A COMPARISON OF SUSTAINED VOWEL AND CONNECTED SPEECH PRODUCTION IN HYPOFUNCTIONAL AND NORMAL VOICES

by

Emily A. Lambert

Bachelor of Science in Communication Sciences & Disorders, 2008
University of Texas at Austin
Austin, Texas

Submitted to the Graduate Faculty of
The Harris College of Nursing & Health Sciences
Texas Christian University
in partial fulfillment of the requirements
for the degree of

Masters of Science in Communication Sciences and Disorders
May 2010

A COMPARISON OF SUSTAINED VOWEL AND CONNECTED SPEECH PRODUCTION IN HYPOFUNCTIONAL AND NORMAL VOICES

Thesis Approved:

Major Professor

Committee Member

Committee Member

College of Nursing & Health Sciences

ACKNOWLEDGMENTS

While completing this project I have been blessed to be surrounded by a group of people who have constantly supported and encouraged me. I would like to thank my husband Jeff for always believing in me. I love you and I could not have done this without you. I want to thank my family for always encouraging me to do my best. I could not have asked for a better group of committee members. I would like to give a big thank you to Kim and Shelby for collecting the acoustic samples and being fantastic mentors and friends. I would especially like to thank Dr. Chris Watts for all of the patience, encouragement, and guidance. This project led to my love of voice and for that I will forever be grateful.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

CHAPTER 1

INTRODUCTION

One major goal of clinical acoustic voice analysis is to obtain measurements which would allow inferences regarding vocal function that represents voice produced in the most natural and habitual manner. It is anecdotally observed (and reported in the literature) that individuals produce a sustained vowel and connected speech in a different manner. The majority of published, clinically-based literature investigating acoustic properties of voice in treatment-seeking populations, such as those with hypofunctional voices which result in a breathy voice quality, have utilized sustained vowel production as the measured stimulus, and have used time-based analyses to determine if differences compared to normal speakers existed. This has been in part due to software limitations which have prevented reliable acoustic analyses of connected speech, due in large part to its transient nature. Additionally, traditional time-based acoustic analyses have been less reliable for the measurement of vocal function (e.g., fundamental frequency, jitter, shimmer) in highly aperiodic signals, such as those produced in a breathy voice. Recently, new software has been developed which allows for more valid acoustic measurements of connected speech using spectral-based methods. The current study will utilize this new software to investigate whether or not detectible differences exist in acoustic measurements of sustained vowels and connected speech in a sample of individuals with a hypofunctional breathy voice quality compared to a control group of normal non-disordered individuals.

CHAPTER II

REVIEW OF LITERATURE

Acoustic Measurement of Voice Production

Traditionally, clinical evaluation of voice type (i.e., breathy, rough, hoarse) and severity (i.e., mild, moderate, severe) have been based on perceptual evaluation via clinical tools (e.g., CAPE-V) and professional judgment (Awan & Roy, 2005). However, due to the subjective nature of perceptual evaluations, the diagnosis may vary across professionals. Acoustic evaluation of voice quality is a noninvasive method applicable to both diagnosis and treatment progress and allows for standardized categorization of voice quality (Awan & Roy, 2006). Results are in an objective, numerical format which can easily be shared with both the patient and other professionals. Two methods of acoustic analyses of voice are time-based measures and spectral/cepstral analysis (Awan & Roy, 2006). Time-based measures depend on accurate identification of cycle boundaries. In a normal voice, phonatory boundaries (i.e., where a cycle begins and ends) are easily identified because of the periodicity and regularity of the sinusoidal wave; however, an aperiodic voice does not have clear phonatory onset and offset due to the influence of significant noise in the voice signal (Awan & Roy, 2006). Cepstral analysis is a method in which the fundamental frequency (F0) and spectral harmonics are extracted from the spectrum of a sound wave. Analysis of a normal voice using this method will result in a distinct harmonic structure with more prominent cepstral peaks (Awan & Roy, 2006). An aperiodic voice will have multiple cepstral peaks with similar amplitudes without a defined pattern of prominence (Heman-Ackan et al., 2003). Previous research studies have observed that greater severity of a voice disorder was associated with an increase in high frequency energy and a decrease in cepstral peak prominence (Awan & Roy, 2006). In this study, cepstral analysis will be used to compare normal and hypofunctional vocal recordings to determine if this analysis method is capable of differentiating disordered and normal voices.

Substrates of Vocal Hypofunction - Vocal Fold Neurophysiology

The larynx is innervated peripherally by Cranial Nerve X, the Vagus, which supplies the larynx via the Recurrent Laryngeal Nerve (RLN) and the Superior Laryngeal Nerve (SLN). The RLN and SLN branch off the Vagus and provide sensory and motor information to specific areas of the larynx (Stemple, Glaze, & Kalben, 2000). The SLN branches off the Vagus nerve in the neck, runs parallel with the carotid artery, and then bifurcates to form an internal and external branch. The internal branch supplies sensory information to the supra-glottal laryngeal structures and the external branch provides motor innervation to the Cricothyroid muscles (Negus, 1949). The RLN extends to the thorax, loops around the aortic arch (on the left side), and then courses superiorly back to the larynx. This branch supplies sensory information to the sub-glottal portion of the larynx and provides the Interarytenoids, Posterior and Lateral Cricoarytenoid, and Thyroarytenoids (the muscles responsible for adduction and abduction) with motor innervations (Negus, 1949). Due to the course of the RLN into the thoracic cavity and around the aortic arch, this pathway is highly susceptible to iatrogenic injury resulting in absent motor innervations to the muscles responsible for adduction and abduction.

The Central Nervous System (CNS) relays efferent (motor) and afferent (sensory) commands from the brain to the body and the larynx to the brain, respectively. Based on current evidence, researchers report that afferent information is transmitted by the internal branch of the SLN to the nucleus tractus solitarious in the medulla, which is essential for vocalization in addition to the control of respiration, swallowing, and other laryngeal movements. However, the CNS relays are not entirely understood due to the difficulty of studying the relevant structures in the larynx in vivo (Stemple et al., 2000).

The CNS control of laryngeal movements is dependent on the integrity of innervation of laryngeal motoneurons from both the laryngeal area of the motor cortex and reflexive patterns from structures in the brainstem, including the periaqueductal grey, nucleus ambiguous, and nucleus retroambiguus (Ludlow, 2005). Speech expression may depend more on voluntary movements that are controlled by the motor cortex while emotional expression is controlled by the structures in the brainstem. Jürgens and Ehrenreich (2007) investigated the descending motorcortical pathway to the laryngeal motoneurons in the squirrel monkey. Based on the data, the researchers concluded that the primate motor cortex contains a laryngeal area that produces bilateral vocal fold adduction when stimulated; however, a direct projection from the motor cortex to laryngeal motoneurons has not been discovered. In contrast, a direct projection from the motor cortex to laryngeal motor neurons in the nucleus ambiguus and to the dorsal reticular nucleus has been found in humans (Jürgens & Ehrenreich, 2007). It is hypothesized that the projection from the motor cortex to the dorsal reticular nucleus is the primary pathway for voluntary control of vocal fold movement as indicated by the complete loss of voluntary control of laryngeal muscles in response to bilateral lesions in the laryngeal area of the motor cortex. Moreover, the projection to the nucleus ambiguous appears to elicit nonverbal emotional vocal utterances and is less robust than the projection to the dorsal reticular nucleus (Jürgens & Ehrenreich, 2007).

Substrates of Vocal Hyperfunction - Etiologies of Vocal Fold Neuropathology

Neuropathologies which affect vocal fold structure and/or movement include laryngeal paralysis, degenerative neurological diseases such as Parkinson's disease, spasmodic dysphonia, tremor, and neurological changes associated with aging (e.g., presbylaryngis). Of these conditions, paralysis, Parkinson's disease, and presbylaryngis are known to result in glottal incompetence (e.g., the inability to completely close the vocal fold from front to back during phonation) and vocal hypofunction and result in a breathy voice quality. The most common type of injury to the RLN is unilateral damage to the left or right branch, wherein efferent information is prevented from being transmitted to the larynx, resulting in dysfunction of the intrinsic muscles on the affected side of the larynx (Wilson, Driscoll, & Quinn, 1995). This results in a compromised ability to adduct the vocal fold on the affected side. The left branch of the RLN is especially vulnerable to injury during thoracic surgical procedures due to its course around the aortic arch.

