PREDICABILITY OF THE VOICE HANDICAP INDEX RELATIVE TO ACOUSTIC MEASURES OF VOICE

By

KAREN MICHELLE WHEELER

A THESIS PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS

UNIVERSITY OF FLORIDA

2003

ACKNOWLEDGMENTS

I would like to extend great appreciation and gratitude to my committee chair, mentor, and friend, Christine Sapienza, Ph.D., for her dedication and help not only with this project but throughout my graduate career. She has spent immense amounts of time guiding me through the development of this project as well as with data analysis. If not for her extensive knowledge of voice disorders and acoustic analysis, this project could not have been successfully completed. Additionally, I would like to thank my other committee member, Rahul Shrivastav, Ph.D., for his guidance and support throughout this project. The opportunity to work with these individuals has enhanced my knowledge and overall graduate experience at the University of Florida greatly. Further acknowledgments are extended to Savita Collins, M.D., and Gayle Woodson, M.D., of the University of Florida Ear, Nose, and Throat Clinic at Ayers for allowing me to record voice samples from patients who were being treated at the clinic. I would like to extend a special thanks to Bari Hoffman-Ruddy, Ph.D., for her contributions to this study as well as to Judith Wingate, M.A., for her assistance with the data measurement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Arts

PREDICABILITY OF THE VOICE HANDICAP INDEX RELATIVE
TO ACOUSTIC MEASURES OF VOICE

By

Karen Michelle Wheeler

May 2003

Chair: Christine M. Sapienza
Major Department: Communication Sciences and Disorders

Currently there is a clinical interest concerning the quantification of the psychosocial affects of voice disorders. The Voice Handicap Index (VHI) is one instrument widely used to quantify the patient's perception of his or her voice handicap. The purpose of this study was to identify the relationship between acoustic measures of the disordered voice and patient responses on the VHI. Fifty voice patients were asked to fill out the VHI questionnaire and provide digitally recorded voice samples consisting of sustained vowels and the reading of a short passage. The results revealed significant positive correlations between scores on chosen VHI items and the acoustic measures of the voice samples. This finding demonstrates that there is a positive correlation between a patient's perception of his or her voice handicap and acoustic analyses of the disordered voice and has implications both clinically and for further research.

CHAPTER 1
INTRODUCTION

The topic of this thesis is concerns the relationship between the *Voice Handicap Index* (VHI) and acoustic measures of disordered voice production. Provided is a review of the classification of voice disorders, general acoustics and acoustic measures of speech, and a review of the literature pertaining to the development, validation, and usage of the VHI as well as other indices used to quantify patient handicap.

Classification of Voice Disorders

The definition of a voice as disordered takes on different parameters depending on the person. Verdolini and Ramig (2001) define voice disorders as "an array of self-reported symptoms and clinically observed signs … the term 'voice disorder' is explicitly or implicitly defined as a condition of sufficient concern for the bearer to report it, register functional disruption because of it, and/or seek treatment because of it" (pg. 26). This definition implies that regardless of the underlying pathology, a voice is defined as disordered, when the person identifies it as aberrant. It is only when the patient recognizes a problem and seeks treatment that a systematic classification of the actual vocal fold pathology can be utilized.

Many different schemes for classifying voice disorders exist. Some schemes classify the disorder based on acoustic properties of the voice, some classify it based on the etiology, some based on symptom presentation, and others classify it based on the pathology associated with the disorder (Boone & McFarlane, 1988; Freeman & Fawcus, 2000; Rammage, Morrison & Nichol, 2001). Rammage et al. (2001) identify six reasons to classify voice disorders: 1. classification helps to identify a specific cause of the disorder, 2. classification allows problems with similar etiology to be grouped together in order to increase an understanding of the dysfunction, 3. classification may help to develop treatment programs based on known factors associated with the disorder, 4. classification may give some idea about the course of the disease with and without treatment, 5. classification may help to facilitate communication among colleagues, and 6. classification may aid in the access of funding both for patient management and research purposes.

Verdolini and Ramig (2001) identify three conditions which affect phonation; 1. structural impairments, meaning no apparent organic injury of structure or function. Stemple, Glaze, and Klaben (2000) also list five separate classifications for voice disorders including structural changes in the vocal folds, neurogenic voice disorders, systemic disease as a contributor to vocal fold pathology, disorders of voice use, and idiopathic voice disorders. As will become more evident, different classification systems often have overlapping subgroups.

Voice disorders have traditionally been grouped into either functional/psychogenic disorders or organic/structural disorders (Rammage et al., 2001). The distinction between hypofunction and hyperfunction has also been used in the classification of voice disorders (Morrison & Rammage, 1994). However, this either/or classification system may be misleading, as voice disorders are often highly complex and multifactorial (Rammage et al., 2001). Following a comprehensive clinical evaluation by an otolaryngologist and a speech pathologist, more than one classification or a more comprehensive classification system may be needed (Rammage et al., 2001).

Colton and Casper (1996) have identified three main categories for the classification of most voice disorders. This outline will be used for the purposes of this paper, however overlapping categories used by other authors will also be identified.

Functional Disorders

Functional disorders are related to how the voice is being used. Functional disorders stem from the utilization of improper vocalization technique, typically done unconsciously on the part of the patient (Benninger, Jacobson, & Johnson, 1994; Boone & McFarlane, 1988; Rammage et al., 2001; Stemple et al., 2000). Boone and McFarlane (1988) describe several factors which may lead to a “faulty manner” of voice production. Some of these include discoordination of respiratory function, use of inappropriate pitch levels, vocal fold hyperfunction, and psychogenic factors such as increased stress and anxiety.

Organic lesions which are considered functional in nature are the result of laryngeal hyperfunction (Colton & Casper, 1996). These include vocal fold nodules, vocal fold polyps, intracordal cysts, laryngitis, sulcus vocalis, and contact ulcers (Boone & McFarlane, 1988; Colton & Casper, 1996; Morrison and Rammage, 1994; Sataloff, 1991).

Hyperfunction can also alter the normal function of the phonatory mechanism to work effectively and efficiently. Terms such as abuse and misuse are used in the literature as well as phonotrauma and repetitive strain injury (Stemple et al., 2000). Stemple et al. (2000) classify these disorders as *disorders of voice use*, where inappropriate voice maladaptations result in voice pathology. According to Stemple et al. (2000), these include psychogenic disorders as well as those related to the functional misuse of the laryngeal muscles in voice production. Vocal behaviors that are categorized as misuse of the phonatory mechanism include: increased tension or strain, inappropriate pitch level, excessive talking, ventricular phonation, and aphonia and dysphonia of psychological origin (Colton & Casper, 1996; Koschkee & Rammage, 1997; Stemple, Glaze, & Gerdeman, 1996). Additionally, Stemple et al. (2000) include muscle tension dysphonia, vocal fatigue, puberphonia, and transgender voice in this category.

Increased tension or strain (Hyperfunction)

Increased tension or strain involves both the intrinsic and extrinsic laryngeal muscles and includes behaviors such as hard glottal attack, high laryngeal position, and anteroposterior squeezing (Benninger et al., 1994; Colton & Casper, 1996; Koschkee & Rammage, 1997; Stemple et al., 1996). Hard glottal attack refers to the rapid and complete adduction of the vocal folds immediately preceding phonation. Hard glottal attack can be produced through medial compression of the vocal folds or prephonatory laryngeal constriction involving the ventricular folds, arytenoids, and epiglottis (Benninger et al., 1994; Colton & Casper, 1996; Stemple et al., 1996).

High laryngeal position refers to the physical raising of laryngeal height, which results in shortening of the vocal tract, stiffness of the vocal fold tissue, and an increased tendency for tight vocal fold closure. Anteroposterior laryngeal squeezing is a condition in which the arytenoids and the epiglottis approach each other during phonation (Colton & Casper, 1996; Stemple et al., 1996).

Muscle tension dysphonia (MTD) is described by Morrison and Rammage (1994) as variable symptoms of voice disruption accompanied by observable tension or stiffness of the neck, jaw, shoulders, and throat. Concomitant issues of psychosocial stress or interpersonal conflicts have been commonly associated with this disorder (Morrison & Rammage, 1994; Stemple et al., 2000).

Inappropriate pitch levels

Inappropriate pitch level includes that which occurs with puberphonia, persistent glottal fry, and/or lack of pitch variability. A person's optimum pitch is difficult to define, however it is known that phonatory range is often diminished in the presence of certain vocal fold pathologies (Benninger et al., 1994; Colton & Casper, 1996, Stemple et al., 1996; Stemple et al., 2000).

Puberphonia refers to the persistence of a high-pitched voice beyond the age at which the voice is expected to change (Colton & Casper, 1996; Rosen & Sataloff, 1997; Stemple et al., 2000). Many causes have been suggested for puberphonia, or mutational falsetto, including attempts by the patient to resist natural growth into adulthood, strong feminine identification, the desire to maintain the childhood soprano singing voice, and embarrassment when the voice lowers dramatically (Rosen & Sataloff, 1997; Stemple et al., 2000).

Glottal fry, or pulse register, is one of three normal voice registers. The other registers include loft (falsetto) and modal (normal speaking range). Glottal fry is the least flexible of the voice registers and is the lowest in fundamental frequency. During glottal fry, the vocal folds close quickly and the closed phase of the glottal cycle is longer compared to the length of the entire period (Aronson, 1990; Benninger et al., 1994;

Colton & Casper, 1996; Sataloff, 1991; Stemple et al., 1996; Stemple et al., 2000). This contributes to laryngeal hyperfunction (Aronson, 1990; Colton & Casper, 1996; Stemple et al., 1996; Stemple et al., 2000).

Lack of pitch variability refers to an individual's inability to vary fundamental frequency. This monotonic voice quality may be a result of neurologic disorder, psychologic depression, or habitual misuse (Colton & Casper, 1996; Stemple et al., 1996). When pitch is not changed, the vocal mechanism rarely varies with regard to the adductory and contact forces. Consequently, these forces consistently occur with the same strength and in the same area on the vocal folds (Colton & Casper, 1996; Stemple et al., 1996).

Excessive talking

Each person has a different physiological limit with regard to the laryngeal structures. Excessive talking may lead to vocal fatigue, which is a common descriptor used to refer to a well-known set of symptoms including decreased endurance, loss of frequency and intensity control, and complaints of effortful, unstable, or ineffective voice production (Benninger et al., 1994; Stemple et al., 1996; Stemple et al., 2000). Clinical complaints include dryness in the throat and neck, pain at the base of the tongue, throat, and neck, feelings of "fullness" or a "lump" in the throat, shortness of breath, and effortful phonation (Benninger et al., 1994; Colton & Casper, 1996; Stemple et al., 1996). Some individuals may talk for hours with minimal hydration, poor nutrition, and physical exhaustion without experiencing phonatory problems. Another person may eat a healthy diet, be well hydrated, and well rested and experience phonatory problems after a moderate amount of talking. Individual differences play a large role in the effect of excessive talking on phonation (Colton & Casper, 1996).

Ventricular phonation

Ventricular phonation, also referred to as plica ventricularis or supraglottic hyperfunction, involves the abnormal constriction of the supraglottis (Benninger et al., 1994; Stemple et al., 2000). Specifically, ventricular phonation occurs when there is greater than expected movement of the ventricular folds towards the midline, and subsequent phonation of those folds (Colton & Casper, 1996; Stemple et al., 2000; Rammage et al., 2001). While typically considered pathologic, phonation of the ventricular folds can occur as compensation for reduced or absent movement of the true vocal folds (Benninger et al., 1994; Colton & Casper, 1996; Stemple et al., 2000).

It is often hard to discriminate between those behaviors that constitute misuse and abuse. Abusive behaviors tend to be harsher with a greater probability of causing trauma to laryngeal tissue (Benninger et al., 1994; Colton & Casper, 1996; Rosen & Sataloff, 1997; Stemple et al., 1996). Vocal behaviors which constitute vocal fold abuse include: excessive and prolonged loudness, strained and excessive use of the voice during periods of swelling, inflammation, or other tissue changes, and excessive coughing and throat clearing. (Benninger et al., 1994; Colton & Casper, 1996; Stemple et al., 1996).

Excessive, prolonged loudness

Excessive and/or prolonged loudness occurs most often in individuals who must speak above environmental noise. These individuals include those with habitual patterns of very loud voice use, for example, teachers or factory workers (Colton & Casper, 1996). Increased loudness can lead to increased subglottal pressure, increased vibratory amplitudes, and increased medial compression of the vocal folds. Irritated, swollen, and inflamed laryngeal mucosa, specifically along the glottal edge of the vocal folds, may result. (Benninger et al., 1994; Colton & Casper, 1996; Stemple et al., 1996).

Swelling, inflammation, or other tissue changes

Vocal fold tissue may become initially inflamed or irritated for numerous reasons; for example, gastroesophageal reflux or laryngitis (Benninger et al., 1994; Sataloff, 1991; Stemple et al., 1996). However, when a person who depends on their voice for occupational demands engages in abusive vocal behaviors in the presence of concurrent vocal fold irritation, the result is an increase in abuse. Further tissue changes may occur and persist beyond the point the original irritation has cleared (Colton & Casper, 1996).

Excessive coughing and throat clearing

All people must cough or clear their throats at some times. However, when these behaviors become compulsive or reactive then they can be abusive to the vocal folds. Because the entire larynx and supraglottal structures are involved in the cough mechanism, multiple structures are at risk for damage as a result of this abusive behavior (Benninger et al., 1994; Colton & Casper, 1996; Stemple et al., 1996; Stemple et al., 2000).

