2018

Are There Differences in Vocal Parameters Across Various Ethnic Groups?

Mason P. Meek
Cleveland State University

ARE THERE DIFFERENCES IN VOCAL PARAMETERS ACROSS VARIOUS ETHNIC GROUPS?

MASON MEEK

Bachelor of Speech and Hearing

Cleveland State University

December 2014

Submitted in partial fulfillment of requirements for the degree

MASTER OF ARTS IN SPEECH-LANGUAGE PATHOLOGY AND AUDIOLOGY

at the

CLEVELAND STATE UNIVERSITY

May 2018

We hereby approve this thesis for

MASON MEEK

Candidate for the Master of Arts in Speech Pathology and Audiology degree

For the

College of Sciences and Health Professions

Program of Speech and Hearing

And

CLEVELAND STATE UNIVERSITY’S

College of Graduate Studies by

____

Thesis Chairperson, Violet O. Cox, Ph.D., MLS., CCC-SLP

____

Department & Date

____

Thesis Committee Member, Myrita Wilhite, Au.D.

____

Department & Date

____

Thesis Committee Member, Elizabeth Domholdt, Ed.D. FAPTA

____

Department & Date

Student’s Date of Defense: April 25, 2018

ARE THERE DIFFERENCES IN VOCAL PARAMETERS ACROSS VARIOUS ETHNIC GROUPS?

MASON MEEK

ABSTRACT

The purpose of this study was to investigate vocal characteristics of various ethnic groups in terms of acoustic, physiological, and perceptual data. Currently there are acoustic parameter norms that have been established for men and women. Prior research suggested that these norms were established on only American Caucasian individuals. This creates the question as to whether or not these norms are reflective of the true nature of vocal fold closure pattern of all ethnic groups. If there is a difference in vocal characteristics then it would be important to establish specific norms for various ethnic groups. Further rationale for this investigation is that the demographic changes in the USA have signaled the need for greater understanding of treatment approaches for SLP and other health professionals. Secondly, establishing empirical evidence to support normative expectations for different groups may be helpful in understanding cultural expectations in treating voice disorders. Broadly speaking, I hypothesized that there would be a difference in fundamental frequency across ethnic groups. Specifically, there will be African-American participants that will demonstrate lower fundamental frequency parameters and that there will be differences in vocal fold behavior across the various ethnic groups. Fifteen individuals participated, one African-American male as well as three African-American females, two Asian males and one Asian female, three Caucasian males and two Caucasian females, one Hispanic male and two Hispanic females. Each participant underwent videostroboscopy for vocal fold observation. Additionally, aerodynamic vocal parameters were measured using the *Visi-Pitch*. All males demonstrated a posterior chink in their glottal closure pattern except for the Caucasian male group. There were no significant differences within acoustic measurements between the participants. When correlating the physiological and the acoustic data, it was demonstrated that the males with a posterior chink demonstrated a higher fundamental frequency compared to the males without a posterior chink. Currently there are no norm studies that have attempted to describe vocal fold closure across ethnic groups. The data identified in this study is relatively unexplored.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

Table Page

NOMENCLATURE

VF- vocal fold

F0 – fundamental frequency

RAP – relative average perturbation (jitter)