Approximately one-third of RLN paralyses result from surgical trauma (Myssiorek, 2004). Disruption of neural transmission can occur from cutting, stretching, compression, thermal damage, and vascular compromise. Surgeries that have been reported to carry the most risk of nerve injury are carotid endarterectomy, anterior approaches to the cervical spine, thyroidectomy, and skull base surgery. Historically, the incidence of RLN damage was greatest for thyroid surgery; however, this incidence has decreased as skull base and anterior approach to the surgical spine procedures have been refined and become more common (Myssiorek, 2004). The risk to the nerve during thyroid surgery increases when surgeons fail to identify the specific location of the RLN during the procedure. Carotid endarterectomy often results in traction injuries, compression injuries (from clamps and vascular loops placed around carotid vessels), and interruption of blood flow to the nerve. RLN damage is the most frequent complication of anterior approach to the cervical spine procedures because the angle of insertion and length of the left branch of the RLN increases the risk of stretch-induced and compression injuries. Skull base surgery often results in transient damage to the Vagus. Despite differences in symptoms observed, damage to the cranial nerve impacts the conduction of neural messages from the Vagus to the RLN to the larynx (Myssiorek, 2004).

The SLN also is susceptible to trauma from surgery although the symptoms of SLN paralysis are less defined than those associated with RLN injury. The SLN is divided into two branches: (1) the external branch (i.e., efferent information) and (2) the internal branch (i.e., afferent information). The SLN courses medial to the internal and external carotid arteries prior to separating into its two branches. The course of the nerve results in increased vulnerability to injury during a thyroidectomy, neck dissection, cricopharyngeal myotomy, carotid endarterectomy, and supraglottic laryngectomy

(Sulica, 2004). The internal branch specifically is important for reflex responses of the intrinsic laryngeal muscles to supraglottic sensory information (Sulica, 2004). Unilateral SLN injury results in both the loss of motor function of the Cricothyroid and sensory information to the supra-glottic larynx (Wilson, Driscoll, & Quinn, 1995). Damage can also be isolated to the external SLN during thyroid surgery and results in motor dysfunction with intact supra-glottic sensory function. Surgical procedures carry a high risk of vocal fold paralysis, specifically unilateral paralysis, due to the pathways of the nerves that innervate laryngeal structures.

Neuromuscular abnormalities have also been associated with normal aging and degenerative diseases. In Presbylaryngis, changes in laryngeal muscle fibers (e.g., atrophy) in addition to neural conduction rates and nerve structure have been identified (Linville, 2001). Two degenerative neurologic diseases that can result in paresis or paralysis of the vocal cords are Parkinson's disease (PD) and Amyotrophic Lateral Sclerosis (ALS). PD is characterized by muscle rigidity, flat facial expression, and resting "pill-rolling" tremor (Woodson, 2008). Out of 2 million Americans with PD, 70% to 90% report some form of Parkinson's related dysphonia (Sewall, Jiang, & Ford, 2006). PD patients often report vocal problems that include reduced vocal loudness, a harsh or hoarse vocal quality, and breathiness (Sapir, Spielman, Ramig, Story, Fox, 2007). Specific to unilateral vocal fold paralysis, laryngeal neuropathy can result in paresis of intrinsic laryngeal muscles, resulting in decreased adduction and abduction movements of the vocal folds (Woodson, 2008). Laryngeal muscles become rigid resulting in impaired phonation and hypophonic vocal quality (Sachin, et al., 2008).

<u>Substrates of Vocal Hypofunction - Vocal Fold Adductor Paralysis</u>

The most common etiology associated with glottal incompetence is unilateral adductor vocal fold paralysis. Vocal impairments such as this which result from incomplete and irregular closure of the glottis against the pulmonary airstream and result in the loss of vocal power and quality are typically diagnosed as hypofunctional voice disorders. Two physical aspects that can cause this change in voice quality are inadequate adduction of the paralyzed vocal fold and loss of vocal fold body and tonicity, which causes bowing, flaccidity, and weakness of the affected fold (Stemple et al., 2000). The degree of dysphonia can be affected by the following factors: (a) the position of the paralyzed vocal fold, (b) the presence and degree of atrophy, and (c) the configuration of the edge of the affected fold (Colton, 1996). There are three positions of paralysis that describe the degree of incomplete glottal closure: (1) midline, (2) intermediate, and (3) abducted (Myerson, 1964). The position of the paralyzed fold depends upon the etiology and nature of denervation (Rubin, Sataloff, & Korovin, 2006). During endoscopy, the clinician should evaluate the degree of atrophy and flaccidity of the paralyzed vocal fold as well as the patient's ability to overadduct the opposite fold and constrict the larynx. According to Richardson and Bastian (2004), other considerations prior to developing a prognosis should include the position and direction of the vocal processes in relation to each other and the presence of scarring.

Untreated UVFP may be associated with changes in acoustic measurements such as voice intensity, fundamental frequency (F0), and frequency range (Schindler et al., 2008). F0 may be decreased after onset of paralysis; however, as the individual begins to compensate for decreased closure (e.g., falsetto voice) F0 may increase (Verdolini, Rosen, & Branski, 2006). The spectral slope may be altered due to changes in the structure and function of the glottis. A consistent decrease in the low frequency range, a decrease in spectral slope, and an increase in the high frequency range is observed in the spectral analysis of breathy voices (Hartl, Hans, Vaissièr, Riquet, & Brasnu, 2001).

Loss of vocal power and quality occurs because of inadequate glottal closure and loss of vocal fold tonicity and bulk (Miller, 2004). According to Eckley, Sataloff, Hawkshaw, Spiegel, and Mandel (1998) the most common presenting symptoms for people with SLN paralysis were vocal fatigue, hoarseness, impaired loudness, loss of high range, and breathiness. Perceptual characteristics of hypophonic voice usually include breathiness and hoarseness with limited pitch and loudness variation, while compensatory behaviors that result in hyperfunction of the larynx can cause diplophonia, roughness, and vocal strain (D'Alatri et al., 2007; Lombard & Steinhauer, 2007). The degree to which each symptom of hypophonic voice is present varies depending on the individual's ability and motivation to compensate for hypofunction, as well as the size, shape, and position of the paralyzed fold (Richardson & Bastian, 2004).

The prognosis for spontaneous recovery is affected by the etiology, type, and time since onset of the paralysis. According to Bryan and Quinn (1994), early return of partial function by six months post-onset has the best prognosis for recovery of laryngeal function, while paralysis persisting for a period longer than 12 months should be considered permanent. However, other researchers encourage patients to delay surgical intervention for a period of 12-18 months post onset (Tsunoda, Kikkawa, Kumada, Higo,

& Tayama, 2003). Vocal fold paralysis resulting from idiopathic etiologies has been reported to spontaneously recover in 58-75% of cases (Tsunoda, Kikkawa, Kumada, Higo, & Tayama, 2003). Greater than 50% of patients diagnosed with idiopathic right true vocal fold can expect to experience some degree of spontaneous recovery within 6 months of onset (Tsunoda, Kikkawa, Kumada, Higo, & Tayama, 2003).

Surgery should be performed immediately on patients with a large glottal gap and subsequent increased risk of aspiration (Havas, Lowinger, & Priestly, 1999). Medical and phonosurgical options include the following: (1) vocal fold injections, (2) medialization laryngoplasty, (3) laryngeal framework surgery, (4) arytenoids adduction, and (5) RLN to RLN reinnervation (Stemple, Glaze, & Klaben, 2000). The goal of medialization thyroplasty is to improve glottic closure, which decreases effort to speak, vocal fatigue, and vocal discomfort. Injection augmentation is often sufficient to manage small glottal gaps; however, an ideal long-lasting material has not been developed (Bielamowicz, 2004). Reinnervation techniques have been developed to manage the loss of vocal fold tonicity and body caused by UVFP and that may follow phonosurgery (Stemple, Glaze, & Klaben, 2000). According to Tucker (1999), reinnervation can be combined with medialization to prevent further muscle wasting and provide improved voice quality and pitch control. According to Havas, Lowinger, and Priestley (1999), management should be tailored to each individual based on the cause of paralysis, resulting disability, and the patient's vocal requirements (e.g., professional voice user).

