A COMPARISON OF ACOUSTIC AND AERODYNAMIC MEASUREMENTS OF LARYNGEAL FUNCTION USING LOW-COST AND HIGH-COST SYSTEMS

by

Paige M. McMullan

Bachelor of Science in Communication Sciences & Disorders, 2014
Texas Christian University
Fort Worth, TX

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 2016

A COMPARISON OF ACOUSTIC AND AERODYNAMIC MEASUREMENTS OF LARYNGEAL FUNCTION USING LOW-COST AND HIGH-COST SYSTEMS

A Thesis for the Degree

Master of Science

by

Paige M. McMullan

Thesis Approved by:

Christopher R. Watts, Ph.D., Major Professor

Danielle Brimo, Ph.D., Committee Member

Ronald Pitcock, Ph.D., Committee Member

Debbie Rhea, Ph.D., Associate Dean
Harris College of Nursing & Health Sciences

May 2016

ACKNOWLEDGEMENTS

I would like to gratefully acknowledge the support and guidance received from Dr. Christopher Watts in the creation of this document. His mentorship not only throughout this project, but throughout my graduate school career, has shaped me into the student, clinician, and person I am today. His student advocacy, constant encouragement, and overwhelming pride in his students and TCU have impacted my classmates and I in a way that will continue to inspire us long after graduation.

I would also like to thank my committee members, Dr. Danielle Brimo and Dr. Ronald Pitcock, for agreeing to be part of this process. Their comments, suggestions, ideas, and multiple perspectives helped shape the document into its final form. I thank them for their counsel, their time, and their interest in my project.

In addition, I would like to thank my father, Henry W. McMullan Jr., for supporting my decision to attend Texas Christian University. His support allowed me to spend not only four, but six years at a prestigious, beautiful university that made my educational and career dreams come true. I look up to him for his leadership in the workplace and family, perseverance during difficult situations, rationality at all times, and kindness when it matters; all of which have shaped me in to the young woman I am today.

A thank you to my mother, Tracy A. McMullan, for being my constant source of joy and encouragement. Her daily support through both good days and bad kept me emotionally healthy throughout the joys, exhaustion, and pains of college and graduate school. Her listening ear, sensible advice, and compassionate heart are just a few of her most admirable qualities that have influenced me. Because of her example, I strive to put others before myself, always seek a new perspective, and spread joy, light, and kindness wherever I go.

Another thank you to my younger sister, Erin N. McMullan, for her endless friendship, support, and company not only through graduate school, but throughout our entire lives. Her constant encouragement, calming presence, honest perspective, and joyful laughter always comfort me in times of stress. Her determination, confidence, kindness, selflessness, and boldness inspire me to be the best clinician, friend, and person that I can be.

A thank you to the TCU Communication Sciences and Disorders graduating class of 2016. I was so blessed to start my graduate school experience with a group classmates and instead leave with lifelong friends. We walked alongside each other throughout everything over the past two years: the good, the bad; the tears, the joy; the work week, the weekends; the finals, the breaks; the discipline of studying, the relief of an “A”. I grew close to these women not only because we lived and breathed similar trials and joys over the past two years, but because they inspired me daily with their kindness, devotion, determination, and wisdom. I am certain that I am a better clinician, person, and friend because of the strong and radiant women that I met through this program.

Finally, a formal thank you to Texas Christian University, the institution that provided both my undergraduate and graduate education and acted as my home away from home for the past six years. At TCU, I was provided with character-building situations, leadership roles, and educational experiences that have grown me in to a more confident, responsible, and ethical citizen.

A specific thank you to the faculty and staff in the Communication Sciences and Disorders department at Texas Christian University for teaching me, preparing me, and shaping me in to the clinician and student I am today. I was blessed to learn from professors who were not only experts in their field, but who genuinely cared about my well-being as an individual and success as a clinician. Their wisdom and professionalism will continue to influence my clinical and personal decision making as I begin my long-awaited journey as a speech-language pathologist.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Introduction

The expression "voice disorder" is a broad term used to describe medical conditions characterized by abnormalities in vocal pitch, loudness, or quality that affect the efficiency with which the vocal folds vibrate, thereby affecting voice production. Voice disorders are common: a study conducted in 2005 reported that the lifetime prevalence of a voice disorder in the general population was 29.9%, whereas 6.6% of participants reported a current voice disorder (Roy, Merrill, Gray, & Smith). When a person experiences a voice disorder, he or she may seek professional help from a speech-language pathologist (SLP). Among the various evaluation options available to an SLP to characterize voice disorders are behavioral and qualitative analysis of voice and resonance, diagnostic laryngoscopy with stroboscopy, and laryngeal function studies. Laryngeal function studies (LFS) consist of acoustic and aerodynamic assessments used to measure air pressure, airflow, frequency, intensity, and/or other spectral characteristics of sound produced by the vocal folds. By quantifying the underlying aerodynamic forces driving phonation and the resulting acoustic energy, LFS provide the clinician with information that can improve evaluation and treatment of voice disorders (Baken & Orlikoff, 2000).

Clinicians have a choice in the instrumentation that they use in order to obtain these acoustic and aerodynamic measurements for LFS, some of which is high-cost and some of which is low-cost. Because companies that produce high-cost systems employ marketing strategies that successfully brand their products as market leaders, some clinicians may be under the impression that it is necessary to purchase high-cost systems in order to perform the instrumental portion of LFS, and that low-cost systems are not a reliable option.

While there is extensive clinical research and data that supports the use of high-cost systems for LFS (Awan, Novaleski, & Yingling, 2013; Zhuang, et al., 2009; Radish, Bhat, &

Mukhi, 2011; Franca & Wagner, 2015), there is also literature that supports the use of low-cost equipment (Maryn, Corthals, Van Cauwenberge, Roy, & De Bodt, 2010; Maryn, De Bodt, & Roy, 2010; Maryn & Weenink, 2014; Awan, 2006; Hirano, Koike, & von Leden, 1968; Iwata & von Leden, 1970; Rau & Beckett, 1984). No contemporary research has compared the relationship between contemporary high and low-cost systems regarding acoustic and aerodynamic measurements. In addition, clinicians do not know if it makes a clinical difference if they use a high-cost versus a low-cost system. The purpose of this study was to investigate the relationship between acoustic and aerodynamic measures of both high and low-cost systems, which may help clinicians make an informed, evidence-based decision about which system is most appropriate for their clinical use.

Review of the Literature

Overview of Laryngeal Anatomy

Laryngeal cartilages. The larynx is a structure composed of cartilages, membranes, ligaments, and muscles located in the midline of the anterior neck. The larynx plays a role in biological functions (respiration, airway protection) and non-biological functions (phonation). There are six laryngeal cartilages including the thyroid, cricoid, epiglottis, and the paired arytenoid, corniculate, and cuneiform cartilages (Tucker, 1994). Of these, the thyroid, cricoid, and arytenoid cartilages are most important for phonation. Phonation can be defined as the physiological processes which lead to vocal fold vibration and sound energy. Specific to phonation, the laryngeal cartilages serve as both a framework for structural support and attachments for the vocal folds, which move based on positioning of the arytenoid and thyroid cartilages.

The thyroid is the largest of the laryngeal cartilages. It is composed of two cartilage plates, called laminae, which come together at a midline angle (Zemlin, 1998). The V-shaped point where they come together is called the thyroid notch, known in layman's terms as the "Adam's apple". There are two projections at the posterior borders of each of the thyroid lamina: the superior thyroid horns project upward and the inferior thyroid horns project downward. The superior horns are attached by a ligament to the corresponding horn of the hyoid bone. The inferior horns point downward and medially, and articulate with the cricoid cartilage (Zemlin, 1998).

The cricoid forms the inferior base for the larynx at the lower portion of the laryngeal framework (Tucker, 1994). It is located directly above and attaches to the first ring of the trachea (Zemlin, 1998). Shaped like a signet ring, the cricoid is thinner anteriorly and thicker posteriorly. Two lateral indentations on either side, called "facets", form the cricothyroid joint, which allows the thyroid to move about the cricoid. On the upper surface of the posterior lamina of the cricoid sit the arytenoid cartilages (Tucker, 1994). The arytenoids are pyramid-shaped with an apex and two projections at their base: the anterior projection is called the vocal process, which serves as the posterior attachment for the vocal folds, and the posterior/lateral projection is called the muscular process, which serves as an attachment for laryngeal muscles. The arytenoid apex is capped by the corniculate cartilage (Zemlin, 1998). The arytenoid cartilages rest on top of the cricoarytenoid joint, which allows the arytenoids to swivel. This swivel allows for the opening (abduction) and closing (adduction) of the vocal folds.

Although not as important to the production of voice, the epiglottis, corniculate cartilages, and cuneiform cartilages must be discussed in order to get a complete overview of laryngeal anatomy. The epiglottis is a leaf-like structure that narrows inferiorly and attaches to the thyroid cartilage (Sataloff, 2005). The epiglottis has an especially important function; during swallowing, it folds over the larynx in order to prevent food from entering the airway, which prevents coughing and/or choking. The corniculate cartilages, which sit atop the arytenoids, and the cuneiform cartilages, located within the aryepiglottic folds, have less-defined functions and are important for neither swallowing nor voice production. However, they may play an important role in structural support of the larynx.

The hyoid bone is a U-shaped bone which forms the superior aspect of the laryngeal skeleton and connects to the thyroid cartilage via the thyrohyoid membrane and ligaments (Sataloff, 2005). It is situated superior to the thyroid cartilage and anterior and inferior to the base of the tongue (Tucker, 1994). Parts of this bone include the body and two pairs of bilateral "horns": the greater horns and the lesser horns. Because the hyoid bone is suspended in the neck from musculature, membranes, and ligaments, it is the only bone in the body that it not attached to another bone (Tucker, 1994).

Membranes and ligaments. The framework of the larynx is held together by a number of membranes and ligaments. Among these include extrinsic membranes which connect the laryngeal framework to the upper neck and torso. Extrinsic laryngeal membranes include the hyothyroid membrane (also known as the thyrohyoid membrane), the cricotracheal membrane, the paired lateral hyothyroid ligaments, and the hyoepiglottic ligament. The hyothyroid membrane is located between the hyoid bone and the superior border of the thyroid cartilage (Zemlin, 1998). This membrane has two parts: the middle thyrohyoid ligament and the lateral thyrohyoid ligament (Tucker, 1994). The cricotracheal membrane attaches the bottom of the cricoid cartilage to the upper portion of the first tracheal ring. The hyoepiglottic ligament extends from the anterior surface of the epiglottis to the superior border of the hyoid bone (Zemlin, 1998).

