeau, 1984; Stell, Stell & Watt, 1982) have detected age-related changes in the mucosa of the male vocal folds and larynx. Following histologic analysis of 64 human larynges (both sexes) ranging from 70 to 104 years old, Hirano et al. determined that while the overall thickness of the mucosa remained relatively constant in males, the mucosal cover tended to increase in thickness until age 70, and then to decrease; edema tended to appear in the superficial layer of the lamina propria; thickness of the intermediate layer tended to decrease, with elastic fibers becoming atrophic and the contour of the layer degenerating; and the deep layer tended to become thicker, with increasing disorder, and sometimes fibrosis, of the collagenous fibers. These changes led to an increase in stiffness of the folds, the authors posited, which, combined with an overall shortening of the folds with age, led to an increase in the older males' fundamental frequency. (It should be noted that, while Hirano et al. did not exclude smokers from their study, no significant differences were observed between the larynges of smokers and non-smokers.)

Stell et al. (1982) examined 105 larynges (both sexes) obtained from non-smoking subjects without chest disease ranging broadly in age. They compared changes in relative coverage by squamous (as opposed to respiratory) epithelium within the larynx, and found that, while the proportion of the supraglottic space covered by squamous epithelium did not change significantly for either sex, subjects (both sexes) in the age range 50-69 were most likely to show vestibular folds partially or totally covered by squamous epithelium, while the proportion of the subglottic space covered by squamous epithelium (when present) increased significantly in men but not in women. Squamous metaplasia was most likely to found in the vestibular folds of men at a mean age of 59.3, and the incidence and extent of subglottic squamous metaplasia increased significantly with age in men.

Perhaps the most widely recognized age-related laryngeal change is vocal fold "bowing" (or the presence of a spindle-shaped glottal configuration), which Sinard and Hall (1998) called "probably the most common benign pathology of the aging voice" (p.

77). Mueller et al. (1984), who did postmortem examinations of 25 larynges of men aged 60 to 88 years with no known diseases or abnormal conditions of the laryngo-pulmonary system, found that 76 percent showed "arrowhead" configurations of the glottis secondary to bowing and atrophy of the vocal cords. None of the ten larynges taken from middle-aged subjects (mean age 44.7 years) showed such a configuration. Other studies (Tanaka, Hirano & Chijiwa, 1994; Woo et al., 1992) have found lower incidence rates of vocal fold bowing in the absence of any clinically noticeable organic lesions. Tanaka et al. noted that, of the 12 patients they encountered who showed bowing (out of 226 dysphonic patients), seven were men over the age of 60. The authors attributed this finding either to subclinical disorders of the vocal fold mucosa, or to a decrease in depth of the vocal fold mucosa due to the aging process. Woo et al., who reviewed the charts of 151 dysphonic patients over the age of 60, found six who demonstrated vocal fold bowing that the authors attributed to the effects of age. This relatively low incidence led the authors to suggest that "age-related dysphonia" (as characterized by bowing, persistent air leakage, and weak voice projection) should be considered a diagnosis of exclusion, to be arrived at only after pathological causes have been eliminated.

Stemple (1998) listed several causes of age-related bowing of the vocal folds, including ossification of the laryngeal cartilages, atrophy or dystrophy of laryngeal muscle and the vocal ligament, fibrosis of the mucous membrane of the vocal folds, and sulcus vocalis. Tanaka et al. (1994) noted that the age and sex distribution both of their patients with "nonorganic" vocal fold bowing and of those with bowing secondary to sulcus vocalis was similar (i.e., in both groups, the largest segment was made up of men over 60). They also noted a reduction in amplitude of vibration and mucosal wave in both conditions, and posited that similar subclinical disorders of the mucosa might exist.

In addition to finding a high incidence of vocal fold bowing and atrophy in their older male specimens, Mueller et al. (1984) also detected fatty degeneration of laryngeal tissues and increased calcification of laryngeal cartilages. Sinard and Hall (1998) reported that laryngeal cartilages and joints may ossify as a consequence of age, possibly limiting adduction of the vocal folds and leading to a weak, breathy voice. Stell et al. (1982) suggested that, as their own investigations showed no significant change in the size of the male larynx with age, observed increases in the weight of the male larynx must be due to changes in its composition, or to increased thickness.

Non-pathological, age-related changes to structures other than the larynx can also account for changes in voice function. Loss of skeletal integrity and poor posture may lead to reduced efficiency of respiratory support (Boone, 1997; Gregg, 1997) and decreased respiratory volumes (Boone, 1997). Hearing loss may be substantial enough to affect characteristics of voice (Boone, 1997). Also, recurrent acid reflux may be associated with dysphonia in some persons (Jones, Lannigan, McCullagh, Anggiansah, Owen & Harris, 1990; Lumpkin, Bishop & Katz, 1989).

Increased use of medications is typically observed in older persons (Boone, 1997). Medication use can be responsible for numerous physical changes that may have consequences for voice function, among them mucosal drying in the larynx and upper respiratory tract (Boone, 1997; Sataloff, Lawrence, Hawkshaw, and Rosen,1994; Woo et al., 1992), poor neuromotor control (Woo et al., 1992), tissue swelling or edema of the vocal folds (Boone, 1997; Sataloff et al., 1994), rebound effects (Boone, 1997; Sataloff et al., 1994), nasopharyngeal irritation (Boone, 1997), gastric irritation, contact inflammation, dehydration, vocal fold hemorrhage, altered pulmonary function, increased salivation, and voice tremor (Sataloff et al., 1994). Long-term use of allergy control medications can cause permanent tissue changes and damage to the larynx (Boone, 1997), while prolonged steroid use can lead to wasting of the vocalis muscle (Sataloff et al., 1994). Additionally, Boone (1997) notes that the interaction effects of multiple medications may cause numerous negative side effects.

Age-related changes in vocal acoustics of the male voice. Following a study of 175 healthy males ranging in age from 20 to 89, Hollien and Shipp (1972) concluded that there was a progressive lowering of the average speaking fundamental frequency (SF_0) in males from pre-adolescence until sometime in the 40-50 year range, followed by a steady rise in SF_0 until old age. The authors attributed this finding to muscle atrophy, reduced vocal fold thickness, and increased vocal fold stiffness. By comparison, Ramig and Ringel (1983) did not find significant differences in fundamental frequency (during reading, spontaneous speech, or sustained vowel phonation) across the ages (26-74) or physical conditions they researched. (Interestingly, Ramig and Ringel found F_0 to be significantly higher during reading than during spontaneous speech across ages and physical conditions.) Some studies of average F_0 of sustained vowels have shown increased F_0 by age in males (Han & Yang, 1990; Mueller et al., 1984; age ranges of 55-74 and 60-92, respectively), while Orlikoff (1990) found no significant variation in mean F_0 across his experimental groups, aged 26-33 (healthy young men), 68-80 (healthy older men) and 60-79 (unhealthy older men), respectively. Orlikoff did note a trend toward increased F_0 with age, however, and suggested that greater differences might have been found if his elderly groups had contained no men younger than 80 and if subjects had been pushed to maximally sustain phonation.

Ramig and Ringel (1983) found that, across age groups, men in good physical condition (in terms of resting heart rate, blood pressure, percentage of body fat, and forced vital capacity) had significantly larger pitch ranges than did subjects in poor condition. They found no statistically significant interaction between age and physical condition for this measure, but noted that the greatest difference in means was found among their elderly subjects. By comparison, Han and Yang (1990) found that the natural voice ranges of their healthy male subjects (i.e. those free of laryngeal or pulmonary complaints) decreased with increasing age, but that falsetto range increased. They attributed this finding either to decreases in laryngeal muscle contraction or to thinning of the vocal fold edge.

Orlikoff (1990) found that all of his subjects were equally able to monitor their vocal sound pressure levels (SPLs) with the help of an analog readout, regardless of age or physical condition. However, compared with younger subjects, significantly greater standard deviations of peak-to-peak amplitude were noted in both healthy and unhealthy (in terms of cardiovascular condition) older subjects. Not surprisingly, measures of mean shimmer and mean percent shimmer also varied significantly between the young group and both groups (healthy and unhealthy) of older men. This matched the findings of Ramig and Ringel (1983), who noted significantly more shimmer in maximally sustained vowels produced by older than younger subjects, regardless of physical condition. However, Ramig and Ringel found no significant age-related differences with regard to jitter, although the greatest variation in mean jitter during maximally sustained phonation occurred within the elderly group. Similarly, Orlikoff found a significant difference, in terms of percentage jitter, between his healthy and unhealthy elderly subjects. Unlike Ramig and Ringel, however, Orlikoff found a significant increase in jitter when the combined older groups were compared to the younger group.

Ramig and Ringel (1983) concluded that their observed acoustic correlates of the aging voice may have reflected a breakdown in fine motor control and/or may have been due to slight degenerative tissue changes. Orlikoff (1990) suggested that the increased variability in acoustic characteristics among his atherosclerotic older subjects may have represented degeneration of laryngeal tissue, alteration in structures that controlled the phonatory mechanism, or vascular changes.

Age-related changes in male aerodynamics. Surprisingly little data is extant concerning aerodynamic characteristics of voice in healthy older males. However, Murty, Carding and Kelly (1991) found a near-doubling of the speed quotient (opening time/closing time) in their older male subjects (who were more than 60 years old) compared with their younger male subjects (all under 25 years old). The authors attributed this finding to a slowing of the opening phase of vocal fold vibration in the older subjects, rather than to a speeding up of the closing phase. By comparison, the authors found no significant difference in open quotient (time open/time of whole cycle) between the two groups, indicating that the proportion of the glottal cycle during which there was airflow was comparable between the two groups. The authors concluded that these findings might account in part for audible changes in the vocal quality of elderly persons, although they did not speculate as to what those changes might be.

Perceptual/self-perceptual characteristics. Following a study comparing SF0 characteristics to the perceived ages of male speakers, Horii and Ryan (1981) concluded that listeners tended to associate higher mean F0s with more advanced age, and vice versa. Morrison and Gore-Hickman (1986) reviewed the literature concerning voice disorders in the elderly and stated that terms used to describe vocal quality changes characterized by altered pitch included "thinned," "reedy," and "breathy." Mueller (1997) reported that naive listeners asked to provide three adjectives describing an older person's voice most frequently supplied the words "shaky," "squeaky," "weak," and "hoarse."

Mueller et al. (1984) noted that a relatively large number of normal elderly persons complained of vocal problems including change of pitch, compressed pitch range, reduced control, and altered pitch quality. As such, they stated, it is important to be able to discern changes caused by normal aging from ones caused by disease, abuse, or psychopathology. Stemple (1998) noted that he had encountered numerous patients suffering from age-related vocal fold bowing who complained of not being able to speak with sufficient loudness to overcome background noise or to be heard in social situations. This problem can be exacerbated, he added, when patients live alone and do little talking, or when their peers are hard of hearing. Stemple stated that senile laryngeous could lead to social isolation that was emotionally unhealthy and unacceptable to the patient.