Patients with adequate glottal closure who are not at risk for aspiration should be considered for voice therapy prior to surgical intervention (Miller, 2004). Primary goals include: development of diaphragmatic breath support for speech and improvement of glottal closure without causing supraglottic hyperactivity (D'Alatri, Gall, Rigante, Antonelli, Buldrini, & Marchese, 2007). Patients with UVFP often attempt to compensate for decreased glottal closure, which results in vocal and laryngeal hyperfunction that may impair voice quality to a greater extent. Voice therapy techniques traditionally used with vocal fold paralysis include: hard glottal attack exercises, pushing exercises, lateral digital pressure, head tilt method, half-swallow boom technique, appropriate tone focus, accent method, tongue or lip trills, and Vocal Function Exercises (Stemple, Glaze, & Klaben, 2000; Miller, 2004). Pushing exercises and hard glottal attack have decreased in use due to the potential for developing laryngeal hyperfunction (Stemple, Glaze, & Klaben, 2000). Patients who are utilizing these techniques should be closely monitored for signs of supraglottic hyperfunction (Miller, 2004).

McFarlane, Holt-Romeo, Lavorato, and Warner (1991) examined vocal quality related to treatment type by presenting voice samples of 16 adult patients with vocal fold paralysis and six adults without laryngeal pathology to three listening groups: otolaryngologists, speech-language pathologists, and lay listeners. Based on the listener groups' perception of vocal quality, the authors determined a conservative approach to treatment using voice therapy in place of surgery in all patients who are not at risk for aspiration should be the standard procedure.

Heuer et al. (1997) performed a retrospective study that included 22 male and 19 female patients with VFP. Of the patients included, 7 females and 8 males elected for surgical intervention while the remainder of the patients chose to participate in voice therapy. Of the patients treated with voice therapy, 92% of the females and 71% of the males reported improved vocal quality. Within the surgery group, 88% of both female and male participants reported an improved voice post-surgery. Colton and Casper (1996) found similar results when they studied 39 patients with unilateral VFP. Twenty-one patients chose surgical intervention, and the remaining 18 patients received voice therapy. Both groups exhibited improved post-treatment voice quality.

Schindler et al. (2008) performed a retrospective study to investigate vocal improvement after voice therapy. The study included 40 subjects with varying etiologies who were assessed by a speech language pathologist and phoniatrician for pretreatment and post treatment data. Each participant received voice therapy to improve glottal closure and avoid or decrease compensatory behaviors that could cause further vocal damage (e.g., falsetto voice, tension of the oral and pharyngeal musculature). The speech language pathologist focused on developing optimal breathing, abdominal support, and improvement of intrinsic muscle strength and agility without hyperfunctional compensation of the supraglottal structures. Humming and resonant voice techniques were taught to reach appropriate tone focus. Hard glottal attacks, pushing, and half-swallow boom were avoided because of the risk of hyperfunctional compensation. Following voice therapy, an improvement in glottal closure was observed in 14 participants. An incomplete glottal closure of the entire glottis persisted in 12 participants. The GIRBAS was administered pre- and post-therapy and a statistically significant reduction in severity was found for all the parameters following voice therapy. Based on the results of this study, researchers concluded that voice therapy can significantly improve voice quality in patients with unilateral VFP; therefore, behavioral intervention may be included in a patient's management plan prior to surgical intervention.

UVFP may be assessed using a multidimensional set of basic measurements including: endoscopic/stroboscopic examination, aerodynamics, acoustics, a subjective rating by the patient (e.g., Voice Handicap Index), and a perceptual evaluation (Schindler et al., 2008). Following the onset of UVFP, patients may experience shortness of breath during speech due to glottal incompetence, increased breathiness or roughness, effort to speak, and higher F0 (Eckley, Sataloff, Hawkshaw, Spiegel, & Mandel, 1998), which can be verified by perceptual evaluation. A consistent decrease in the low frequency range, a decrease in spectral slope, and an increase in the high frequency range is observed in the spectral analysis of breathy voices (Hartl, Hans, Vaissièr, Riquet, & Brasnu, 2001). Indicators of recovery may include changes in the patient's and listener's perception of their speech, general increase in glottal closure, and improved quality of life as reported by the patient. The set of measurements listed above serve as an assessment protocol as well as indicators of recovery following therapy or a period of spontaneous recovery (Schindler et al., 2008).

Traditional methods of diagnosing and determining improvement of UVFP have been contingent on visual examination of the gross and fine movement of the vocal folds as well as perceptual changes (Patel & Parsram, 2005). Acoustics can provide objective evaluation of the components commonly observed in hypofunctional voice disorders (e.g., increased jitter and shimmer). However, analysis of UVFP has been limited to prolonged vowels due to the lack of software capable of analyzing the complexity of time- and spectral-based measures of the vocal signal during connected speech reliably (Patel & Parsram, 2005).

CHAPTER III

STATEMENT OF PURPOSE

The purpose of this study is to determine if differences exist in selected acoustic measurements of voice production in individuals with hypofunctional voices compared to those with normal voices. A second purpose of this study is to determine if selected acoustic measurements of voice production in hypofunctional and normal speakers are influenced by the speaking context, when comparing sustained vowel production to connected speech. Specifically, the following questions will be addressed:

1) Will acoustic measurements of Cepstral peak differ between hypofunctional and normal speakers during sustained vowel and connected speech production?
2) Will acoustic measurements of the standard deviation in the Cepstral peak differ between hypofunctional and normal speakers during sustained vowel and connected speech production?
3) Will acoustic measurements of the discrete Fourier transform ratio differ between hypofunctional and normal speakers during sustained vowel and connected speech production?
4) Will acoustic measurements of the standard deviation in the discrete Fourier transform ratio differ between hypofunctional and normal speakers during sustained vowel and connected speech production?

CHAPTER IV

METHODOLOGY

Participants

Two types of participants were recruited for this study: (1) Speaking participants and (2) Perceptual judges.

Speaking participants consisted of 14 total individuals. 7 participants had a confirmed laryngeal pathology associated with a hypofunctional voice disorder and presented with a confirmed breathy voice quality (disordered speakers), and 7 individuals had no history of a voice disorder (control speakers). Disordered speakers were recruited from the local community via treatment seeking populations (e.g., those seeking evaluation and/or treatment from a speech-language pathologist for a self-perceived voice problem, n = 1). In addition, existing recordings of disordered speakers from the KayPentax Disordered Voice Database (a commercially available recording of a variety of disordered voices) were used (n = 6). Inclusion criteria for all recorded disordered voices were (1) diagnosed laryngeal pathology by an otolaryngologist, (2) an etiological diagnosis of unilateral recurrent laryngeal nerve paresis/paralysis, Parkinson's disease, or presbylaryngis, and (3) predominantly moderate-to-severe breathy voice quality which was confirmed by 8/10 judges via auditory-perceptual analysis by perceptual judges. Inclusion criteria for control speakers was (1) no history of diagnosed voice disorder (2) no history of voice therapy, (3) no history of neurological impairment, (4) no history of habitual smoking, and (5) no current voice problems. All recordings of normal speakers (n = 7) were taken from the KayPentax Disordered Voice Database.

Perceptual judges consisted of 10 college students recruited from the graduate student population at Texas Christian University. The 10 judges rated voice quality type (breathy or normal) and severity of recorded signals in order to determine inclusion eligibility of those signals (see inclusion criteria #3 above for disordered voices). Inclusion criteria for perceptual judges were (1) passed hearing screening at 25dBSPL at 500Hz, 1KHz, 2Khz, and 4KHz, (2) no history of hearing disorder, (3) no current reported hearing problems, and (4) no history of neurological disorder.