Organic Disorders

Unlike functional disorders, organic disorders are unrelated to the manner in which the voice is used. Stemple et al. (2000) group these disorders as *structural changes in the vocal folds.* Preliminary diagnosis of these disorders is typically based upon visual examination of the vocal folds. Changes in the mucosal layers or the thyroarytenoid muscle will affect the mass, size, stiffness, flexibility, tension, glottal closure pattern, and phase duration of the vibrating mechanism (Benninger et al., 1994; Sataloff, 1991; Stemple et al., 2000). Stemple et al. (2000) group vocal fold nodules polyps, contact ulcers, laryngitis and sulcus vocalis in this category.

Treatment for these disorders is typically pharmacological or surgical, although voice therapy may be helpful with regard to patient education concerning the disorder or in order to ensure that the patient avoids developing abusive vocal habits as a result of difficulty producing voice (Aronson, 1990; Benninger et al., 1994; Colton & Casper, 1996). Conditions which are classified as organic disorders include: keratosis, granulomas, ankylosis of the cricoarytenoid joint, papillomas, carcinoma and other malignancies, blunt or penetrating trauma, chemical or heat trauma, congenital and acquired webs, presbylaryngus, congenital cysts, Reinke’s edema or polypoid degeneration, and vascular lesions (including vocal hemorrhage and varix) (Aronson, 1990; Colton & Casper, 1996; Rammage et al., 2001; Stemple et al. 2000).

Neurologic Disorders

“Neurogenic voice pathologies are those voice disorders directly caused by and interruption of the nervous innervation supplied to the larynx, including both central and peripheral insults” (Stemple et al., 2000, pg. 114). Aronson (1990) groups neurologic voice disorders in the “organic” group of disorders. The nervous system is divided subdivided into the central nervous system (CNS) and the peripheral nervous system (PNS). The CNS is that part of the nervous system which resides in the cranial cavity and is responsible for the initiation and coordination of function. The PNS resides outside of the skull, throughout the body, and carries the instructions of the CNS to the organs or muscles (Aronson, 1990; Colton & Casper, 1996; Duffy, 1995; Webster, 1999).

The act of producing phonation and speech is highly complex, involving many different subsystems (Aronson, 1990; Benninger et al., 1994; Colton & Casper, 1996). In order for speech to occur normally, the nervous system must oversee the coordination between the respiratory, laryngeal, velopharyngeal, and articulatory subsystems. This involves the chestwall, the larynx and pharynx, the velum, lips, tongue, teeth, and mandible. The CNS must provide initiation and coordination of these functions. The PNS sends innervation to those muscles and organs, which are needed for the proper functioning of these subsystems. Laryngeal function in the production of voice depends on this coordination. Injuries to the laryngeal nerves, neurologic disorders or diseases, or abnormal growths may disrupt nervous system function and, consequently, phonatory or speech function (Aronson, 1990; Benninger et al., 1994; Colton & Casper, 1996; Duffy, 1995).

Some neurogenic voice disorders are confined to voice and laryngeal manifestations, while others reflect a larger deterioration of many motor control systems. These can include impairments of respiration, resonance, swallowing, and other functions beyond the head and neck (Benninger et al., 1994; Duffy, 1995; Stemple et al., 1996; Stemple et al., 2000). Hallmark symptoms of some of these systemic neurologic disorders are based on clusters of perceptual attributes and deficits from the speech pattern. Some of these systemic neurologic disorders include Parkinson’s disease, Huntington’s chorea, multiple sclerosis, and amyotrophic lateral sclerosis (Benninger et al., 1994; Duffy, 1995; Koschkee & Rammage, 1997; Stemple et al., 2000).

Vocal fold paralysis is identified as the most common neurogenic voice disorder (Benninger et al., 1994; Morrison & Rammage, 1994; Stemple et al., 2000). Vocal fold paralysis is associated with lesions to the vagus nerve (CNX) and may be bilateral or unilateral. Lesions of CNX at any point along the pathway from nucleus ambiguous to the brainstem to the musculature may result in paresis or paralysis of laryngeal muscles (Aronson, 1990; Sataloff, 1991). However, vocal fold paralysis is most typically caused by peripheral involvement of the recurrent laryngeal nerve (RLN) and, to a lesser extent, the superior laryngeal nerve (SLN) (Aronson, 1990; Morrison & Rammage, 1994; Sataloff, 1991; Stemple et al., 2000;).

A more proximal lesion to the CNX may affect muscles innervated by both the SLN and the RLN. This would result in problems both abducting and adducting the vocal folds (Brown, Vinson, & Crary, 1996; Stemple et al., 2000). The extent of muscle weakness and the degree of voice impairment depend largely upon the location of the lesion along the pathway of the nerve and whether the lesion is bilateral or unilateral (Aronson, 1990).

Spasmodic dysphonia (SD) is another neurogenic voice disorder, although its etiology remains uncertain (Stemple et al., 2000). Morrison and Rammage (1994) described SD as related to stress or psychological factors, although more recent evidence has suggested neurologic origin (Stemple et al., 2000). Spasmodic dysphonia refers to a family of symptoms which include strained, strangled, or effortful voice production. It is characterized by a normal appearing larynx when at rest but abnormal involuntary movements that are action-induced and task specific (Morrison & Rammage, 1994; Stemple et al., 2000).

Spasmodic dysphonia can be identified as either the abductor or adductor type. Adductor SD (ADSD) is more common and results in a strained-strangled voice quality, multiple pitch breaks, and occasional voicing blocks of tension or effort which may block continuous phonation (Sataloff, 1991; Stemple et al., 1996; Stemple et al., 2000). Laryngeal behavior while phonating is characterized by intermittent, tight adduction of the vocal folds. Muscles which may be involved include the thyroarytenoid muscle, lateral cricoarytenoid and the interarytenoid muscles. This behavior creates the strained, forced voice quality (Sataloff, 1991; Stemple et al., 2000)

Abductor SD (ABSD) is characterized by involuntary spasms of vocal fold abduction resulting in a period of aphonia followed by a burst of air. This primarily involves the posterior cricoarytenoid muscle. The voice is characterized by prolonged bursts of breathy phonation and intermittent aphonia (Sataloff, 1991; Stemple et al., 2000).

Essential tremor (or benign tremor) is another neurogenic voice disorder characterized by rhythmic tremors various body parts, including the larynx. Other body parts involved may include the arms, head, neck, tongue, palate, and face (Aronson, 1990; Stemple et al., 1996; Stemple et al., 2000). Essential tremor is often familial and typically begins in middle age (Aronson, 1990; Morrison & Rammage, 1994; Stemple et al., 2000). Vocal tremor is most noticeable during prolonged vowels, however connected speech may be negatively affected as well. Laryngeal tremor is visible during prolonged phonation during the visual examination. The rate of essential tremor is typically between 4 and 7Hz. Benign essential tremor is exclusive of tremor in other neurologic disease processes (Aronson, 1990; Stemple et al., 1996; Stemple et al., 2000).

Acoustics

Perceptual terms for describing voiced sounds have been examined by several different researchers. Titze (1994) describes four dimensions which may be used to describe sounds: pitch, loudness, vowel (or voice consonant), and quality. *Quality* is perhaps the most poorly defined of these categories. Terms such as *harsh, rough, breathy,* and *pressed* have been used to describe voice quality (Baken & Orlikoff, 2000; Titze, 1994).

In order to bring a greater degree of objectivity to the study of voice and voice disorders, acoustic correlates of perceptual measures have been explored (Wolfe & Martin, 1997). These acoustic measures have been developed in order to aid in the quantification of voice characteristics. These measures, when proven reliable and reproducible, provide a means of following changes in the voice over time or between subjects (Heman-Ackah, Michael, & Goding, 2002).

Overview

The term *acoustics* refers to the study of sound. Because speech is continuous sound, understanding the nature of sound is essential to understanding the production of speech (Borden, Harris, & Raphael, 1994).

The simplest sound pattern is the pure tone. By definition, a pure tone has only one frequency of vibration (Borden et al., 1994). *Frequency* refers to the number of cycles per second. A pure tone results from a vibration that repeats itself at a constant number of vibrations per second (Borden et al., 1994). One instrument which produces a pure tone is a tuning fork. When struck, a tuning fork vibrates in *simple harmonic motion*; the prongs move back and forth at a fixed rate regardless of how hard the tuning fork is struck (Borden et al., 1994; Titze, 1994). Simple harmonic motion is the projection of circular motions at a constant speed onto one axis, resulting in a sine wave. It is referred to as simple harmonic motion because it is the simplest smoothly connected back and forth movement possible (Titze, 1994).

The properties of *elasticity* and *intertia* are responsible for keeping the tuning fork prongs in motion. After being struck, the prongs are brought back to their original position via elastic force (Borden et al., 1994; Colton & Casper, 1996). The prongs are kept in motion because of inertia, which is the tendency of an object in motion to remain in motion (or an object at rest to remain at rest). These forces of elasticity and inertia are nearly always simultaneously at work, although one may dominate at a particular moment in time. The simultaneous interplay of the two forces is therefore responsible for the continuous cyclic motion of the prongs (Borden et al., 1994).

Most sound sources (including speech) produce complex vibrations which produce more than one frequency, resulting in a complex tone. This is a result of vibrating in a complex manner instead of simple harmonic motion (Borden et al., 1994). Complex tones can be classified as either *periodic* or *aperiodic*. Periodic tones are those which the pattern of vibration repeats itself regardless of its level of complexity (Borden et al., 1994, Kent & Read, 2002). Pure tones are periodic in nature. Aperiodic sounds are the result of random vibration with no repeatable pattern (Borden et al., 1994).

Fundamental Frequency

The component frequencies of complex periodic signals are integral measures of the lowest frequency component, or the fundamental frequency (Borden et al., 1994). The fundamental frequency (F_0) of a sound source is the lowest frequency of a complex, periodic wave (Borden et al., 1994; Colton & Casper, 1996). Fundamental frequency is derived from the rate at which the sound source is vibrating. When applied to speech, F_0 is an acoustic measure that directly reflects the rate of vocal fold vibration (Colton & Casper, 1996). Pitch is described as the perceptual correlate of the F_0 (Baken & Orlikoff, 2000; Colton & Casper, 1996; Borden et al., 1994; Kent & Read, 2002).

The measurement of the F_0 depends largely on the assumption that the signal is approximately periodic (Baken & Orlikoff, 2000). Normal vocal signals are described as "nearly periodic" because they have no tones of strictly constant pitch and are constantly changing in frequency and quality (Simon in Baken & Orlikoff, 2000; Kent & Read, 2002). Titze (1995) described three types of vocal signals:

<u>Type 1</u> – Nearly periodic (or, ideally, periodic). These waveforms do not undergo qualitative changes during the time intervals being analyzed.

<u>Type 2</u> – Signals that have sudden qualitative changes, or bifurcations, in the interval to be analyzed. These signals have no single F_0 that characterizes the entire segment.

<u>Type 3</u> – Signals which have no apparent periodicity.

The evaluation of speaking fundamental frequency (SF_0) during connected speech may give information regarding whether one speaker's vocal frequency is very different from comparable speakers (Baken & Orlikoff, 2000). Vocal pitch (the perceptual correlate of SF_0) is subject to expectations based on age, sex, body type, social situation, emotional state, and other factors (Wolfe, Ratusnik, Smith & Northrop, 1990).

Speaking fundamental frequency may be attained by collecting speech samples in different ways, including spontaneous speech and reading a passage (Baken & Orlikoff, 2000). While each method has advantages, passage reading allows for the same materials to be used repeatedly or between subjects. This allows for comparison between different subjects or between sessions with the same subject (Baken & Orlikoff, 2000). Fundamental frequency has been shown to vary with age and sex, however F_0 alone does not yield a sufficiently detailed picture of vocal fold vibratory patterns to effectively differentiate between normal and disordered voices (Ferrand, 2002). It is important to note that most disorders of the larynx do not have, by themselves, a consistent effect on F_0 (Baken & Orlikoff, 2000).

Vocal F_0 is reflective of the biomechanical characteristics of the vocal folds as they interact with glottal airflow. The laryngeal structure and muscle forces of the larynx determine the biomechanical properties of the vocal folds (Baken & Orlikoff, 2000; Titze, 1994). A combination of reflexive, affective, and learned voluntary behaviors result in an adjustment of the muscle forces. The ability of a speaker to adjust his/her F_0 gives information regarding the mechanical adequacy of the laryngeal structures, and about the precision of laryngeal control (Baken & Orlikoff, 2000; Titze, 1994).

Jitter

Overall stability of phonatory adjustment is reflected in the amount of short-term variability of the speech signal. Frequency perturbation, or jitter, provides an index concerning the stability of the laryngeal system (Baken & Orlikoff, 2000; Borden et al., 1994; Colton & Casper, 1996). It measures small, cycle-to-cycle changes of period that occur during phonation (Baken & Orlikoff, 2000; Borden et al., 1994; Colton & Casper, 1996; Kent & Read, 2002; Titze, 1994). Jitter is a measure of variability which is not accounted for by voluntary changes in frequency (Baken & Orlikoff, 2000). Measures of perturbation (both frequency and amplitude) need to be taken from sustained vowels. Connected speech confounds the measure due to linguistically produced variations in amplitude and frequency which cannot be separated from the biomechanical characteristics of the vocal folds (Colton & Casper, 1996). This measure may reflect small differences in mass, tension, biomechanical properties, or neural control of the vocal folds (Baken & Orlikoff, 2000; Baer in Colton & Casper, 1996).

When evaluating jitter %, the more a measure deviates from zero, the more it correlates with erratic vibratory patterns of the vocal folds (Baken & Orlikoff, 2000). The vibratory cycles of all speakers are erratic to some extent; however an abnormal voice would be expected to be more erratic than a normal voice. While jitter is considered to be sufficiently sensitive to pathologic changes in the phonatory system, it is in no way a guide to the type or classification of dysphonia the patient presents with (Baken & Orlikoff, 2000; Titze, 1994).