NHR – noise-to-harmonic ratio

VTI – voice turbulence index

AAM - African-American male

AAF – African-American female

ASM- Asian male

ASF- Asian female

CM - Caucasian male

CF - Caucasian female

HM - Hispanic male

HF - Hispanic female

CHAPTER I

INTRODUCTION

The acoustic-perceptual parameters of the human voice have been studied for varying purposes over many years. From a medical perspective, the human voice has been studied as an indicator of the state of health as well the severity/progression of disease. From an intellectual point of view, the human has also been investigated as a linguistic source of information. Furthermore, culturally speaking, many studies have examined the physical characteristics of the human voice from the point of view of conveying a person's cultural and even dialectical background. (Titze,1995). Physiologically the process of converting the air pressure from the lungs into audible vibration is called phonation. Specifically, phonation occurs when air is expelled from the lungs through the glottis – the space between the vocal folds (VF). During this expulsion of air, the pressure at the glottis changes causing the VFs to vibrate. It is this vibration that produces the sound or phonation. The muscles of the larynx change the elasticity and tension of the VFs to determine the pitch of phonation. A voice is perceived as sounding normal providing that it meets certain expectations associated with a person's community, society, culture, age, gender, and profession. A voice disorder or voice difference exists when quality, pitch, loudness, and flexibility of a person's voice differ from the voice of others of a similar age, gender, and cultural group (Aronson, 1990). As mentioned above, it is possible to recognize a person's voice based on characteristics such as pitch, culture, age gender and even profession. But the question that appears to be relatively unanswered is whether these differences rise to the level of ethnicity. Are there certain vocal fold characteristics that are unique to specific ethnicities? Cox & Mueller (2004) noted that there is evidence from both perceptual and acoustic studies that seem to suggest a possible difference in fundamental frequency across certain ethnic groups, but it is unclear as to whether the difference is linguistically or physiologically influenced.

Laryngeal Physiology

Laryngeal physiology examines the functional dynamics of the larynx in relation to the control mechanisms that influence the production of voice. The VFs are located within the larynx which is anatomically positioned adjacent to the cervical vertebrae C3 through C6. Healthy vocal folds are required for airway protection, respiration, swallowing, and phonation. Speech results from the production of a fundamental tone produced at the larynx and is modified by resonating chambers of the upper areodigestive tract. The production of the fundamental tone is due to the mucosal vibration of the VFs, generated by the passage of air between them. According to Hirano (1987) the vibratory cycle is described as having three phases: opening, closing, and closed. Within the opening phase as subglottic pressure increases, the VFs are forced apart from an inferior to superior direction until the glottis opens, letting air escape thus releasing subglottic pressure. As the elastic recoil of the VFs forces them back together, the superior portion of the vocal folds are the last to close. Therefore, the VFs close from bottom to top. The folds then remain closed until the subglottic pressure builds up enough to force them open again. Anatomically, the movement of the mucosal wave depends on the soft and compliant lamina propria and healthy layered structure of a normal VF. VF shaping and positioning are under neural regulation. The high innervation ratio of the human laryngeal muscles, estimated at 100 to 200 cells per motor unit, render the laryngeal muscles capable of a great degree of precise control as required for adjustment of speaking frequency and intensity (Kempster, 1988). It has been shown that the laryngeal muscles start to contract about 100 to 200 milliseconds prior to the onset of phonation. (Buchtal, 1964). The muscle that appears to be critical to varying VF adjustments and thus influencing phonation styles or frequency variation is the thyroarytenoid muscle since it forms the mass of the VFs. The frequency of vibration depends on the vibratory mass of both VFs, anterior to posterior tension, functional damping at high pitch, and subglottic pressure (Hirano, 1964) The cricothyroid muscle increases fundamental frequency (F0) by tensing the VF. The VF is stretched, elongated, thinned, and slightly adducted adjacent to midline as the VF is lowered within the larynx. These changes result in tightening the mucosal cover and increasing F0. The thyroarytenoid muscle generates the opposite effect as it thickens the VF and loosens the mucosal cover. (Gay, 1972). Additionally, it increases glottal resistance, contributing to vocal intensity as subglottic pressure increases. Therefore, vocal control is achieved by the coordinated efforts of respiratory, laryngeal, and articulatory muscles capable of producing great variations of tonal capabilities that characterize the human voice. These important physiological VF changes are readily appreciated under video stroboscopic viewing.

Video Stroboscopic Examination of the Larynx

Laryngeal video stroboscopy provides viewing of several vocal parameters. The stroboscobe flashes a light at a rate equal to or approximating the vibrating rate of the vocal folds (VFs) (Alberti, 1978; Kitzing, 1985). When the flash rate equals the vibrating rate of the VFs they will appear to stand still. This is because they are illuminated at the same phase as the vibratory cycle. On the other hand, if the flash rate of the strobe light is slightly different from the vibration rate of the VFs, then they would be illuminated at different phases of their cycle. This creates an illusion of a slow vibratory motion of the VFs thus allowing the human eye to observe vibratory details of the VFs. Through the process described above, one is able to stroboscopically view various parameters of the VFs such as, the degree of glottal closure, amplitude, VF margins and symmetry, mucosal wave, phase symmetry, periodicity and arytenoid and ventricular fold movement.