The intrinsic laryngeal membranes connect the laryngeal cartilages to one another and control the direction of their movements. This group consists of the conus elasticus (cricovocal membrane), the medial cricothyroid ligament, the paired lateral cricothyroid membranes, the paired quadrangular membranes, and the aryepiglottic folds (Zemlin, 1998). These membranes and ligaments are part of a single sheet of connective tissue called the elastic membrane, which lines the majority of the larynx. At the inferior portion of the elastic membrane is the conus elasticus, which connects the thyroid, cricoid, and arytenoid cartilages. This membrane is further divided into a medial cricothyroid ligament and two lateral cricothyroid membranes. It extends from the superior border of the cricoid cartilage to the superior portions of the vocal folds and can be considered to connect laryngeal structures below the level of the glottis (Zemlin, 1998). The upper portion of the elastic membrane is the quadrangular membrane, named for its shape (Zemlin, 1998). The paired quadrangular membranes begin at the inner portion of the epiglottis and attach posteriorly to both the arytenoid and corniculate cartilages. Inferiorly, these membranes thicken to form the false vocal folds and extend around the ventricle to the point of attachment of the upper portion of the true vocal folds (Tucker, 1994). The superior portions of the quadrangular membranes are also thickened and form the aryepiglottic folds, in which the cuneiform cartilages are embedded. One other intrinsic laryngeal membrane is the medial cricothyroid ligament. This is a thick band of elastic tissue that stretches from the cricoid arch to the inferior border of the thyroid cartilage (Zemlin, 1998).

Laryngeal muscles. The extrinsic muscles of the larynx, sometimes referred to as the “strap muscles”, are those that have one attachment to structures outside of the larynx (Tucker, 1994; Zemlin, 1998). This group of muscles can be divided into those above the hyoid bone (suprahyoid muscles) and those below (infrahyoid muscles; Sataloff, 2005). Infrahyoid muscles depress the larynx and include the omohyoid, sternothyroid, thyrohyoid, and sternohyoid muscles (Sataloff, 2005; Tucker, 1994). The "inferior belly" of the omohyoid originates from the scapula and inserts into the tendon of the omohyoid muscle in the neck. From this same tendon originates the "superior belly", which connects to the greater horn of the hyoid bone. This muscle pulls the hyoid bone downward (Sataloff, 2005). The sternothyroid originates at the sternum and inserts at the thyroid cartilage. The thyrohyoid, which is also located in the anterior neck, originates from the thyroid lamina and inserts at the greater horn of the hyoid bone. When this muscle contracts it decreases the distance between the hyoid bone and thyroid cartilage (Zemlin, 1998). The sternohyoid muscle originates at the clavicle and inserts at the hyoid bone. Contraction of this muscle lowers the hyoid bone (Sataloff, 2005).

Suprahyoid muscles elevate the larynx and include the digastric, mylohyoid, geniohyoid, and stylohyoid muscles (Sataloff, 2005). The digastric muscle has two parts: the anterior belly and the posterior belly. The anterior belly originates at the mandible and inserts into the digastric intermediate tendon. The posterior belly originates at the mastoid process of the temporal bone and also inserts at the digastric intermediate tendon. Both the anterior and posterior bellies raise the hyoid bone, which is connected to the digastric intermediate tendon (Zemlin, 1998; Sataloff, 2005). The mylohyoid muscle originates from the mandible and inserts at the hyoid. The geniohyoid originates from the inferior mental spine of mandible and inserts at the hyoid bone. Lastly, the stylohyoid originates from the styloid process and inserts into the hyoid bone. The interaction and coordination of these extrinsic laryngeal muscles controls the vertical position of the larynx in the neck (Sataloff, 2005).

The intrinsic muscles of the larynx have both their point of origin and point of insertion within the larynx. The purpose of these muscles is to adduct (close), abduct (open), tense (elongate), or relax (shorten) the vocal folds. The intrinsic muscles include the thyroarytenoid (TA, including the vocalis muscle and the muscularis muscle), posterior cricoarytenoid (PCA), lateral cricoarytenoid (LCA), interarytenoids (IA), and the cricothyroid (CT; Zemlin, 1998). The TA can act as an adductor, tensor, or relaxer of the vocal folds. More specifically, the vocalis division is a tensor or relaxer while the muscularis division is an adductor. Both the vocalis and muscularis originate at the anterior commissure of the thyroid. The vocalis inserts at the vocal process of the arytenoid and the muscularis inserts at the muscular process of the arytenoid (Zemlin, 1998). The PCA, the sole abductor of the vocal folds, originates at the posterior lamina of the cricoid and inserts at the muscular process of the arytenoid. The LCA, an adductor, originates at the anterolateral arch of the cricoid and inserts at the muscular process of the arytenoid (Sataloff, 2005). The IA muscles (2) originate at the arytenoids, insert at the opposite arytenoid, and act as adductors of the vocal folds. Finally, the CT, a tensor of the vocal folds, originates from the anterolateral arch of the cricoid cartilage and inserts at the muscular process of the arytenoid (Zemlin, 1998).

Vocal fold structure. The vocal folds consist of five structural layers. The epithelium is the outermost layer which consists of squamous cell tissue. Below this layer is the lamina propria, which consists of three separate layers. The first of these is the superficial layer of the lamina propria, also known as Reinke's space, which consists of loose fibrous components. Next is the intermediate layer of the lamina propria, which is mainly composed of elastic fibers. The last layer of the lamina propria is the deep layer, which contains more collagenous fibers than any other layer. The fifth and last layer of the vocal folds is the vocalis muscle, which is considered the main body of the vocal fold and is part of the thyroarytenoid muscle (Colton, 1994; Zemlin, 1998).

Although there are five structural layers, the vocal folds are commonly divided into three functional layers. The first of these is the cover, which consists of the epithelium and the superficial layer of the lamina propria. During most instances of phonation it is the vocal fold cover which moves during oscillation. The next structural layer is the vocal ligament, which is made of the intermediate and deep layers of the lamina propria. This layer connects the cover to the last functional layer, the body, which consists solely of the vocalis muscle (Colton, 1994; Sataloff, 2005).

Overview of Phonation

Phonation refers to the physiological processes which lead to vocal fold vibration and the production of acoustic energy. Phonation requires two types of forces, muscular and aerodynamic, acting on the vocal fold tissue. Extrinsic muscular forces are only utilized to position the larynx in the midline of the neck with minor contribution to vocal fold vibration. For example, when a speaker raises their fundamental frequency to high levels it is often accompanied by increased contraction in the laryngeal elevators with a subsequent rise in the vertical position of the larynx. Intrinsic laryngeal muscular forces are crucial to phonation, as their levels of activation influence glottal configuration (e.g., abducted, adducted) and the degree of vocal fold tension.

To bring the vocal folds to midline (adduction) three intrinsic laryngeal muscles contract including the interarytenoids, the muscularis division of the thyroarytenoid, and the lateral cricoarytenoid. These three adductor muscles set the medial compression force of the vocal folds. At the same time, tension in the vocal folds is set by the cricothyroid and vocalis muscles, which determines the fundamental frequency of the voice that is produced. In addition, aerodynamic forces act on the vocal folds. The respiratory muscles contract and send air flowing upward until it meets a closed (adducted) glottis. The blocked air flow results in an increase in air pressure below the vocal folds. Eventually this increased subglottic pressure overcomes the medial compression forces of the vocal folds and air separates the cover of the vocal folds, from bottom to top. As air flows through the glottis, the bottom edge opens first, then the top edge opens; next the bottom edge closes before the top edge closes. While aerodynamic forces act on the vocal folds, the tensor and adductor muscles maintain their contraction to hold the vocal folds in an adducted position throughout phonation (Baken, 2005; Colton, 1994; Zemlin, 1998).

Once phonation is initiated, vocal fold vibration is maintained through multiple forces including intraglottal pressure differences, vortices, elasticity, and the Bernoulli Effect. Each of these factors help the vocal folds return to midline after the air flow blows them apart. Intraglottal pressure differences sustain oscillation and vibration of the vocal folds. Because the bottom and top lip of the vocal folds open and close at difference times, there are pressure differences at the top and bottom lip of the folds. The air pressure is greater where the vocal folds are closer together. This inverse relationship between air pressure and volume between the vocal folds help to open and close them during vibration. Vortices, or areas of extreme negative pressure, are created along the superior edges of a divergent glottis (top edges apart, bottom edges approximating each other). This negative pressure helps the top edges of the vocal folds snap back together during the closing phase of vocal fold vibration. The elastic nature of the vocal folds, or the ability for the vocal folds to distend and return to their original position, also brings the vocal folds together during phonation. Lastly, the Bernoulli Effect acts on the vocal folds. As air flows up through the glottis, it leaves behind a negative pressure perpendicular to the air flow, which brings the vocal folds together. Together, these four factors act on the vocal folds and help sustain oscillation during phonation (Baken, 2005; Colton, 1994; Zemlin, 1998).

Laryngeal Function Studies

By quantifying the underlying aerodynamic forces driving phonation and the resulting acoustic energy, LFS provide the clinician with information that can improve evaluation and treatment of voice disorders (Baken & Orlikoff, 2000). Aerodynamic assessments associated with LFS typically measure air volume, airflow, air pressure, and/or vocal efficiency (CPT Assistant, January 2006 – American Medical Association). Airflow measures have included vital capacity, maximum phonation time, phonation quotient, s/z ratio, as well as average and peak flow rates (Dejonckere, 2001; Zraick, Smith-Olinde, Shotts, 2012). Air pressure is typically measured using estimates of subglottal pressure during habitual and soft phonation, the latter from which measures of phonation threshold pressure have been derived (Zhuang et al., 2013). To date, no individual or set of aerodynamic measures has been demonstrated to be clinically more important or more cost effective than others.

Acoustic assessments associated with LFS have typically included measures of frequency, intensity, perturbation, signal-to-noise ratio, and spectral analyses (Dejonckere, 2001). Classical acoustic measurements have utilized time-based measurements including fundamental frequency, frequency perturbation, and amplitude perturbation which have relied on software that detects vocal fold vibratory cycle boundaries. A disadvantage of time-based measurements is that they lose accuracy when applied to connected speech and voices which are highly dysphonic. However, recent technology has allowed clinicians to efficiently collect cepstral and spectral acoustic measurements, which hold advantages to the traditional time-based measurements of jitter and shimmer in that they can be validly applied to moderate-to-severe dysphonic voices in addition to connected speech (Awan et al., 2006; Heman-Ackah et al., 2003). The dominant rhamonic (an anagram of “harmonic”) of the cepstrum, referred to as the cepstral peak prominence (CPP), is an indirect measure of phonation periodicity and spectral noise, and is reduced in dysphonic voices (Heman-Ackah et al., 2003). Measurements of CPP have also correlated well to auditory-perceptual measures of dysphonic severity (Awan et al., 2006). Similar to aerodynamic measures, the literature illustrates a large variation in the clinical application and reporting of specific acoustic voice measurements (Dejonckere, 2001; Heman-Ackah et al., 2003; Awan et al., 2006).

Available technology for conducting LFS includes precision high-tech options sold by corporations specializing in computerized analysis of vocal function in addition to low-tech options available for free or at a low cost. High-tech options include products sold by PENTAX Medical, such as the Computerized Speech Lab (CSL) for acoustic voice analysis and the Phonatory Aerodynamic System (PAS) for aerodynamic voice analysis. Advantages of these systems include excellent measurement precision, professional customer support, and a large body of literature validating their use in clinical populations (Awan, Novaleski, & Yingling, 2013; Zhuang, et al., 2009; Radish, Bhat, & Mukhi, 2011; Franca & Wagner, 2015**).** One disadvantage of these products is their high cost, which can equal or exceed thousands of dollars. While these products are not the only high-tech/high-cost systems available for conducting LFS, they stand among the most sold acoustic and aerodynamic systems in the United States.