Some potential sources of jitter include the following:

1. Neurogenic problems (ie – vocal fold paralysis, “spasms” of spasmodic dysphonia)
2. Rapid aerodynamic changes (changes in glottal airflow)
3. Biomechanical alterations in the properties of the vocal folds. This can include changes in vocal fold mass associated with vocal fold pathology (Baken, 2000).
4. Stylistic changes (or artistic fluctuations), typically thought of when dealing with the performing arts; for example, vibrato.
5. Chaotic oscillation, which assumes the vocal apparatus is in part, a chaotic system.

Voice onset and termination have greater frequency perturbation than the midportion of sustained vowels (Baken & Orlikoff, 2000; Titze, 1994). Therefore, when measuring this parameter, a midsection of the vowel production should be analyzed.

Several studies have examined, among other acoustic measures, jitter and its potential perceptual correlates. Normal voices should have little cycle-to-cycle variability in frequency, i.e., jitter, whereas "hoarse" or "breathy" voices would be expected to have higher degrees of jitter. Heman-Ackah et al. (2002) conducted a study which aimed to evaluate the ability of specified acoustic measures, including jitter, to predict overall dysphonia within the perceptual categories of breathiness and roughness in pathological voices. Jitter was found to be a poor predictor of dysphonia at levels which would be clinically applicable (Heman-Ackah et al., 2002).

Ferrand (2002) aimed to evaluate the ability of different acoustic measures (including harmonics-to-noise ratio, jitter, and fundamental frequency) to provide information regarding the integrity of the vocal mechanism in women with normal voices. Harmonics to noise ratio is the relative amount of additive noise in a voice signal. The women were divided into three groups: young adults, middle-aged adults, and elderly adults. Significant differences were found in the harmonics-to-noise ratio (HNR) between the group of elderly women and the two younger groups. Differences in F_0 were also found between the elderly and younger groups (Ferrand, 2002). However there were no significant differences in jitter between the three groups. Consequently, HNR was judged to be a more sensitive indicator of vocal function than jitter (Ferrand, 2002).

An additional study by Wolfe and Martin (1997) explored the acoustic discrimination and graded severity of three clinical voice types. Evidence suggests that the perceptual categorization of voice qualities is associated with the interaction of acoustic parameters, specifically F_0, in conjunction with spectral slope (Wolfe & Martin, 1997). Two trained listeners classified 102 samples of dysphonic voices as one of three voice types: breathy, hoarse, or strained. The speech sample consisted of the vowels /a/ and /i/. The vowels were analyzed acoustically with cepstral peak prominence (CPP), jitter standard deviation, F_0, and signal-to-noise ratio (SNR) standard deviation. Findings revealed that voice type is indeed associated with the interaction of spectral noise, F_0, and signal irregularity (jitter and SNR). Results also suggested that dysphonic severity is associated with similar parameters (Wolfe & Martin, 1997).

Intensity

The measurement of vocal intensity correlates with the perception of loudness (Baken & Orlikoff, 2000; Borden et al., 1994; Colton & Casper, 1996; Kent & Read, 2002; Titze, 1994). Intensity is defined as power per unit area in watts (Baken & Orlikoff, 2000; Borden et al., 1994). Vocal intensity is dependent on the interaction of subglottal pressure, biomechanics and aerodynamics at the level of the vocal fold as well as the status of the vocal tract (Baken & Orlikoff, 2000; Borden et al., 1994; Titze, 1994). The range of intensities at which the voice can be produced is an indication of the limits of adjustment of the phonatory system (Baken & Orlikoff, 2000). This makes intensity a potentially important measure in the assessment of voice disorders (Baken & Orlikoff, 2000).

Intensity can be measured using sustained vowels or connected speech samples (Colton & Casper, 1996). Measuring the intensity of connected speech requires different considerations than sustained vowels (Baken & Orlikoff, 2000; Colten & Casper, 1996). Connected speech shows very large fluctuations over short time intervals. This is due in part to silences contained in connected speech, varying of intensity for syllable and word stress, varying intensity among phonemes characterized by different acoustic powers (Baken & Orlikoff, 2000).

Many different voice disorders can result in a decrease in loudness. Speech intensity is notably reduced in disorders of the central nervous system, or laryngeal or ventilatory pathology (Baken & Orlikoff, 2000). For patients with these types of disorder, increasing loudness is a common treatment goal.

In a study conducted by Angerstein & Neuschaefer-Rube (1998) it was found that hyperfunctional voice disorders had little effect on the intensity of sustained vowels at comfortable loudness levels. However, findings did suggest that a decrease of intensity during "loud" sustained vowel did relate to the severity of the voice disorder (Angerstein & Neuschaefer-Rube, 1998).

Shimmer

Shimmer is defined as small, cycle-to-cycle changes of amplitude which occur during phonation (Baken & Orlikoff, 2000; Borden et al., 1994; Colton & Casper, 1996; Kent & Read, 2002; Titze, 1995). Shimmer values quantify short-term amplitude instability (Titze, 1995). Shimmer is thought to contribute to the perception of hoarseness, however the relationship of shimmer to specific abnormalities of vocal fold function remains unclear (Baken & Orlikoff, 2000; Titze, 1995). Ideally, the signal being analyzed would be Type 1 (near-periodic) in order to acquire a reliable measure of shimmer (Baken & Orlikoff, 2000; Titze, 1995).

Morsomme, Jamart, Wery, Giovanni, and Remacle (2001) attempted to establish relevant objective parameters for evaluating dysphonia following unilateral vocal fold paralysis. The study compared objective and perceptual measures of voice from 40 subjects. The subjects were divided into either the dysphonic group (28 subjects) or the control group (12 subjects). The perceptual measures were taken from the GIRBAS-scale (grade, instability, roughness, breathiness, asthenia, and strain) and compared to the acoustic measures. Findings suggested that measures pertaining to the aperiodicity of the phonatory signal (including F_0 coefficient and jitter) correlated well with the GIRBAS scale's criteria of grade, breathiness, and asthenia. However, mean F_0 and shimmer did not significantly correlate with the subjective data (Morsomme et al., 2001).

Signal-to-Noise Ratio

Signal-to-noise ratio as defined by Milenkovic (1987) is the ratio of the total energy of a voice signal to the energy of the aperiodic component of the voice signal. Because most speaking situations (except those in a controlled environment, such as a sound booth) occur in the presence of noise, listeners have to pick which sounds are important to attend to (Kent, 1997). Background noise refers to noise in the environment, or ambient sound. Speech must compete with background sound created in the environment in order to be perceived by the listener (Kent, 1997). Voice signals which have a significant aperiodic component may be more difficult to hear when competing with background noise. Large positive values of the signal-to-noise ratio are indicative of a strong voice signal relative to aperiodic noise. Smaller positive values, or even negative values, are less preferable because they indicate that aperiodic components and background noise rival or exceed the speech signal in intensity (Kent, 1997, Milenkovic, 1987).

Amplitude Spectrum and Sound Spectrogram

The acoustic analysis of speech is highly dynamic in nature. The acoustic signals of speech change rapidly and nearly continuously; any change in the acoustic characteristics of the speech signal represent movement in the structures of speech production (Baken & Orlikoff, 2000; Ferrand, 2001; Kent, 1997).

Sound spectrography provides a dissection of an acoustic signal into its most basic components (Baken & Orlikoff, 2000; Kent, 1997; Titze, 1994). The basis of sound spectrography is the Fourier theorem (Baken & Orlikoff, 2000; Kent & Read, 2002;). The Fourier theorem states that "any periodic wave can be expressed as the sum of an infinite series of sine waves of different amplitudes, whose frequencies are in integer ratio to each other, and which have different phase angles with respect to each other" (Baken & Orlikoff, 2000 pg. 227). An *amplitude spectrum* is the result of a Fourier analysis (Baken & Orlikoff, 2000). An amplitude spectrum is the plot of the relative amplitudes versus frequencies of all components of the signal (Titze, 1994).

The slope of the amplitude spectrum (or spectral slope) is a measure of how the amplitudes of signal components decrease with increasing harmonic number. Spectral slope is typically given in dB/octave (an octave is a doubling or halving of frequency; Titze, 1994). Spectral slopes relate to the quality of a sound. Sounds with many high frequencies would have a more shallow, or small, spectral slope. A smaller spectral slope implies that the second and third harmonics have relatively large amplitudes relative to the fundamental frequency (Titze, 1994). Waveforms with a smaller spectral slope relate to a more pressed voice quality. Waveforms with many lower frequency sounds might have a more steep, or larger, slope. A larger spectral slope relates to a more breathy voice quality. This spectrum is filled with non-harmonic components, or noise, between the harmonic lines (Titze, 1994).

A sound spectrogram achieves a short-term running spectrum, or a dynamic analysis, which can reveal spectral features in a nearly continuous fashion (Kent, 1997; Kent & Read, 2002). A spectrogram includes three dimensions: time, frequency, and intensity (Baken & Orlikoff, 2000; Kent, 1997; Titze, 1994). Time is represented along the horizontal axis (read from left to right), frequency is along the vertical axis (increases as it goes up) and intensity is represented according to the darkness or lightness of the signal (dark indicates more intensity) (Ferrand, 2001; Kent, 1997).

There are two types of spectrograms, wide band and narrow band (Kent, 1997; Kent & Read, 2002; Titze, 1994). Wide band spectrograms have a relatively wide analyzing filter (typically 300-500 Hz) whereas narrow band spectrograms have a narrow analyzing filter (45-50 Hz) (Kent, 1997; Kent & Read, 2002; Titze, 1994). Wide band spectrograms can display formant energy because of their rather widespread acoustic energy (Kent, 1997). Wide band spectrograms also offer good time resolution; individual periods of vibration can be seen as striations in the waveform. This makes it possible to determine F_0 by simply counting striations and dividing by a unit of time (Titze, 1994).

Narrow band spectrograms, because of their low pass filtering, typically pass only one harmonic at a time. Consequently, narrow band spectrograms provide a finer resolution in frequency and clearly display harmonics (Kent, 1997; Kent & Read, 2002). Dark horizontal lines indicate the intensity of individual harmonics in the signal (Titze, 1994). Narrow band spectrograms do not have vertical striations which give vibratory and formant information, however, because the time resolution is not sufficient to respond to individual periods of vibration (Titze, 1994).

Vowels

When studying vowels, it is often desirable to analyze formant pattern, harmonics, duration, and F_0, therefore a wide band spectrogram is used (Ferrand, 2001; Kent & Read, 2002). Areas of increased vowel resonance, or formants, are the immediately obvious feature in the spectrogram (Baken & Orlikoff, 2000). Formants are a result of increased resonance around harmonic frequencies by the vocal tract (Ferrand, 2001, Kent & Read, 2002). On wide band spectrograms, formants appear as wide, dark horizontal strips which reflect the concentration of acoustic energy at those frequencies. Because the vocal tract is widely tuned, many harmonics are amplified near vocal tract formants, which accounts for the width of the bands (Ferrand, 2001). Formant frequencies can be modified by the size and shape of the vocal tract (Titze, 1994). Modification of the vocal tract can be achieved by raising or lowering the larynx or by protruding or retracting the lips (Titze, 1994). Placement of the vowel (front, central, back) as well as dialectical differences can also influence the formant frequencies of vowels (Kent, 1997; Kent & Read, 2002). Spectrographically, vowels are most saliently characterized by their first three formants (Ferrand, 2001; Kent & Read, 2002). However, formants for specific vowels do not remain constant given anatomical variation across speakers. The important identifying factor is not the formant itself, but the relationship of relative formant frequencies and amplitudes (Baken & Orlikoff, 2000).

The differences between formant amplitudes can give valuable information regarding glottal configuration and closure patterns (Hanson, 1996; Hanson & Chaung, 1999). Hanson (1996) as well as Hanson and Chaung (1999) conducted research aimed at formulating acoustic parameters of the voicing source in order to differentiate between individuals and female and male speakers. The configuration of the glottis varies among speaker, and, notably, between females and males (Hanson, 1996, Hanson & Chaung, 1999). Persons who do not achieve complete closure even during the closed phase of the glottal vibratory cycle experience constant airflow through the glottis. The effects of this DC flow on the glottal waveform increase with its size, providing a source of variability in voicing characteristics (Hanson, 1996). For instances in which the glottis closes completely, if a speaker modifies production such that it results in a larger open (abducted) quotient, the spectrum of the source undergoes a change at low frequencies only. However, if there is significant variation, the amplitude of the first harmonic relative to that of the second (H1-H2) changes by approximately 10dB (Hanson, 1996). The spectrum is ultimately influenced by the abruptness of closure (the rate at which the airflow is cut off when the membranous part of the vocal folds close during the cycle; Hanson, 1996). Sodersten (in Hanson, 1996) observed glottal closure patterns via fiberscopy and then related the degree of closure to perceptions of breathiness. It was found that significant correlations exist between perceived breathiness and the relative amplitude of the first harmonic (H1) to the first formant peak (A1) Sodersten in Hanson, 1996).

The relationship of the amplitude of the first harmonic (H1) relative to that of the second harmonic (H2) is used as an indication of the open quotient, or the ratio of the open phase of the glottal cycle to the total period (Hanson & Chaung, 1999). The amplitude of the first harmonic (H1) relative to the amplitude of the first formant (A1) reflects the source spectral tilt (a more tilted shape of the amplitude spectrum waveform) (Hanson, 1996). A longer open phase and a more tilted shape, or larger slope, of the amplitude spectrum is associated with a more breathy voice quality (Titze, 1994). Several different vocal fold pathologies, including vocal fold nodules, polyps, granulomas, and cysts, can result in incomplete vocal fold closure. Consequently, the open phase of vibration persists and a larger spectral slope characterized by a larger positive difference between H1 and H2 as well as H1 and A1 exists (Hanson, 1996; Hanson & Chaung, 1999).