Glottal Closure

Glottal closure is a term that describes the extent and configuration of how the VFs adduct. The terms most commonly used to describe glottal closure include: complete closure, posterior chink, anterior chink, bowed, hour glass configuration, irregular closure and incomplete closure. According to Abou-Elsaad (2007), in normal habitual phonation, complete glottal closure is present along the vocal edges in males whereas in females there is a closure along the vibrating edges with a small posterior chink. Each vibration represents an opening and closing phase of the VFs. During conversational speech the VFs of the typical male vibrate on an average of 100-150 times a second, and for the typical female, the VFs vibrate an average of 180-250 times a second (Hollien, Dew, &

Phillips, 1971). These vibrations are acoustically known as fundamental frequency (F0). This allows one to assess how well the VFs are vibrating and to see any problems in their functioning. Imaging of VF vibratory function during phonation offers a vital role in diagnostic, therapeutic and surgical decisions during the management and treatment of most laryngeal dysfunctions (Mehta & Hillman, 2012).

Amplitude

Abou-Elsaad (2007) defines the movements of the VF musculature in the horizontal plane as amplitude. In very simple terms, amplitude refers to the maximum displacement or distance moved by a point on a vibrating body. In the case of VF movement, amplitude refers to the maximum displacement of the vocal folds from their mid-line. The amplitude of vibration is affected by VF stiffness and subglottal pressure. It is the horizontal excursion of the VFs that is directly assessed during stroboscopy.

Vocal Fold Margins and Symmetry

During phonation, the margins of the VFs may vary from normal smooth and straight to abnormally rough and irregular. Symmetry is the degree to which the pair of VFs are synchronized in vibratory pattern. Harmonization of VF vibration is indicative of VFs containing similar mechanical properties. This is evaluated during stroboscopy in terms of amplitude, horizontal excursion, and timing phases.

Mucosal Wave

The VF vibration is categorized into two types of movement. The transverse movement of the muscular body and the more vertical movement of the overlying layer. Stroboscopic examination makes it possible to observe both the transverse opening and closing of the folds as well as the vertical sliding motion of the mucosal cover. Normally, this rippling motion travels medio-laterally across the surface of the VFs and is known as the mucosal wave. Abou-Elsaad (2007) explains that the mucosal wave is a defining diagnostic phenomenon that can be viewed only stroboscopically. When evaluating the mucosal wave, a rating is assigned to whether it is present (normal) or absent. If the mucosal wave is defined as present, this is indicative that the wave travels approximately 50% across the upper fold when at F0. Absence of a mucosal wave can be the result of stiffness or immobility of the VF. Stroboscopy can be helpful in observing unique vibratory patterns within respective groups of people.

Periodicity

Periodicity is the measurement of stability and regularity of VF vibration while aperiodicity is the measurement of volatility and abnormality of the VF vibration. Aperiodicity can be observed secondarily since normal periodic vibration appears frozen in video recording, whereas aperiodic vibration demonstrates movement. Acoustic measurement can then symbolize aperiodicity and equate it to normative data. Periodicity is rated by the investigator as regular to always regular. Oscillation in a synchronized, rhythmic, metronome manner should be present in the VFs.