Low-tech/low-cost products are also available to practicing clinicians. These options include the computer program “PRAAT” for acoustic analysis and hand-held spirometers (different manufacturers) for aerodynamic analyses. PRAAT is freeware available for download via the Internet, and has a substantial body of research literature validating its use for acoustic analyses (Maryn, Corthals, Van Cauwenberge, Roy, & De Bodt, 2010; Maryn, De Bodt, & Roy, 2010; Maryn & Weenink, 2014). Hand-held spirometers are available in analog or digital varieties, and their use is also backed by substantial research literature (Awan, 2006; Hirano, Koike, & von Leden, 1968; Iwata & von Leden, 1970; Rau & Beckett, 1984). A primary advantage to these low-tech systems is their cost, which ranges from free to a few hundred dollars. Disadvantages include a lack of customer support for their application and possible variability in measurement compared to high-tech systems. However, contemporary research evidence comparing the measurement reliability between high-cost and low-cost systems for LFS is non-existent.

Statement of the Problem

Clinicians performing voice evaluations, including Laryngeal Function Studies (LFS), have options regarding the types of equipment they will use to obtain measurements including both high-cost and low-cost options. The majority of published studies on LFS use high-cost equipment; therefore, some clinicians feel the need to use expensive, similar equipment in order to perform and bill for LFS. Many clinicians are not performing LFS at all because they do not have the high-cost equipment that they believe is necessary to perform the exam (Christopher R. Watts, personal communication, March 6, 2015). The purpose of this study was to investigate the strength of the relationship (parallel forms) of different measurements obtained from both high-cost and low-cost equipment, specifically measurements obtained for the acoustic and aerodynamic portions of LFS. Specific questions include: (1) Do high-cost and low-cost equipment used to obtain measures for the acoustic portion of LFS (e.g., CPP) exhibit a strong degree of relationship? (2) Do high-cost and low-cost equipment used to obtain measures for the aerodynamic portion of LFS (e.g., VC, airflow rate) exhibit strong degree of relationship? Additional research questions included: (3) Do acoustic measurements obtained from high-cost and low-costs systems differ as a function of sex? (4) Do aerodynamic measurements obtained from high-cost and low-cost systems differ as a function of sex?

Methodology

Subjects

This study included 40 participants recruited from the student body population at Texas Christian University and the local community. Participants were divided into two groups of 20 males and 20 females. Inclusion criteria were as follows: (a) age between 18-25, (b) no current or past history of diagnosed voice disorder, (c) no history of neurological disease and (d) no current complaints of voice or laryngeal problems.

Instrumentation

Two categories of instruments were used to conduct this study: high-cost and low-cost. The high-cost category was defined as equipment costing more than $1,000, and the low-cost category was defined as equipment costing less than $1,000. The choice for this cutoff was based on equipment budgets for speech-language pathology programs in hospital-based rehabilitation centers in the local region. For acoustic analyses, hardware and software from the Computerized Speech Lab (CSL; Kay Pentax; Montvale, NJ) represented high-cost instrumentation. The Analysis of Dysphonia in Speech and Voice (ADSV) software component of CSL was used to collect cepstral peak prominence in vowel (cppV) and cepstral peak prominence in speech (cppS) from this system. The computer program “PRAAT” (http://praat.en.softonic.com/) represented low-cost instrumentation and was used to collect the same dependent variables as CSL. Both CSL and PRAAT were connected to the same desktop PC.

For aerodynamic analyses, measures of vital capacity (VC) and airflow (EMFR/TarAirflow) were obtained. The Phonatory Aerodynamic System (PAS; Kay Pentax;

Montvale, NJ) represented the high-cost instrumentation. The PAS consists of a pneumotachograph connected via cables to a PC computer. A Contec SP10 hand-held digital spirometer (Contec Medical Systems; Qinhuangdao, China) represented the low-cost option. The hand-held spirometer consists of a flow tube leading to a turbine which is driven by pulmonary air forced through the flow tube by the participant.

Procedures

All participants underwent consenting procedures prior to participation in the study approved by the university Institutional Review Board. All recording and measurement occurred in the Laryngeal Function Laboratory in the Miller Speech and Hearing Clinic on the campus of Texas Christian University. The order in which any participant provided measurements using different instrument options was counterbalanced across all participants. For acoustic analyses, participants wore a head-mounted microphone (AKG Acoustics; Vienna, Austria) with the microphone head positioned approximately three centimeters from the left corner of the participant's mouth. For recordings on CSL, the microphone was directly connected to the CSL preamplifier. For recordings using PRAAT, the microphone was connected to an Audiogram Preamplifier (Yamaha Corp; Buena Park, CA). Gain settings on CSL and Yamaha were set at a constant level for all participants, and only the microphone head distance from the corner of the mouth was adjusted if signal clipping occurred.

Dependent variables for this study included acoustic and aerodynamic measurements. Acoustic measures included: cepstral peak prominence (CPP in dB) in (1) sustained vowel and (2) connected speech. Aerodynamic measures included: (1) vital capacity (VC in mL), (2) estimated mean flow rate (EMFR in mL/s = 77+ [.236 x phonation quotient]; phonation quotient [PQ] = [spirometer VC/MPT]), and (3) target airflow (TarAirflow in mL/s).

For acoustic analyses, measures of cepstral peak prominence in both vowel (cppV) and speech (cppS) were collected. To obtain cppV, participants were recorded sustaining the vowel /ɑ/. To obtain cppS, participants were recorded reading the sentence "we were away a year ago". Three tokens of each stimulus were recorded using both PRAAT (low-cost) and CSL (high-cost), resulting in twelve total acoustic tokens per participant. Instructions to participants for vowel and connected speech recording on both systems were as follows:

(Vowel): "Take a nice easy breath and say the vowel /ɑ/ at a comfortable pitch and loudness as steady as you can until I tell you to stop. I will tell you to stop after about four to five seconds."

(Connected speech): "Take a nice easy breath and say the sentence 'we were away a year ago' at a comfortable pitch and loudness just as if you were speaking to me in conversation."

For high-cost aerodynamic analyses, participants completed two protocols using the PAS: vital capacity (VC) and target airflow. Three tokens of each stimulus were recorded, for a total of six tokens per participant. Instructions to participants for measures using the PAS were as follows:

(Vital Capacity): "Take two easy breaths. On the third breath, take a deep breath in, place the mask over your nose and mouth, creating a tight seal, and blow out all of your air until there is nothing left in your lungs."

(Target Airflow): "I will now insert a small tube in to the mask. Take an easy breath, place the mask on your face, creating a tight seal, and position the small tube so it is resting just inside your of your mouth. When you have the mask positioned correctly, say the syllable /pɑ/ six times slowly, like this (clinician demonstrated). There should be about a 1.5 second pause in between each syllable. Complete all six syllables on one breath, without inhaling in between."

Low-cost aerodynamic measures were obtained by measuring vital capacity (VC) with the hand-held spirometer and maximum phonation time (MPT) with a digital timer on an iPhone. Three tokens of each recording were acquired, for a total of six tokens per participant. While using the hand-held spirometer, the participant wore a nose clip to prevent air leakage from the nostrils. Instructions for MPT and VC from the hand-held spirometer were as follows:

(Maximum Phonation Time): "I want you to take a deep breath and say the vowel /ɑ/ at a comfortable pitch and loudness for as long as you possibly can."

(Vital Capacity): "I want you to take two easy breaths. On the third breath I want you to breathe in as deeply as possible, place your lips and teeth around the spirometer tube tightly and blow out all of your air through the tube, until you have nothing left in your lungs."

Analyses

There were two classifications of independent variables in this study: (1) equipment type, of which there were two levels: high-cost and low-cost; and (2) sex, which consisted of male and female levels. There were four dependent variables in this study: (1) vital capacity, defined as the volume of gas exhaled from the lungs after maximum inhalation, measured in mL; (2) airflow rate, defined as the quantity of airflow through the glottis over a period of time, measured in mL/s; (3 and 4) cepstral peak prominence in A) vowel and B) speech, defined as the peak rhamonic in the cepstrum, representing periodicity and stability of phonation.

For all statistical analyses, the level of significance was set at alpha = 0.025, which is an adjusted level of significance to correct for Type 1 error. To measure the strength of relationship between physiological measurements obtained from high-cost and low-cost systems, Pearson Product Moment Correlations were applied to the data in order to determine linear relationships between the following variables, with data collapsed across sex:

- High-cost vital capacity (VCpas) and Low-cost vital capacity (VCspir)
- High-cost target airflow rate (TarAirflow) and Low-cost estimated mean flow rate (EMFR)
- High-cost CPP in vowels (ADSVcppV) and Low-cost CPP in vowels (PRAATcppV)
- High-cost CPP in speech (ADSVcppS) and Low-cost CPP in speech (PRAATcppS)

It was expected that sex would affect some physiological measurements due to structural physical differences between males and females. By further analyzing this factor across the dependent variables of this study, we were able to determine whether or not sex influenced all aerodynamic and acoustic measurements as a function of instrumentation type. It was also expected that different instrumentation would result in significantly different measurements because formulas used to calculate aerodynamic and acoustic measurements vary between high-cost and low-cost systems. Because of this, instrumentation type was not a factor when investigating the effect to sex. To investigate the effect of sex on acoustic and aerodynamic measurements obtained from different types of instrumentation, one-way multivariate analyses of variance (MANOVA) were applied to the data with sex as the independent variable and VC, EMFR/Target Airflow, CPP in vowels and CPP in speech as dependent variables. Because data points of CPP in speech were missing from two males and tw0 females, a separate MANOVA was applied to that data.

Reliability

Intra- and inter-measurement reliability was obtained by re-analyzing 10% of all recorded digital files. For inter-measurement reliability, a second laboratory assistant trained in the acquisition and analysis of acoustic and aerodynamic measurements was utilized. Because some types of measures used to acquire the dependent variables were not in digital form (e.g., phonation time measured from a stopwatch used to calculate phonation quotient), reliability calculations for all dependent variables were not possible. The correlation coefficients for intra- and inter-reliability data are displayed in Tables 1 and 2. This data translated to very good reliability in how these dependent variables were measured.

Table 1. **Correlation coefficients for intra-reliability data.**

VCpas	TarAirflow	PRAATcppV	PRAATcppS	ADSVcppV	ADSVcppS
1	0.99	1	1	0.99	0.99

Table 2. **Correlation coefficients for inter-reliability data.**

VCpas	TarAirflow	PRAATcppV	PRAATcppS	ADSVcppV	ADSVcppS
0.99	0.99	1	1	0.99	0.99

Results

Descriptive Statistics

Overall, there were four measures collected in this study: vital capacity (VC), an aerodynamic measure of airflow (transglottal airflow =TarAirflow; estimated mean flow rate = EMFR), cepstral peak prominence in vowel (cppV), and cepstral peak prominence in speech (cppS). Each measure was acquired with a high-cost instrument (PAS or CSL) and with a low-cost instrument (hand-held spirometer and iPhone or PRAAT). The means, standard deviations, and ranges of these measures are displayed in Table 3. In this table, both males and females were pooled together (N = 40). Data gathered using low-cost equipment is listed first, followed by the same (or similar) measure gathered using high-cost equipment. It should be noted that adequate data was not collected for CPP in speech on neither PRAAT nor ADSV for four participants (two males and two females) and therefore these participants were excluded (n = 36).