Instrumentation

Data for this study were computed from recordings of vocal productions produced by individuals in the disordered and control groups. All recordings were digitized using hardware and/or software produced by KayPentax (Lincoln Park, NJ), which included the Computerized Speech Lab (CSL) and Sonaspeech. All recordings were acquired with participants wearing a head-mounted microphone positioned approximately 3cm off the corner of the mouth.

Two custom software applications, developed by Dr. Shaheen Awan at Bloomsburg University were used for initial ratings by perceptual judges and to analyze the recorded vocal productions for data collection, respectively. Anchors software was used for auditory-perceptual measurements of breathiness and vocal severity (that is, the perceptual judges used this software to rate voices) for stimulus inclusion. The validity of the software for use in auditory-perceptual ratings has been established in previous studies (Awan & Roy, 2005; Awan & Roy, 2006; Awan & Lawson, in press, Watts, et al., 2008). A complete description of the auditory-perceptual rating task is described further below. Spectral-Cepstral Analysis of Running Speech (SCARS) software, which utilizes digital signal processing to analyze multiple spectral-based measures of the recorded vocal signals, was used. For perceptual training sessions and experimental data collection sessions, a standard desktop computer wired to circumaural headphones was used to present auditory stimuli to the perceptual judges at a self-reported comfortable listening level, as established by individual judges

Procedures

Prior to initiation of the study, the general purpose of the study were described to the participants (e.g., to compare the voices of speakers with hypofunctional voice disorders to those of individuals without disorders). The participants were given the consent form to read, provided an opportunity to ask questions, and then asked to sign the consent form.

Vocal Recordings: Speaking participants who were recruited locally were recorded in a quiet room at the D. Wayne Tidwell Voice & Swallowing Center located in Baylor All-Saints Hospital. The head-mounted microphone was worn and placed with the microphone head off-center at the right mouth corner, with a mouth-to-microphone distance of approximately 3-8cm. The microphone had a direct line input to the KayPentax digital recording system. This procedure matched that used to acquire recordings on the KayPentax Voice Disordered Database. Speaking participants were asked to produce two types of vocal productions: (1) a sustained vowel, /a/, produced for a 3 second duration at a self-reported comfortable pitch and loudness, held as steady as possible, and (2) a reading of a standard paragraph, "The Rainbow Passage" (Fairbanks, 1960), produced at a self-reported comfortable pitch and loudness.

A total of 8 disordered speaking participants were recruited locally, and pre-recorded signals from an additional 6 disordered speaking participants were taken from the KayPentax Voice Disordered Database. Pre-recorded disordered speaking participants were selected from the database based on the first 6 who matched the inclusion criteria of hypofunctional etiology (paralysis, presbylaryngis, or Parkinson's disease). All recordings of disordered speakers were applied to the pre-experimental judgment of breathiness task (described below), which yielded a total of 7 recordings that matched inclusion criteria. A total of 14 recordings of control speakers were selected from the pre-recorded database. All recordings of control speakers were applied to the pre-experimental judgment of breathiness task to ensure that no control speaker was judged as breathy.

Pre-experimental judgments of breathiness: Prior to experimental data collection, ratings of breathiness or normal voice quality were confirmed in the recorded hypofunctional voices by 10 perceptual judges. An explanation of the study procedures and associated risks was given to the participants and the consent form was signed. Prior to the judgment task, judges participated in a 20-minute training session. During the training session, instructions regarding use of the Anchors software and a review of the operational definitions for voice quality types (normal and breathy) and severity (normal, mild, moderate, and severe) were provided. The perceptual judgment made by the judges was a two-item forced choice task. Voice quality types were defined for each judge as they appeared in the Anchors software: <u>Normal</u> - “The voice does not differ substantially from your expectations in terms of parameters such as quality”, and <u>Breathy</u> – “Commonly perceived as a whispery or airy voice; associated with hypoadduction (Awan & Roy, 2005). Each judge then listened to recorded voice samples which were representative for the range of voice types (normal and breathy) and severities (normal, mild deviation from normal, severe deviation from normal) possible for individuals sustaining the vowel /a/. These representative stimuli have been validated in three published studies (Awan & Roy, 2005; Awan & Roy, 2006, Watts, et al., 2008).

Once the training session was complete, judges then listened to the recorded sustained /a/ vowels of the hypofunctional and normal voices, one stimulus at a time, while seated and wearing the circumaural headphones. The Anchors software allowed the user to select samples for playback, in an order that was randomized for each judge. After listening to each stimulus, the judges were asked to make ratings regarding the quality and severity of that voice. For voice quality ratings, judges could choose “normal” or “breathy. For severity ratings, judges placed a curser along a 100mm visual analog scale, anchored with labels of “mild”, “moderate”, and “severe”. Judges were allowed to replay each sample as many times as necessary before rating it. In addition, judges were allowed to compare each voice sample with two pre-selected “normal” voice samples (i.e., normal male and normal female) used in the training session. These samples served as an external auditory standard (e.g., a perceptual anchor). The use of perceptual anchors in this way have been shown to be beneficial to auditory-perceptual judgments of voice quality and severity as they help to reduce listener-related variability (Awan & Lawson, in press; Awan & Roy, 2006; Eadie & Baylor, 2006). The same two normal perceptual anchors were used for all judgments. In addition to the perceptual anchor, judges also had access to the operational definitions for each rating category, as reviewed in the training session. For purposes of measuring intra-judge reliability, every judge listened to and rated each voice sample twice (first rating and second rating).

For hypofunctional voices to be included in the study, it was a requirement for them to be rated as “breathy” for voice quality by 80% (8/10) of the perceptual judges. For control voices to be included in the study, it was a requirement for them to be rated as “normal” by 100% (10/10) of the perceptual judges. Any voices failing to meet these criteria were excluded from the study.

Digital Signal Analysis: The SCARS software was utilized to analyze all recorded signals. The SCARS software uses built-in algorithms to automatically compute the following measures: (A) sustained vowels and (B) connected speech. Detailed descriptions of the analyses produced by this software have been previously published (Awan & Roy, 2006; Awan & Roy, 2009). For sustained vowels, the middle 1-second steady state portion of the vowel (e.g., from 1.0 to 2.0 second marks of the recordings, which represents the middle of the 3-second recorded sample) was isolated using the computer and applied to signal analysis processing using the SCARS software. For connected speech, the 2nd sentence of the rainbow passage was isolated using the computer and applied to signal analysis processing using the SCARS software. The specific measurements collected from each vowel and each connected speech sample using the SCARS software will be as follows:

(1) Cepstral peak-expected peak ratio (CPP/EXP - ratio of the amplitude of the cepstral peak (CPP) to the expected (EXP) amplitude of the cepstral peak, as derived via linear regression). This measure reflects the degree of regularity or periodicity in the voice signal. Higher values reflect greater periodicity.

(2) Standard Deviation of the cepstral peak-expected peak ratio (CPP/EXPSD). This measure reflects the variability in periodicity for each recorded signal. Higher values reflect greater variability.

(3) Discrete Fourier Transform Ratio (DFTR – the ratio of low [<4000Hz] versus high [>4000Hz] frequency energy. This measure reflects the relationship between energy at high frequencies in the voice signal compared to energy at low frequencies. This measure is commonly referred to as spectral tilt. Higher values reflect greater low frequency energy, which is commonly seen in normal speakers.

(4) Standard Deviation of the discrete Fourier transform ratio (SD DFTR). This measure reflects the variability in DFTR for each recorded signal. Higher values reflect greater variability.

Until recently, these measures in connected speech lacked internal validity due to limitations in the algorithms of the software used for analysis. The SCARS software has been developed specifically to overcome these validity concerns for measuring spectral-based parameters of the vocal signal.