Arytenoid and Ventricular Fold Movement

Arytenoid movement is evidently observed stroboscopically and is rated as normal movement or poor movement by the investigator. Arytenoids should move symmetrically within both phases of adduction and abduction. Ventricular fold should be noted as normal to full compression. Within phonation, the ventricular folds should be reasonably immobile and apart. Videostroboscopy is an advanced technique that provides a magnified, slow motion view of the vocal cords in action. Rosen (2005) describes videostroboscopy as a dynamic voice assessment and evaluation tool that allows for a natural in vivo evaluation of the entire vocal tract during rest, vegetative activities, and phonation. Furthermore, stroboscopy allows the examiner insight into key vocal fold vibratory activity, specifically the physiologic activities. Earlier studies have attempted to perceptually and acoustically quantify differences in the vocal parameters of different ethnic groups, but with no conclusive evidence. Many of these studies lacked the more advanced technological assessment tools such as video stroboscopy, hence any conclusions drawn were merely speculative. Cox & Mueller (2002) in their extensive review of the available literature reported that although inconclusive, the data at that time pointed to the possibility of a difference in F0 between African American (AA) speakers and Caucasian speakers. For example, Awan & Mueller (1992) examined the speech production of African American kindergarteners and found no significant difference in speaking fundamental frequency between the Hispanic and Caucasian speakers, but the AA speakers demonstrated lower F0s compared to the other two groups. These findings contrasted with those of Xue and Mueller (1996), who found no difference in speaking fundamental frequency (SSF) of institutionalized AA elderly speakers compared to the established norms for a Caucasian cohort group. Most of these studies however lacked the full utilization of combined acoustic, perceptual and physiological data. For example, perceptual studies of ethnic speech characteristics suggest that it may be possible to identify ethnicity based on vocal quality (Irwin 1977; Lass, Mertz and Kimme 1978; Lass, Trapp, Baldwin, Sherbick and Wright (1983). One study that attempted to quantify perceptual data with acoustic findings was carried out by Walton and Orlikoff (1994). In this study, 50 Caucasian and 50 AA males recorded sustained phonation of the vowel /a/. The samples of both groups were paired and listeners were asked to identify the ethnicity of each speaker. Acoustic results showed that AA speakers had a greater frequency perturbation, and greater amplitude perturbation when compared to their counterparts. Perceptually, the judges identified the AA speech based on the vocal perturbation. This suggests that the listeners in this study were cued into ethnic identity on perceived vocal quality. Clearly perceptual and acoustic data may prove beneficial helping to understand possible differences that may be present in the vocal characteristics of different ethnic groups.

Acoustic and Perceptual Parameters

According to Boone, McFarland, Von Berg, and Zraick (2014), there are five aspects of voice that determine dysfunctionality. These are (1) loudness (volume of speaker’s voice), (2) hygienic production (whether or not a speaker’s voice sounds “healthy”), (3) pleasantness (smooth and easy or broken and harsh), (4) flexibility (ability to use appropriate intonation), and (5) representative (does the voice fit the speaker, such as man vs. woman). If a person’s voice violates any one or more of these five aspects, a voice disorder exists. The criteria for judging a voice as normal, abnormal, or belonging to a particular cultural group depend on the orientation of the person making the judgment, the manner in which the measurement is obtained, and the normative thresholds or "norms" to which the acquired acoustic data are compared. For example, if a speaker's pitch is not representative of their age, gender, or culture, then the listener's expectations of normality in voicing are violated. Vocal pitch is the perceptual representation of the VF vibration rate. The acoustic correlate of this perception is known as fundamental frequency (F0). As the speaker varies his or her intensity or loudness, F0 will also vary. F0 is acoustically represented as the number of vocal fold phases, which represents openings and closings that occur in one second otherwise known as Hertz (Vargo, 2012). The F0 of a person's voice is determined by a number of different factors such as the length, size, and tension of the vocal folds. To oscillate, the VFs are brought near enough together such that air pressure builds up beneath the larynx (Figure 1.1). The folds are pushed apart by this increased subglottal pressure, with the inferior part of each fold leading the superior part. Such a wave-like motion causes a transfer of energy from the airflow to the fold tissues (Lucero 1995). According to Zhang (2016), this propagates pitch through the vocal tract and is selectively amplified at different frequencies. This selective modification of the voice produces perceptible contrasts, which are used to convey different linguistic sounds and meaning.