Woo et al. (1992) noted that some patients may be satisfied with fairly dysphonic voices, while other may suffer damage to their self-image with even mild vocal symptoms. As such, the authors proposed, therapeutic decisions regarding dysphonia should be made based upon the perceptions of the patient, rather than those of the voice care professional.

Need for the Current Study

A great part of the research currently available on vocal fatigue has been done with women 50 years old or younger (e.g. Buekers, 1998; De Bodt, Wuyts, Van de Heyning, Lambrechts, & Abeele, 1998; Gelfer et al., 1991, 1996; Kostyk & Rochet, 1998; Linville, 1995; Neils & Yairi, 1987; Rantala et al., 1998; Scherer et al., 1991; Stemple et al., 1995; Verstraete, 1993). Studies that have involved both sexes have typically assessed characteristics of vocal fatigue in persons of working age who belonged to specific professions (e.g. Gotaas & Starr, 1993; Kitch & Oates, 1994; Koufman & Blalock, 1988; Novak et al., 1991). Eustace et al. (1996) retrospectively analyzed the acoustic, aerodynamic, and videostroboscopic measures of persons of both genders, up to the age of 79, who complained of chronic laryngeal fatigue. To date, however, no study had systematically attempted to *induce* vocal fatigue in older men with normal-sounding voices, and then to describe the consequences for the voice, if any, that resulted.

As numerous studies cited in the preceding section on age-related characteristics of the older male voice have shown, anatomical changes are likely to occur with advancing age even in healthy men. It seems reasonable to theorize that age-related changes to the vocal mechanism could affect older men's vulnerability to sustaining vocal fatigue, even if the men's voices are usually perceived as normal. It also seems sensible to suggest that use of a fatiguing task, as utilized in numerous studies cited in the above section on characteristics of vocal fatigue, might help to create fatigue (whose characteristics could then be analyzed) in the voices of vocally normal older men.

Use of "vocal endurance tests" to analyze vocal fatigue has recently come under criticism (Buekers, 1998; De Bodt, et al., 1998). However, the study described below proposed neither to attempt to *detect* vocal fatigue in symptomatic persons (like Buekers) nor to try to *predict* its future occurrence in persons with diverse vocal physiologies (like De Bodt et al.). Rather, the purpose of the current study was to attempt to *create* fatigue of the vocal mechanisms of healthy subjects without a history of voice problems, and then to describe any changes in vocal characteristics that followed the fatiguing task.

No prior study had assessed the character of vocal fatigue when intentionally induced in healthy older men free of any history of vocal disorder. Based on the findings of the study described here, it is hoped, a determination can begin to be made as to whether healthy older men with normal voices, in the midst of a period of physical change, demonstrate distinct physiologic, acoustic, perceptual, and/or self-perceptual responses to a vocally fatiguing task.

Research Hypotheses

The following hypotheses were tested during the current investigation.

H_1: When mean reading fundamental frequencies (F_0s) are calculated, pre- and post-task, there will be a significant difference between the changes in the means of the experimental and control groups.

H_2: When mean reading fundamental intensities (I_0s) are calculated, pre- and post-task, there will be a significant difference between the changes in the means of the experimental and control groups.

H_3: When mean maximum phonation times (MPTs) are calculated, pre- and post-task, there will be a significant difference between the changes in the means of the experimental and control groups.

H_4: When mean highest attainable pitches are calculated, pre- and post-task, there will be no significant difference between the changes in the means of the experimental and control groups.

H_5: When mean lowest attainable pitches are calculated, pre- and post-task, there will be no significant difference between the changes in the means of the experimental and control groups.

H_6: Expert perceptual ratings of overall vocal quality will be significantly lower, pre- to post-task, for the experimental subjects, but not for the control subjects.

H_7: When mean self-ratings of subjects' own voice quality are calculated, pre- and post-task, there will be a significant difference between the changes in the means of the experimental and control groups.

H_8: When mean self-ratings of subjects' own effort to speak are calculated, pre- and post-task, there will be a significant difference between the changes in the means of the experimental and control groups.

$H_{9:}$ Expert raters of videoendoscopic results will find glottic closure patterns to be more open, pre- to post-task, for a majority of the experimental subjects, and unaffected for a majority of the control subjects.

$H_{10:}$ Expert raters of videoendoscopic results will find increased supraglottic activity, pre- to post-task, for a majority of the experimental subjects, but not for a majority of the control subjects.

$H_{11:}$ Expert videoendoscopic ratings of vertical level approximation will be notably different, pre- to post-task, for a majority of the experimental subjects, but not for a majority of the control subjects.

$H_{12:}$ Expert videoendoscopic ratings of condition of the vocal fold edge will be notably different, pre- to post-task, for a majority of the experimental subjects, but not for a majority of the control subjects.

$H_{13:}$ Expert videoendoscopic ratings of vocal fold mobility will be notably different, pre- to post-task, for a majority of the experimental subjects, but not for a majority of the control subjects.

Methods

Subjects

Fourteen male subjects (eight experimental, six control) were recruited for the study. Subjects were randomly assigned to either the experimental or the control group. Experimental subjects ranged in age from 60 to 76. Control subjects ranged in age from 64 to 73. The mean age of the experimental group was 67.75 years (SD = 5.90). The mean age of the control group was 68.50 years (SD = 3.15).

Subjects were asked to complete questionnaires regarding general health, educational/vocational history, and vocal characteristics and usage. Prerequisites for the study included no current or prior history of voice disorder, laryngeal pathology, prolonged recent intubation, significant hearing loss, central nervous system disorder, cardiac/circulatory problems requiring restriction of activities, significant uncontrolled reflux, or significant respiratory ailments. Subjects were non-smokers and had no recent (<5 year) history of smoking. Subjects had voices judged to be of normal perceptual quality by an experienced speech-language pathologist. No subjects had received extensive vocal training (e.g. as would befit a professional singer or actor). Subjects were able to read at least at an elementary level, which was necessary for them to complete the reading task.

All subjects but one were recruited from the greater Cincinnati area, through volunteer organizations, senior centers, and religious congregations. One subject came from the state of Georgia.

Procedures

Informed consent was obtained from all subjects prior to their participation in the study. A verbal and written explanation of experimental procedures was provided to each subject. A number was assigned to each subject for the analysis and reporting of data to assure confidentiality.

For the 24 hours before his appointment, each subject was asked to refrain from excessive voice use (e.g. yelling, singing, whispering, or harsh throat clearing), to avoid caffeine, and to drink six to eight glasses of water.

Pre-task assessments: Self-perceptual. Before beginning the prolonged reading task, each subject was asked to make a rated judgment of how his voice sounded (vocal quality) during reading of the Rainbow Passage (Fairbanks, 1960), using a seven-point (1 = normal, 7 = very severely abnormal) equal interval scale (Aronson, McCaffrey, Litchy & Lipton, 1993). Each subject was also asked to rate the physical effort required to produce voice (phonatory effort) during the reading, again using a seven-point (1 = normal, 7 = very severely abnormal/complete fatigue) equal interval scale.

Pre-task assessments: Acoustic. Measures of reading fundamental frequency (F_0), reading fundamental intensity (I_0), maximum phonation time (MPT), and pitch range were taken instrumentally. A Shure Prologue Dynamic Microphone was held by the examiner at a microphone-to-mouth distance of approximately four inches for all tasks. Subjects were allowed to practice each task, with demonstration by the clinician as necessary.

Acoustic data were collected using Real Time Pitch software packaged with

Computerized Speech Laboratory (CSL) (Kay Elemetrics: Lincoln Park, New Jersey). Reading fundamental frequency (F_0) and reading fundamental intensity (I_0) were collected during reading of the Rainbow Passage. Subjects were asked to read the passage twice at their normal pitch and loudness.

To assess maximum phonation time, a counting technique was first used to help subjects find their modal pitch. Subjects were next asked to take a deep breath and then to sustain the vowel /a/ for as long as possible at modal pitch using a conversational level of intensity. Subjects were given three opportunities to perform this task, and their longest phonation time was recorded for analysis.

Frequency range was assessed using pitch trials during which subjects glided upward or downward in pitch while sustaining the vowel /u/ or /o/. Using three successive trials, subjects first glided from their mid-range to their highest possible tone, then from mid-range to their lowest possible tone. The sample with the greatest range in each direction was documented.

Pre-task assessments: Perceptual. While subjects were reading the Rainbow Passage for acoustic assessment, simultaneous audio recordings were made using a JVC TD-W354 Double Cassette Deck audio recorder. A desk-mounted Realistic Unidirectional Dynamic Microphone, placed at a microphone-to-mouth distance of approximately 18 inches, was used to record the readings. Audio recordings were subsequently analyzed for overall voice quality by two speech-language pathologists with at least 15 years each of professional experience.

Pre-task assessments: Laryngeal structure and function. Laryngeal structure and glottic closure were assessed using a Nagashima LS-3A laryngostroboscope with stroboscopic light source, a CCD Color Camera (model MN 401 E), a Panasonic AG-1980 SVHS VCR, and a Zenith color television monitor. The principal investigator, who had more than ten years' experience in performing videoendoscopic laryngeal examinations, advanced the rigid scope to the posterior oral cavity. Subjects were then requested to sustain the vowel /i/ at a comfortable pitch so that the vocal folds and surrounding structures could be clearly visualized. At least two clear one- to two-minute samples were collected. The samples were subsequently randomized, and were then rated by the principal investigator plus a second speech-language pathologist with more than ten years of experience in rating laryngeal function endoscopically.

Vocal endurance task. Before beginning the prolonged loud reading task, subjects were asked to drink six ounces of water and to avoid using their voices for five minutes. Subjects were then seated comfortably in a quiet room containing a sound level meter (SLM) and reading material of their choice. The SLM, which was placed at a distance of approximately 18 inches from the speaker's mouth, was used to measure subjects' vocal intensity throughout the reading task. Subjects were asked to maintain an intensity level of between 75 and 80 dB SPL throughout the reading task, which lasted for two hours.

During the reading task, a research assistant regularly monitored the speaker's loudness level, and cued subjects to change their intensity level if their voices fall outside the prescribed range. Every fifteen minutes during the reading task, the research assistant asked subjects to make subjective ratings of their vocal quality and phonatory effort, using the 7-point scales used during the pre-task assessment period.

Prior to beginning the reading task, subjects were informed that they were free to terminate their participation in the experiment at any time if they began to experience serious discomfort. Additionally, the clinicians listened for signs of undue dysphonia and discomfort in the subjects' voices, with the intent to terminate the participation of any subject whose voice seemed unduly affected by the fatiguing task. All subjects were able to complete the study.