The Impact of Voice Disorders

Approximately ¼ of the working population in the United States depends on their voice as a critical tool for their occupation (National Center for Voice and Speech, 1993). This equates to approximately 28,000,000 people who need their voice in order to do their job (Verdolini & Ramig, 2001).

The impact of a disordered voice varies greatly from person to person. Occupation, environment, family members, and overall personality are all variables that can affect the way a voice disorder affects a specific person. In general, people with dysphonia tend to encounter problems that include psychological, emotional, social (family and friends), and employment related difficulties (Scott, Deary, Wilson, & MacKenzie in Wilson, Deary, Millar, & MacKenzie, 2002). Dysphonia has also been found to have an impact the overall health status of a patient (Wilson et al., 2002).

In a recent study by Wilson et al. (2002), dysphonia was found to have a marked impact on patients' reports regarding health status. Wilson's purpose was two-fold: 1.) to compare self-rated general heath status in a large cohort of dysphonic patients to those from control groups, and 2.) to examine the differential impact of dysphonia on various health status domains (Wilson et al., 2002). The study included 163 patients; 38 men and

125 women. The subjects were required to complete the Short Form 36 (SF36), a 36-item questionnaire assessing quality of life, and the Voice Handicap Index (VHI). The SF36 consists of eight subscales including physical functioning, social functioning, role limitations due to physical problems, role limitations due to emotional problems, mental health, energy/fatigue, bodily pain, and general health perceptions. Results revealed that patients with dysphonia had significantly lower scores than age-matched controls on all eight subscales of the SF36. These results emphasize the importance of including a quality of life measure in an otolaryngologic assessment (Wilson et al., 2002).

The study of self-perceived handicap in relation to voice disorders has gained much attention recently (Jacobson, Johnson, Grywalski, Silbergleit, Jacobson, & Benninger, 1997; Murry & Rosen 2000; Rosen & Murry, 2000; Rosen, Murry, Zinn, Zullo, & Sonbolian, 2000). Scales with indices aimed at quantifying an individual's quality of life and handicap related to voice difficulty have been developed and are becoming widely used in voice and otolaryngologic clinics. These tools provide insight as to why two people with similar vocal pathologies experience varying levels of handicap and disability (Jacobson et al., 1997).

The fields of audiology and medicine frequently utilize surveys as a part of disability, handicap, and outcomes assessment (Benninger et al., 1998; Jacobson et al., 1997). For example, the *Dizziness Handicap Inventory* is employed by audiologists to assess the effect dizziness has on patient daily living (Jacobson & Newman, 1990). Additionally, the *Hearing Handicap Inventory for Adults* is used to measure the effects of hearing loss on quality of life (Newman, Jacobson, Weinstein, & Hug, 1990). The *Medical Outcomes Trust 36-Item Short Form General Health Survey* (SF-36) is a quality of life survey used in the field of medicine (Benninger et al., 1998). All of these surveys offer health care providers information regarding the patient's self-perception of his or her handicap both before and after treatment.

Otolaryngologists (ENT's) and speech-language pathologists (SLP's) use the GRBAS (**g**rade, **r**oughness, **b**reathiness, **a**sthenic, **s**trained quality) rating scale. This system allows members of a voice team to quantify their perception of the overall severity of a patient's voice disorder (DeBoldt, Wuyts, Van de Heyning, & Croux, 1997). The GRBAS scale has the added advantage of practical use in the clinical setting because it has only 5 parameters (grade, roughness, breathiness, asthenic, and strained quality) and 4 rating categories (normal, slight, moderate, and severe), (DeBoldt et al., 1997). However, because this scale is completed by the SLP or ENT, and not by the patient, it does not reflect patient perception of voice handicap.

The *Voice Handicap Index* (VHI) and the *Voice-Related Quality of Life* survey (V-RQOL) are two self-administered patient questionnaires utilized for the quantification of handicap and quality of life (respectively) related to voice disorders (Hogikyan & Sethurman, 1999; Jacobson et al., 1997). Both scales have strong test-retest reliability as well as construct validity and are applicable to a wide range of voice disorders (Hogikyan & Sethurman, 1999; Jacobson et al., 1997; Rosen et al., 2000). While these instruments are similar in many ways, the fundamental difference between the two is that the VHI measures voice handicap and the V-RQOL measures quality of life (Hogikyan & Sethurman, 1999).

Jacobson et al. (1997) developed the VHI in order to quantify a patient's handicap resulting from his or her voice disorder. Sixty-five subjects were asked to complete a preliminary 85-item scale. These original questions were developed according to previous patients' reports of the psychosocial effects of voice disorders. The initial 85 items were divided into functional, physical, and emotional facets of voice disorders. Following a statistical analysis for internal consistent validity, the preliminary 85-item version was reduced to 57 items. The final version consists of a functional, a physical, and an emotional subscale, each comprised of 10 items (Jacobson et al., 1997). Final version test-retest reliability for both subscale and total scores was found to be strong for the functional (r=.84), physical (r=.86), and emotional (r=.92) and total scores (r=.92). The relationship between the functional, emotional, and physical subscales were moderate-strong. Pearson product-moment correlations ranged from r=.70 to r=.79 (Jacobson et al., 1997).

Scoring of the VHI is based on an ordinal scale. The patient rates each question between "0" and "4." Zero represents a response of "never," 1 represents "almost never," 2 represents "sometimes," 3 represents "almost always," and 4 represents "always" (Jacobson et al., 1997). A total score of 120 represents a maximum perceived handicap resulting from the voice disorder. The VHI was found to be useful in assessing the patient's judgment in relation to the effect the voice disorder has on daily living. According to its developers, the VHI is also useful in measuring functional outcomes of behavioral, medical, and surgical treatment of voice disorders (Jacobsen et al., 1998).

In a study by Murry and Rosen (2000), the VHI was used to assess changes in the degree of handicap patients experience following voice treatment. This study revealed that patients reliably identify the degree of handicap they are experiencing as well as the significant changes in that handicap after treatment. In addition, a study by Rosen et al.

(2000) demonstrated that patients from three different diagnostic groups (unilateral vocal cord paralysis, muscle tension dysphonia, and vocal fold polyp or vocal cord cyst) showed a decrease in average VHI score following treatment. This study suggests that while the absolute score on the VHI is important, the percentage of change between the pre-treatment and post-treatment score is the more critical measure when assessing treatment outcome (Rosen et al., 2000).

Spector, Netterville, Billante, Clary, Reinisch, and Smith (2001) conducted a study that also found the VHI to be sensitive to change in patient perception of voice handicap following surgical treatment of unilateral vocal cord paralysis (UVCP). In this study the VHI provided information regarding functional, physical, and emotional changes when each subscale was examined individually (Spector et al., 2001). Spector et al. advocated using the VHI when planning a course of treatment. For example, a low VHI score (indicating minimal handicap) might suggest a conservative approach to treatment would be best (as opposed to an invasive surgery), while a high score might suggest an aggressive treatment is necessary (Murry & Rosen, 2000).

The VHI has also been used with the singing population (Rosen & Murry, 2000). Rosen and Murry examined the degree of handicap expressed by singers (both professional and amateur) and non-singers presenting with a voice complaint. The VHI was administered to assess the patients' perception of the handicapping effects of their voice disorder. Results revealed that the VHI scores of the singers were, on average, lower than those of non-singers. Rosen and Murry (2000) hypothesized that the reason for this difference could be multifactorial: the questions on the VHI may not address problems related specifically to singing, singers may be more sensitive to changes in voice and therefore present earlier in the course of their voice problem, and non-singers may not present until a time when their voice reaches a more handicapping level.

The VHI has proven to be a valuable tool in assessing self-perceived handicap in a diverse population of voice patients. It has also proved to be effective in the evaluation of treatment outcome in a wide range of voice disorders (Benninger et al., 1998; Jacobsen et al., 1998; Murry & Rosen, 2000, Rosen et al., 2000, Spector et al., 2001). However, to date there are no studies which examine the relationship between patient self-perceived handicap and objective measures of voice. It is known that patients are generally adept at assessing the degree of handicap they experience in daily life using the VHI (Murry & Rosen, 2000), but how does this measure relate to measures of vocal dysfunction? It is important to know the relationship between acoustic measures used in the diagnosis of voice disorders and patients' self-perception of their voice handicap as quantified by the VHI. Once known, the relationship between the VHI and acoustic measures will provide information concerning the course of treatment that will best suit the patient as well as be able to more definitively document treatment outcome. This study was conducted for the purpose of determining the relationship between acoustic measures of the disordered voice and patient responses on the VHI. It was hypothesized that a significant positive correlation would exist between specific patient responses on the VHI and acoustic measures of voice.

CHAPTER 2
METHODS

Participants

Fifty patients from the Ayers Outpatient Ear, Nose, and Throat Clinic in Gainesville, Florida participated in this study. Potential subjects were identified based on the presence of a voice complaint. Criteria for inclusion in the study consisted of 1) primary complaint regarding the quality of the voice, and 2) over the age of 18 years. Once identified, the potential participants were asked to voluntarily participate in the study. Thirty-eight females and 12 males, ranging in age from 19 to 80 years, with a mean age of 49 years were selected as participants.

Table 1 provides information about each subject including age, sex, and diagnosis.

Equipment and Procedures

All participants completed a VHI questionnaire. It was explained to each participant that the VHI would ask questions regarding how the voice problem affected different aspects of their life. The questionnaire was completed during their outpatient office visit to the otolaryngologist, either before or after a sample of their voice was recorded. The VHI can be found in Appendix B. Total time to complete the questionnaire was 10 minutes.

Table 1. Participant information regarding age, sex, and diagnosis of voice condition

Participant #	Age	Sex	Diagnosis
1	28	M	GERD
2	56	F	prenodules
3	62	F	contact ulcer/granuloma
4	48	F	GERD
5	27	M	polyp/reflux
6	66	M	GERD
7	24	F	prenodules
8	57	M	polyp
9	62	F	dysphonia
10	20	F	prenodules
11	31	F	edema/prenodules
12	64	F	dislocated arytenoid
13	45	F	nodules
14	31	F	nodules
15	22	F	prenodules
16	19	F	prenodules
17	48	F	dysphonia
18	62	M	leukoplakia
19	22	F	nodules
20	53	F	MTD
21	54	M	VF paralysis
22	56	F	PD
23	60	M	VF edema
24	42	F	MTD
25	72	F	mild presbylaryngis
26	25	F	vocal nodules
27	47	F	MTD
28	71	F	GERD
29	65	F	reduced VF movement
30	73	M	PD
31	30	F	MTD
32	60	F	unilateral VF paralysis
33	33	F	dysphonia
34	42	F	paradoxical VF function
35	43	M	contact ulcer/granuloma
36	59	F	MTD
37	48	F	GERD
38	80	F	tremor, atrophy
39	33	F	VF hemorrhage
40	72	F	age-related changes
41	40	F	prenodules/reflux
42	74	F	Tremor
43	51	F	SD
44	75	F	VF paralysis
45	45	F	R. VF paralysis
46	43	M	MTD
47	43	M	dysphonia
48	35	F	GERD
49	71	M	paralysis
50	59	F	paralysis

Voice samples were collected using a high quality condensor type unidirectional stand microphone. The participant stood in front of the microphone and the microphone head was placed 8.8cm from the mouth. In some circumstances a unidirectional headset microphone was used and placed 2.2cm from the left corner of the mouth. Voice samples were preamplified using DBX microphone preamplifier (model 760-X) or phantom power, and recorded to a Sony digital audio tape recorder (model ZA5ES).

For the voice recordings, all participants were asked to produce vocalizations at either soft, comfortable, or loud effort levels. Three trials of the sustained vowel /a/ at soft, comfortable, and loud effort levels, and *The Zoo Passage* (Fletcher, 1972) at a comfortable effort level were produced by each participant. *The Zoo Passage* was chosen because the majority of the segment is produced with voicing. A total of 212 samples were collected (53 recordings x 3 vowels + 53 zoo passage). Acoustic segments were viewed and analyzed using *Cool Edit 2000* (Syntrillium Software Corporation, 2000) and *TF32* (Milenkovic, 2001). The following measures were made.

Measures

Vowels

Measures of fundamental frequency (F_0), jitter%, shimmer%, signal-to-noise ratio (SNR), H_1-H_2, H_1-A_1, and H_1-A_3 were obtained using the TF32 program from the middle 500ms of the vowel. H_1-H_2, H_1-A_1, and H_1-A_3 were obtained by identifying the first and second harmonics as well as first and third formants on an amplitude spectrum generated by TF32. (Figure 1). The corresponding amplitudes were then recorded onto an Microsoft excel spreadsheet. Each measure was then calculated by subtracting the appropriate amplitudes from one another. The RMS intensity of each sustained /a/ was automatically calculated using the Cool Edit software.