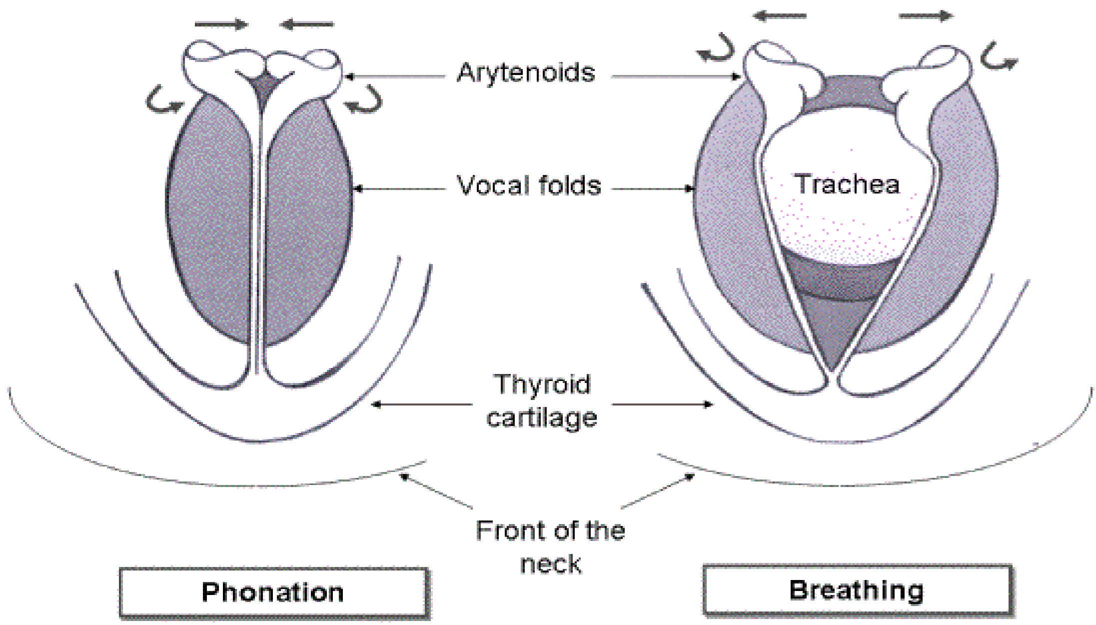

Figure 1.1 Vocal folds in the adducted position for phonation and the abducted position for breathing (Wolfe, Garnier, & Smith 2009)

Since phonation is fundamentally a product of the VFs, the F0 is primarily determined by the frequency of the vocal fold behaviors. VF natural vibration patterns depend on its geometry, including length, depth, thickness, stiffness, and stress conditions of the VFs. Shorter VFs tend to have low pitch vibration pattern while longer VFs tend to have a higher pitch vibration pattern. Thus, because of the small VF size, children tend to have the highest F0. The VFs of children are smaller than those of adult males and females. The current norms state that the F0 of a child's voice should be approximately 265 Hz (Williamson, 2006). According to Williamson the normal range for an adult female's F0 is approximately 225-250 Hz and an adult male's F0 is approximately 128-150 Hz. Adult male voices are usually lower pitched due to longer and thicker folds. According to Titze (1995) the male vocal folds are between 1.75 cm and 2.5 cm in length, while female vocal folds are between 1.25 cm and 1.75 cm in length. Over the years, there have been reports of differences in F0 (Hudson & Holbrook, 1981, 1982) across different ethnic groups. Walton and Orlikoff (1992) found that listeners were able to identify race at levels significantly greater than chance merely on the basis of one second portions excised from natural vowel productions. These results suggest that there are acoustic differences that may be attributable to race. These acoustic differences appear to be present in the signal and that listeners are able to recognize these differences in determining a speaker's race. Walton and Orlikoff (1992) also found shimmer and harmonic-to-noise ratio measures to differ significantly by race in the participants of their study. Prior studies uncovered significantly more prevoicing on average on the part of African American individuals. Prevoicing, in phonetics, is voicing or phonation before the onset of a consonant or the beginning with the onset of the consonant but ending before its release.