Post-task assessments. At the conclusion of the two-hour reading task, subjects were asked to remain silent until they began post-task assessments. Post-task measures were first taken during a re-reading of the Rainbow Passage, which the subjects were instructed to read at their normal pitch and loudness. This reading was used to acquire an immediate post-task measure of reading fundamental frequency (F_0) and reading fundamental intensity (I_0). Audio recordings were made, which were subsequently randomized with the pre-task readings and rated for perceptual characteristics. When subjects finished the immediate post-task reading of the Rainbow Passage, they were asked to make subjective ratings of their vocal quality and phonatory effort, using the 7-point scales used throughout the earlier parts of the experiment.

Post-task measurements of pitch range, maximum phonation time, and laryngeal structure/glottic closure were also taken, using procedures identical to those used in the pre-task period. Additionally, subjects re-read the Rainbow Passage four more times (at five, ten, fifteen, and twenty minutes after the conclusion of the prolonged reading task), and new measures of F_0 were taken. This was done to track patterns of vocal recovery in the immediate aftermath of the fatiguing event.

Post-task directions to subjects. Subjects were asked to avoid excessive voice use during the 24-hour period following the conclusion of the experimental task. Subjects were also advised to drink six to eight glasses of water and to avoid caffeine and passive smoke.

At the end of the 24-hour period, a research assistant called subjects to document any after effects of their participation in the study. During the follow-up telephone conversation, subjects were asked whether they felt any residual effects of the reading task, and if so, what symptoms they were experiencing. If a subject no longer felt any effects, he was asked to state how much time had passed following the experiment before he felt completely normal. If after 24 hours had elapsed a subject had complained of vocal symptom not evident prior to his participation in the study, referral to ENT and/or Voice Lab would have been offered. All subjects reported that they felt fully recovered by the morning following their participation.

Control group procedures. Subjects randomly assigned to the control group participated in the same protocol as the experimental subjects, with one exception: Rather than reading loudly for two hours, controls sat silently for most of the two-hour period. Once every 15 minutes, a research assistant asked control group subjects to read aloud for three to five minutes using their normal levels of pitch and intensity. This was done in order to simulate normal voice use across a two-hour period.

Data analysis. Pre- and post-task measures of reading fundamental frequency (F_0), reading fundamental intensity (I_0), maximum phonation time, and pitch range were compared using unpaired t-tests to test equality of mean differences between the experimental and control groups. Pitch range was analyzed using separate scores for lowest F_0 and highest F_0, to account for the possibility that changes in range could occur at either or both extremes of pitch.

Following randomization of pre- and post-task audio recordings, overall quality of voice was rated by two speech language pathologists with more than 15 years of experience in analyzing voice disorders. Similarly, video recordings of stroboscopic exams were randomized, and observed changes in the structure or function of the larynx, pre- to post-task, were noted.

Additionally, subjects' self-ratings of voice quality and phonatory effort, immediately pre-task and immediately post-task, were compared using unpaired t-tests to test equality of mean differences between the experimental and control groups.

Results

The research hypotheses tested in this study were based on data documented in the literature review and on clinical evidence.

Acoustic Analysis

- H_1: When mean reading fundamental frequencies (F_0s) are calculated, pre- and post-task, there will be a significant difference between the changes in the means of the experimental and control groups.

Reading fundamental frequency (F_0). Table 1 reports mean F_0s and standard deviations, immediately pre- and post-task, for the experimental and control groups. Post-task minus pre-task scores were calculated, and unpaired t-tests were performed to test equality of mean differences between the two groups. No significant difference was found between the mean differences of the two groups ($p > .05$). Therefore, the hypothesis was rejected.

Although significant differences were not found, seemingly different trends were noted between groups in the relative effects of the reading tasks, pre- to post-task and during the subsequent 20 minute recovery period. As shown in Table 2, and illustrated in Figure 1, mean F_0 for the experimental group climbed from immediately pre-task to immediately post-task (mean difference = 6.11 Hz, SD = 14.60) and remained elevated throughout the 20 minute recovery period. By comparison, mean F_0 for the control group declined from immediately pre-task to immediately post-task (mean difference = -4.59 Hz, SD = 5.46) and remained depressed throughout the recovery period.

TABLE 1: Pre- and post-task means and standard deviations (SD) for acoustic variables of reading fundamental frequency, reading average intensity, maximum phonation time, and the highest and lowest pitches in the total pitch range. Differences between pre and post test values were calculated, and the probability of significant differences between experimental and control groups is presented.

VARIABLE	PRE-TASK MEAN	PRE-TASK SD	POST-TASK MEAN	POST-TASK SD	DIFFERENCE (POST - PRE)		*p*
Reading Fundamental Frequency (Hz)							
Control	118.23	12.14	113.64	11.97	-4.59	5.46	
Experimental	109.99	12.85	116.09	16.92	6.11	14.60	0.09
Reading Fundamental Intensity (dB SPL)							
Control	58.99	3.60	60.33	2.29	1.33	3.89	
Experimental	59.79	3.64	61.49	3.15	1.71	3.94	0.86
Maximum Phonation Time (s)							
Control	23.92	4.53	23.46	6.30	-0.45	5.02	
Experimental	18.60	6.65	18.43	6.37	-0.17	5.64	0.92
Highest Pitch (Hz)							
Control	552.62	57.81	545.69	63.12	-6.93	64.06	
Experimental	592.66	50.06	573.90	54.66	-18.76	96.02	0.80
Lowest Pitch (Hz)							
Control	81.91	13.79	70.61	11.90	-11.29	12.74	
Experimental	75.70	14.62	77.76	18.24	2.07	17.15	0.14

Additionally, it was noted that, while the ***highest*** mean F_0 shown by the experimental group occurred five minutes after the completion of the fatiguing task (mean = 123.83 Hz, SD = 15.10), the *lowest* mean F_0 for the control group was also measured five minutes post-task (mean = 109.07, SD = 10.81). Indeed, while five of the eight experimental subjects demonstrated raised F_0s immediately post-task (compared to their pre-task F_0s), by five minutes post-task seven showed raised F_0s, while the remaining subject's F_0 was within .03 Hz of his pre-task F_0. At 20 minutes post-task, five of the eight experimental subjects still registered F_0s that were higher than immediately pre-task, two had F_0s that were lower, and one demonstrated a F_0 that was within one Hz of his immediate pre-task score.

By comparison, while five of the six control subjects showed lower immediate post-task F_0s (compared to their pre-task F_0s), by five minutes post-task all six controls demonstrated reduced F_0s compared to pre-task. Twenty minutes post-task five of the six controls continued to show reduced F_0s, while one showed a higher F_0 than at pre-task.

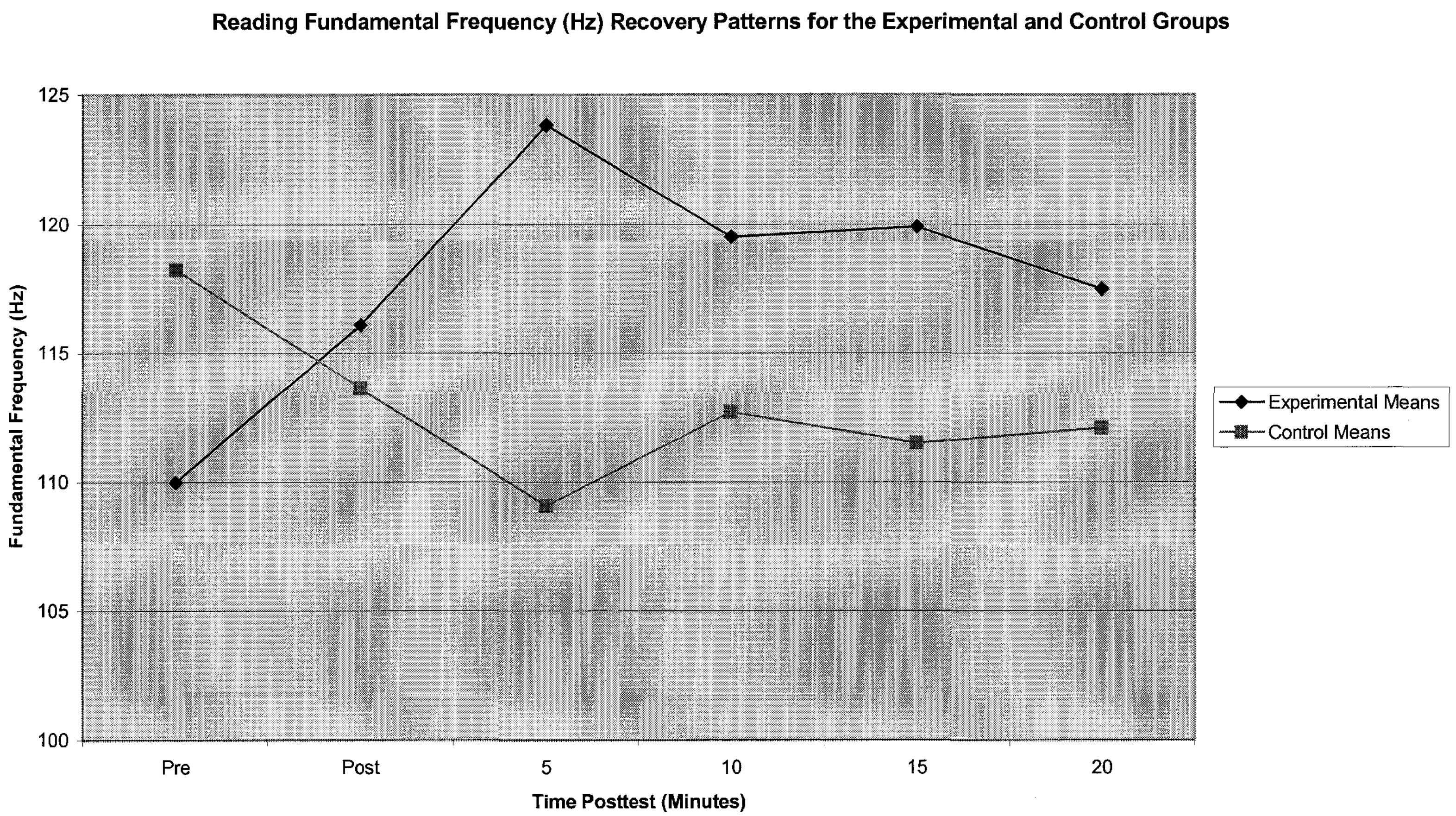
- H_2: When mean reading fundamental intensities (I_0s) are calculated, pre- and post-task, there will be a significant difference between the changes in the means of the experimental and control groups.