Analyses

Four separate dependent variables (see above) were measured using the SCARS software. The study design utilized one between-subjects independent variable (group – hypofunctional vs. normal) and one within-subjects independent variable (speaking condition – sustained vowel vs. connected speech). Two separate analyses were utilized to investigate the effect of group on the dependent variables when (1) reading the Rainbow passage and (2) sustaining a vowel. Multivariate Analyses of Variance (MANOVA) were applied to the data sets for these two speaking conditions. The main effect for both MANOVA's was group, and the level of significance was set at .05 for both statistical tests.

Reliability

100% percent of all recorded samples were re-measured using the SCARS software. A Pearson product-moment correlation coefficient was used to compare these measures to the original measures to assess the degree of relationship. The resulting correlation coefficient was used to assess the reliability of the initial measurements. The correlation coefficient was r=1.0, which demonstrated a perfect correlation and indicated 100% measurement reliability with the software.

CHAPTER V

RESULTS

The design of this study utilized two independent variables, one between-subjects (group) and one within-subjects (speaking condition) and four dependent variables (Cepstral peak-expected peak ratio, standard deviation of the Cepstral peak-expected peak ratio, discrete Fourier transform ratio, & standard deviation of the discrete Fourier transform ratio). A two-way MANOVA was performed for both independent variables (group and speaking condition). For the MANOVA and any subsequent post-hoc testing, an alpha level of .05 was used as the criterion for statistical significance. Fourteen participants were included in this study. Table 1 lists the participants' age, group (hypo vs. normal), and disorder. Descriptive statistics summarizing the resulting raw data is summarized in Table 2. Trends in this data were as follow:

<u>Cepstral peak-expected peak ratio (CPP/EXP):</u> Figure 1 illustrates group performance on the measure of CPP/EXPSD in the two speaking conditions Comparison of CPP/EXP between groups manifested noticeable differences. For the sustained vowel, the hypofunctional group's mean was 2.57% with a standard deviation of 1.99 while the normal group's mean was 13.91% with a standard deviation of 2.42. During connected speech, the mean of the hypofunctional group was 2.52% with a standard deviation of 1.74 while the normal group's mean was 6.47% with a standard deviation of .74.

<u>Standard deviation of the cepstral peak-expected peak ratio (CPP/EXPSD):</u> Figure 2 illustrates group performance on the measure of CPP/EXPSD in the two speaking conditions. Analysis revealed the hypofunctional group had an average CPP/EXPSD

during a sustained vowel of .48% with a standard deviation of .38 while the normal group had an average of .54% with a standard deviation of .25. The CPP/EXPSD for connected speech had greater variability between groups. The mean for the hypofunctional group was 1.54% with a standard deviation of 0.93 while the normal group had an average of 3.45% with a standard deviation of 0.46.

Table 1. Demographic Information

Subject	Group	Age	Diagnosis
1	Hypo	42	UVFP
2	Hypo	59	UVFP
3	Hypo	52	UVFP
4	Hypo	42	UVFP
5	Hypo	49	UVFP
6	Hypo	43	UVFP
7	Hypo	42	UVFP
8	Normal	52	N/A
9	Normal	31	N/A
10	Normal	37	N/A
11	Normal	40	N/A
12	Normal	40	N/A
13	Normal	39	N/A
14	Normal	40	N/A

Table 2. Descriptive statistics (mean & standard deviation) summarizing the raw data for each dependent variable (DV) and independent variable (group and condition).

DV (measure)	***IV (Group)***	***IV (Condition)***	***Mean***	***Std. Deviation***
CPP/EXP	Hypofunctional	Rainbow	2.52	1.74
		Ah	2.57	1.99
	Normal	Rainbow	6.47	0.73
		Ah	13.91	2.42
CPP/EXPSD	Hypofunctional	Rainbow	1.54	0.93
		Ah	0.48	0.38
	Normal	Rainbow	3.45	0.46
		Ah	0.54	0.25
DFTR	Hypofunctional	Rainbow	19.65	6.69
		Ah	19.57	7.98
	Normal	Rainbow	30.21	1.19
		Ah	32.92	5.84
DFTRSD	Hypofunctional	Rainbow	8.99	3.00
		Ah	1.06	0.32
	Normal	Rainbow	10.94	1.12
		Ah	1.39	0.87

Discrete Fourier transform ratio (DFTR): Figure 3 illustrates group performance on the measure of DFTR in the two speaking conditions. Both groups in the two speaking contexts had large standard deviations for this particular measure with the exception of the normal group during connected speech. The hypofunctional group was measured to have a lower DFTR during a sustained vowel with a mean of 19.57% and standard deviation of 7.98 while the normal group had an average of 32.92% with a standard deviation of 5.84. During the analysis of connected speech, the hypofunctional group had an average of 19.65% with a standard deviation of 6.69 while the normal group had a mean of 30.21% with a standard deviation of 1.19.

Standard deviation of the discrete Fourier transform ratio (DFTRSD): Figure 4 illustrates group performance on the measure of DFTRSD in the two speaking conditions. Both groups manifested similar measures of DFTRSD in both speaking contexts although a larger difference was observed during connected speech. Analysis of a sustained vowel manifested an average of 1.06% with a standard deviation of .32 for the hypofunctional group and an average of 1.39% with a standard deviation of .86 for the normal group. For the connected speech measurements, the hypofunctional group had an average of 8.99% with a standard deviation of 3.00 and the normal group had an average of 10.95% with a standard deviation of 1.12.

Figure 1. Averaged group (Normal, Hypofunctional) performance on the measure of CPP/EXP in the two speaking conditions (/a/, Rainbow).

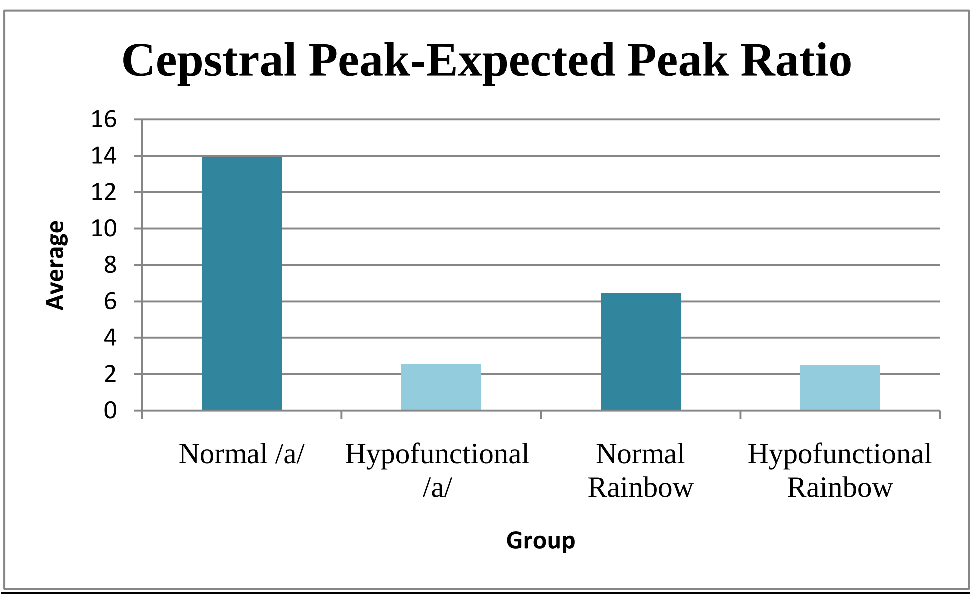


Figure 2. Averaged group (Normal, Hypofunctional) performance on the measure of CPP/EXPSD in the two speaking conditions (/a/, Rainbow).