***Table 3*. Descriptive statistics for acoustic and aerodynamic measures.**

	Mean	Std. Deviation	Range	N
VCspir	4093.25	907.641	3930	40
VCpas	4894.00	1232.558	5410	40
EMFR	122.23	17.427	71	40
TarAirflow	156.75	63.140	280	40
PRAATcppV	26.77	3.597	13.68	40
ADSVcppV	12.18	2.565	9.16	40
PRAATcppS	23.25	1.952	8.04	36
ADSVcppS	9.24	1.280	6.3	36

VCspir = vital capacity in mL from hand-held spirometer; VCpas = vital capacity in mL from PAS system; EMFR = estimated mean flow rate in mL/s derived from spirometer PQ; TarAirflow = target airflow in mL/s from PAS system; PRAATcppV = cepstral peak prominence in dB (vowel) from PRAAT; ADSVcppV = cepstral peak prominence in dB (vowel) from ADSV program in CSL; PRAATcppS = cepstral peak prominence in dB (connected speech) from PRAAT; ADSVcppS = cepstral peak prominence in dB (connected speech) from ADSV program in CSL.

Vital capacity: Hand-held spirometer and Phonatory Aerodynamic System. Based on descriptive statistics, the mean, standard deviation, and range of vital capacity derived from the hand-held spirometer data (VCspir) was less than that derived from the PAS (VCpas) data. Figure 1 visually displays the VC data trends across the 40 participants (female participants = 1 – 20; male participants = 21 – 40).

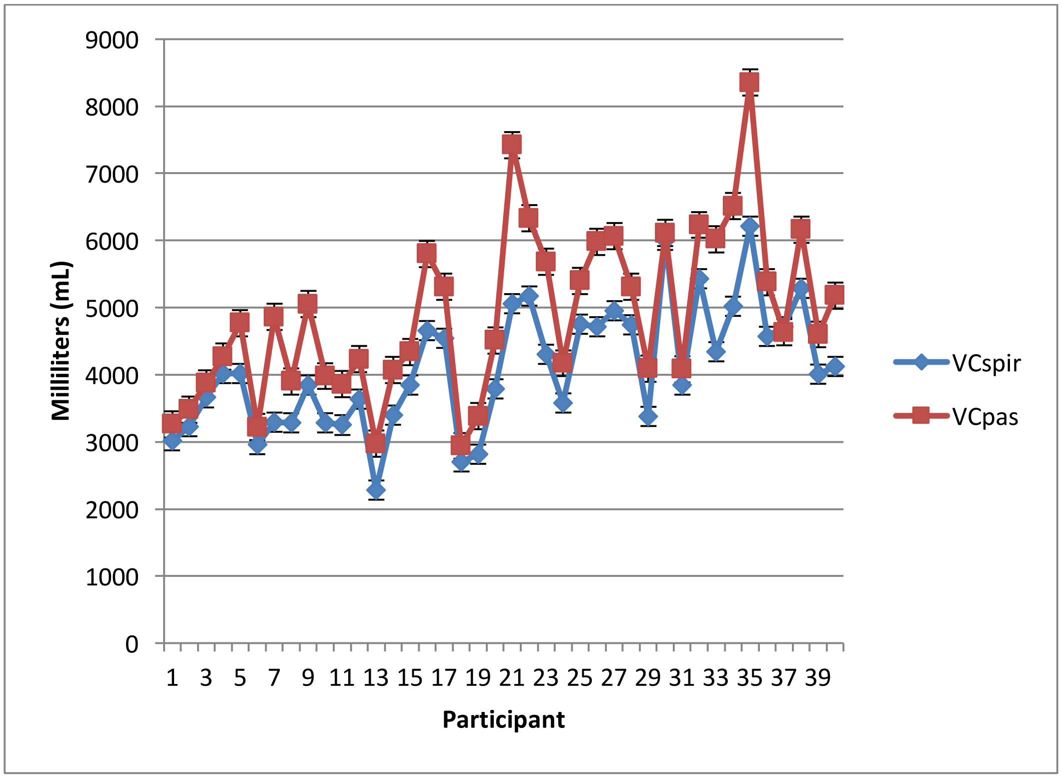

Figure 1. **Vital capacity obtained from the low-cost hand-held spirometer (VCspir) and vital capacity obtained from the high-cost PAS system (VCpas).**

Although VCspir was consistently lower than VCpas, when VCspir increased from one participant to the next, so did VCpas. Overall, the line depicting VCspir closely follows the line depicting VCpas, which indicates the possibility of a strong relationship of VC measured from the two systems.

Estimated mean flow rate (spirometer) and target airflow (PAS). As can be observed in Table 3, the mean, standard deviation, and range derived from EMFR data across all participants was less than those derived from Target Airflow data. Figure 2 visually displays the data trends across all participants (females = 1 – 20; males = 21 – 40).

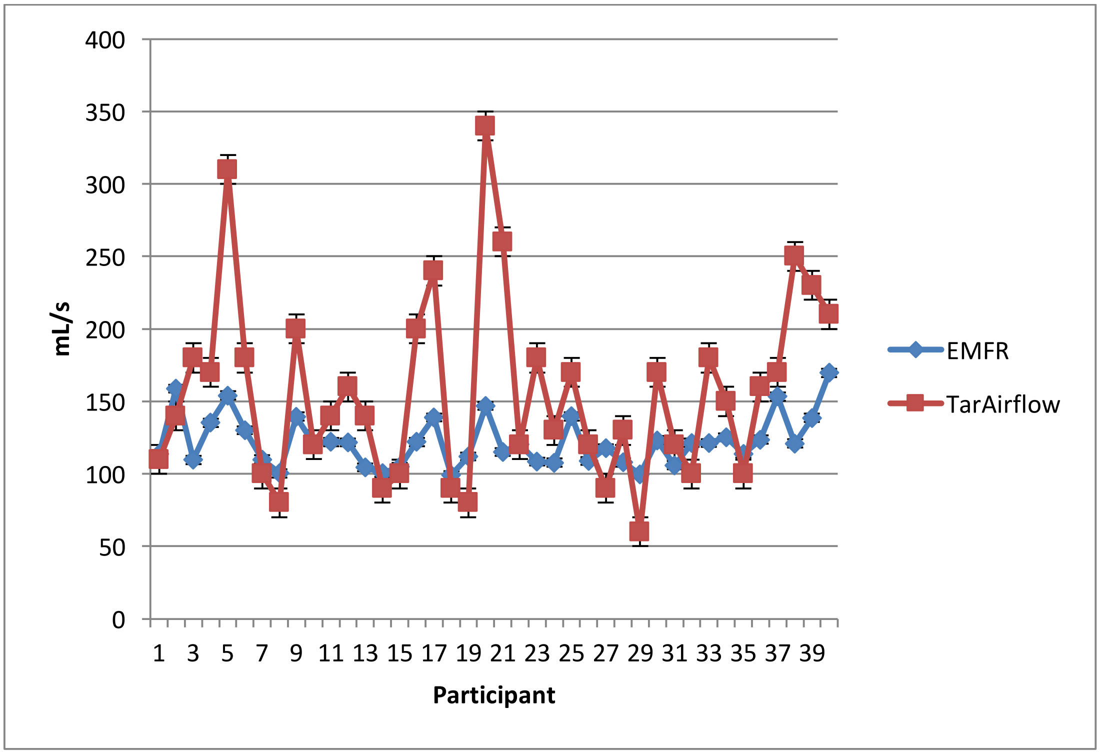

Figure 2. **Estimated mean flow rate (EMFR) derived from the low-cost hand-held spirometer and target transglottal airflow (TarAirflow) derived from the high-cost PAS system.** All measurements are in milliliters per second (mL/s) as indicated on the Y-axis.

Although the trend is less apparent, when Target Airflow increased, EMFR also increased. The line depicting EMFR roughly follows the line depicting Target Airflow, which indicates the possibility of a moderate relationship of EMFR and Target Airflow measured from the two systems. Visually, the airflow data appeared to manifest more variability than the VC data, especially with the high-cost system.

Cepstral peak prominence in vowel: PRAAT and ADSV. Figure 3 visually displays the data trends across all participants (females = 1 – 20; males = 21 – 40). Based on descriptive statistics, the mean, standard deviation, and range of CPP in vowel derived from PRAAT (PRAATcppV) data was greater than those derived from ADSV (ADSVcppV) data (see Table

3). Mean differences were expected because these two programs use different algorithms to generate data.

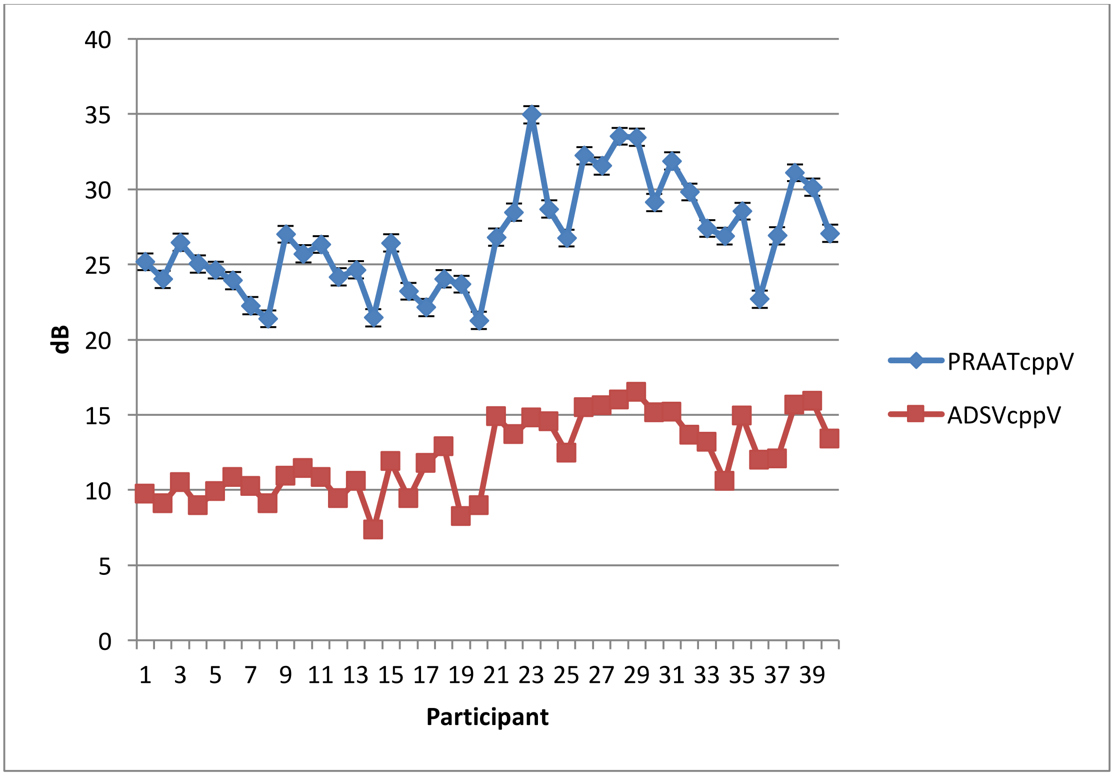

***Figure 3*. Cepstral peak prominence in vowel measured from the low-cost PRAAT (PRAATcppV) and high-cost ADSV (ADSVcppV) systems.**

However, a general trend still occurred in this data set; when PRAATcppV increased, ADSVcppV also increased. The line depicting ADSVcppV roughly follows the line depicting ADSVcppV, which indicates the possibility of a strong relationship in CPP in vowel measured by PRAAT and ADSV.