Reading fundamental intensity (I_0). Table 1 reports mean I_0s and standard deviations, immediately pre- and post-task, for the experimental and control groups. Post-task minus pre-task scores were calculated, and unpaired t-tests were performed to test equality of mean differences between the two groups. No significant difference was

TABLE 2: Mean reading fundamental frequencies (F_o), immediately pre-task, immediately post-task, and during the 20-minute recovery period.

F_o (Hz)	Immediately Pre-Task	Immediately Post-Task	5 Minutes Post-Task	10 Minutes Post-Task	15 Minutes Post-Task	20 Minutes Post-Task
Control	118.23	113.64	109.07	112.71	111.51	112.1
Experimental	109.99	116.09	123.83	119.51	119.91	117.48

FIGURE 1: Mean pre- and post-task fundamental frequencies, and mean fundamental frequencies at five, ten, fifteen, and twenty minutes post-task, experimental and control groups.



found between the mean differences of the two groups ($p > .05$). Therefore, the hypothesis was rejected.

As shown in Table 1, mean change in I_0 was by less than two dB SPL for both the experimental group (mean difference = 1.71 dB SPL, SD = -3.94) and control group (mean difference = 1.33 dB SPL, SD = 3.89).

- H_3: When mean maximum phonation times (MPTs) are calculated, pre- and post-task, there will be a significant difference between the changes in the means of the experimental and control groups.

Maximum Phonation Time (MPT). Table 1 reports mean MPTs and standard deviations, immediately pre- and post-task, for the experimental and control groups. Post-task minus pre-task scores were calculated, and unpaired t-tests were performed to test equality of mean differences between the two groups. No significant difference was found between the mean differences of the two groups ($p > .05$). Therefore, the hypothesis was rejected.

It should be noted that, both pre- and post task, the mean MPT of the control group was approximately five seconds longer than that of the experimental group. In both circumstances, the MPT scores of the experimental group were slightly more variable than were those of the control group.

- H_4: When mean highest attainable pitches are calculated, pre- and post-task, there will be no significant difference between the changes in the means of the experimental and control groups.

Pitch range: Highest pitch. Table 1 reports mean highest pitches and standard deviations, immediately pre- and post-task, for the experimental and control groups. Post-task minus pre-task scores were calculated, and unpaired t-tests were performed to test equality of mean differences between the two groups. No significant difference was found between the mean differences of the two groups ($p > .05$). Therefore, the hypothesis was accepted.

- H_5: When mean lowest attainable pitches are calculated, pre- and post-task, there will be no significant difference between the changes in the means of the experimental and control groups.

Pitch range: Lowest pitch. Table 1 reports mean lowest pitches and standard deviations, immediately pre- and post-task, for the experimental and control groups. Post-task minus pre-task scores were calculated, and unpaired t-tests were performed to test equality of mean differences between the two groups. No significant difference was found between the mean differences of the two groups ($p > .05$). Therefore, the hypothesis was accepted.

<u>Perceptual analysis: Expert raters</u>

- H_6: Expert perceptual ratings of overall vocal quality will be significantly lower, pre- to post-task, for the experimental subjects, but not for the control subjects.

Voice quality. For the experimental group, mean overall voice quality using a seven-point scale (1 = normal, 7 = aphonic) was rated to be 1.38 (SD = .5175) pre-task, and 1.25 (SD = .4629) post-task. A paired samples t-test for related measures found no significant change in perceived quality, pre- to post-task ($p > .05$).

There was no change in perceived overall voice quality for the control group, as all controls received perceptual voice quality ratings of 1 both pre- and post-task (mean = 1, SD = .00). Therefore, the hypothesis was rejected.

<u>Perceptual analysis: Subject self-ratings</u>

- H_7: When mean self-ratings of subjects' own voice quality are calculated, pre- and post-task, there will be a significant difference between the changes in the means of the experimental and control groups.

Voice quality. Table 3 reports mean self-ratings of vocal quality (1 = normal, 7 = severely abnormal) immediately pre- and post-task for the experimental and control groups. Post-task minus pre-task scores were calculated, and unpaired t-tests were performed to test equality of mean differences between the two groups. The comparison of mean differences for the experimental and control groups, pre- to post-task, was significant ($p = .03$). Therefore, the hypothesis was accepted.

For the control group, the mean self-rating for vocal quality, pre- to post-task, was unchanged (mean difference = 0.00, SD = .89). By comparison, the experimental group's mean score for vocal quality increased (i.e. self-ratings of quality declined) by a mean amount of 1.31 (SD = 1.03), indicating they felt mildly dysphonic following the task.

Among the control group's members, two subjects concluded that their voice quality had worsened, two felt that it had improved, and two gave their voices the same rating, pre- to post-task. No control changed his rating of vocal quality by more than one point in either direction. By comparison, seven of the eight experimental subjects concluded that their voice quality had worsened by the end of the reading task, while one

TABLE 3: Pre- and post-task means and standard deviations for the subject self ratings of vocal quality and phonatory effort. Both ratings were based on a 7-point scale, where 1=normal and 7=severely abnormal. Differences between pre and post test values were calculated, and the probability of significant differences between experimental and control groups is presented.

VARIABLE	PRE-TASK MEAN	PRE-TASK SD	POST-TASK MEAN	POST-TASK SD	DIFFERENCE (POST - PRE)		*p*
Vocal Quality							
Control	1.83	.98	1.83	.75	0.00	.89	
Experimental	1.81	1.51	3.19	1.41	1.38	1.16	0.03
Phonatory Effort							
Control	1.50	.55	1.83	.75	0.33	0.52	
Experimental	1.69	.80	3.00	1.31	1.31	1.03	0.06

gave his voice the same rating both pre- and post-task. Three of the experimental subjects changed their self-ratings for vocal quality by two points or more.

- H_8: When mean self-ratings of subjects' own effort to speak are calculated, pre- and post-task, there will be a significant difference between the changes in the means of the experimental and control groups.

Effort. Table 3 reports mean self-ratings of effort to speak (1 = normal, 7 = severely abnormal/complete fatigue) immediately pre- and post-task for the experimental and control groups. Post-task minus pre-task scores were calculated, and unpaired t-tests were performed to test equality of mean differences between the two groups. No significant difference was found between the mean differences of the two groups. Therefore, the hypothesis was rejected. However, near-significance was observed ($p = .06$).

Effort to speak was rated as increasing, pre- to post-task, by six of the eight experimental subjects, while the remaining two claimed to be exerting the same effort post-task as pre-task. Two of the control subjects felt their effort to speak had increased post-task, while the remaining four reported no change in their effort to speak. As noted in Table 3, the mean difference in effort to speak was greater, pre- to post-task, for the experimental group (mean difference = 1.31, SD = 1.03) than for the control group (mean difference = .33, SD = .52).

<u>Videostroboscopy: Expert ratings</u>

- $H_{9:}$ Expert raters of videoendoscopic results will find glottic closure patterns to be more open, pre- to post-task, for a majority of the experimental subjects, and unaffected for a majority of the control subjects.

Glottic closure. Table 4 includes data describing glottic closure patterns, pre- and post-task, for the experimental and control groups.

Four experimental subjects were rated as having complete glottic closure pre-task, although one of those four showed “touch” closure with a slight spindle shape anteriorly. Of the same four, one continued to show complete glottic closure at post-task evaluation, while the other three demonstrated small anterior gaps, including the one who had shown “touch” closure pre-task.

One experimental subject was rated during the pre-task assessment as having a slight anterior gap, which had enlarged to a small anterior gap at post-task. Another experimental subject showed a small anterior gap during both pre-task and post-task evaluations.

One experimental subject was rated as having a medium anterior gap pre-task, and a small anterior gap post-task. Similarly, the remaining experimental subject showed a large anterior gap pre-task, and a medium anterior gap post-task.

In all, of the eight experimental subjects, four showed more open glottic configurations, pre- to post-task, two showed less open patterns, and two showed no detectable change in glottic closure.

By comparison, only one of the six control subjects demonstrated complete glottic closure pre-task; following the reading task, the same subject demonstrated a slight anterior gap, or “anterior slit.” Another control subject showed a slight, or “minimal,” anterior gap pre-task; at post-task, the anterior gap had grown and was rated as “small.”

TABLE 4: Expert ratings of glottic closure patterns. Measures of laryngeal function, including glottic closure, supraglottic activity, vertical level approximation, vocal fold edge, and vocal fold mobility.

LARYNGEAL ACTIVITY	FREQUENCY OF OCCURRENCE, EXPERIMENTAL		FREQUENCY OF OCCURRENCE, CONTROL	
	PRE	POST	PRE	POST
Glottic Closure				
Complete:	4	1	1	2
Anterior Gap:				
Slight:	1	0	1	1
Small:	1	6	3	2
Medium:	1	1	1	0
Large:	1	0	0	1
Supraglottic Activity				
Latero-Medial Compression:				
None:	3	1	2	1
Mild:	1	2	3	3
Mild-Moderate:	1	1	1*	2
Moderate:	3*	4	0	0
	(*1 left side only)		(*left side only)	
Supraglottic Activity				
Antero-Posterior Compression:				
None:	5	3	4	4
Mild:	1	2	2	2
Mild-Moderate:	1	1	0	0
Moderate:	1	2	0	0

TABLE 4, continued: Expert ratings of glottic closure patterns.

LARYNGEAL ACTIVITY	FREQUENCY OF OCCURRENCE, EXPERIMENTAL		FREQUENCY OF OCCURRENCE, CONTROL	
	PRE	**POST**	**PRE**	**POST**
Vertical Level Approximation				
Equal:	8	8	6	6
Vocal Fold Edge				
Smooth, Straight:				
Right:	8	8	6	6*
Left:	8	8	6	6*
			(*1 control showed slight bowing, both folds, post-task)	
Vocal Fold Mobility:				
Normal:				
Right:	8	8	6	6
Left:	8	8	6	6

Three control subjects showed small anterior gaps pre-task, one with a tendency toward "spindle shape," or bowing. Of those three, the subject who had begun with a spindle-shaped glottis continued to demonstrate both a small anterior gap and a tendency toward bowing at post-task; one showed a large anterior gap at post-task analysis; and one demonstrated complete glottic closure, although with a tendency to come apart anteriorly.

The remaining control demonstrated a medium anterior gap, with a tendency to bow anteriorly, at pre-task; following the reading task, he demonstrated complete glottic closure. No control subjects had large anterior gaps pre-task.

Of the six control group members, three showed more open glottic configurations at post-task than at pre-task, two showed less open patterns, and one showed the same type of glottic closure both pre- and post-task. Therefore, the hypothesis was rejected.

- $H_{10:}$ Expert raters of videoendoscopic results will find increased supraglottic activity, pre- to post-task, for a majority of the experimental subjects, but not for a majority of the control subjects.