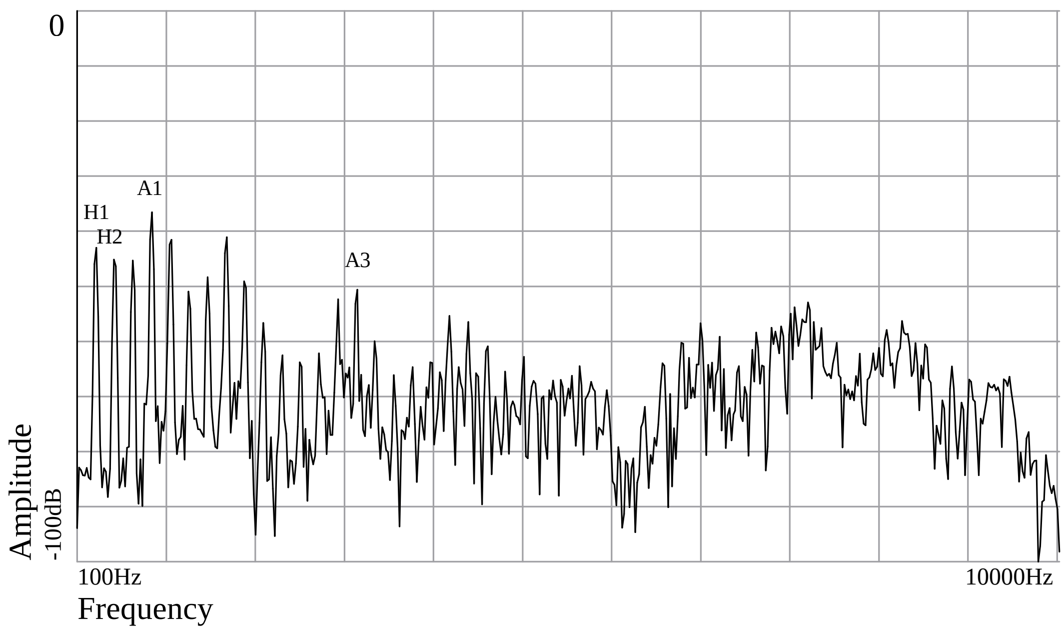


Figure 1. Schematic of sound spectrum indicating measurement points for H1, H2, A1, and A3

Connected Speech

The connected speech samples were analyzed using TF32 and Cool Edit. Cool edit was used to calculate the mean intensity level of the passage. TF32 was used to generate a pitch trace, which produced measures of mean frequency and frequency standard deviation during speaking.

Aphonic periods and phrase duration were also measured in the connected speech samples using Cool Edit. Aphonic periods were identified as those in which there was total absence of voicing during a word or group of words that would typically be voiced. Once identified both aurally and visually from the waveform, the length of each aphonic period was measured in milliseconds by placing cursers around the aphonic period.

To calculate phrase duration, the number of breaths each subject took during the *Zoo Passage* was calculated by listening to the passage and identifying each breath sound both visually and or aurally. Once the total number of breaths was ascertained, the total number of syllables in the zoo passage (83) was divided by the number of breaths. The resulting number was the phrase duration in syllables per breath.

Statistical Analysis

Intrameasurer and intermeasurer reliability was completed on 10% of the data. Pearson r correlation was used to measure intrameasurer and intermeasurer reliability. As well, pair-wise t-tests were performed to determine the direction of the differences between measurement one and two and their significance. Results of the intrameasurer reliability are in Table 2. Results of the intermeasurer reliability are in Table 3.

In order to determine whether significant correlations existed between participant responses on specific VHI questions and acoustic measures, a Pearson r correlation statistic was used. A significance level of 0.05 was set. Multiple linear regression was used to examine the relationship between the overall VHI and acoustic measures. Univariate analysis of variance (ANOVA) was used to identify the effects of group or sex effects on overall VHI score or acoustic measures.

CHAPTER 3
RESULTS

Table 2 and Table 3 indicate the results of the intrameasurer and intermeasurer reliability for each of the dependent measures. Results indicated significant correlation within and between measurements as well as non-significant differences as indicated by the t-test results.

Appendix A shows participant information regarding age, diagnosis, overall VHI score, and mean data for the dependent variables. The results of Pearson r correlations for the chosen items on the VHI and specific acoustic measures can be found in Table 4. Figures 2-19 show scatter-plots for those correlations which were significant.

A multiple linear regression was completed using the dependent variables of interest for each task (vowel and zoo passage) for predicting overall VHI score. The results showed that none of the dependent variables analyzed from the vowel sample were significant predictors of the overall VHI score (Table 5). The results also showed that none of the dependent variables analyzed from the connected speech sample were significant predictors of the overall VHI score (Table 6).

Table 2. Correlation results of intrameasurer reliability

Vowel

	F_0	Jit %	Shim %	Intensity	H1-H2	H1-A1	H1-A3	SNR
r	1.000	0.974	0.925	1.000	0.936	0.889	0.638	0.859
p	0.000	0.000	0.000	0.000	0.000	0.000	0.011	0.000
t	-0.373	1.203	0.994	-0.269	0.660	-1.109	1.745	-1.069
p	0.715	0.249	0.337	0.792	0.520	0.286	0.103	0.303

Zoo Passage

	F_0	F_0 SD	Instensity	# of breaths	# of aphonic periods	Duration of aphonic periods
r	1.000	1.000	1.000	0.922	0.944	1.000
p				0.026	0.001	0.000
t				-2.449	1.000	1.408
p				0.070	0.374	0.232

Table 3. Correlation results of intermeasurer reliability

Vowel

	F_0	Jit %	Shim %	Intensity	H1-H2	H1-A1	H1-A3	SNR
r	0.992	0.956	0.997	0.986	0.935	0.891	0.812	0.848
p	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
t	0.898	1.350	0.648	-0.944	0.981	-1.033	-0.155	-1.429
p	0.384	0.198	0.527	0.361	0.343	0.319	0.879	0.175

Zoo Passage

	F_0	F_0 SD	Instensity	# of breaths	# of aphonic periods	Duration of aphonic period
r	0.998	0.994	1.000	0.877	0.999	1.000
p	0.000	0.000		0.051	0.000	0.000
t	-1.000	1.000		-1.633	1.238	1.408
p	0.374	0.374		0.178	0.284	0.232

Table 4. Correlation results for individual VHI questions (F=functional, P=physical) and acoustic measures. Significant correlations are identified with an asterisk (*).

	F-1	F-2	F-3	F-4	P-1	P-4	P-5	P-10
Vowel								
F_0	0.190	-0.006	0.020	0.047	-0.157	-0.147	0.046	-0.010
Intensity	-0.270	*-.365	-0.216	-0.162	-0.109	-0.105	-0.016	-0.250
Shim %	0.175	*.291	0.225	0.260	0.197	0.100	0.208	0.097
Jit %	0.190	*.320	0.211	*.358	0.206	-0.008	0.239	0.212
SNR	-0.257	*-.320	-0.164	*-.299	-0.267	-0.209	*-.337	-0.131
H1-H2	0.138	0.008	0.136	0.055	0.095	0.146	0.182	0.228
H1-A1	0.023	-0.058	0.004	-0.004	-0.004	-0.016	0.115	0.185
H1-A3	0.089	0.233	0.073	0.165	0.144	0.100	0.000	0.068
Zoo Passage								
F_0	0.057	-0.034	0.117	0.088	0.076	0.020	0.157	0.083
F_0 SD	0.209	0.267	0.158	0.248	0.062	0.123	*.328	0.249
Intensity	-0.141	-0.177	-0.082	-0.117	-0.148	-0.096	0.044	-0.065
# of Breaths	0.172	*.317	0.153	*.324	0.130	0.263	*.423	0.201
Phrase duration	-0.267	*-.353	-0.242	*-.332	-0.162	-0.238	*-.485	*.285
Aphonic periods	0.132	*.334	0.142	*.306	0.152	0.209	0.263	*.283

Further analysis was done by taking the composite VHI score (maximum 120) and dividing it by four to provide handicap severity levels. A mild handicap level was defined as 0-30, a moderate handicap level was defined as 31-60, a severe handicap level was defined as 61-90, and a profound handicap level was defined as 91-120. A univariate analysis of variance was run on the database with between subject factors of handicap level, sex, age, and task. Results for the sustained vowel task showed no significant difference as a function of sex (F=.000, df=1,49, p=.996), age (F=1.329, df=3,49, p=.293), or an interaction between handicap level and sex (F=95.343, df=3,49, p=.369) or handicap level and age (F=1.171, df=8,49, p=.363). Results for the connected speech task (Zoo Passage) revealed no significant difference as a function of sex (F=.472, df=1,49, p=.499), age (F=.587, df=3,49, p=.630), or an interaction between handicap level and sex (F=87.583, df=3,49, p=.430) or handicap level and age (F=1.097, df=8,49, p=.402).

Table 5. Multiple linear regression results for the dependent variables analyzed from the vowel

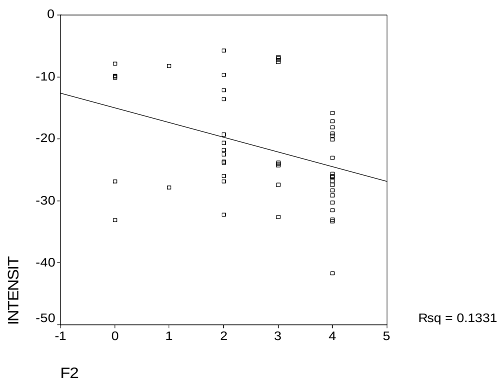
Model Summary				
Model	R	R square	Adjusted R Square	Std. Error of the Estimate
1	0.535	0.286	0.186	25.5299

Coefficients						
		Unstandardized Coefficients		Standardized Coeffieients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	57.395	65.579		0.875	0.286
	F_0	1.22E-01	0.118	0.149	1.034	0.307
	F_0 SD	-4.31E-02	0.253	-0.028	-0.17	0.866
	Intensity	-0.915	0.502	-0.249	-1.823	0.075
	# of Breaths	-0.189	3.418	-0.02	-0.055	0.956
	Phrase duration	-2.657	2.369	-0.419	-1.121	0.268
	Aphonic periods	1.372	1.174	0.186	1.168	0.249

Table 6. Multiple linear regression results for the dependent variables analyzed from the Zoo Passage

Model Summary				
Model	R	R square	Adjusted R Square	Std. Error of the Estimate
1	0.535	0.286	0.186	25.5299

Coefficients						
		Unstandardized Coefficients		Standardized Coeffieients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	57.395	65.579		0.875	0.286
	F_0	1.22E-01	0.118	0.149	1.034	0.307
	F_0 SD	-4.31E-02	0.253	-0.028	-0.17	0.866
	Intensity	-0.915	0.502	-0.249	-1.823	0.075
	# of Breaths	-0.189	3.418	-0.02	-0.055	0.956
	Phrase duration	-2.657	2.369	-0.419	-1.121	0.268
	Aphonic periods	1.372	1.174	0.186	1.168	0.249



Figrure 2. Scatter plot of item 2 from the Functional subscale and vowel Intensity

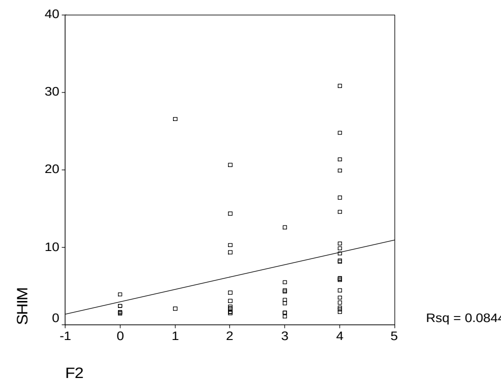


Figure 3. Scatter plot of item 2 from the Functional subscale and Shimmer %

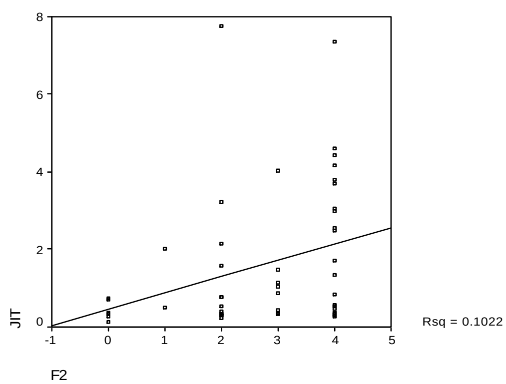


Figure 4. Scatter plot of item 2 from the Functional subscale and Jitter %

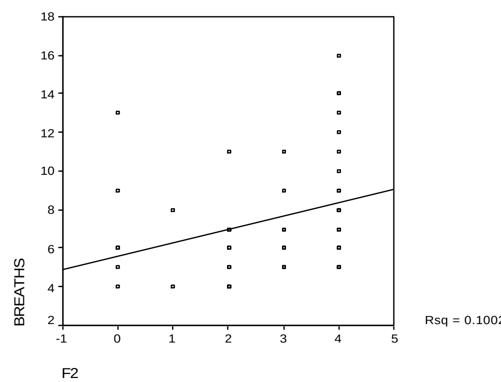


Figure 5. Scatter plot of item 2 from the Functional subscale and # of breaths

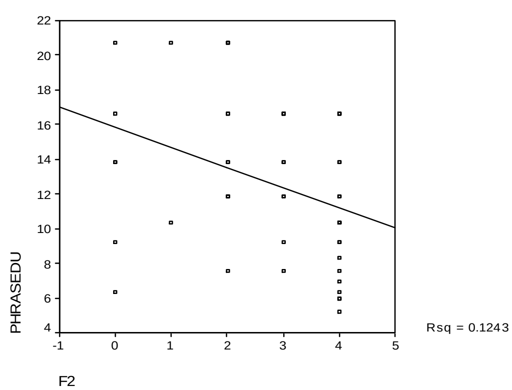


Figure 6. Scatter plot of item 2 from the Functional subscale and Phrase Duration