Cepstral peak prominence in speech: PRAAT and ADSV. Figure 4 visually displays the data trends across all participants (females = 1 – 20; males = 21 – 40). As expected, the mean, standard deviation, and range of CPP in speech derived from PRAAT (PRAATcppS) data was greater than those derived from ADSV (ADSVcppS) data (see Table 3).

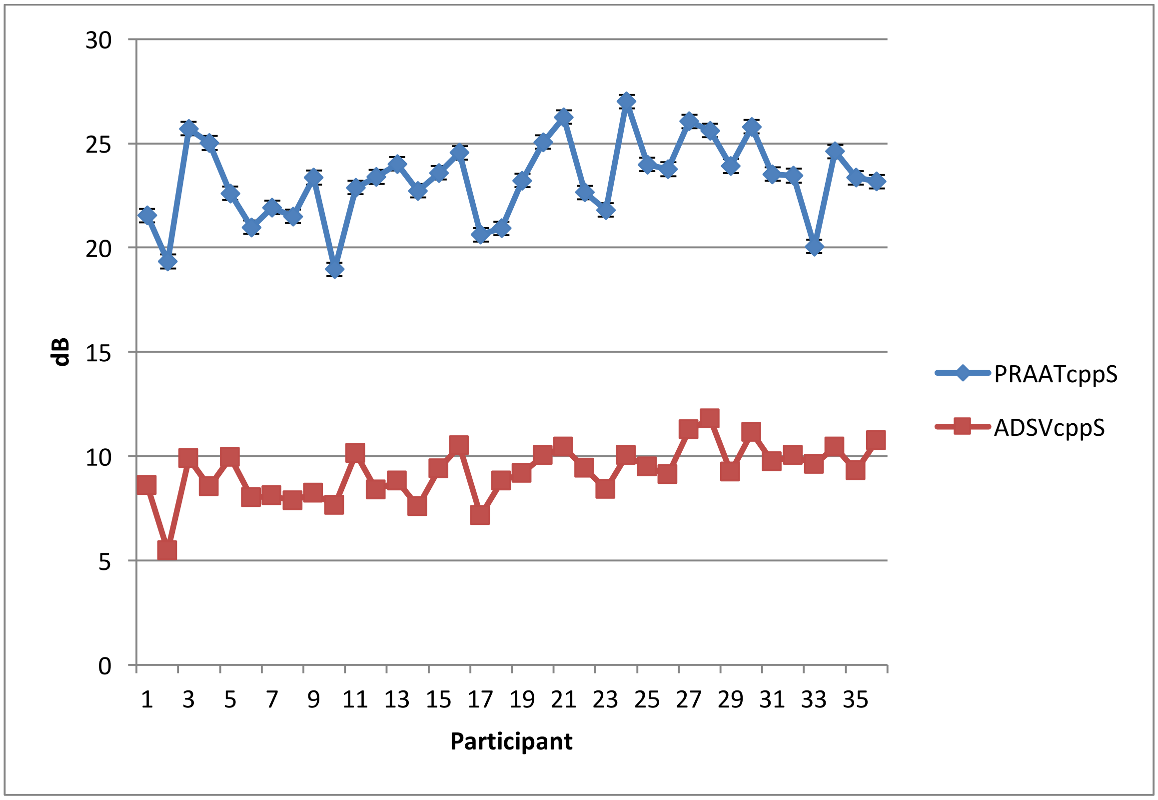

***Figure 4*. Cepstral peak prominence in connected speech measured from the low-cost PRAAT (PRAATcppS) and high-cost ADSV (ADSVcppS) systems.**

As was true for CPP in vowel, PRAAT also consistently produced a higher CPP in speech than ADSV for each participant. A similar trend occurred in this data set; when PRAATcppS increased, ADSVcppS also increased. The line depicting ADSVcppS generally follows the line depicting ADSVcppS, which indicates the possibility of a moderate-to-strong relationship in CPP in speech measured by PRAAT and ADSV.

Correlation Analyses

To quantify the degree of relationship between data derived from high-cost instruments and their low-cost counterparts, a bivariate Pearson correlation analysis was performed on each data set. The results are displayed below in Table 4.

***Table 4.* Correlation data.**

		VCspir	VCpas	EMFR	TarAirflow	PRAATcppV	ADSVcppV	PRAATcppS	ADSVcppS
VCspir	Pearson Correlation		.920*						
	Sig. (2-tailed)		.000						
	N		40						
VCpas	Pearson Correlation	.920*							
	Sig. (2-tailed)	.000							
	N	40							
EMFR	Pearson Correlation				.628*				
	Sig. (2-tailed)				.000				
	N				40				
TarAirflow	Pearson Correlation			.628*					
	Sig. (2-tailed)			.000					
	N			40					
PRAATcppV	Pearson Correlation						.854*		
	Sig. (2-tailed)						.000		
	N						40		
ADSVcppV	Pearson Correlation					.854*			
	Sig. (2-tailed)					.000			
	N					40			
PRAATcppS	Pearson Correlation								.730*
	Sig. (2-tailed)								.000
	N								36
ADSVcppS	Pearson Correlation							.730*	
	Sig. (2-tailed)							.000	
	N							36	

*Correlation is significant at the 0.01 level (2-tailed). VCspir = vital capacity in mL from hand-held spirometer; VCpas = vital capacity in mL from PAS system; EMFR = estimated mean flow rate in mL/s derived from spirometer PQ; TarAirflow = target airflow in mL/s from PAS system; PRAATcppV = cepstral peak prominence in dB (vowel) from PRAAT; ADSVcppV = cepstral peak prominence in dB (vowel) from ADSV program in CSL; PRAATcppS = cepstral peak prominence in dB (connected speech) from PRAAT; ADSVcppS = cepstral peak prominence in dB (connected speech) from ADSV program in CSL.

Vital capacity: Hand-held spirometer and Phonatory Aerodynamic System.

Correlation analysis between VCspir and VCpas revealed a correlation coefficient of r = .920, which was statistically significant (P = 0.01). This indicated a strong relationship between VC measured from high-cost and low-cost systems. Figure 5 visually displays this relationship in a scatterplot. Each dot on the scatterplot represents the VCspir and VCpas data collected from a single participant. The coefficient of determination for this relationship was $R^2 = 0.846$, which indicated that approximately 85% of the variability in VCpas and VCspir measures could be explained by instrument type.

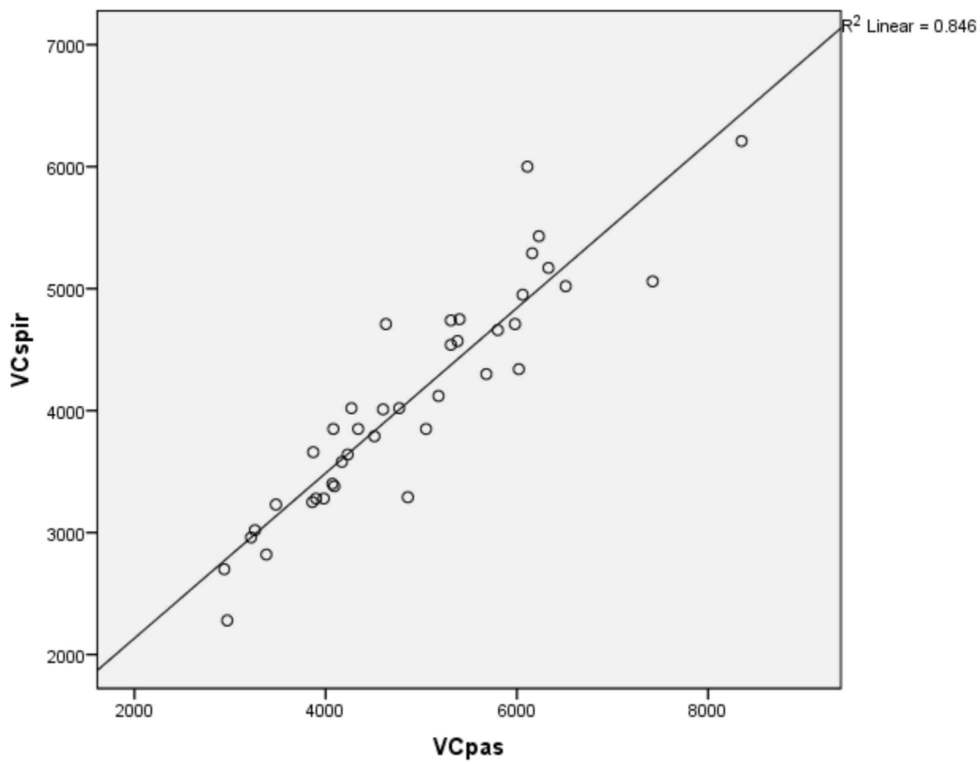

***Figure 5*. Correlation scatterplot between vital capacity obtained from the low-cost spirometer (VCspir) and the high-cost PAS system (VCpas)**. One dot represents one participant.

Estimated mean flow rate (spirometer) and target airflow (PAS). Correlation analysis between EMFR and Target Airflow revealed that r = .628, which was also statistically significant (P = 0.01) and deemed to represent a moderate-to-strong relationship. Figure 6 visually displays the relationship between EMFR and Target Airflow. The coefficient of determination was $R^2 = 0.394$, which indicated that approximately 39% of the variability in Target Airflow measures and EMFR measures could be explained by instrument type.