Supraglottic activity. During pre-task assessment, three of the eight experimental subjects showed no latero-medial or antero-posterior compression. Of those three, one continued to show no latero-medial or antero-posterior compression at post-task, one showed mild latero-medial compression but no antero-posterior compression, and one demonstrated both mild-moderate latero-medial compression and mild antero-posterior compression at post-task.

One experimental subject showed mild latero-medial compression and no antero-posterior compression both pre- and post-task. Another experimental subject demonstrated mild-moderate latero-medial compression and moderate antero-posterior compression pre-task, which changed to moderate latero-medial compression and moderate antero-posterior compression by post-task examination.

One experimental subject had moderate latero-medial compression and mild antero-posterior compression pre-task, which by post-task became moderate latero-medial compression combined with moderate antero-posterior compression. Another showed moderate latero-medial compression without antero-posterior compression, pre-task; at post-task, he showed moderate latero-medial compression and mild antero-posterior compression. The remaining experimental subject was rated at pre-task as having moderate latero-medial compression of the left side only, combined with mild-moderate antero-posterior compression; at post-task, he was found to demonstrate moderate latero-medial compression on both sides, and he continued to show mild-moderate antero-posterior compression.

Overall supraglottic activity was judged to have increased, pre-task to post-task, in six of the eight experimental subjects. In the remaining two experimental subjects, supraglottic activity was rated as being the same both pre- and post-task. None of the eight experimental subjects demonstrated a reduction in hyperfunction from pre-task to post-task assessment.

Among the control subjects, one showed no latero-medial or antero-posterior compression at pre-task evaluation; at post-task he demonstrated mild-moderate latero-medial compression, but continued to show no antero-posterior compression. Another control was rated at pre-task as having no latero-medial compression, combined with mild antero-posterior compression; he continued to show mild antero-posterior compression at post-task, but was found to have mild-moderate latero-medial compression.

Two control subjects demonstrated mild latero-medial compression and no antero-posterior compression at pre-task. Of those two, one continued to demonstrate mild latero-medial compression with no antero-posterior compression at post-task, while the other was found to have mild-moderate latero-medial compression, still without antero-posterior compression, at post-task.

An additional control subject showed mild latero-medial compression and mild antero-posterior compression at pre-task evaluation. At post task, however, it was found that he no longer demonstrated true hyperfunction, although "raised tissue" within the larynx was noted.

The remaining control subject was found to show mild-moderate latero-medial compression on the left side only, combined with mild antero-posterior compression, during pre-task assessment. At post-task, he was found to show mild latero-medial compression (both sides), still with mild antero-posterior compression.

For the six members of the control group, overall supraglottic activity was judged to have increased in two of the subjects, pre- to post-task; to have been reduced in two of the subjects; and to have remained the same in two of the subjects. Therefore, the hypothesis was accepted.

- $H_{11:}$ Expert videoendoscopic ratings of vertical level approximation will be notably different, pre- to post-task, for a majority of the experimental subjects, but not for a majority of the control subjects.

Vertical level approximation. All eight experimental subjects and all six control subjects had equal vertical level approximation both pre- and post-task. Therefore, the hypothesis was rejected.

- $H_{12:}$ Expert videoendoscopic ratings of condition of the vocal fold edge will be notably different, pre- to post-task, for a majority of the experimental subjects, but not for a majority of the control subjects.

Vocal fold edge. All eight experimental subjects and five of the six control subjects had smooth, straight vocal fold edges (left and right folds) both pre- and post-task; one control subject showed slight bowing of both folds, post-task. Therefore, the hypothesis was rejected.

- $H_{13:}$ Expert videoendoscopic ratings of vocal fold mobility will be notably different, pre- to post-task, for a majority of the experimental subjects, but not for a majority of the control subjects.

Vocal fold mobility. All eight experimental subjects and all six control subjects had normal vocal fold mobility (left and right folds) both pre- and post-task. Therefore, the hypothesis was rejected.

Additional observations. The two experienced speech-language pathologists who rated the videostroboscopic exams observed mucus, pre-task, upon the laryngeal structures of two of the experimental subjects. They also noted that secretions appeared within the pyriform sinuses of a third experimental subject during his pre-task assessment. None of those subjects was noted to show an abnormal collection of mucus or other secretions during post-task assessment.

One control subject was rated as showing an abnormal accumulation of mucus within the larynx both before and after the reading task. Another control showed no such accumulation pre-task, but did demonstrate increased laryngeal mucus during post-task assessment. A third control was observed, both pre- and post-task, to have a small varix upon the right vocal fold.

Self-reported physical symptoms during reading task: During the course of the two-hour reading tasks undertaken by both groups, subjects were asked every 15 minutes whether they were experiencing any abnormal physical symptoms. During the first hour of the reading task, one experimental subject said his mouth felt dry, while another spoke of thirstiness and mild throat soreness; a third complained of boredom. During the second hour, two experimental subjects said they felt an overall tiredness, one said his throat felt dry, and another said his throat felt "raspy." One experimental subject said he felt eye fatigue and had a sore back, while the subject who had earlier complained of a sore throat stated that the soreness had increased from mild to moderate.

Four of the control subjects described no abnormal physical effects at any point during the reading task. One control observed during the second hour that he would like to take a nap, but added that he "could do this [reading task] all day." A second control stated during the first hour that he was experiencing "scratchiness" in his throat, as well as mild hoarseness; however, he reported by the end of the first hour that the "scratchiness" was going away. During the second hour, the same subject complained of having tired eyes and slightly increased hoarseness, but added the "scratchiness" was continuing to subside. Near the end of the second hour this subject began to note a slight soreness at the back of his throat.

Long-term recovery time

Following the fatiguing task, three experimental subjects reported feeling fully recovered right away. Three said that they felt fully recovered within two to three hours of the end of their participation, while the remaining two said that they felt normal by the following morning.

Of the control subjects, two said they felt normal right away, and one reported feeling fully recovered by the time he had driven home. One control reported a recovery time of less than one hour, while another needed two to two and one-half hours to feel fully normal. One control subject stated that he felt fully recovered by the following morning.

Discussion

None of the subjects in this study demonstrate vocal fatigue or other functional voice disorders pre-task. The larynges of the subjects were free of pathology, and reflect the normal aging of the laryngeal tissues. Results of the present investigation indicated some effects of the reading task on the post-test measures taken. It is believed that those effects reflected the responses of normal, rather than disordered, larynges of men between the ages of 60 and 80. Those effects are examined in the present chapter.

Acoustic findings

Fundamental Frequency (F_0). Although a significant pre- and post-test difference was not observed in mean reading F_0 of the experimental and control groups, five out of the eight experimental subjects demonstrated raised F_0s immediately post-task. This trend may have been expected to reach significance with a larger number of subjects. Additionally, seven subjects reached their highest F_0s five minutes post-task, which was an unexpected finding because it had not been noted in previous literature. The five-minute fundamental frequency was not included in the statistical analysis comparing the group F_0s before and after the experimental task, which may partially explain the lack of significant differences. This effect was observed in the recovery patterns. The tendency for elevation of F_0 following loud reading may support the findings of other studies (Gelfer et al., 1991; Stemple et al., 1995; Toner, 2002) that found significant increases in F_0 following fatiguing tasks. Some studies (Neils and Yairi, 1987; Novak et al, 1991) found no such increases in F_0 due to vocal fatigue. However, as noted earlier, both Novak et al. and Neils and Yairi studied changes in F_0 across tasks that may not have been as fatiguing for the voice as were those in the studies that found significant changes in F_0. The reading task used in the present task was most similar to that used by Stemple et al. and Toner.

Both Stemple et al. (1995) and Toner (2002) theorized that the increases in F_0 that they observed may have resulted from straining of the thyroarytenoid (TA) muscles during the course of the fatiguing task. The subjects studied were relatively young in all three of the above studies that showed significant changes in F_0 (i.e. Stemple et al. and Gelfer et al. studied young women, while Toner studied adolescent males). The subjects evaluated during the current study were older males. Given that changes in the vocal fold mucosa and atrophy of the TA have been found to be common characteristics of male larynges even in older men who are healthy (Hirano et al., 1989; Mueller et al., 1984), the results obtained here have added interest in that they should reflect the results of placing prolonged strain upon vocal folds that may already be compromised by structural degradation.

This observation (that aging vocal folds may respond uniquely to stress) may be particularly relevant when one considers the recovery patterns (in terms of F_0) shown by the two groups in this study. It may be that two members of the experimental group, whose F_0s were found to rise above pre-task levels in the five minutes between the end of the fatiguing task and the second post-task measurement of F_0, showed a delayed effect of the fatiguing task upon their vocal folds. Alternatively, it may be that other post-task measures that were taken in the interim (including maximum phonation time and pitch range) added additional fatigue for these two subjects and caused their F_0s to increase.

Five of the control subjects showed lowered F_0s immediately post-task, possibly indicating a "warm-up" effect upon the voice as a result of their less-fatiguing reading task. It is also possible that the single control subject whose F_0 declined later than the others, but was below pre-task level at five minutes post-task, may have shown a sort of delayed warm-up effect. Alternatively, it may be that the assessments performed upon this subject during the five minutes after his first post-task measurement of F_0 may have served to warm up his vocal mechanism.

The notion of a "warm-up effect" following a relatively unstressful vocal task may be supported by the findings of Linville (1995), who noted that her subjects generally showed improved glottic closure following only 15 minutes of loud reading. On the other hand, Scherer et al. (1991) found that their untrained subject began to feel her voice deteriorating after just fifteen minutes of loud reading. As such, it may be that "warming up" of the voice, if it does happen as a consequence of loud speech, may be an event that varies across persons.

Some of the older adults in this study did not appear to recover from the fatiguing task as quickly as younger subjects; i.e. their larynges did not appear to be as resilient. Compared to their pre-task scores, four of Toner's (2002) adolescent male experimental subjects showed F_0s at 20 minutes post-task that remained elevated, while six were below their pre-task frequencies. All of Toner's experimental subjects demonstrated a decline in F_0 between the end of the fatiguing task and the end of the 20-minute recovery period. By comparison, five of the eight experimental subjects in the current study still showed F_0s at 20 minutes post-task that were higher than their pre-task scores. Three of those five had F_0s at 20 minutes post-task that remained above their F_0s immediately post-task. It seems possible that this may reflect a decrease in the resilience of the TA (compared to younger males) in some healthy older males, with the result that the muscle continues to show signs of strain for a longer period following stress, and is, for a time, unable to overcome the pull of the cricothyroid muscles sufficiently to lower F_0 to its habitual level.