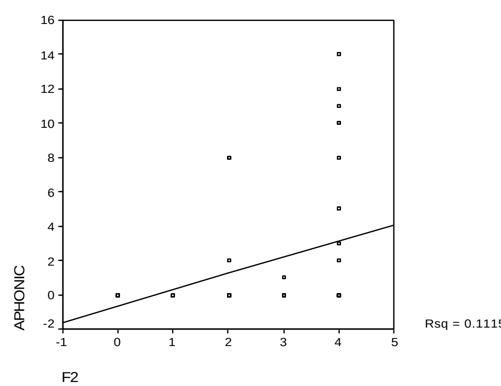
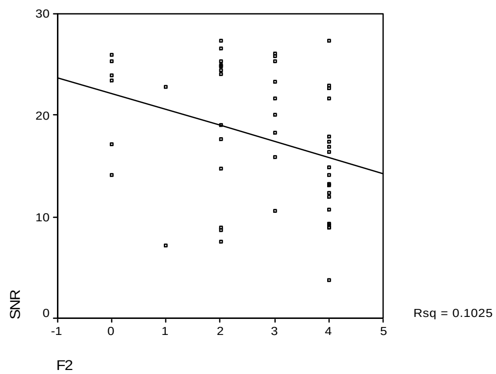


Figure 7. Scatter plot of item 2 from the Functional subscale and Aphonic Periods



Firgure 8. Scatter plot of item 2 from the Functional subscale and SNR

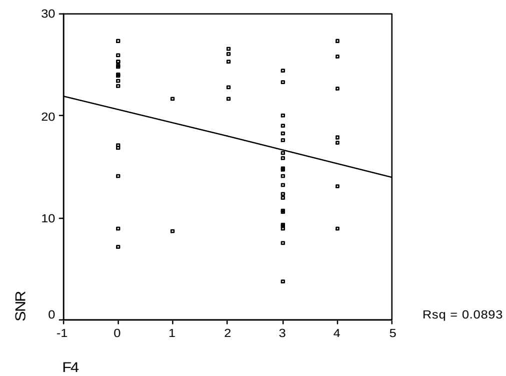


Figure 9. Scatter plot of item 4 from the Functional subscale and SNR

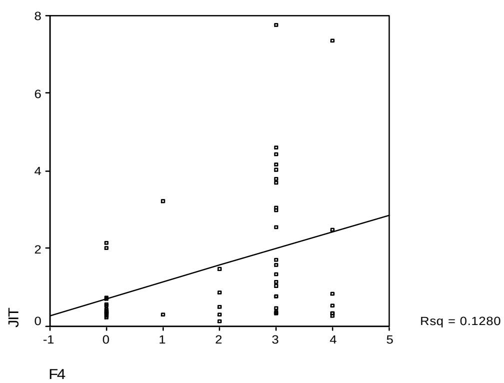


Figure 10. Scatter plot of item 4 from the Functional subscale and Jitter %

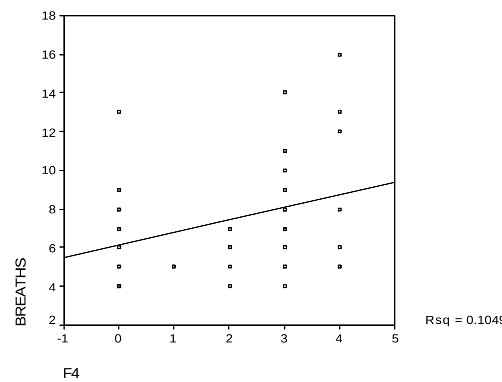


Figure 11. Scatter plot of item 4 from the Functional subscale and # of breaths

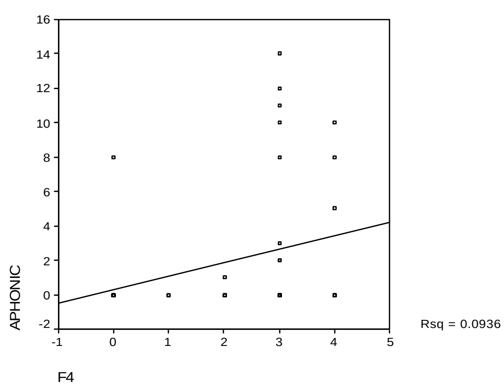


Figure 12. Scatter plot of item 4 from the Functional subscale and Aphonic Periods

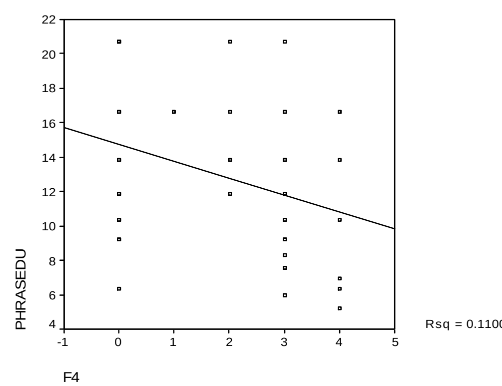


Figure 13. Scatter plot of item 4 from the Functional subscale and Phrase Duration

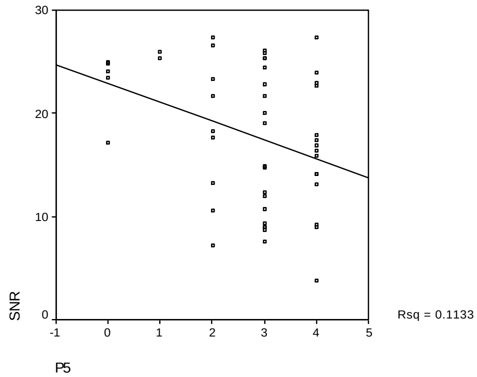


Figure 14. Scatter plot of item 5 from the Physical subscale and SNR

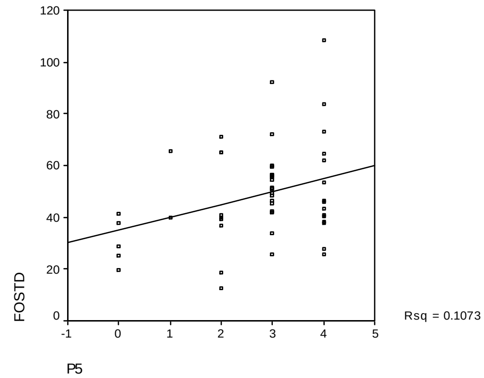


Figure 15. Scatter plot of item 5 from the Physical subscale and F_0 SD

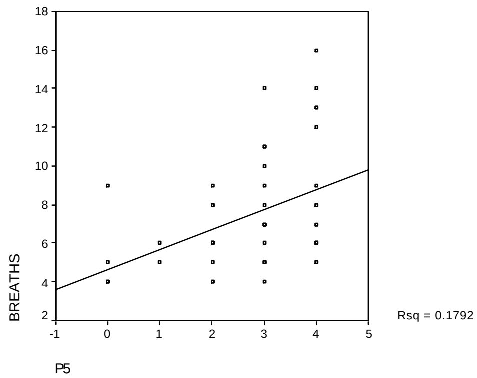


Figure 16. Scatter plot of item 5 from the Physical subscale and # of breaths

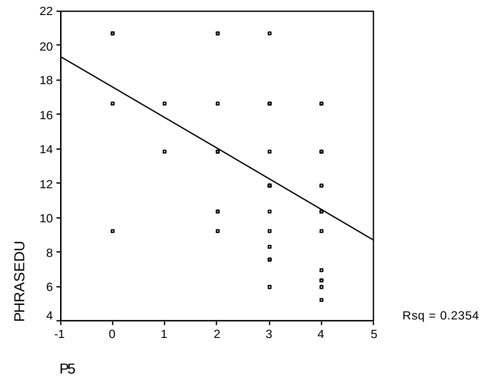


Figure 17. Scatter plot of item 5 from the Physical subscale and Phrase

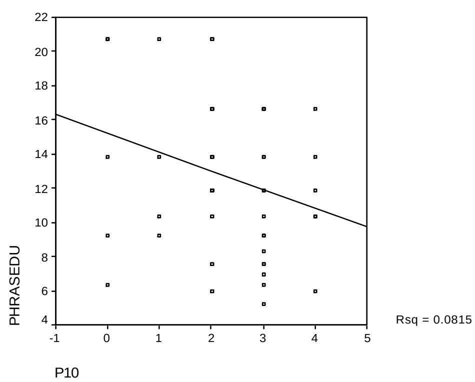


Figure 18. Scatter plot of item 10 from the Physical subscale and Phrase duration

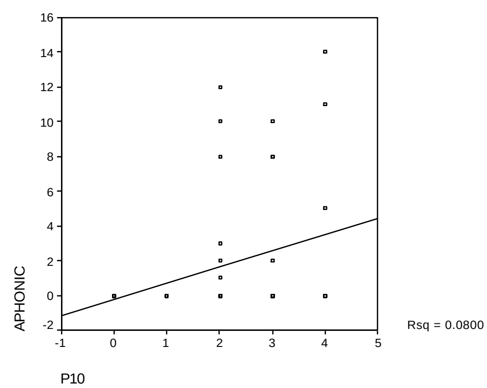


Figure 19. Scatter plot of item 10 from the Physical subscale and Aphonic periods

CHAPTER 4
DISCUSSION

The purpose of this study was to examine the relationship between specific questions on the *Voice Handicap Index* (VHI) and acoustic measures made of disordered voice samples. It was hypothesized that a positive correlation would exist between specific VHI responses from patients with voice disorders as well as the overall VHI score to these acoustic measures.

The VHI is currently used as a tool for assessing patient handicap as a result of a voice problem by many speech-language pathologists (SLP's) as well as otolaryngologists. Several studies have shown that the VHI is useful in measuring functional outcomes of behavioral, medical, and surgical treatment of voice disorders (e.g., Jacobson et al., 1998; Rosen et al., 2000; Spector et al., 2001). As well, the VHI has been used to assess the affect voice disorders have on patient daily living (Jacobson et al., 1998). The overall VHI score, as well as the percentage change between VHI scores pre- to post-intervention, and scores on the individual subscales of the VHI can be important for assessing treatment options and treatment outcome (Murry & Rosen, 2000; Rosen et al., 2000; Spector et al., 2001).

The present study adds information regarding how acoustic measures relate to the degree of handicap a patient experiences (as measured by the VHI) as a result of their voice disorder. As health insurance companies require more objective measures to assess treatment outcome, identifying a clear relationship between patient level of handicap and acoustic measures of the voice becomes more important. Identifying a relationship between the VHI and acoustic measures can provide the SLP and the patient with additional information when weighing treatment options. For example, if a patient whose VHI score is indicative of higher degree of handicap also has acoustic measures which vary from expected norms, a more aggressive treatment option may be more appropriate (i.e., surgery versus therapy).

The data set was analyzed with regard to both individual acoustic measures and the overall VHI score as they related to specific acoustic measures. It was found that existing acoustic measures are not predictive of the overall VHI score reported by patients. This is most likely because as a whole, the VHI does not query an individual about the behavior of the vocal folds but rather asks how the disordered voice affects general social, economic, and emotional aspects of their life. It was for this reason that specific items from the *functional* and *physical* subscales of the VHI were targeted and a more detailed analysis was completed to determine the relationship between these questions and the acoustic measures of interest. These questions were chosen based on their potential ability to correlate with acoustic measures. Responses to the following questions were used to complete the correlations:

Functional subscale items:

F-1. My voice makes it difficult for people to hear me.

F-2. People have difficulty understanding me in a noisy room.

F-3. My family has difficulty hearing me when I call them throughout the house.

F-4. I use the phone less often than I would like to.

Physical subscale items:

P-1. I run out of air when I talk.

P-4. My voice sounds creaky and dry.

P-5. I feel as though I have to strain to produce voice.

P-10. My voice “gives out” on me in the middle of speaking.

Results showed that items F-2, F-4, P-5, and P-10 did have significant positive correlations with some of the acoustic measures. Item F-2 was found to correlate positively with measures of intensity, shimmer %, jitter %, SNR, number of breaths, phrase duration, and aphonic periods. The feeling that a patient’s voice makes it difficult for people to understand them in a noisy environment may be due to decreased loudness (acoustic correlate, intensity), a voice that is variable (% jitter, % shimmer, SNR), such as occurs with a hoarse, breathy, or rough voice quality, or due to decreased subglottal pressure which may manifest itself as increased number of breaths or decreased phrase duration (Baken & Orlikoff, 2000; Borden et al., 1994; Colton & Casper, 1996; Kent & Read, 2002; Titze, 1994). Additionally, it is reasonable that patients who experience aphonic periods, or periods when no voicing occurs when it is expected, would feel that their voice is less understandable when competing with a noisy environment.

Item F-4 positively correlated with acoustic measures of jitter %, SNR, number of breaths, phrase duration, and aphonic periods. Item F-4 states, “I use the phone less often than I would like to.” Over the telephone voices undergo filtering which causes distortion to the voice signal. People often need to speak louder and more clearly in order to be understood. Because increased jitter % and SNR both correspond to the perception of breathy, hoarse, or rough voices, which are perceived as less clear, it is expected that more problems would exist when trying to communicate effectively via the telephone (Baken & Orlikoff, 2000; Borden et al., 1994; Colton & Casper, 1996; Kent & Read, 2002; Titze, 1994). Patients who exhibit increased number of breaths or decreased phrase duration may feel fatigued when trying to speak in a clear and loud enough voice to effectively use the telephone. Additionally, patients who have significant aphonic periods when speaking may feel it is too difficult to communicate effectively using the telephone.