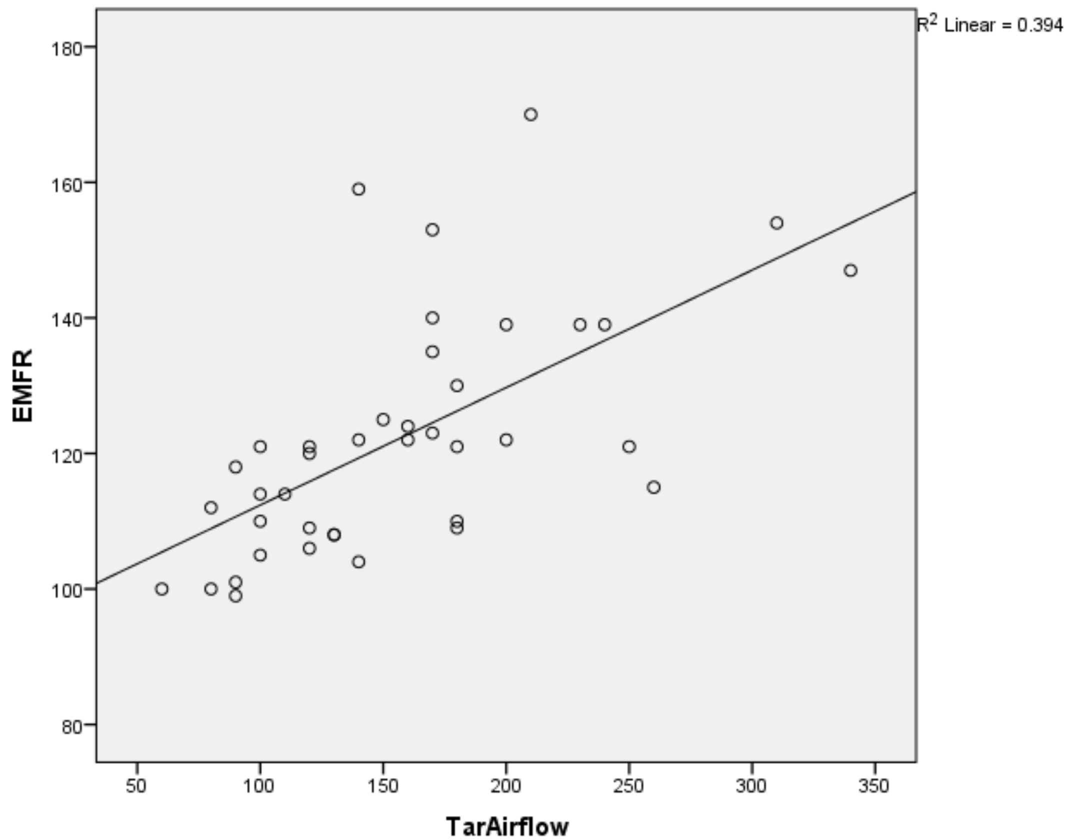

***Figure 6*. Correlation scatterplot between estimated mean flow rate derived from the low-cost spirometer (EMFR) and target transglottal airflow derived from the high-cost PAS system (TarAirflow).** One dot represents one participant.

Cepstral peak prominence in vowel: PRAAT and ADSV. Correlation analysis between PRAATcppV and ADSVcppV revealed that r = .854, which was also statistically significant (P = 0.01) and interpreted as a strong relationship. Figure 7 visually displays the relationship between PRAATcppV and ADSVcppV. The coefficient of determination was $R^2 = 0.730$, indicating that approximately 73% of the variability in ADSVcppV and PRAATcppV measures could be explained by instrument type.

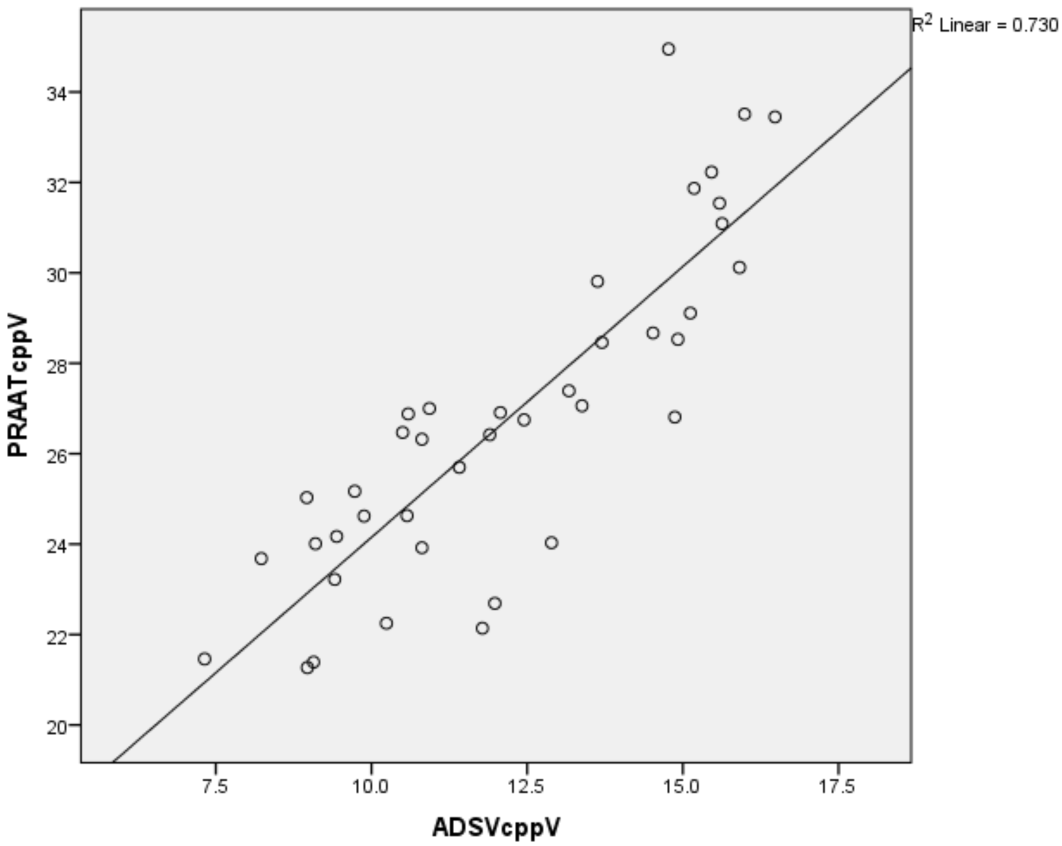

***Figure 7*. Correlation scatterplot between cepstral peak prominence in vowel obtained from the low-cost PRAAT system (PRAATcppV) and the high-cost ADSV system (ADSVcppV).** One dot represents one participant.

Cepstral peak prominence in speech: PRAAT and ADSV. Correlation analysis between PRAATcppS and ADSVcppS revealed that r = .730, which was also statistically significant (P = 0.01) and interpreted as a strong relationship. Figure 8 visually displays the relationship between PRAATcppS and ADSVcppS. The coefficient of determination was $R^2 = 0.533$, indicating that approximately 53% of the variability in ADSVcppS and PRAATcppS measures could be explained by instrument type.

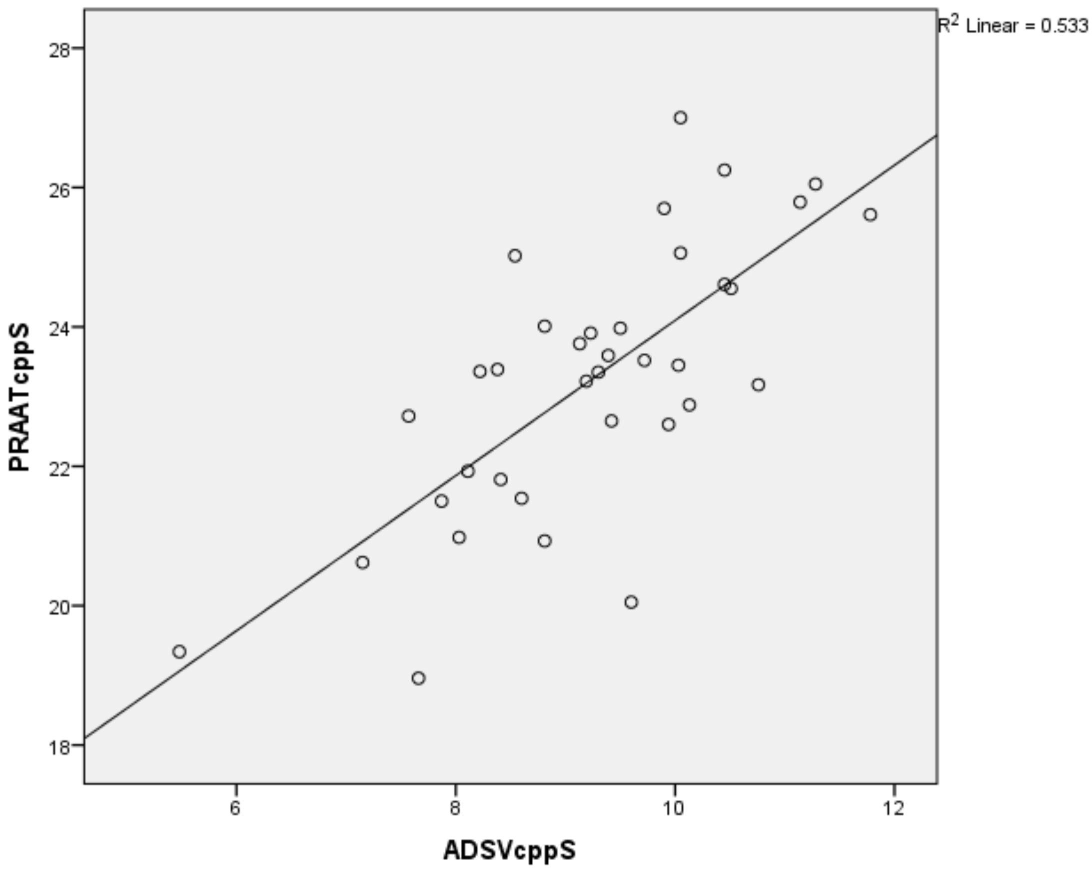

***Figure 8*. Correlation scatterplot between cepstral peak prominence in connected speech obtained from the low-cost PRAAT system (PRAATcppS) and the high-cost ADSV system (ADSVcppS).** One dot represents one participant.

Males vs. Females

The first purpose of this study was to compare high-cost and low-cost equipment used in voice analyses, which is described above. The second purpose was to examine acoustic and aerodynamic data gathered from males and females to evaluate the effect of sex on these measures. Descriptive statistics for male and female data for each dependent variable are displayed in Table 5.

***Table 5.* Descriptive Statistics for Males vs. Females.**

	Gender	Mean	Std. Deviation	N
VCspir	Female	3477.00	595.192	20
	Male	4709.50	732.957	20
	Total	4093.25	907.641	40
VCpas	Female	4103.50	785.576	20
	Male	5684.50	1088.884	20
	Total	4894.00	1232.558	40
EMFR	Female	122.25	18.278	20
	Male	122.20	17.010	20
	Total	122.23	17.427	40
TarAirflow	Female	158.50	72.712	20
	Male	155.00	53.754	20
	Total	156.75	63.140	40
PRAATcppV	Female	24.15	1.777	20
	Male	29.39	2.986	20
	Total	26.77	3.597	40
ADSVcppV	Female	10.10	1.350	20
	Male	14.27	1.587	20
	Total	12.18	2.565	40
PRAATcppS	Female	22.42	1.862	18
	Male	24.07	1.715	18
	Total	23.25	1.952	36
ADSVcppS	Female	8.51	1.215	18
	Male	9.97	.872	18
	Total	9.24	1.280	36

VCspir = vital capacity in mL from hand-held spirometer; VCpas = vital capacity in mL from PAS system; EMFR = estimated mean flow rate in mL/s derived from spirometer PQ; TarAirflow = target airflow in mL/s from PAS system; PRAATcppV = cepstral peak prominence in dB (vowel) from PRAAT; ADSVcppV = cepstral peak prominence in dB (vowel) from ADSV program in CSL; PRAATcppS = cepstral peak prominence in dB (connected speech) from PRAAT; ADSVcppS = cepstral peak prominence in dB (connected speech) from ADSV program in CSL.

Mean measures of vital capacity, cepstral peak prominence in vowel (cppV), and cepstral peak prominence in speech (cppS) were greater for males when gathered with both high-cost and low-cost equipment. Mean airflow rate was similar across males and females using the two equipment types (EMFR and TarAirflow), differing by no more than 4 mL/s between sexes.