Acoustic Studies

Awan & Muller (1996) studied speech samples from groups of Caucasian, African American, and Hispanic kindergarten-age children and compared measures of mean speaking fundamental frequency (F0), maximum and minimum speaking F0, pitch sigma, and speaking range (in semitones). The results of these comparisons indicate that Caucasian kindergarten-age children do not differ significantly in terms of mean speaking F0 from their African American or Hispanic counterparts. The mean speaking F0 distribution for each group shows that the Caucasian group had a wide dispersion of mean speaking F0 productions that substantially overlap both the African-American and Hispanic distributions. In contrast, both Hispanic and African American groups present greater homogeneity in mean speaking F0, with a relatively large number of subjects within each group clustered about significantly different mean speaking F0 values

(236.40 Hz for African American vs. 248.71 Hz for Hispanic). Another finding of this study was that the Hispanic participants were observed to have significantly reduced speaking ranges compared to both Caucasian and African American subjects. It may be speculated that these findings can be attributed to differences in the intonational patterns used by the Hispanic children as compared to the Caucasian and African American children. Deutsch, Jinghong, Shen & Henthorn (2009) suggest that pitch differences may be due to cultural rather than racial factors. Different linguistic communities have expectations for their members and an individual's vocal characteristics can adapt to the surrounding community, even sounding similar in pitch. Deutsch et al (2009) examined pitch levels of females from two Chinese villages, each community being homogenous ethnically and culturally. The dialects of Mandarin spoken in the villages were also similar. The F0 values were clustered within each village but differed by approximately three semitones. These data support the claim that F0 is influenced by a representation acquired through long-term exposure to the speech of others (i.e., one's linguistic community) and suggests a cultural, rather than a physiological, influence on pitch. In contrast, Yamazawa & Hollien (1992) consider pitch variability to be directly related to different language structures. The speaking fundamental frequency (SFF) mean levels, variability and patterns of Japanese females were measured and compared with those of Caucasian speakers of American English. In this study, fifty-six young women, thirty-two Japanese and twenty-four Caucasian Americans read standard passages in either or both the Japanese and English languages. It was found that the Japanese speakers exhibited higher F0 than did the Americans for all speaking conditions and this contrast was statistically significant. Additionally and unlike the American speakers, most Japanese women exhibited bimodal SFF distribution patterns. It is judged that the observed differences in level and distribution result primarily from differences in the structure of the two languages. It is unclear as to the physiological characteristics of participants across both groups. It is well established that height, weight and age have significant bearings on F0. Perhaps studies that attempt to quantify F0 across ethnic groups may yield more compelling results if data reflected comparisons of perceptual, acoustic, and physiologic data.

Physiological Study

One physiological study conducted by Xue and Hao (2006) attempted to determine vocal tract measurements of speakers from different races. One hundred and twenty participants composed of Caucasian American, African American, and Chinese were involved in this study. Each racial group was comprised of twenty females as well as males. Race was found to be a significant variable for oral volume and total vocal tract volume. The Caucasian American speakers were found to have significantly smaller mean volume and mean total tract volume than Chinese speakers. There were no significant differences in the oral volume and total tract volumes between Caucasian American and African American speakers. Within the male speakers, Chinese males had significantly larger oral volume than both Caucasian American and African American male speakers. Chinese male speakers were also found to have significantly larger total vocal tract volume than both Caucasian American and African American male speakers. Within the female group results displayed that Chinese female speakers had significantly shorter vocal tract length than the other two groups. Caucasian American female speakers had significantly larger pharyngeal volume than both African American and Chinese female speakers. If race is a determining factor in pitch, the implication is that physiological differences may be the cause. An earlier physiological study done by Boshoff (1945) of 102 cadaveric larynges found that larynges of black South African males were stronger and more complex organs than the larynges of their Caucasian counterparts. Boshoff observed that the intrinsic laryngeal muscles were broader, stronger, and had more complicated points of attachment than those of Caucasian males. The finding was significant, as intrinsic laryngeal musculature has a direct impact on the F0 of the voice. Contraction and relaxation of these muscles allow the speaker to abduct, adduct, tense, and relax the vocal folds. Broader and stronger laryngeal muscles may increase the mass of the laryngeal framework, thus contributing to a lower pitch.