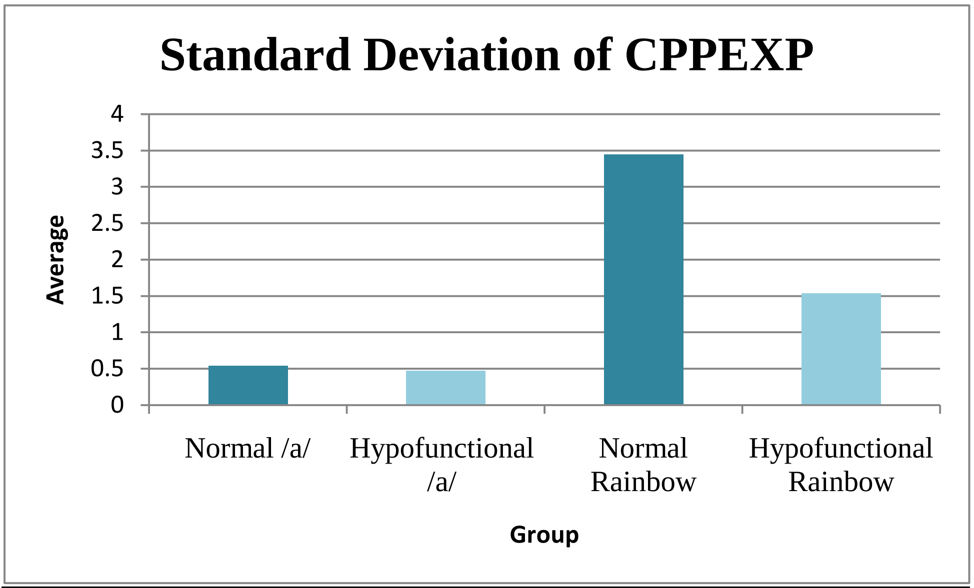


Figure 3. Averaged group (Normal, Hypofunctional) performance on the measure of DFTR in the two speaking conditions (/a/, Rainbow).

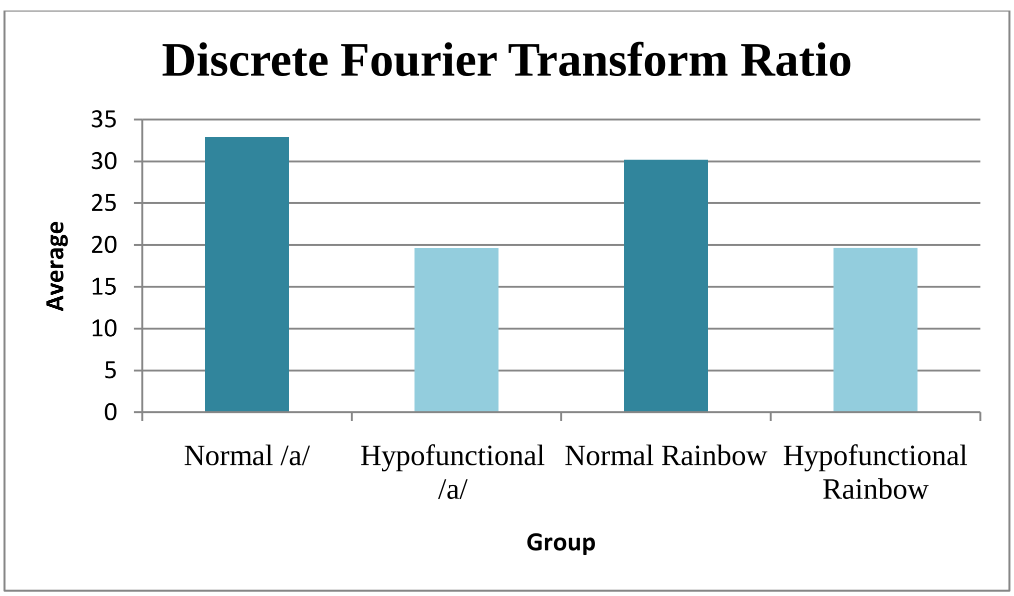
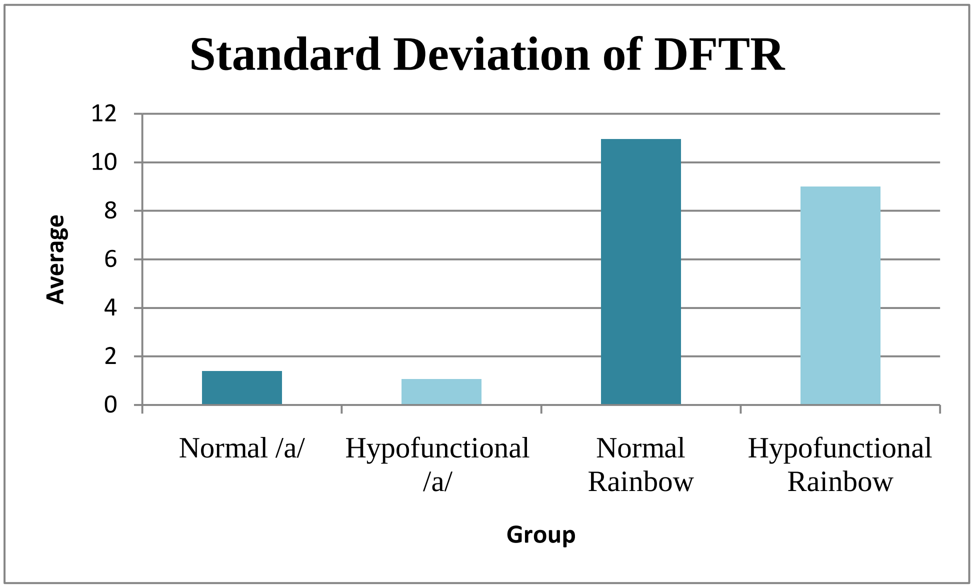


Figure 4. Averaged group (Normal, Hypofunctional) performance on the measure of DFTRSD in the two speaking conditions (/a/, Rainbow).



A two-way (group x condition) Multivariate Analysis of Variance (MANOVA) was completed to determine the effect of the independent variables on the dependent variables (CPP/EXP, CPP/EXPSD, DFTR, and DFTRSD) and any potential interaction effects. The MANOVA results are illustrated in Table 3. The results of the MANOVA revealed main effects for group (Pillai's Trace [1, 21] = .838, p <.001)and condition (Pillai's Trace [1,21] = .928), p <.001), as well as a significant interaction between the two (Pillai's Trace [1,21] = .787, p<.001). Significant main effects for group were observed on CPP/EXP (F[1,24] = 122.09, p<.001), CPP/EXPSD (F[1,24] = 21.09, p<.001), and DFTR (F[1,24] = 27.75, p<.001). However, no significant difference were manifested for DFTRSD (F[1,24] = 3.30, p=.081). Significant main effects for condition were observed on CPP/EXP (F[1,24] = 29.29, p<.001), CPP/EXPSD (F[1,24] = 85.73, p<.001), and DFTRSD (F[1,24] = 192.25, p<.001). No significant difference on DFTR

(F[1,24] = 0.33, p=.569) was observed for speaking condition. A significant interaction effect was also present for CPP/EXP (F = 28.48, p<.001) and CPP/EXPSD (F = 18.52, p<.001).

Table 3. MANOVA table showing probability values for each main effect in the study.

Source	***Dependent Variable***	***Sig. (p)***
Group	CPP/EXP	<0.01
	CPP/EXPSD	<0.01
	DFTR	<0.01
	DFTRSD	=0.08
Condition	CPP/EXP	<0.01
	CPP/EXPSD	<0.01
	DFTR	=0.56
	DFTRSD	<0.01
Group*Condition	CPP/EXP	<0.01
	CPP/EXPSD	<0.01
	DFTR	=0.54
	DFTRSD	=0.21

The MANOVA results for DFTR revealed that, irrespective of speaking condition (which did not have an effect on this measure), measures of DFTR were always significantly higher in the normal group. In addition, the MANOVA results for DFTRSD showed that, irrespective of group (which did not have an effect on this measure), measures of DFTRSD were always significantly higher in the speaking (Rainbow condition). Because of the significant interaction between group and condition on the CPP/EXP measure and the CPP/EXPSD measure, post-hoc tests utilizing independent t-tests were conducted on those variables to determine the locations of significant differences. Four separate t-tests were conducted (Hypo vs. normal on CPP/EXP during

Rainbow, Hypo vs. normal on CPP/EXP during vowel, Hypo vs. normal on CPP/EXPSD during Rainbow, Hypo vs. normal on CPP/EXPSD during vowel). The descriptive statistics for CPP/EXP and CPP/EXPSD are reported in Table 2.