In addition to descriptive statistics, a MANOVA was used to evaluate group differences between males and females. Results from the MANOVA revealed a significant effect of sex (Pillai's trace, $F = 21.23$, $p < 0.001$). To determine where differences existed, follow up F-tests were applied. These tests revealed a significant effect of sex for vital capacity (VCpas: $F = 27.73$, $p < 0.001$; VCspir: $F = 34.079$, $p < 0.001$) and cepstral peak prominence in vowel (ADSVcppV: $F = 80.228$, $p < 0.001$; PRAATcppV: $F = 45.594$, $p > .001$). There were no significant group differences between males and females for measures of airflow on high-cost (TarAirflow: $F = .030$, $p = .863$) nor low-cost (EMFR: $F = .000$, $p = .993$) equipment.

Because there were only 18 participants in each group for cepstral peak prominence in speech (cppS), a second MANOVA was applied to cppS data. Results from this MANOVA revealed a significant effect of sex (Pillai's trace, $F = 8.394$, $p = .001$). To determine where differences existed in the cppS data, follow up F-tests were applied. These tests revealed significant differences between males and females for both high (ADSVcppS: $F = 17.295$, $p < .001$) and low-cost equipment (PRAATcppS: $F = 7.604$, $p = .009$) on measures of cppS.

Discussion

The purpose of this study was to investigate the strength of relationship in different measurements obtained from both high-cost and low-cost equipment, specifically measurements obtained for the acoustic and aerodynamic portions of LFS. The first research question asked: Do high-cost and low-cost equipment used to obtain measures for the acoustic portion of LFS (e.g., CPP) exhibit a strong degree of relationship? The answers was "yes"; this study found that measures obtained using high-cost (ADSV) and low-cost (PRAAT) equipment exhibited a strong degree of relationship for cepstral peak prominence measured in both vowels and speech. However, there was a stronger relationship between the high and low-cost systems for CPP in vowel. This finding is consistent with previous studies which have found that when CPP is compared to perception of severity, correlations are stronger in vowels than in speech (Peterson et al., 2013; Brinca, Batista, Tavares, Goncalves, & Moreno, 2014; Watts & Awan, 2011). Peterson et al. (2013) compared listener perceived severity and CPP in both vowels and speech and found that correlations were stronger between perceived severity and CPP in sustained vowel (r = 0.81) than in CPP in connected speech (r = 0.67). Brinca, Batista, Tavares, Goncalves, and Morena (2014) used PRAAT software to investigate the use of CPP in vowel (/ɑ/) and CPP in speech to differentiate dysphonic from nondysphonic voices. They found that correlations were significant and strongest between CPP in vowel and auditory-perceptual measures such as grade (overall voice abnormality), roughness, breathiness, and asthenia (weakness). Watts and Awan (2011) evaluated the diagnostic value of CPP in vowels and in speech to differentiate dysphonic and nondysphonic voices. The researchers found that although both measures of CPP in vowels and CPP in speech are sensitive in identifying a dysphonic voice, CPP in vowels was a more sensitive measure than CPP in speech.

A possible reason for the stronger relationship of CPP in vowels is because the acoustic signal in vowels does not vary substantially, whereas it is highly variable in connected speech. This creates a situation where speakers might produce a vowel multiple times in a similar manner, but produce connected speech with slight differences in prosody which will influence the subsequent acoustic measurements recorded with different equipment. This can introduce greater variability into the acoustic signal and influence the correlation measurements. The fact that separate recordings of vowels and speech were obtained for analysis via PRAAT and ADSV, rather than analyzing the same file on different systems, may have also introduced measurement variability that influenced correlational analyses. In other words, the same vowel recording was not analyzed twice by two separate systems; rather, two different vowel recordings were obtained to be analyzed separately by each of the two systems. This was also true for the connected speech data.

Although different algorithms are used to calculate CPP in ADSV and PRAAT and correlations were stronger for vowels than connected speech, the results of this study revealed that both systems will indicate when phonation is more periodic or when it is less periodic. Because the measures from both systems had strong correlations for cppV and cppS, we concluded that both ADSV and PRAAT will provide similar information regarding periodicity of the voice, and therefore either is a valid option for use during the acoustic portion of laryngeal function studies.

The second research question asked: Do high-cost and low-cost equipment used to obtain measures for the aerodynamic portion of LFS (e.g., VC, airflow rate) exhibit a strong degree of relationship? The answer was also "yes"; this study found that there was a strong degree of relationship between vital capacity measures and a moderate-to-strong relationship between airflow measures obtained from high-cost (PAS) and low-cost (spirometer) equipment. A study by Rau and Beckett (1984) revealed similar conclusions. The researchers compared aerodynamic measures obtained from four types of instrumentation: a Collins P-900 respirometer, a Calculair Portable Electronic Spirometer, a Propper Compact Spirometer, and a Ventilation Monitor. In the study, the Collins P-900 could be considered high-cost and the other three instruments could be considered low-cost. When investigating relationships between vital capacity data collected from all four instruments, correlation coefficients were significant ($p = .001$) and ranged between 0.94 and 0.98. When comparing vital capacity data collected from high and low-cost instrumentation in the current study, the correlation coefficient was 0.92. This is slightly below but very close to the findings in the Rau and Beckett study; in both studies, vital capacity measures obtained from high and low-cost instrumentation exhibited a strong degree of relationship. Rau and Beckett also examined relationships among measures of phonation quotient (PQ) and mean flow rate (MFR) calculated from the same instrument for all four spirometers. Phonation quotient is the ratio between vital capacity and maximum phonation time and is an aerodynamic measure that is comparable to the Target Airflow measure used in the current study. Mean flow rate was calculated by the equation (mean flow rate = 77 + [.236 x Phonation Quotient]) which was the same equation used to calculate EMFR in the current study. Rau and Beckett found strong relationships between measures of PQ and MFR among all four instruments with correlation coefficients ranging from 0.74 to 0.87. The current study compared similar data (Target Airflow and EMFR) obtained from high and low-cost instrumentation and found a moderate-to-strong relationship ($r = 0.628$). The correlation coefficient found in this study was slightly below but approached that of the Rau and Beckett study. The differences between Rau and Beckett's findings and the findings in the current study could be explained by the use of different equipment and different subjects. Both Rau and Beckett and the current study concluded that vital capacity and airflow measures obtained from high and low-cost instrumentation exhibited moderate-to-strong relationships.

In the current study, although VC measures were not the same when measured by PAS and the low-cost spirometer, both systems indicated when VC was high and when it was low. Similarly, both systems indicated when airflow through the glottis was high or low. Although EMFR and Target Airflow measures were not the same, there was a moderate-to-strong relationship between the high and low-cost systems. Because VC measures exhibited a strong relationship and airflow measures exhibited a moderate-to-strong relationship when obtained with PAS and a hand-held spirometer, we concluded that both systems will provide similar information, and therefore either is a valid option for use during the aerodynamic portion of laryngeal function studies.

The third research question asked: Do acoustic measurements obtained from high-cost and low-costs systems differ as a function of sex? The answer was “yes”; this study found a significant effect of sex for both high-cost measures of CPP in vowel and CPP in speech as well as low-cost measures of CPP in vowel and CPP in speech. In other words, males and females were significantly different on both high and low-cost measures of both cppV and cppS.

Multiple studies have found differences between males and females on other acoustic measures such as fundamental frequency, maximum phonation time, jitter, and shimmer. Fundamental frequency (F_0) is a measure of the lowest frequency in habitual speaking pitch and is commonly recognized as lower in adult males than in adult females in both speech (Nittrouer, McGowan, Milenkovic, & Beehler, 1990; Klatt & Klatt, 1990) and sustained vowels (Sussman & Sapienza, 1994; Dehqan, Ansari, & Bakhtiar, 2008). Maximum phonation time (MPT in seconds) is a measure that represents respiratory-phonatory coordination and airflow control and is obtained by asking the participant to sustain a vowel (typically /ɑ/), at a comfortable pitch and loudness, for as long as he possibly can. Multiple studies have revealed that MPT is, on average, greater in males than females (Dehquan, Ansari, & Bakhtiar, 2010; Maslan, Leng, Rees, Blalock, & Butler, 2011).

Jitter and shimmer are measures of cycle to cycle variation in vocal frequency and amplitude, respectively. These measures have been widely used both clinically and in research to describe acoustic characteristics of voice. However, these time-based measures have been found to have poor measurement reliability, sensitivity, and specificity, especially in severely dysphonic voices (Brockmann, Drinnan, Storck, & Carding, 2011). Cepstral peak prominence (CPP), an indirect measure of phonation periodicity, gives similar, but more reliable, acoustic information about a voice (Heman-Ackah et al., 2003). The current study found that CPP in vowel and CPP in speech differ based on sex; similarly, other studies have found that measures of jitter and shimmer differ based on sex. For example, in a study comparing acoustic measurements of men's and women's voices, Nittrouer et al. (1990) found a between-subject main effect of speaker sex for both jitter and shimmer. Multiple studies have showed that female voices generally display less shimmer but more jitter than male voices (Horii, 1980; Sorensen & Horii, 1983). In a study examining gender effects on acoustic measures, gender had a significant effect on jitter, but not on shimmer (Brockmann et al., 2011).

Although the acoustic measures mentioned above ($F_{0,}$ MPT, jitter, and shimmer) are not the same as the ones examined in the current study (cppV and cppS), this research supports the fact that males and females differ in various acoustic measures. This is consistent with the current study, which found that males and females differed in both cppS and cppV obtained from high and low-cost instrumentation. Physiologically, these differences might be explained by larger structures in male speakers which cause differences in how the vocal folds oscillate. One obvious example of this is the difference in fundamental frequency between males and females – males have greater vocal fold mass which vibrates at a slower rate.

The fourth and final research question asked was: Do aerodynamic measurements obtained from high-cost and low-cost systems differ as a function of sex? The answer was "yes" and "no"; males and females differed significantly on measures of vital capacity obtained with the PAS and spirometer, but not on measures of EMFR/Target Airflow. Existing literature supports these findings. Biersteker and Biersteker (1985) examined vital capacity in healthy young adults and found that males had greater inspiratory vital capacity than females. Although this study examined expiratory rather than inspiratory vital capacity, one can conclude that a greater inspiratory ability would result in a greater expiratory ability. Regarding general pulmonary function, prediction equations for lung function have shown a significant sex difference in adults (as cited by Harms, 2006). It is also generally accepted that men have larger lung volumes than women, which would explain the larger vital capacity volumes found in men in this study (Harms, 2006).

Regarding flow rate, Goozee et al. (1998) found that males and females were not significantly different in measures of phonatory (mean) flow rate in vocal efficiency, vowel, and sentence speaking tasks. However, there was a significant effect of sex for flow rate in the most comfortable phonation task (Goozee et al., 1998). In this study, the vocal tasks used to measure EMFR and TarAirflow were the sustained vowel /ɑ/ and the syllables /pɑ pɑ pɑ/, respectively. Although syllables were not tested in the Goozee study, the current study supports Goozee's findings because males and females were not significantly different in a vowel task. Goozee tested vowels /i/ and /u/; the current study tested the vowel /ɑ/. Both studies agreed that males and females were not significantly different in flow rate for vowels.