Definitions

It has been unclear from the prior studies described, how terms such as “race” and “ethnicity” for research purposes have been defined. Such terms can be fluid and may mean different concepts to different people. Therefore it is important to distinguish terms such as “race,” “culture,” and “ethnicity.” According to Bauman-Waengler (2012), race is considered an organic label that is defined in terms of observable physical features and biological characteristics, such as genetic composition. Culture is a way of life consisting of values, norms, beliefs, attitudes, behavioral styles, and traditions that have been developed by a group of individuals to meet psychosocial needs. Ethnicity refers to commonalities such as religion, nationality, and region. For the purpose of this study, the term “race” will be operationally defined as African American, Asian, Caucasian American, or Hispanic American and will be used interchangeably with the term “ethnicity.” African American (AA) will be used to designate an individual who was born in the United States and whose biological parents are both of African descent. The term Caucasian American (CA) will be used for those individuals who were born in the United States and whose biological parents are of European descent. In this study, Asian is defined as individuals whose biological parents were of Chinese, Filipino, and Myanmar. Hispanic American was defined as any individual whose biological parents were from Puerto Rico or Mexico and who were not of African or Caucasian descent.

Purpose

Not many studies have attempted to quantify possible vocal differences across racial boundaries and fewer studies have attempted to document physiological differences. The purpose of this study is to investigate the vocal characteristics of different ethnic groups in terms of acoustic and physiologic data. Specific questions that will be investigated are: (1) Are there acoustic differences in the vocal parameters across different ethnic groups? (2) Do the physiologic and perceptual data reflect these differences? Specific aspects that will be considered from an aerodynamic perspective are: F0 and perturbation and from a physiologic perspective: glottal closure patterns, amplitude and arytenoid and ventricular movements. This study will analyze descriptive data provided by individuals from different ethnicities.

Hypothesis

Broadly speaking, it is hypothesized there would be a difference in fundamental frequency across ethnic groups. Specifically, there will be African-American participants who demonstrate a lower fundamental frequency and there will be differences in vocal fold behavior across the various ethnic groups.

Rationale

As stated above, not many published studies have attempted to establish physiologic, acoustic, or perceptual differences in terms of vocal characteristics as a function of ethnicity. To be able to distinguish differences from a physiologic or acoustic perspective, can be beneficial from at least two perspectives. Firstly, the demographic changes in the United States have signaled the need for a greater understanding of treatment approaches for the speech language pathologist (SLP) and other health professionals based on evaluation, diagnosis, and intervention (Battle,1998). For example it has been well accepted in the field of speech-language pathology that the average F0 for females approximates 225 – 250 Hz range and for males 100 -150 Hz range. These norms have been established mostly for Caucasian individuals consequently, it is not known that they are necessarily accurate for all ethnicities. The significance of this is that it may very well be possible that different ethnic groups are being held to an inaccurate acoustic standard. If this turns out to be the case, it could significantly impact diagnosis and expectations for intervention. Secondly, establishing empirical evidence to support vocal normative expectations for different groups may be helpful in understanding cultural expectations and may also be assistive in diminishing some social communicative barriers.

CHAPTER II

METHODOLOGY

IRB Approval and Consent Form

The investigation, materials, and procedures for this study were approved by the Institutional Review Board (IRB) of Cleveland State University. The investigator recruited, screened, and collected data for all participants. All data were collected in the voice lab of the Cleveland State University Speech and Hearing Clinic. Informed consent was obtained from each participant following a detailed explanation of the nature of the study. (Appendix A).

Participants

Fifteen participants between the ages of 18 and 35 from Cleveland State University (CSU) in Cleveland, Ohio served as speakers for this study. Participants were recruited through word of mouth by the investigator. Gift cards worth $10 to Rascal House and Starbucks were provided as incentives for participation. The participants consisted of eight groups. The groups were as follows: one African American male (AAM), three African American females (AAF), two Asian males (ASM), one Asian female (ASF), three Caucasian males (CM), two Caucasian females (CF), one Hispanic male (HM), and two Hispanic females (HF). The physiological data for one Caucasian male, one African American female, as well as one Asian female participant had to be eliminated from the study because of inadequate visual images extracted from the videostrobe but their acoustic data were retained for the study. Two individuals were recruited as listening participants for the perceptual task.

Inclusionary Criteria

To meet the inclusionary criteria, prospective participants had to be between the ages of 18 – 35 and had to be members of one of the designated ethnic groups. Within similar height and weight categories. For purposes of the investigation, each ethnic group is defined as having both parents of the same descent as previously stated.