The results of the current study also differed with those found by Toner (2002) with regard to post-task changes in the control group's F_0s. While half (i.e. two) of Toner's controls showed decreased F_0's 20 minutes post-task, the F_0s of the other two had increased. For the present study, by contrast, five of the six controls continued to show reduced F_0s at twenty minutes post-task. Hence, if a warm-up effect may be said to have occurred for the controls in the present study, its effects appear to have been durable. Since adolescent males and men in their sixties and seventies are both members of groups with transitioning vocal systems, further study of the comparative responses of these two groups to mildly fatiguing vocal events might prove enlightening.

Intensity (I_0). With regard to changes in mean fundamental intensity (I_0) due to a fatiguing vocal task, the findings of the current study contradicted those of Gelfer et al. (1991, 1996) in that the current study showed minimal, nonsignificant change for either group, pre- to post-task. Indeed, the current study may have reinforced the findings of Kostyk and Rochet (1998), who noted no significant change in intensity across time for either their fatigued or control groups. Kostyk and Rochet theorized, however, that their two groups may have used different combinations of laryngeal and respiratory function, one efficient and one inefficient, in order to maintain consistent airway resistence. This observation may be relevant to the current investigation, and will be addressed further shortly.

Maximum Phonation Time. The findings of the current study regarding maximum phonation time are congruent with those of Stemple et al. (1995) and Toner (2002), who found no significant changes in MPT from pre- to post-task. Maximum phonation times (MPTs) were highly consistent, in the current study, for both the experimental and control groups. Moreover, in the present study, all of the control subjects and all but one of the experimental subjects demonstrated pre-task MPTs that fell within the normal range for "aged" males found by Kent, Kent and Rosenbeck (1987 [cited in Colton and Casper, 1990]); all subjects fell within the normal range at post-task. This contradicts the findings of Eustace et al. (1996), who noted abnormally reduced maximum phonation times for all but their youngest subjects. An important difference between the study by Eustace et al. and the current one, however, was that Eustace et al. studied subjects whose voices were chronically fatigued, whereas (like Stemple et al. and Toner) the current investigation attempted to induce vocal fatigue in subjects who presented with normal voices. It may be surmised, then, that the MPTs measured by Eustace et al. reflected the long-term effects of functional misuse upon the vocal mechanism (including highly abnormal glottal configurations and increased airflow rates).

Pitch Range. The findings of the current study were inconclusive with regard to the effects of the fatiguing task upon the pitch ranges of the two groups. Both the experimental and control groups showed mean highest and lowest pitches that surpassed the means noted by Ptacek, Sander, Maloney, and Jackson (1966 [cited in Colton &

Casper, 1990]) and Colton (1982 [cited in Colton & Casper, 1990]). In terms of the mean overall pitch ranges of the two groups, neither group's range changed substantially. This makes it difficult to suggest either that the fatiguing task significantly affected (i.e. reduced) the mean pitch range of the experimental subjects, or that a "warm-up" effect occurred that notably expanded the pitch ranges of the controls.

These results appear to support the findings of Stemple et al (1995), who found that their experimental subjects were ultimately able to maintain their pre-task pitch ranges despite observed difficulty in reaching their earlier lowest pitches. Similarly, although Eustace et al. (1996) did not employ a fatiguing task, their subjects, despite being chronically fatigued, were found to have normal frequency ranges. Unlike with Toner (2002), the present study showed no clear trend in the effects of the fatiguing task upon the lowest attainable pitches of the experimental subjects, as four of the present study's experimental subjects showed an increase in lowest pitch, three saw a decrease, and one demonstrated essentially the same lowest pitch both pre- and post-task. The current study also appears to contradict the results of Koufman and Blalock (1988), who found that their subjects with "Bogart-Bacall" syndrome showed significantly reduced vocal ranges. In results analogous to those found by Stemple et al., the findings of the current study appear to suggest that, whatever the weakening effects upon the TA muscle of the prolonged loud reading task, they were not sufficient to significantly alter the pitch range of the healthy older men who participated.

<u>Physiologic (videostroboscopic) ratings</u>

The findings of the current study with regard to pre- and post-task glottal configuration appear to be unique within the literature. In that half (four) of the experimental subjects and all but one (five) of the controls showed at least slight anterior gaps pre-task, the findings of the present study might at first glance seem analogous to those of Eustace et al (1996). However, unlike in the Eustace study, none of the subjects who participated in the current study were believed to be vocally fatigued pre-task. Hence, while Stemple et al. (1995) proposed that the presence of an anterior glottal gap might be considered a diagnostic sign of vocal fatigue, this did not appear to have held true for these subjects.

However, it may be suggested, based on the observation of Stemple et al. (1995) that the presence of an anterior chink may point to TA involvement, that the subjects observed in this study demonstrated glottic configurations indicative not of vocal fatigue, but of structural changes within the TA muscle. That is to say, the "abnormal" glottic configurations observed pre-task in the majority of subjects may have been diagnostic not of pathology or functional misuse, but of normal aging.

Hence, it is difficult (and perhaps, given the relative difference in subject ages, inappropriate) to attempt to place this study's findings regarding change in glottic configuration in response to vocal fatigue on the continuum proposed by Stemple et al. (1995). On one hand, while Linville (1995) found that her subjects were more likely to show glottic configurations that were either smaller or unchanged following a 15-minute reading task, half of the experimental subjects in the present study demonstrated more open glottal patterns post-task -- as did half of the controls, who, somewhat like Linville's subjects, undertook a relatively non-fatiguing vocal task. On the other hand, while seventy percent of Stemple et al.'s experimental subjects developed anterior chinks as a result of the two-hour fatiguing task, half of the experimental subjects in the current study showed glottal configurations that either stayed the same or became smaller, pre- to- post-task. And while neither Gelfer et al. (1996) nor Toner (2002) observed significant changes in glottic configuration as a result of reading tasks (one hour and two hours long, respectively), six of the eight experimental subjects in the current study, and five of the six controls, showed altered glottal configurations, pre- to post-task.

Given the diversity of responses to the experimental and control tasks, in terms of change in glottic configuration, observed during this study, it may be appropriate to recall Orlikoff's (1990) warning to be cautious when comparing the laryngeal responses of individual males past the sixth decade of life with group means. Perhaps the divergent laryngeal physiologies observed in response to a fatiguing task (or even the relatively undemanding control task) reflected the sort of "elderly individualism" that Orlikoff suggested could be expected by researchers.

Results were more consistent, during the present study, when estimates of supraglottic hyperfunction were taken. Supraglottic activity, both lateral and medial, was observed in a majority of the subjects, both experimental and control. While this appears to support the findings of Colton and Casper (1996) and Linville (1995), it must be cautioned that, in the case of the current study, supraglottic compression was observed *before* the reading task for five of the eight experimental subjects and five of the six controls. Hence, if we are to accept that the subjects in this study were not vocally fatigued pre-task, it may perhaps be assumed that the supraglottic configurations shown by these subjects represented compensatory postures, meant to accommodate for glottal weakness, of the sort described by Neils and Yairi (1987).

Sander and Ripich (1983) stated that laryngeal hyperfunction, as an inappropriate means for maintaining loudness in the absence of adequate breath support, may explain vocal fatigue in those who employ it. During the current study, six of the eight experimental subjects (versus only two of the control subjects) showed increased overall supraglottic activity, pre- to post-task. Notably, however, mean intensity barely changed, pre- to post-task, for either group. These findings appear to correspond to those of Kostyk and Rochet (1998), who found that their non-fatigued controls tended to increase their translaryngeal airway pressure, during the course of a speaking day, while maintaining constant airflow. By contrast, subjects who tended to fatigue were found to show decreasing airflow as the day progressed, while maintaining constant pressure. Intensity levels remained constant for both groups, leading the authors to suggest that the two groups adjusted in different ways to the need to maintain habitual laryngeal airway resistance. The non-fatigued controls met this need, the authors theorized, by coordinating greater respiratory effort with laryngeal adjustments, while the fatigued subjects were thought to have relied exclusively on adjustments at the laryngeal valve, a method known to be less efficient and more stressful for the laryngeal muscles.

Since it is believed that half of the experimental subjects were already compensating for somewhat open glottic configurations pre-task (perhaps due to atrophy of the TA muscle, or to changes in the vocal fold mucosa), it may be theorized that the degree of hyperfunction they showed, both pre- and post-task, would have depended upon how well they compensated through respiratory support for imperfect glottal closure. As in the study by Kostyk and Rochet (1998), fundamental intensity (I_0) remained essentially the same, pre- to post-task, for the experimental group. Given the facts that (a) only one of eight experimental subjects showed normal glottic closure by the end of the reading period, and (b) six of eight showed at least some increase in hyperfunction, it would seem that a majority of the experimental subjects were compensating for a somewhat open glottal configuration by post-task. In fact, of the six experimental subjects who showed increased hyperfunction, five were found to demonstrate a somewhat open glottal configuration at post-task. Also of interest was that, of the five experimental subjects who showed raised average F_0s at 20 minutes post-task, four showed increased supraglottic hyperfunction, again suggesting that laryngeal compression had been increased to compensate for tiring of the TA muscles.

As Neils and Yairi (1987) suggest, it is not loud voice, per se, that produces vocal fatigue, but *overuse* of loud voice; moreover, the fatiguing process may be accelerated, they say, by producing loud voice in a hyperfunctional manner. As in the current study, Neils and Yairi observed little change in the perceptual quality of their subjects' voices, pre- to post-task. The authors acknowledged that their relatively brief (45 minute) reading task may not have been sufficiently stressful to produce consistent symptoms of vocal fatigue in their subjects. However, regarding both that study and the current one, it may be surmised that, had the fatiguing task been longer, dysphonia would eventually have been observed due to strain upon the laryngeal musculature; if Neils and Yairi are correct, this dysphonia would have appeared first, in both studies, in those subjects who attempted to compensate for fatigue of the glottal muscles via laryngeal hyperfunction.

This conclusion is supported, with regard to the present study, by the findings of Woo et al. (1992). In their study of 151 dysphonic patients over the age of 60, the authors noted that a common finding was poor breath support during sustained phonation. The authors observed that their subjects tended to expend too much air in the early part of the phonatory cycle (due, perhaps, to open glottal configurations); toward the latter part of expiration, the authors state, the subjects' larynges assumed a "squeezed, hyper-adducted posture" (p. 142). It may be suggested, then, that among those subjects in the current study who tended to compensate for open glottic configurations with supraglottic hyperfunction, the risk of eventually developing dysphonic voices may have been greater compared to those who were able to increase respiratory support in order to maintain their normal speaking intensity.