Limitations

There are several limitations in this study which warrant guarded generalizations. First, although the sample size was moderate (N = 40), only participants aged 18-25 were included in this study. Although we suspect that the study would have produced the same results with a wider age range of participants, generalizations should be made with caution and findings support the need for larger studies with more variable participant characteristics. Also, only healthy participants were included this study. Future directions may include replication of this study with a larger sample size of a wider age range of participants including both healthy and voice disordered participants. Sufficient acoustic data for connected speech was not collected for all 40 participants. Subsequently, two males and two females were excluded from cppS data analysis for both high-cost and low-cost equipment (n = 36). This smaller sample size should be noted when interpreting cppS data.

Conclusion

Acoustic and aerodynamic measurements are useful components of laryngeal function studies that assist a speech-language pathologist in evaluation and treatment of voice disorders. Existing literature illustrates a large variation in the clinical application and reporting of specific measurements such as vital capacity, target airflow, and cepstral peak prominence. Both high and low-cost equipment is commercially available to obtain these measures, but no recent studies have compared equipment types. In this study, high and low-cost equipment was used to compare measures of vital capacity (VC), air flow (EMFR/TarAirflow), cepstral peak prominence in vowels (cppV), and cepstral peak prominence in connected speech (cppS). We demonstrated that measures obtained from high and low-cost equipment had a strong relationship for VC, cppV, and cppS, and a moderate-to-strong relationship for EMFR/TarAirflow. This study also evaluated differences between sexes in the same dependent variables across high and low-cost equipment. Significant differences between males and females were found for measures of vital capacity, cppV, and cppS, which supported existing studies. The results of this study support the use of either high or low-cost equipment as a reliable option to obtain acoustic and aerodynamic measures for evaluation and management of voice disorders.

References

Awan SN. The aging female voice: acoustic and respiratory data. Clin Linguist Phon. 2006 Apr-May; 20(2-3):171-80.

Awan SN, Novaleski CK, Yingling JR. Test-retest reliability for aerodynamic measures of voice. J Voice. 2013 Nov; 27(6):674-84. doi: 10.1016/j.jvoice.2013.07.002. Epub 2013 Oct 8.

Awan SN, Roy N. Toward the development of an objective index of dysphonia severity: a four-factor acoustic model. Clin Linguist Phon. 2006 Jan-Feb; 20(1):35-49.

Baken RJ, Orlikoff RF. Clinical Measurement of Speech and Voice. 2000, San Diego: Singular Publishing.

Baken, R. J. (2005). An overview of laryngeal function for voice production. In R. T. Sataloff (Ed.), Voice science (53-87). San Diego: Plural Publishing.

Biersteker MW, Biersteker PA. Vital capacity in trained and untrained healthy young adults in the Netherlands. Eur J Appl Physiol Occup Physiol. 1985;54(1):46-53.

Brinca LF, Batista AP, Tavares AI, Goncalves IC, Moreno ML. Use of cepstral analyses for differentiating normal from dysphonic voices: a comparative study of connected speech versus sustained vowel in European Portuguese female speakers. J Voice. 2014 May;28(3):282-6. doi: 10.1016/j.jvoice.2013.10.001. Epub 2014 Feb 1.

Brockmann M, Drinnan MJ, Storck C, Carding PN. Reliable jitter and shimmer measurements in voice clinics: the relevance of vowel, gender, vocal intensity, and fundamental frequency effects in a typical clinical task. J Voice. 2011 Jan;25(1):44-53. doi: 10.1016/j.jvoice.2009.07.002. Epub 2010 Apr 8.

Colton, R. H. (1994). Physiology of phonation. In M. S. Benninger, B. H. Jacobson & A. F. Johnson (Eds.), Vocal arts medicine: the care and prevention of professional voice disorders (30-60). New York: Thieme Medical Publishers, Inc.

Dehqan A, Ansari H, Bakhtiar M. Objective voice analysis of Iranian speakers with normal voices. J Voice. 2010 Mar;24(2):161-7. doi: 10.1016/j.jvoice.2008.07.005. Epub 2009 Feb 20.

Dejonckere PH, Bradley P, Clemente P, Cornut G, Crevier-Buchman L, Friedrich G, Van De Heyning P, Remacle M, Woisard V; Committee on Phoniatrics of the European Laryngological Society (ELS). A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. Guideline elaborated by the Committee on Phoniatrics of the European Laryngological Society (ELS). Eur Arch Otorhinolaryngol. 2001 Feb; 258(2):77-82.

Franca MC, Wagner JF. Effects of vocal demands on voice performance of student singers. J Voice. 2015 May; 29(3):324-32. doi: 10.1016/j.jvoice.2014.07.004. Epub 2014 Dec 12.

Goozée JV, Murdoch BE, Theodoros DG, Thompson EC. The effects of age and gender on laryngeal aerodynamics. Int J Lang Commun Disord. 1998 Apr-Jun;33(2):221-38.

Harms, CA. Does gender affect pulmonary function and exercise capacity? Respir Physiol Neurobiol. 2006 Apr 28;151(2-3):124-31. Epub 2006 Jan 10.

Heman-Ackah YD, Heuer RJ, Michael DD, Ostrowski R, Horman M, Baroody MM, Hillenbrand J, Sataloff RT. Cepstral peak prominence: a more reliable measure of dysphonia. Ann Otol Rhinol Laryngol. 2003 Apr; 112(4):324-33.

Hirano M, Koike Y, Von Leden H. Maximum phonation time and air usage during phonation. Clinical study. Folia Phoniatr (Basel). 1968; 20(4):185-201.

Horii Y. Vocal shimmer in sustained phonation. J Speech Hear Res. 1980;23:202-209.

Iwata S, von Leden H. Phonation quotient in patients with laryngeal diseases. Folia Phoniatr (Basel). 1970; 22(2):117-28.

Klatt D & Klatt L. Analysis, synthesis and perception of voice quality variations among female and male talkers. J Acoust Soc Am. 1990 Feb; 87(2): 820-57.

Maryn Y, Corthals P, Van Cauwenberge P, Roy N, De Bodt M. Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels. J Voice. 2010 Sep; 24(5):540-55. doi: 10.1016/j.jvoice.2008.12.014. Epub 2009 Nov 2.

Maryn Y, De Bodt M, Roy N. The Acoustic Voice Quality Index: toward improved treatment outcomes assessment in voice disorders. J Commun Disord. 2010 May-Jun; 43(3):161-74. doi: 10.1016/j.jcomdis.2009.12.004. Epub 2009 Dec 23.

Maryn Y, Weenink D. Objective dysphonia measures in the program PRAAT: smoothed cepstral peak prominence and acoustic voice quality index. J Voice. 2015 Jan; 29(1):35-43. doi: 10.1016/j.jvoice.2014.06.015. Epub 2014 Dec 9.

Maslan J, Leng X, Rees C, Blalock D, Butler SG. Maximum Phonation Time in Healthy Older Adults. Journal of voice : official journal of the Voice Foundation. 2011;25(6):709-713. doi:10.1016/j.jvoice.2010.10.002.

Nittrouer S, McGowan RS, Milenkovic PH, Beehler D. Acoustic measurements of men's and women's voices: a study of context effects and covariation. J Speech Hear Res. 1990 Dec;33(4):761-75.

Peterson EA, Nelson R, Awan SN, Merrill RM, Banks R, Tanner K. Toward validation of the Cepstral Spectral Index of Dysphonia (CSID) as an objective treatment outcomes measure. J Voice. 2013 Apr; 27(4): 401-410.

Radish Kumar B, Bhat JS, Mukhi P. Vowel harmonic amplitude differences in persons with vocal nodules. J Voice. 2011 Sep; 25(5):559-61. doi: 10.1016/j.jvoice.2010.06.009. Epub 2010 Oct 6.

Rau D, Beckett RL. Aerodynamic assessment of vocal function using hand-held spirometers. J Speech Hear Disord. 1984 May; 49(2):183-8.

Roy N, Merrill RM, Gray SD, Smith EM. Voice disorders in the general population: prevalence, risk factors, and occupational impact. Laryngoscope. 2005 Nov; 115(11):1988-95.

Sataloff, R. T. (2005). Clinical anatomy and physiology of the voice. In R. T. Sataloff (Ed.), Voice science (53-87). San Diego: Plural Publishing.

Sorensen D, Horii Y. Frequency and amplitude perturbation in the voices of female speakers. J Commun Disord. 1983;16:57-61.

Sussman JE, Sapienza C. Articulatory, developmental, and gender effects on measures of fundamental frequency and jitter. J Voice. 1994 Jun;8(2):145-56.

Tucker, H. M. (1994). Gross and microscopic anatomy of the larynx. In M. S. Benninger, B. H. Jacobson & A. F. Johnson (Eds.), Vocal arts medicine: the care and prevention of professional voice disorders (11-29). New York: Thieme Medical Publishers, Inc.

Watts CR, Awan SN. Use of spectral/cepstral analyses for differentiating normal from hypofunctional voices in sustained vowel and continuous speech contexts. J Speech Lang Hear Res. 2011 Dec;54(6):1525-37. doi: 10.1044/1092-4388(2011/10-0209).

Zemlin, W. R. Speech and Hearing Science: Anatomy and Physiology. 1998; Boston: Allyn & Bacon.

Zhuang P, Sprecher AJ, Hoffman MR, Zhang Y, Fourakis M, Jiang JJ, Wei CS. Phonation threshold flow measurements in normal and pathological phonation. Laryngoscope. 2009 Apr; 119(4):811-5. doi: 10.1002/lary.20165.

Zhuang P, Swinarska JT, Robieux CF, Hoffman MR, Lin S, Jiang JJ. Measurement of phonation threshold power in normal and disordered voice production. Ann Otol Rhinol Laryngol. 2013 Sep; 122(9):555-60.

Zraick RI, Smith-Olinde L, Shotts LL. Adult normative data for the KayPENTAX Phonatory Aerodynamic System Model 6600. J Voice. 2012 Mar; 26(2):164-76.

ABSTRACT

Purpose: This investigation compared high-cost and low-cost options for obtaining acoustic and aerodynamic voice measures and evaluated differences between males and females for each of the dependent variables.

Methodology: Forty healthy participants were recorded performing various acoustic (vowel, speech) and aerodynamic (vital capacity, airflow) tasks with both high and low-cost equipment. Recorded voices were analyzed with PRAAT and CSL and aerodynamic measures were obtained with PAS and a hand-held spirometer.

Results: Strong, statistically significant relationships were found between high and low-cost equipment for measures of vital capacity, cppV, and cppS, and a moderate-to-strong relationship was found for airflow measures (EMFR/TarAirflow). Significant differences between males and females were found for measures of vital capacity, cppV, and cppS.

Conclusions: Because of the strong reliability found between equipment types, we concluded that either high or low-cost equipment is a valid option to conduct laryngeal function studies. The significant effects of sex found in this study are supported by the existing body of literature. The results suggest that the high-cost and low-cost instrumentation used in this study are both valid options for use in clinical voice evaluations.

Key Words: voice, acoustic analysis, aerodynamic analysis, laryngeal function studies, vital capacity, cepstral peak prominence, air flow