Item P-5, “I feel as though I have to strain to produce voice,” correlates positively with SNR, F_0 SD, number of breaths, and phrase duration. Increased noise in the voice sample may indicate breathiness or hoarseness, which may give the speaker the feeling that they need to increase effort in order to produce a voice. Additionally, variability in F_0 may indicate an inability to control voicing, as would be expected with voice disorders, and perhaps lead to the feeling of vocal fatigue (Baken & Orlikoff, 2000; Borden et al., 1994; Colton & Casper, 1996; Kent & Read, 2002; Titze, 1994). Increased breaths or decreased phrase duration may be indicative of inadequate respiratory drive, which results in the need for increased effort in order to produce voice. It is not unusual for voice disorders to be associated with laryngeal hyperfunction as a means to compensate for disordered vibratory mechanics (Benninger et al., 1994; Colton & Casper, 1996; Stemple et al., 1996).

For item P-10, “my voice ‘gives out’ on me in the middle of speaking,” significant positive correlations were found when compared to phrase duration and aphonic periods. It is reasonable that measurement of aphonic periods, which are times when no voicing occurs when it is expected, is positively correlated to this question.

When an aphonic period occurs during speech, the vocal folds cease to vibrate, which may give the speaker the feeling that the voice has "given out." Additionally, decreased phrase duration may indicate that voice may "give out" secondary to inadequate subglottal pressure to maintain vocal fold vibration.

A closer look at the specific item results indicated that F-2, F-4, P-5, and P-10 all positively correlated with the measure of phrase duration. Three of these four items (F-2, F-4, and P-5) were positively correlated with the measures of number of breaths and SNR. Additionally, three of these four items (F-2, F-4, and P-10) positively correlated with the number of aphonic periods. With the exception of SNR, these measures were all made from the connected speech sample (Zoo Passage). It may be that these measures are more sensitive measures to examine clinically when evaluating the patient's level of handicap.

There were four selected VHI items which did not correlate significantly with the acoustic measures. These were items F-1, F-3, P-1, and P-4. Reasons these VHI items did not correlate may be multifactorial. Item F-1 may be too vague to correlate to specific acoustic measures. Perhaps the patient's response to this question was based on an ideal speaking situation; quiet environment, one-on-one speaking, facing the speaking partner, etc. In this situation it may not be difficult for the listener to hear the patient's voice. Additionally, many voice disorders do not affect the patient's ability to be adequately loud. For example, a disorder of hyperfunction may result in a voice that is too loud. The overall voice quality of that disordered voice may be perceived as undesirable, but not difficult to hear.

Patients often make adaptations when in their own home and with their family which may influence the response to item F-3, which states “my family has difficulty hearing me when I call them throughout the house.” Patients may develop strategies with family members for communicating throughout their home which make this statement less handicapping for them. Additionally, the VHI assumes that all individual filling out the scale have a family or are required to interact with family members. The size of the patient’s home could potentially influence the rating given to this question as well. Item F-3 is also closely related to item F-1 in that they both relate to decreased loudness. Again, many disordered voices are not necessarily inadequately loud.

Items P-1 (I run out of air when I talk) and P-4 (my voice sounds creaky and dry) also did not significantly correlate with any acoustic measures. It is surprising that item P-1 did not correlate with increased number of breaths or phrase duration. However, during the study, many patients asked questions concerning items P-1 and P-4 because they were unsure of the appropriate responses. Many patients were unclear as to what “creaky and dry” meant, or they were not sure if their voice was “creaky and dry.” Additionally, many patients commented that they never felt they were gasping for air when speaking, and consequently chose not to respond with a more severe rating on item P-1. During the study, although some patients requested further definition, no assistance was given in order to ensure consistency between participants.

The acoustic measures of F_0 (vowel), H1-H2, H1-A1, H1-A3, F_0 (Zoo Passage), and intensity (Zoo Passage) were not significantly correlated with any of the selected VHI items. While some disorders of pitch, such as puberphonia or persistent glottal fry, are related to abnormal frequency, many voice disorders are not necessarily associated with abnormal F_0 (Colton & Casper, 1996; Rosen & Sataloff, 1997; Stemple et al., 2000). Because the voice samples which were analyzed are disordered, many fell into Titze's Type 3 category (mostly aperiodic) and therefore did not lend themselves to accurate F_0 extraction (Titze, 1995). The measurement of the F_0 depends largely on the assumption that the signal is approximately periodic (Baken & Orlikoff, 2000). When a voice signal is mostly aperiodic (as is the case in many disordered voices), an accurate measurement of F_0 is not possible.

With regard to the spectral measures, they may have been less than accurate due to small variations that occurred in the distance the microphone was placed relative to the patient's mouth. Often times, patients found it difficult to stand perfectly still when performing the speech tasks and, consequently, may have varied in position slightly. Variations in distance from the microphone distort the signal and may result in less accurate measures of amplitude. These variations may also explain the poor correlation of intensity measures to VHI items.

Finally, not all items from the VHI were compared to acoustic measures. It is possible that correlations existed outside of the specific VHI items analyzed. However, for the purposes of this study, only those items which were identified based upon their potential ability to correlate with acoustic measures were examined. Comparing the remaining items with acoustic measures may be valuable in a future study.

It may also be relevant with regard to future study to examine the relationship between item responses on the VHI relative to each other. For example, determining the relationship between items on the functional subscale and items on the physical subscale. This may help narrow down which questions are the most sensitive and avoid redundancy in the patient's responses as well as indicate which responses relate most closely to the quantitative outcome measures.

Results of this study indicate that the specific questions from the VHI did show positive correlation to acoustic measures and as such may be the more clinically useful items to examine when tracking pre- and post- treatment measures. Instead of having the patient answer all 120 VHI items, it may only be necessary with regard to outcomes measures to answer specific items which correlate significantly to objective acoustic measures. This would substantially reduce the amount of time the patient is required to fill out questionnaires. Additionally, when a clinician is selecting acoustic measures to make of a voice sample, those which showed significant correlations to VHI items according to this study may be the more important measures to examine.

APPENDIX A
PARTICIPANT INFORMATION

Appendix A. Data pertaining to each participant's outcome for the VHI along with the means for each acoustic dependent variable as a function of the vowel task and the production of the *Zoo Passage.* F = Functional subscale. P = Physical subscale.

							VHI response			
Participant	Diagnosis	VHI Score	F-1	F-2	F-3	F-4	P-1	P-4	P-5	P-10
1	GERD	6	0	0	0	0	0	0	0	2
2	prenodules	10	1	2	0	0	0	1	0	0
3	contact ulcer/granuloma	10	0	0	0	0	0	2	0	0
4	GERD	12	2	2	1	0	0	0	0	0
5	polyp/reflux	13	0	0	0	0	0	0	1	0
6	GERD	20	2	2	2	0	0	1	0	3
7	prenodules	21.5	0	0	0	2	0	0	1	2
8	polyp	23	2	0	0	0	0	3	4	0
9	dysphonia	38	3	3	4	0	3	2	3	3
10	prenodules	39	1	2	1	2	0	0	2	1
11	edema/prenodules	39	2	2	3	0	0	3	3	4
12	dislocated arytenoid	42	2	2	2	3	0	0	3	2
13	nodules	46	4	4	4	3	0	2	3	3
14	nodules	46	2	2	1	3	0	1	2	1
15	prenodules	47	0	0	0	0	3	4	4	3
16	prenodules	47	0	1	1	2	1	2	3	0
17	dysphonia	50	2	2	2	0	2	3	2	2
18	leukoplakia	54	3	3	3	2	0	1	3	2
19	nodules	57	1	1	0	0	2	2	2	1
20	MTD	58	2	3	2	3	0	3	2	2
21	VF paralysis	59	2	2	2	3	0	0	3	2
22	PD	60	4	4	3	0	0	3	4	3
23	VF edema	61	2	2	1	0	4	4	3	3
24	MTD	62	2	3	0	3	2	0	2	1

VHI response

Participant	Diagnosis	VHI Score	F-1	F-2	F-3	F-4	P-1	P-4	P-5	P-10
25	mild presbylaryngis	63	3	3	2	2	2	3	2	2
26	vocal nodules	65	2	2	2	3	2	2	3	3
27	MTD	66	3	4	3	3	3	2	3	2
28	GERD	70	4	4	4	3	4	4	4	3
29	reduced VF movement	70	3	4	4	3	1	3	3	2
30	PD	71	2	4	4	3	3	1	4	3
31	MTD	72	2	2	3	3	2	2	3	2
32	unilateral VF paralysis	77	3	4	4	3	3	3	3	2
33	dysphonia	80	4	4	4	0	2	2	4	3
34	paradoxical VF function	82	3	4	3	1	4	4	3	4
35	contact ulcer/granuloma	83	3	3	3	3	1	4	2	3
36	MTD	83	2	2	2	1	2	3	3	2
37	GERD	82	3	4	4	4	3	0	4	3
38	tremor, atrophy	85	2	4	4	3	3	3	2	4
39	VF hemorrhage	88	3	4	3	3	3	3	3	3
40	age-related changes	89	3	3	3	3	3	2	3	3
41	prenodules/reflux	89	3	3	3	4	4	3	3	3
42	Tremor	90	2	4	0	3	0	4	4	4
43	SD	90	3	4	4	3	2	4	4	2
44	VF paralysis	91	4	3	4	3	3	0	4	3
45	R. VF paralysis	94	2	4	4	4	3	4	4	3
46	MTD	95	4	4	4	4	2	4	4	4
47	dysphonia	99	2	4	3	3	3	4	3	2
48	GERD	102	4	4	4	4	4	4	4	2
49	paralysis	104	3	4	3	4	2	4	4	3
50	paralysis	118	3	4	4	4	4	4	4	4

Vowel

Participant	F_0	Intensity	Jitter%	Shimmer%	SNR	H1-H2	H1-A1	H1-A3
1	133.7	-33.1	0.34	1.57	23.5	1.63	-6.57	3.37
2	143.9	-26.0	0.31	2.08	24.9	-1.50	-9.00	10.77
3	186.6	-7.9	0.73	2.41	17.1	7.50	-0.97	6.13
4	179.6	-23.7	0.41	2.17	24.1	-1.63	-10.17	6.00
5	195.6	-10.1	0.28	3.94	26.0	1.07	-1.53	8.83
6	139.9	-20.6	0.53	2.35	24.8	-3.93	-11.57	12.30
7	183.3	-26.9	0.14	1.44	25.3	0.57	-6.27	5.40
8	189.0	-9.8	0.71	2.42	14.1	8.73	3.03	16.47
9	177.6	-6.8	0.44	1.56	25.3	2.43	-6.67	-2.13
10	428.5	-26.9	0.31	1.48	26.6	3.83	0.90	16.70
11	234.0	-23.8	0.33	1.63	25.3	2.00	-4.13	12.07
12	189.8	-5.7	0.35	1.62	24.5	-2.10	-0.40	12.00
13	188.3	-26.0	4.16	19.92	8.9	6.17	0.50	4.83
14	136.6	-32.2	0.76	4.12	17.6	-1.63	-8.40	7.17
15	194.2	-9.9	0.37	1.66	23.9	6.40	0.43	14.27
16	210.0	-8.2	0.49	2.04	22.8	2.33	-0.50	8.33
17	202.8	-13.6	0.23	1.63	27.4	4.80	-2.77	17.43
18	216.3	-24.3	0.86	2.81	26.1	3.03	-2.73	9.77
19	76.1	-27.9	2.03	26.55	7.2	-2.90	-16.97	-10.70
20	201.3	-7.0	0.37	1.59	23.3	2.47	-6.03	12.23
21	187.6	-12.1	0.76	3.04	19.0	6.63	-1.90	9.27
22	133.4	-15.8	0.36	1.63	23.0	3.53	-5.40	7.90
23	82.4	-22.5	2.16	10.31	9.0	9.33	3.27	3.47
24	160.4	-24.0	4.04	12.54	10.6	-6.77	-20.50	-16.00
25	169.6	-7.6	1.47	4.28	21.7	2.67	-0.87	8.33
26	218.1	-9.6	7.77	20.64	7.6	6.17	1.97	10.77
27	100.7	-26.1	2.56	9.85	9.3	4.70	-4.40	3.47
28	186.3	-23.0	1.71	6.02	16.4	-4.63	-13.50	-2.37
29	224.8	-33.0	3.80	24.82	12.0	8.60	9.30	14.67
30	147.9	-18.1	2.98	10.49	14.1	-3.13	-3.97	13.20
31	153.3	-21.8	1.58	9.37	14.7	1.30	-13.23	-4.87
32	223.7	-31.5	1.34	9.18	10.7	7.17	-2.30	8.60
33	196.4	-27.4	0.56	3.51	16.9	-0.30	-4.40	11.30
34	234.2	-41.7	0.30	2.86	21.7	10.63	8.77	18.73
35	90.4	-32.6	0.34	3.19	18.3	11.80	-3.27	12.70
36	148.1	-19.3	3.24	14.33	8.7	4.53	-5.00	3.90
37	157.9	-20.1	7.36	21.33	8.9	8.57	1.87	14.73
38	229.9	-25.6	4.60	16.43	13.2	12.77	5.23	11.30
39	179.3	-26.2	3.71	14.55	12.3	-2.43	-6.07	-1.30
40	241.3	-23.8	1.13	5.50	20.0	1.77	-8.70	13.43
41	140.8	-7.2	0.35	1.09	25.9	-4.53	2.67	16.10
42	224.4	-30.3	3.07	8.12	9.2	3.90	3.73	3.40
43	168.8	-28.3	4.43	30.83	3.8	0.13	-25.33	-13.10
44	192.4	-27.4	1.05	4.43	15.9	15.10	5.27	14.33
45	188.5	-26.8	0.27	2.00	27.3	-1.20	-11.33	16.47
46	101.2	-33.3	0.84	5.96	17.4	-3.30	-18.53	-10.97
47	125.3	-19.5	0.46	5.76	14.9	-15.53	-35.20	-30.93
48	168.6	-29.1	0.35	2.19	22.7	3.13	-6.30	9.30
49	166.6	-17.1	2.50	8.29	13.1	15.57	9.70	25.60
50	209.6	-19.1	0.54	4.41	17.9	7.33	-0.50	29.50