The lack of significant change, in the current study, of subjects' measures of vertical level approximation, vocal fold edge, and vocal fold mobility, accord with the findings of Toner (2002), who found little change for these parameters, and of Eustace et al. (1996), who found limited abnormality of vertical approximation or vocal fold edge even in the chronically fatigued subjects they studied. The finding of no pre-task vocal fold bowing for any subject agreed with the findings of Tanaka et al. (1994) and Woo et al. (1992), who found low incidences of bowing even among dysphonic subjects. The finding that one control subject showed mild bowing at post-task was surprising, but may reflect pre-existing atrophy or dystrophy of laryngeal muscle, which Stemple (1998) stated was one cause of age-related vocal fold bowing. That this subject may have demonstrated atrophy of the TA muscles is reinforced by the fact that he presented pre-task with a small anterior gap and mild latero-medial and antero-posterior supraglottic compression; the anterior gap had become large, concurrent with reduced hyperfunction, at post-task.

With regard to additional observations made during videostroboscopy, it is not clear whether dehydration may have been the primary reason for the presence of abnormal secretion levels in some subjects. Subjects who demonstrated mucus or other secretions fell along a broad continuum that included both the oldest (76) and one of the youngest (64) subjects in the study. Although two of the three experimental subjects who demonstrated laryngeal secretions pre-task reported drinking at around the mean amount of caffeine per day (2.28 cups, SD=2.00), and one reported drinking substantially more than the mean amount of alcohol per week (3.38 drinks, SD = 3.89), the third subject reported that he drank no caffeine or alcohol and consumed a great deal of water. Among the controls, the subject who showed mucus both pre- and post-task reported drinking three cups of coffee per day, while the control who only showed mucus pre-task claimed to drink two to three alcoholic beverages daily.

The control subject who presented with a small varix upon the right vocal fold, both pre- and post-task, had also complained of “scratchiness” in his throat during his first hour of reading. In that the subject reported that the “scratchiness” had largely subsided by the end of the two hour reading task, this may lend weight to the suggestion that some of the controls may have experienced a “warm-up” effect, leading to improved function of voice, while participating in the relatively non-strenuous control task.

Expert ratings of overall voice quality

That expert ratings of overall voice quality did not change significantly pre- to post-task agrees with the findings of Neils and Yairi (1987) and Toner (2002). Although both Koufman and Blalock (1988) and Gotaas and Starr (1993) found increased dysphonia in the voices of their fatigued subjects, those two studies were performed using subjects with histories of chronic fatigue. As in the present study, it appears that the loud speaking tasks employed by Neils and Yairi (45 minutes) and Toner (two hours) were not sufficiently fatiguing to produce detectable increases in dysphonia among experimental subjects.

In light of the current investigation, the observation by Eustace et al. (1996) that their chronically vocally fatigued subjects consistently demonstrated high airflow rates, decreased maximum phonation times, and incomplete glottal closure, is of interest. As the authors observed, high airflow rates would support the presence of breathiness, which would presumably be interpreted as dysphonic by listeners. Given the frequency, in the current study, with which somewhat open glottic configurations were found among both the experimental and control groups, the lack of perceived breathiness seems to lend support to the idea that the subjects were preventing, via hyperfunctional laryngeal postures, what might otherwise have been perceptible levels of breathiness. This may have been true in particular during post-task, when a substantial majority of the experimental subjects showed increased overall supraglottic activity, presumably due to encroaching fatigue of the laryngeal muscles. If correct, would seem to lend further support to the observation of Kostyk and Rochet (1998) that fatigued persons may use inefficient valving of the larynx to maintain subglottic pressure while reducing airflow. Future studies including aerodynamic measurements appear warranted.

Hence, a possible difference between chronic laryngeal fatigue and temporary, experimentally induced fatigue may be that, short-term, the larynx may be able to compensate for an open laryngeal configuration, while sustaining a loud voice for an extended period of time, without audible changes to the voice. In a similar vein, Neils and Yairi (1987) theorized that, while a singer may claim to be able to "sing over" a speaking voice disorder (p. 110), a longer sustained effort would be likely to result in an inability to counter the symptoms of fatigue. Thus, the hypothesis can again be posited that, while some of the subjects in the current study may have employed compensatory hyperfunctioning of the larynx to make their voices sound normal post-task, it could be expected that, had the fatiguing task been more extended in length, those subjects might have been among the first to demonstrate audible dysphonia.

Perceptual analysis: Subject self-ratings

For the current study, subjects were asked to rate their perceptions of the quality of their own vocal output, and of the effort they were expending in producing voice, compared to normal. Gelfer et al. (1996) observed that many of their subjects complained of feeling vocally fatigued, following a one-hour period of loud reading, despite the fact that few had demonstrated notable laryngeal alterations post-task. The authors suggested that a person's perceptions of fatigue may occur substantially before any laryngeal damage can be observed. Given that notable laryngeal changes, including increased supraglottic hyperfunction, were found in a majority of the experimental subjects during the current study, it should perhaps not be surprising that mean self-ratings of voice quality changed significantly compared to those of the controls, while self-ratings of change in vocal effort approached significance.

The findings of the current study regarding self-perception of voice change agreed with those of Toner (2002), who noted significantly increased ratings of physical effort, and significantly diminished ratings of voice quality, by her subjects. She suggested, however, that some of her subjects, knowing the nature of the study in which they were participating, may have allowed their expectations as to how their ratings would be expected to change to influence their self-ratings. Given the risk of such an effect with subjective data, the same could have held true for the current study.

On the other hand, Scherer et al. (1991) found that their vocally trained subject's self-evaluation of the quality of her voice proved to correlate significantly with the shimmer level present in her voice. Eustace et al. (1996) noted that the increased effort required to produce loud voice in the presence of incomplete glottal closure may, through the process of weakening or straining the laryngeal muscles, cause some patients to complain of feelings of tension or a "lump in the throat" (p. 153). Moreover, Kostyk and Rochet (1998) stated that their subjects' ratings of the severity level of their vocal fatigue appeared to correspond well to changes in their airflow means, i.e. as airflow rates declined, subjects tended to rate their fatigue as being more severe. This finding could be of particular relevance to the current study, given the hypothesis, stated above, that a majority of the experimental subjects who showed increased hyperfunction, pre- to post-task, were attempting to compensate for at least somewhat open glottal configurations by employing supraglottic hyperfunction to reduce their airflow rates. Indeed, of the six experimental subjects who demonstrated increased hyperfunction, pre- to post-task, five rated their effort as increasing, and six rated their quality as declining.

It must ultimately be acknowledged, however, that subjective ratings are, by their nature, limited in the consistency with which they can be expected to be employed across subjects. This point is illustrated by the difference in results found by Stemple et al. (1995) and the current study, both of which employed a two-hour reading task, and both of which requested that subjects rate their own level of effort. While eight of ten subjects in the study by Stemple et al. claimed to be employing maximum effort to speak by the end of the two hour task, only one experimental subject in the current study rated his level of effort as being higher than a four (out of seven); as reported previously, the mean rating of effort for the experimental group, post-task, was 3.00 (SD = 1.31), while the mean for the control group was 1.83 (SD = .75). It is, of course, possible that these apparent discrepancies represent a real difference between the physical symptoms felt, at the end of a two-hour reading task, by the female graduate students studied by Stemple et al. and the men who participated in the current study. Alternatively, it may be that the men from the current study held to different internal standards, compared to Stemple et al.'s subjects, as to what constituted "effort. A third possibility is that some portion of the men did not wish to admit that they were expending great effort. (Indeed, one experimental subject rated his levels of both effort and vocal quality as normal throughout the two hours of the fatiguing task, despite being observed to show an increasing amount of coughing over time.) It may be that, as suggested by Toner (2002), optimum use of a subjective scale may demand a detailed explanation to subjects of standard meanings entailed by different scores, rather than dependence upon the internal referents of the subjects.

Self-reported symptoms of the fatiguing task (including mouth dryness, thirstiness, throat soreness, hoarseness, and overall fatigue), as described by experimental subjects in the current study, appear to have accorded well with subjective complaints described by other authors (e.g. Scherer et al., 1991; Stemple et al., 1995; Toner, 2002). Self-reported recovery times (i.e. the amount of time needed to feel completely normal, post-task) also appear to have been in line with those reported by these authors.

Toner (2002) noted a mild-to-moderate correlation between speaking fundamental frequency (F_0) and self-ratings of effort. Based on the results of the current study, this observation may, perhaps, be added to: Self-ratings of effort appear, for this group of subjects, to have related to an attempt, on the part of the speaker, to compensate for tiring of the TA muscle (as demonstrated by increased F_0), and for the presence of a consistently open glottic configuration (due to weakening of the TA muscle and/or degeneration of the vocal fold cover), by increasing supraglottic hyperfunction. This conclusion may be a fruitful one for testing by future researchers.

Conclusions

Despite Orlikoff's (1990) warning to be wary of "elderly individualism" when generalizing about healthy older men, based upon the findings of this study, it may be possible to state that the population of healthy men between the ages of 60 and 80 displays some distinctive responses to a fatiguing vocal task. Fundamental frequency appears to rise, perhaps due to tiring of the thyroarytenoid muscles. Moreover, laryngeal recovery time, as reflected in the presence of a raised fundamental frequency post-task, may be extended compared to those of younger groups.

In terms of physiology, members of this group seem likely to present with somewhat open glottal configurations, including anterior gaps, even in the absence of audible dysphonia. Such a configuration may pose no difficulty for the speaker during normal voice use. However, when required to undergo a vocally stressful activity, the speaker is likely to attempt to compensate by employing supraglottic hyperfunction, presumably to maintain a consistent fundamental speaking intensity. While this compensatory posture may prove effective at masking breathiness over the short term, if prolonged, it may lead to dysphonia. Also, declining self-ratings of voice quality and vocal effort may reflect increasing laryngeal hyperfunction.

All of these results make it seem likely that subtle degradation of muscular support will be a common finding even among healthy men of this age group. In particular, age-related changes to the vocal fold, including dystrophy or atrophy of the thyroarytenoid, may be found. It is hoped that future research will further refine these conclusions.

References

Boone, D. R. (1997). The singing/acting voice in the mature adult. Journal of Voice, 11, 161-164.

Buekers, R. (1998). Are voice endurance tests able to assess vocal fatigue? Clinical Otolaryngology, 23, 533-538.

Colton, R. H., & Casper, J. K. (1990). Understanding voice problems. Baltimore: Williams and Wilkins.

Colton, R. H., & Casper, J. K. (1996). Understanding voice problems: A physiological perspective for diagnosis and treatment (2nd ed.). Baltimore: Williams and Wilkins.

De Bodt, M. S., Wuyts, F. L., Van de Heyning, P. H., Lambrechts, L., & Abeele, D. V. (1998). Predicting vocal outcome by means of a vocal endurance test: A 5-year follow-up study in female teachers. The Laryngoscope, 108, 1363-1367.

Eustace, C. S., Stemple, J. C., & Lee, L. (1996). Objective measures of voice production in patients complaining of laryngeal fatigue. Journal of Voice, 10, 146-154.