Results from the independent t-tests revealed significant differences between groups for both CPP/EXP ($p<.001$) and CPP/EXPSD ($p<.001$) while reading the Rainbow passage, so that both measures were significantly higher in the normal group. However, although a t-test for the sustained vowel revealed significant differences between groups for CPP/EXP ($p<.001$), there was no significant group difference on CPP/EXPSD ($p=.727$). This indicated that the normal group manifested significantly higher measures of CPP/EXPSD when reading the Rainbow passage, but there was no measureable group difference on this variable when sustaining a vowel.

CHAPTER VI

DISCUSSION

The purpose of this study was to determine if differences exist in selected acoustic measurements of voice production in individuals with hypofunctional voices compared to those with normal voices. A second purpose of this study was to determine if selected acoustic measurements of voice production in hypofunctional and normal speakers are influenced by the speaking context, when comparing sustained vowel production to connected speech. The results of the study revealed significant differences between hypofunctional and normal speakers in measurements of cepstral peak, discrete Fourier transform ratio, and the standard deviation of the discrete Fourier transform ratio during connected speech and sustained vowel production; however, the standard deviation of Cepstral peak in hypofunctional voices during sustained vowel production does not differ significantly from normal speakers.

Measurement of the cepstral peak is an indication of the periodicity of the spectrum of the voice. Analysis of a normal voice produces a more dominant (higher) Cepstral peak than that of a disordered voice (Awan & Roy, 2006). The CPP/EXP ratio is most likely affected by the increase in high-frequency noise that is present in hypofunctional voices. Due to this increase in noise, the ratio of Cepstral peak prominence to expected peak prominence is reduced (Awan & Roy, 2005). This pattern was also observed in the results of this study. The consistently lower Cepstral peaks observed in the hypofunctional group across both speaking conditions appears to discriminate between hypofunctional and normal voices. Previous studies conducted by

Awan and Roy (2006) and Heman-Ackah et al. (2003) also indicated that measures CPP/EXP correlate well with dysphonia. The greater degree of variation in measurements of the CPP/EXPSD measurement during connected speech when compared to sustained vowel production fulfill reasonable expectations that the variability of Cepstral peak is limited during sustained vowel production due to the consistency of the voice signal, particularly in normal voices.

Discrete Fourier transform ratio (DFTR), the ratio of low versus high spectral energy, was observed by Awan & Roy (2006) to contribute to predict the severity of breathy voice, which has been characterized by spectral noise in the signal particularly at high frequencies. Significant differences manifested in comparison of group (normal vs. hypofunctional) signify that DFTR may be used to classify a voice as breathy and can distinguish those with hypofunctional voice pathologies from normal voices. However, the degree of variability between normal and hypofunctional voices may be limited and the use of this measurement to analyze production of sustained vowels may not be reliable.

Current methods of acoustic analysis of the voice signal have not been as reliable in the ability to quantify voice type and severity. CPP, an acoustic measure used in the current study to differentiate between hypofunctional and normal voices, has been researched to determine the reliability of this measure for diagnosis of voice type (Heman-Ackah et al., 2003). Acoustic analyses of dysphonic voices revealed that CPP seems to be able to discriminate between normal and dysphonic voices (Awan & Roy, 2005). Specific to breathy voices, the increase of high frequency noise in the voice signal resulted in reduced DFTR and a reduced ratio of the CPP/EXP. According to Heman-Ackah et al. (2003), CPP is more reliable in predicting dysphonic voices than jitter, shimmer, and noise-to-harmonic ratio (time-based measures) because of better specificity and sensitivity. Awan & Roy (2006) found that DFTR may be useful in predicting the severity of hypofunctional voices.

CHAPTER VII

CONCLUSIONS

Four types of acoustic analyses (CPP/EXP, CPP/EXPSD, DFTR, & DFTRSD) were performed to analyze differences in vocal recordings based on group (hypofunctional vs. normal) and speaking context (connected speech vs. prolonged vowel). Within the context of group, CPP/EXP & CPP/EXPSD were significantly higher in normal speakers. However, an interaction effect was observed on both CPP/EXP & CPP/EXPSD. Results of independent t-tests revealed a measureable difference in CPP/EXP between groups during both speaking tasks, but for CPP/EXPSD there was only a group difference when reading the Rainbow passage. Measures of DFTR were statistically significant regardless of speaking condition, so that this measure was always lower in the hypofunctional group. These results indicate that selected cepstral analysis of vocal recordings for prolonged vowel and connected speech can differentiate between hypofunctional and normal voices.

Further research of the ability to identify dysphonic voices using cepstral analysis of the voice signal should include a larger sample size and a variety of type and severity of dysphonia (rough and hoarse). Since most acoustic research up to this point focuses on prolonged vowels, this speaking context should be included in order to compare the results of analyses of both prolonged vowel and connected speech. Care should also be taken in training perceptual judges to identify dysphonic voices more consistently.

CHAPTER VIII

STUDY LIMITATIONS

Several limitations in the methodology of this study and suggestions for further research should be noted.

1. This study assessed characteristics of normal and *hypofunctional* voices only. It is unclear whether the acoustic variables found to be statistically significant in differentiating between hypofunctional and normal voices (CPP/EXP, CPP/EXPSD, & DFTR) would have the similar results when used to assess other types of dysphonia (e.g., roughness, hoarseness) in comparison to normal voices. It is also unclear if the acoustic analyses found to be significant in the context of speaking condition (CPP/EXP, CPP/EXPSD, & DFTRSD) would remain the same for other types of dysphonia.
2. Perceptual judges appeared to have difficulty discriminating between normal and hypofunctional voices during a prolonged vowel. Although training was provided to each judge prior to participating in the listening task, the time constraint on this project prevented training from being sufficient enough for the judges to identify the breathiness which characterizes a hypofunctional voice. It was observed that judges had difficulty making a distinction between "breathy" and "hoarse," which resulted in the variability in the categorization of voice type. This resulted in only 7 speakers being rated as "breathy" by 8/10 judges. Further research should include more intensive training for voice

quality and severity in order to ensure that a larger number of samples can be included for analysis.

REFERENCES

Abboud, B., Tabchy, B., Jambart, S., Hamad, W.A., & Farrah, P. (1999). Benign disease of the thyroid gland and vocal fold paralysis. *The Journal of Laryngology and Otology, 113,* 473-474.

Awan, N.S., Roy, N. (2005). Acoustic prediction of voice type in women with functional dysphonia. *Journal of Voice, 19(2),* 268-283.

Awan, N.S., Roy, N. (2006). Toward the development of an objective index of dysphonia severity: A four-factor acoustic model. *Clinical Linguistics & Phonetics, 20(1),* 35-49.

Awan, N.S., Roy, N. (2009). Outcomes measurement in voice disorders: Application of an acoustic index of dysphonia severity. *Journal of Speech, Language, and Hearing Research, 52(2),* 482-500.

Bando, H., Nishio, T., Bamba, H., Uno, T., Hisa, Y. (2006). Vocal fold paralysis as a sign of chest diseases: A 15-year retrospective study. *World Journal of Surgery, 30(3),* 293-298.

Bielamowicz, S. (2004). Perspectives on medialization laryngoplasty. *Otolaryngologic Clinics of North America, 37 (1),* 139-160.

Colton, R.H. (1996). Physiology of voice quality. In W.S. Brown, Jr., B.P. Vinson, & M.A. Crary (Eds.). Organic voice disorders assessment and treatment, 49-88 San Diego: Singular Publishing Group, Inc.

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

D'Alatri, L., Galla, S., Rigante, M., Antonelli, O., Buldrini, S., Marchese, M.R. (2007). Role of early voice therapy in patients affected by unilateral vocal fold paralysis. *The Journal of Laryngology & Otology, 122,* 936-941.

Department of Otoloaryngology, UTMB, Grand Rounds. (1994). *Vocal cord paralysis.* Galveston, TX: Bryan, M.D., & Quinn, F.B. (1994).