Zoo Passage

Participant	F_0	F_0 SD	Intensity	Number of Breaths	Phrase duration	Aphonic Periods
1	106.4	19.7	-28.51	4	20.75	0
2	133.2	28.9	-28.15	4	20.75	0
3	172.4	41.4	-7.67	9	9.22	0
4	153	37.9	-28.85	4	20.75	0
5	175.7	40	-16.7	6	13.83	0
6	114.1	25.3	-26.25	5	16.60	0
7	167	65.3	-23.58	5	16.60	0
8	132.5	40.2	-19.59	13	6.38	0
9	192	55.6	-8.26	5	16.60	0
10	203.4	40.7	-36.01	6	13.83	0
11	172.9	50.7	-24.22	7	11.86	0
12	176.8	54.3	-10.66	6	13.83	0
13	173.4	59.4	-17.27	9	9.22	2
14	130.9	18.6	-31.74	4	20.75	0
15	188.8	46.6	-14.79	6	13.83	0
16	199.2	56.3	-9.98	4	20.75	0
17	140.9	36.7	-21.82	4	20.75	0
18	158.1	56.3	-15.54	7	11.86	1
19	106.5	39.5	-25.9	8	10.38	0
20	211.6	64.9	-8.65	6	13.83	0
21	155.4	46.6	-14.15	7	11.86	2
22	111.7	43.2	-14.28	8	10.38	0
23	139.6	42.1	-29.16	7	11.86	8
24	119.3	71.2	-26.47	9	9.22	0
25	159.9	39.9	-10.24	6	13.83	0
26	214.3	72	-16.56	11	7.55	0
27	131.2	92.2	-26.78	8	10.38	12
28	169.7	62	-23.01	5	16.60	0
29	176.9	48.4	-33.66	7	11.86	0
30	145.2	27.6	-22	6	13.83	0
31	231.7	51.3	-25.92	7	11.86	8
32	163.7	59.9	-19.37	11	7.55	0
33	192.5	64.4	-16.83	9	9.22	0
34	171.4	45.3	-34.77	5	16.60	0
35	89.6	12.7	-29.31	5	16.60	0
36	159.9	56	-21.97	5	16.60	0
37	98.4	53.4	-12.7	12	6.92	10
38	168.3	64.8	-22.59	8	10.38	11
39	196.4	49.37	-22.47	10	8.30	0
40	202.6	42.4	-26.22	11	7.55	0
41	152.2	33.9	-13.63	5	16.60	0
42	166.6	83.7	-20.26	14	5.93	14
43	173.9	73.1	-29.83	7	11.86	10
44	183.8	37.8	-29.55	7	11.86	0
45	139	38.2	-29.06	16	5.19	0
46	109.9	108.3	-32.63	8	10.38	5
47	116.2	25.8	-22.32	14	5.93	3
48	128.1	25.7	-36.26	5	16.60	0
49	144	40.8	-23.47	13	6.38	8
50	227.2	45.9	-18.64	6	13.83	0

APPENDIX B
VOICE HANDICAP INDEX

Voice Handicap Index – Appendix B

Instructions: These are statements that many people have used to describe their voices and the effect of their voices on their lives. Circle the response that indicates how frequently you have the same experience.

Key: 0 = never
1 = almost never
2 = sometimes
3 = almost always
4 = always

Part I – Functional

*1. My voice makes it difficult for people to hear me.	0 1 2 3 4
*2. People have difficulty understanding me in a noisy room.	0 1 2 3 4
*3. My family has difficulty hearing me when I call them throughout the house.	0 1 2 3 4
*4. I use the phone less often than I would like to.	0 1 2 3 4
5. I tend to avoid groups of people because of my voice.	0 1 2 3 4
6. I speak with friends, neighbors, or relatives less often because of my voice.	0 1 2 3 4
7. People ask me to repeat myself when speaking face-to-face.	0 1 2 3 4
8. My voice difficulties restrict personal and social life.	0 1 2 3 4
9. I feel left ouit of conversations because of my voice problem.	0 1 2 3 4
10. My voice problem causes me to lose income.	0 1 2 3 4

Part II - Physical

*1. I run out of air when I talk.	0 1 2 3 4
2. The sound of my voice varies throughout the day.	0 1 2 3 4
3. People ask, “What is wrong with your voice?”	0 1 2 3 4

4. My voice sounds creaky and dry. 0 1 2 3 4

*5. I feel as though I have to strain to produce voice. 0 1 2 3 4

6. The clarity of my voice is unpredictable. 0 1 2 3 4

7. I try to change my voice to sound different. 0 1 2 3 4

8. I use a great deal of effort to speak. 0 1 2 3 4

9. My voice sounds worse in the evening. 0 1 2 3 4

*10. My voice “gives out” on me in the middle of speaking. 0 1 2 3 4

Part III – Emotional

1. I am tense when talking to others because of my voice. 0 1 2 3 4

2. People seem irritated with my voice. 0 1 2 3 4

3. I find that other people don’t understand my voice problem. 0 1 2 3 4

4. My voice problem upsets me. 0 1 2 3 4

5. I am less outgoing because of my voice problem. 0 1 2 3 4

6. My voice makes me feel handicapped. 0 1 2 3 4

7. I feel annoyed when people ask me to repeat. 0 1 2 3 4

8. I feel embarrassed when people ask me to repeat. 0 1 2 3 4

9. My voice makes me feel incompetent. 0 1 2 3 4

10. I am ashamed of my voice problem. 0 1 2 3 4

* indicates items chosen for comparison to acoustic measures

REFERENCES

Angerstein, W. & Neuschaefer-Rube, C. (1998). Sound pressure level examinations of the calling and speaking voice in healthy persons and in patients with hyperfunctional dysphonia. *Logopedics Phoniatrics Vocology, 23*, 23-25.

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

Benninger, M.S., Ahuja, A.S., Gardner, G., & Grywalski, C. (1998). Assessing outcomes for dysphonic patients. *Journal of Voice, 12 (4),* 540-550.

Benninger, M., Jacobson, B., & Johnson, A. (1994). *Vocal Arts Medicine: The Care and Prevention of Professional Voice Disorders.* New York: Thieme Medical Publishers.

Baken, R.J. & Orlikoff, R.F. (2000). *Clinical Measurement of Speech and Voice,* 2nd ed. San Diego, CA: Singular.

Boone, D. & McFarlane, S. (1988). *The Voice and Voice Therapy.* Englewood Cliffs, NJ: Prentice Hall.

Borden, G., Harris, K. & Raphael, L. (1994). *Speech Science Primer: Physiology, Acoustics, and Perception of Speech.* Baltimore, MD: Lippincott, Williams & Wilkins.

Brown, W.S., Vinson, B.P. & Crary, M.A. (1996). *Organic Voice Disorders: Assessment and Treatment.* San Diego, CA: Singular Publishing Group.

Colton, R. & Casper, J. (1996). *Understanding Voice Problems: A Physiological Perspective for Diagnosis and Treatment,* 2nd ed. Baltimore, MD: Williams & Wilkins.

DeBodt, M.S., Wuyts, F.L., Van de Heyning, P.H. & Croux, C. (1997). Test-retest study of the GRBAS scale: Influence of experience and professional background on perceptual rating of voice quality. *Journal of Voice, 11(1),* 74-80.

Duffy, J. (1995). *Motor Speech Disorders: Substraits, Differential Diagnosis, and Treatment.* St. Louis, IL: Mosby.

Ferrand, C. (2001). *Speech Science: An Integrated Approach to Theory and Clinical Practices.* Boston, MA: Allyn & Bacon.

Ferrand, C. (2002), Harmonics-to-noise ratio: An index of vocal aging. *Journal of Voice, 16(4),* 480-487.

Fletcher, S.G. (1972). Contingencies for bioelectric modification of nasality. *Journal of Speech and Hearing Disorders, 37,* 329-346.

Freeman, M. & Fawcus, M. (2000). *Voice Disorders and Their Management,* 3rd ed. London: Whurr.

Hanson, H.M. (1997). Glottal characteristics of female speakers: Acoustic correlates. *Acoustical Society of America, 101(1),* 466-481.

Hanson, H.M. & Chuang, E.S. (1999). Glottal characteristics of male speakers: Acoustic correlates and comparison with female data. *Acoustical Society of America, 106(2),* 1064-1077.

Heman-Ackah, Y., Micheal, D., & Goding, G. (2002). The relationship between cepstral peak prominence and selected parameters of dysphonia. *Journal of Voice, 16(1),* 20-27.

Hogikyan, N. & Sethuraman, G. (1999). Validation of an instrument to measure voice-related quality of life (V-RQOL). *Journal of Voice, 13 (4),* 557-569.

Jacobson, B., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., & Benninger, M. (1997). The Voice Handicap Index (VHI): Development and validation. *American Journal of Speech-Language Pathology, 6(3),* 66-69.

Jacobson, G.P. & Newman, C.W. (1990) The development of the Dizziness Handicap Inventory (DHI). *Archives of Otolaryngology-Head and Neck Surgery, 116,* 424-427.

Kent, R. (1997). *The Speech Sciences.* San Diego, CA: Singular Publishing Group, Inc.

Kent, R.C. & Read, C. (2002). *Acoustic Analysis of Speech,* 2nd ed. Albany, NY: Singular.

Koschkee, D, & Rammage, L. (1997). *Voice Care in the Medical Setting.* San Diego, CA: Singular Publishing.

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

Morrison, M. & Rammage, L. (1994). *The Management of Voice Disorders*. San Diego, CA: Singular Publishing Group, Inc.

Morsomme, D., Jamart, J., Wery, C, Giovanni, A., & Remacle, M. (2001). Comparison between the GIRBAS scale and the acoustic and aerodynamic measures provided by EVA for the assessment of dysphonia following unilateral vocal fold paralysis. *Folia Phoniatrica et Logopaedica, 1(53),* 317-325.

Murry, T., & Rosen, C.A. (2000). Outcome measurements and quality of life in voice disorders. *Otolaryngologic Clinics of North America, 33(4),* 905-916.

Newman, C.W., Jacobson, G.P., Weinstein, B.E. & Hug, G.A. (1990). The hearing handicap inventory for adults: Psychometric adequacy and audiometric correlates. *Ear and Hearing, 11,* 430-433.

Rammage, L., Morrison, M. & Nichol, H. (2001). *Management of the Voice and Its Disorders.* San Diego, CA: Singular Thomson Learning.

Rosen, C.A. & Murry, T. (2000). Voice Handicap Index in singers. *Journal of Voice, 14(3),* 370-377.

Rosen, C.A., Murry, T., Zinn, A., Zullo, T., & Sonbolian, M. (2000). Voice Handicap Index change following treatment of voice disorders. *Journal of Voice, 14(4),* 619-623.

Rosen, D. & Sataloff, R. (1997). *Psychology of Voice Disorders.* San Diego, CA: Singular Publishing Group.

Sataloff, R. (1991). *Professional Voice: The Science and Art of Clinical Care.* San Diego, CA: Singular Publishing Group

Spector, B.C., Netterville, J.L., Billante, C., Clary, J., Reinisch, L., & Smith, T.L. (2001). Quality of life assessment in patients with unilateral vocal cord paralysis. *Otolaryngology, Head and Neck Surgery, 125(3)*, 176-182.

Stemple, J.C. (1984). *Clinical Voice Pathology: Theory and Management*. Columbus, OH: Charles E. Merrill Publishing Co.

Stemple, J., Glaze, L., & Gerdeman, B. (1996). *Clinical Voice Pathology: Theory and Management*, 2rd ed. San Diego, CA: Singular Publishing Group.

Stemple, J.C., Glaze, L.E., & Klaben, B.G. (2000). *Clinical Voice Pathology: Theory and Management*, 3rd ed. San Diego, CA: Singular Publishing Group.

Titze, I.R. (1994). *Principles of Voice Production.* Englewood Cliffs, NJ: Prentice Hall, Inc.

Titze, I. (1995). *Workshop on Acoustic Voice Analysis: Summary Statement.* Iowa City, IA: National Center for Voice and Speech.

Verdolini, K. & Ramig, L. (2001). Review: Occupational risks for voice problems. *Logoped Phoniatrica Vocol, 26(1)*, 37-46.

Weatherly, C., Worrall, L., & Hickson, L. (1997). The effect of hearing impairment on the vocal characteristics of older people. *Folia Phoniatrica et Logopaedica, 49,* 53-62.

Webster, D. (1999). *Neuroscience of Communication.* San Diego, CA: Singular Publishing Group.

Wilson, J., Deary, I., Millar, A. & MacKenzie (2002). The quality of life imact of dysphonia. *Clinics of Otolaryngology, 27(3)*, 179-182.

Wolfe, V., Ratusnik, D, Smith, F., & Northrop, G. (1990). Observation of perturbation in a lumped-element model of the vocal folds with application to some pathological cases. *Journal of Speech and Hearing Disorders, 55,* 43-50.

Wolfe, V. & Martin, D. (1997). Acoustic correlates of dysphonia: type and severity. *Journal of Communication Disorders, 30,* 403-416.

BIOGRAPHICAL SKETCH

Karen Michelle Wheeler received her Bachelor of Arts degree in communication sciences and disorders from the University of Florida in 2001. She will complete the requirements for the Master of Arts degree in speech pathology at the University of Florida as well. After graduation, Karen plans to pursue a doctoral degree in speech pathology.