Gelfer, M. P., Andrews, M. L., & Schmidt, C. P. (1991). Effects of prolonged loud reading on selected measures of vocal function in trained and untrained singers. Journal of Voice, 5, 158-167.

Gelfer, M. P., Andrews, M. L., & Schmidt, C. P. (1996). Documenting laryngeal change following prolonged loud reading: A videostroboscopic study. Journal of Voice, 10, 368-377.

Gotaas, C., & Starr, C. D. (1993). Vocal fatigue among teachers. Folia Phoniatrica, 45, 120-129.

Gregg, J. W. (1997). The singing/acting mature adult -- Singing instruction perspective. Journal of Voice, 11, 165-170.

Han, Z., & Yang, S. (1990). Analysis of voice of healthy aged persons with narrow band spectrum analyzer. Auris Nasus Larynx, 17, 45-48.

Hirano, M., Kurita, S., & Sakaguchi, S. (1989). Ageing of the vibratory tissue of human vocal folds. Acta Oto-Laryngologica, 107, 428-433.

Hollien, H., & Shipp, T. (1972). Speaking fundamental frequency and chronologic age in males. Journal of Speech and Hearing Research, 15, 155-159.

Horii,, Y., & Ryan, W. J. (1981). Fundamental frequency characteristics and perceived age of adult male speakers. Folia Phoniatrica, 33, 227-233.

Jackson, C. (1940). Myasthenia laryngis. Archives of Otolaryngology, 32, 434-463.

Jones, N. S., Lannigan, F. J., McCullagh, M., Anggiansah, A., Owen, W., & Harris, T. M. (1990). Acid reflux and hoarseness. Journal of Voice, 4, 355-358.

Kitch, J. A., & Oates, J. (1994). The perceptual features of vocal fatigue as self-reported by a group of actors and singers. Journal of Voice, 8, 207-214.

Kostyk, B. E., & Rochet, A. P. (1998). Laryngeal airway resistance in teachers with vocal fatigue: A preliminary study. Journal of Voice, 12, 287-299.

Koufman, J. A., & Blalock, P. D. (1988). Vocal fatigue and dysphonia in the professional voice user: Bogart-Bacall Syndrome. Laryngoscope, 98, 493-498.

Linville, S. E. (1995). Changes in glottal configuration in women after loud talking. Journal of Voice, 9, 57-65.

Lumpkin, S. M. M., Bishop, S. G., & Katz, P. O. (1989). Chronic dysphonia secondary to gastroesophageal reflux disease (GERD): Diagnosis using simultaneous dual-probe prolonged pH monitoring. Journal of Voice, 3, 351-355.

Morrison, M. D., & Gore-Hickman, P. (1986). Voice disorders in the elderly. The Journal of Otolaryngology, 15, 231-234.

Mueller, P. B. (1997). The aging voice. Seminars in Speech and Language, 18, 159-169.

Mueller, P. B., Sweeney, R. J., & Baribeau, L. J. (1984). Acoustic and morphologic study of the senescent voice. Ear, Nose and Throat Journal, 63, 71-75.

Murty, G. E., Carding, P. N., & Kelly, P. J. (1991). Combined glottographic changes in the elderly. Clinical Otolaryngology, 16, 532-534.

Neils, L. R., & Yairi, E. (1987). Effects of speaking in noise on vocal fatigue and vocal recovery. Folia Phoniatrica, 39, 104-112.

Novak, A., Dlouha, O., Capkova, B., & Vohradnik, M. (1991). Vocal fatigue after theater performance in actors. Folia Phoniatrica, 43, 74-78.

Orlikoff, R. F. (1990). The relationship of age and cardiovascular health to certain acoustic characteristics of male voices. Journal of Speech and Hearing Research, 33, 450-457.

Ramig, L. A., & Ringel, R. L. (1983). Effects of physiological aging on selected acoustic characteristics of voice. Journal of Speech and Hearing Research, 26, 22-30.

Rantala, L., Paavola, L., Korkko, P., & Vilkman, E. (1998). Working-day effects on the spectral characteristics of teaching voice. Folia Phoniatrica et Logopaedica, 50, 205-211.

Sander, E. K., & Ripich, D. E. (1983). Vocal fatigue. Annals of Otology, Rhinology and Laryngology, 92, 141-145.

Sataloff, R. T., Lawrence, V. L., Hawkshaw, M. J., & Rosen, D. C. (1994). Medications and their effects on the voice. In M. S. Benninger, B. H. Jacobson, & A. F.

Johnson (Eds.), Vocal arts medicine: The care and prevention of professional voice disorders (pp. 216-225). New York: Thieme Medical Publishers, Inc.

Sataloff, R. T., Rosen, D. C., Hawkshaw, M., & Spiegel, J. R. (1997). The aging adult voice. Journal of Voice, 11, 156-160.

Scherer, R. C., Titze, I. R., Raphael, B. N., Wood, R. P., Ramig, L. A. , & Blager, R. F. (1991). Vocal fatigue in a trained and an untrained voice user. In T. Baer, C. Sasaki, & K. S. Harris (Eds.), Laryngeal function in phonation and respiration (pp. 533-555). San Diego, CA: Singular Publishing Group, Inc.

Sinard, R. J., & Hall, D. (1998). The aging voice: How to differentiate disease from normal changes. Geriatrics, 53, 76-79.

Stell, P. M., Stell, I. M., & Watt, J. (1982). Age changes in the epithelial lining of the human larynx. Gerontology, 28, 208-214.

Stemple, J. C. (1998). Some thoughts related to voice therapy for senile laryngeous. Voice and Voice Disorders (December), 6-7.

Stemple, J. C., Stanley, J., & Lee., L. (1995). Objective measures of voice production in normal subjects following prolonged voice use. Journal of Voice, 9, 127-133.

Tanaka, S., Hirano, M., & Chijiwa, K. (1994). Some aspects of vocal fold bowing. Annals of Otology, Rhinology and Laryngology, 103, 357-362.

Toner, M. (2002). Effects of prolonged loud reading on normal adolescent male voices: A study of vocal fatigue. Unpublished master's thesis, University of Cincinnati, Cincinnati, OH.

Verstraete, J., Forrez, G., Mertens, P. & Debruyne, F. (1993). The effect of sustained phonation at high and low pitch on vocal jitter and shimmer. Folia Phoniatrica, 45, 223-228.

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

APPENDIX I: SUBJECT CASE HISTORY FORM

SUBJECT CASE HISTORY

1. PERSONAL INFORMATION:

Initials: **Date:**

Age:

Smoking History: **Never** **Past Use:**

2. MEDICAL HISTORY

Current/past neurological problems: No Yes/explain:___________

Current/past GI (reflux) problems: No Yes/explain:___________

Current/past breathing/lung problems: No Yes/explain:___________

Current/past heart/circulatory problems: No Yes/explain:___________

Current/past high blood pressure: No Yes/explain:___________

Current/past diabetes: No Yes/explain:___________

Any past debilitating illnesses: No Yes/explain:___________

____

Current/past allergies: No Yes/explain:___________

Current/past dysphagia (swallowing difficulty) No Yes/explain:_________

Current/past physical trauma (to head / neck): No Yes/explain:_________

Past surgeries with dates (related to head and neck area): ___________________

____

____

Past history of intubation?____

Current medications (prescription and over-the-counter, including aspirin and anti-histamines): ____

APPENDIX I: SUBJECT CASE HISTORY FORM

____

____

Medication side effects?: ____

____

____

Current/past hearing problems: No Yes/explain:___________

Current/past voice problems: No Yes/explain:___________

Current/past disease of larynx/throat? No Yes/explain:___________

Previous speech/voice therapy: No Yes/explain:___________

General Health per subject description: poor fair good excellent

3. EDUCATIONAL/VOCATIONAL HISTORY

Years of education: _____ Highest degree obtained: _____

Occupation: ___________ Years in longest-held position: _____

Total years worked: _____ Are you retired? _____ If so, for how long? _____

APPENDIX II: VOICE QUESTIONNAIRE

VOICE QUESTIONNAIRE

Initials: Date:

Age:

Briefly describe your voice:____________________________________
____

Have you noticed any changes to your voice? _____

If so, please describe: ____
____

When did the changes begin? ____

How frequently do you exercise? _____ Please describe your typical workout:
____

Do you do any vocal performing (e.g. sing in a choir)? _____ If so, please describe: ____

What is the average amount of water you consume each day?________

What is the average amount of caffeine you consume each day?______

Do you drink alcoholic beverages? _____ If so, frequency/amount: ___________

Does your voice tire? No Yes If so, when:

Does your voice get hoarse? No Yes If so, when:

Does your voice get breathy? No Yes If so, when:

Does your voice sometimes become weaker? No Yes If so, when:

Does speaking sometimes seem to require great effort? No Yes If so, when:

APPENDIX II: VOICE QUESTIONNAIRE

Do you sometimes experience physical discomfort (such as throat pain) during or after speaking? No Yes If so, please describe: ______________________

____

Do you sometimes experience tension or tightness in the throat and/or chest during or after speaking? No Yes If so, when:

Do you notice changes in your voice during the course of the day? No Yes

If so, please describe:

____

____

Briefly describe how you typically use your voice on a daily basis (e.g. do you sometimes sing, shout, etc.) and during which activities (e.g. during choir, social activities, etc.) ?____

____

____

____

____

____

____

APPENDIX III: ACOUSTIC DATA FORM

SUBJECT #____

PRE / POST - TEST

ACOUSTIC MEASURES

Reading Fundamental Frequency (SFo) ___________Hz

Reading Intensity (SIo) ____________dB

Frequency Range

Gliding Mid-range to High: Trial #1_______
Trial #2_______
Trial #3_______

Gliding Mid-range to Low: Trial #1_______
Trial #2_______
Trial #3_______

Greatest Frequency Range _______________Hz

Max. Phonation Time: _____________________ Sec.

F0: 5 min.: _____________ Hz

10 min.: _____________ Hz

15 min.: _____________ Hz

20 min.: _____________ Hz

APPENDIX IV: SUBJECT SELF-RATING FORM (VOICE QUALITY, VOCAL EFFORT)

SUBJECT SELF RATING

Sound of voice – impairment of voice quality.

1	2	3	4	5	6	7
normal	mild	mild-mod	mod	mod-sev.	sev.	very sev.

Physical effort needed to produce voice.

1	2	3	4	5	6	7
normal	mild	mild-mod	mod	mod-sev.	sev.	very sev./ complete fatigue

APPENDIX V: FOLLOW-UP FORM, 24 HOURS POST-TASK

24-HOUR FOLLOW UP QUESTIONS:

Subject Number: __________

Initials: _________

Date: ___________

1. Do you feel anything abnormal at this point? If so, please describe.

2. If not, how long did it take for you to feel totally recovered from the experimental task?