Department of Otoloaryngology, UTMB, Grand Rounds. (1995). *Vocal cord paralysis.* Galveston, TX: Wilson, D., Driscoll, B., & Quinn, F.B.

Eclkey, C.A., Sataloff, R.T., Hawkshaw, M., Spiegel, J.R., & Mandel, S. (1998). Voice range in superior laryngeal nerve paresis and paralysis. *Journal of Voice, 12(3),* 340-348.

Fairbanks, G. (1960). Voice and articulation drillbook. (2nd ed.). New York: Harper & Row, 124-139.

Hartl, D.M., Hans, S., Vaissière, J., Riquet, M., Brasnu, D.F. (2001). Objective voice quality analysis before and after onset of unilateral vocal fold paralysis. *Journal of Voice, 15 (3),* 351-361.

Havas, T., Lowinger, D., Priestley, J. (1999). Unilateral vocal fold paralysis: Causes, options, and outcomes. *Australia and New Zealand Journal of Surgery, 69,* 509-513.

Heman-Ackah, Y.D., Heuer, R.J., Michael, D.D., Ostrowski, R., Horman, M., Baroody, M.M., Hillenbrand, J., Sataloff, R.T. (2003). Cepstral peak prominence: A more reliable measure of dysphonia. *Annals of Otology, Rhinology, & Laryngology,112,* 324-333.

Heuer, R.J., Sataloff, R.T., Emerich, K., Rulnick, R., Baroody, M., Spiegel, J.R., Durson, G., & Butler, J. (1997). Unilateral recurrent laryngeal nerve paralysis: The importance of "preoperative" voice therapy. *Journal of Voice, 11(1),* 88-94.

Hillenbrand, J., Houde, R.A. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *Journal of Speech and Hearing Research, 39,* 311-321.

Jürgens, U., & Ehrenreich, L. (2007). The descending motorcortical pathway to the laryngeal motoneurons in the squirrel monkey. *Brain Research*, *1148*, 90-95.

Linville, S.E. (2001). Vocal aging. San Diego, CA: Singular/Thomson Learning.

Lombard, L.E., & Steinhauer, K.M. (2007). A novel treatment for hypophonic voice: Twang therapy. *Journal of Voice, 21(3),* 294-299.

Ludlow, C.L. (2005). Central nervous system control of the laryngeal muscles in humans. *Respiratory Physiology & Neurobiology* 147 (2), 205-222.

McFarlane, S.C., Holt-Romeo, T.L., Lavorato, A.S., & Warner, L. (1991). Unilateral vocal fold paralysis: Perceived vocal quality following three methods of treatment. *American Journal of Speech Language Pathology,1*, 45-48.

Miller, S. (2004). Voice therapy for vocal fold paralysis. *Otolaryngologic Clinics of North America, 37 (1),* 105-119.

Myerson, M.C. (1964). *The human larynx.* Springfield, IL: Charles C Thomas Publisher

Myssiorek, D. (2004). Recurrent laryngeal nerve paralysis: Anatomy and etiology. *Otolaryngologic Clinics of North America*, 37(1), 25-44.

Negus, V.E. (1949). The comparative anatomy and physiology of the larynx. New York, NY: Hafner Publishing Company, Inc.

Patel, R., Parsram, K.S. (2005). Acoustic analysis of subjects with vocal cord paralysis. *Indian Journal of Otolaryngology and Head and Neck Surgery, 57(1),* 48-51.

Richardson, B.E., & Bastian, R.W. (2004). Clinical evaluation of vocal fold paralysis. *Otolaryngologic Clinics of North America,37(1),* 45-58.

Robert, D., Pouget, J., Giovanni, A., Azulay, J., Triglia, J. (1999). Quantitative voice analysis in the assessment of bulbar involvement in amyotrophic lateral sclerosis. *Acta Oto-Laryngologica, 119 (6),* 724-731.

Rubin, J., Sataloff, R.T., & Korovin, G. (2006). Diagnosis and treatment of voice disorders (2nd ed.). San Diego: Plural Publishing.

Sachin, S., Shukla, G., Goyal, V., Singh, S., Aggarwal, V., Gureshkuma, Behari, M. (2008). Clinical speech impairment in Parkinson's disease, progressive supranuclear palsy, and multiple system atrophy. *Neurology India, 56 (2),* 122-126.

Sapir, S., Spielman, J.L., Ramig, L.O., Story, B.H., & Fox, C. (2007). Effects of intensive voice treatment (the Lee SilvermanVoice Treatment [LSVT]) on vowel articulation in dysarthric individuals with idiopathic parkinson's disease: Acoustic and perceptual findings. *Journal of Speech, Language, & Hearing Research, 50(4),* 899-912.

Sewall, G.K., Jiang, J., & Ford, C.N. (2006). Clinical evaluation of Parkinson's-related dysphonia. *Laryngoscope, 116(10),* 1740-1744.

Schindler, A., Bottero, A., Capaccio, P., Ginocchio, D., Adorni, F., & Ottaviani, F. (2008). Vocal improvement after voice therapy in unilateral vocal fold paralysis. *Journal of Voice, 22(1),* 113-118.

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

Sulica, L. (2004). The superior laryngeal nerve: Function and dysfunction. *Otolaryngologic Clinics of North America*, 37(1), 183-201.

Tsunoda, K., Kikkawa, Y.S., Kumada, M., Higo, R., Tayama, N. (2003). Hoarseness caused by unilateral vocal fold paralysis: How long should one delay phonosurgery? *Acta Oto-Laryngologica, 123*, 555-556.

Tucker, H.M. (1999). Long-term preservation of voice improvement following surgical medialization and reinnervation for unilateral vocal fold paralysis. *Journal of Voice 13(2)*, 251-256.

Verdolini, K., Rosen, C.A., Branski, R.C . (2006). Classification manual for voice disorders – 1. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Woodson, G. (2008). Management of neurologic disorders of the larynx. *Annals of Otology, Rhinology, & Laryngology*, *117(5)*, 317-326.

VITA

Emily Lambert was born on July 21, 1987 in Houston, Texas. She received a Bachelor of Science degree with a major in Communication Sciences and Disorders from the University of Texas in 2008. While at the University of Texas, Emily participated in the Austin Center for Stuttering Intervention and Research with both preschool children and a young adult.

In August of 2008, she enrolled for graduate study at Texas Christian University. She is a member of the National Student Speech-Language-Hearing Association and Special Interest Division 3, Voice and Voice Disorders.

She is married to Jeffrey Lambert of Waco, Texas.

ABSTRACT

A COMPARISON OF SUSTAINED VOWEL AND CONNECTED SPEECH PRODUCTION IN HYPOFUNCTIONAL AND NORMAL VOICES

by Emily A. Lambert, M.S., 2010

Department of Communication Disorders

Texas Christian University

Thesis Advisor: Chris Watts, Ph.D.

The purpose of this study was to: 1) to determine if differences exist in acoustic measurements of voice production in individuals with hypofunctional and normal voices and 2) to determine if measurements of voice production are influenced by the speaking context, when comparing sustained vowel production to connected speech. Disordered speaking participants were recruited from Baylor All Saints in Fort Worth, TX in addition to existing recordings. Perceptual judges were recruited from Texas Christian University (TCU) to rate voice quality (breathy vs. normal) of prolonged vowels. A MANOVA revealed main effects on group and condition, and a significant interaction between the two. Significant main effects for group (hypo. vs. normal) were observed on Cepstral peak/expected peak (CPP/EXP) ratio ($p<.001$), the standard deviation of Cepstral peak/expected peak ratio (CPP/EXPSD) ($p<.001$), and discrete Fourier transform ratio (DFTR) ($p<.001$). Significant main effects for condition were observed on CPP/EXP ($p<.001$), CPP/EXPSD ($p<.001$), and DFTRSD($p<.001$). A significant interaction effect was found for CPP/EXP ($p<.001$) and CPP/EXPSD ($p<.001$). These results indicate that cepstral analysis can differentiate between hypofunctional and normal voices. The consistency of significant differences observed in this study suggests that acoustic analyses are capable of discriminating between normal and hypofunctional voices reliably.