Exclusionary Criteria

Prospective participants who considered themselves biracial were excluded from the study. Based on individual account, prospective participants were screened for conditions and daily living habits known to alter the pitch of the voice. Specifically, those who reported a history of asthma, acid reflux, or laryngeal pathologies, and current smoking were excluded from the study. Individuals who currently had a head cold, upper respiratory infection, or chronic sinusitis were excluded. These precautions were taken as a measure to ensure the validity and reliability of the results obtained from the study. (See Screening tool Appendix B).

Acoustic Procedures

All experimental procedures performed and all data collected took place in the voice laboratory of the Speech and Hearing Clinic at CSU. The following quantitative measures were collected on each participant using the Visi-Pitch IV, Model 3950 issued by KayPentax: Fundamental frequency, jitter, shimmer, voice turbulence index (VTI), and noise-harmonic ratio (HNR). The participant was positioned approximately 2 inches from a standard microphone attached to the Visi-pitch and asked to sustain the vowel /a/ for 4 seconds. Participants' speech samples were recorded using the Visi-Pitch Multi-Dimensional Voice Program (MDVP). The MDVP extracts four different acoustic parameters from a single voiced sample. These are jitter, shimmer, VTI, and HNR. This program also allows the investigator to retrieve the fundamental frequency (F0). Participants were asked to sustain the vowel /a/ for four seconds. Instructions to the participants can be found in Appendix C. The vowel /a/ was chosen because of the open position of the vocal tract. Sustained /a/ is a commonly used vowel in research and clinical practice to determine habitual pitch and fundamental frequency (Fairbanks, 1940). The participants were asked to perform this task twice to ensure proper recording of the sustained vowel.

Physiological Procedures

KayPentax Rhino-Laryngeal Stroboscope™ model RLS 9100B was used to conduct videostroboscopy in order to complete a physiological examination of the vocal folds. Videostroboscopy is a routine procedure used by otolaryngologists and speech-language pathologists as a tool for assessment of vocal fold damage, dysfunction, and vocal fold oscillation. A rigid endoscope attached to a camera from the videostroboscopy instrument was placed at the base of the participants tongue while the participant vocalized /a/ and then /ee/. The endoscope uses a synchronized, flashing light passed through a rigid telescope. The equipment in turn saved and analyzed data obtained with audiovisual recordings for each participant. The investigator directly observed the participants' laryngeal structures, as well as the vocal fold parameters.

Perceptual Procedures

One-second samples were extracted from the middle portion of each of the 15 speakers' vowel prolongations. The middle portion of the sample was used in order to minimize any frequency variability that occurs during voice onset time and any possible glottal fry at the end of the vowel sample. The samples were randomly paired, with two samples from each race. A one-second interval separated the paired samples of voices and the order of the voices in each pair was randomized. Listening participants were asked to make a perceptual judgement of the race of each speaker with the understanding that there were two different ethnic groups in each sample. This open-choice method of presentation and data collection was chosen so the listening participants were able to compare the vowel samples with minimal delay between the samples. It was reasoned by Walton and Orlikoff (1994) that if the listening participants guessed the speaker's race after each vowel sample, they would be correct approximately 50% of the time and that a forced-choice method of presentation helps eliminate this variable. Additionally, this method of pair comparison was chosen because this study attempts to focus on the listeners' ability to make comparisons between voices to determine the race of the speaker, not the ability to rely on judgments from memory of linguistic qualities of various races.

Specific directions provided to the listening participants can be found in Appendix D Listeners were provided with a form and instructed to write the order of presentation of the paired voice samples with African American male denoted by "AAM", African American Female denoted by "AAF", Asian Male by "ASM", Asian

Female by “ASF”, Caucasian Male by “CM”, Caucasian Female by “CF”, Hispanic Male by “HS”, and Hispanic Female by “HF.”

CHAPTER III

RESULTS

This investigation into patterns of vocal fold (VF) closure across four ethnic groups yielded results from stroboscopic assessment, acoustic, and perceptual analysis. The results were analyzed in a descriptive manner.

Videostroboscopy Results

AAM #1

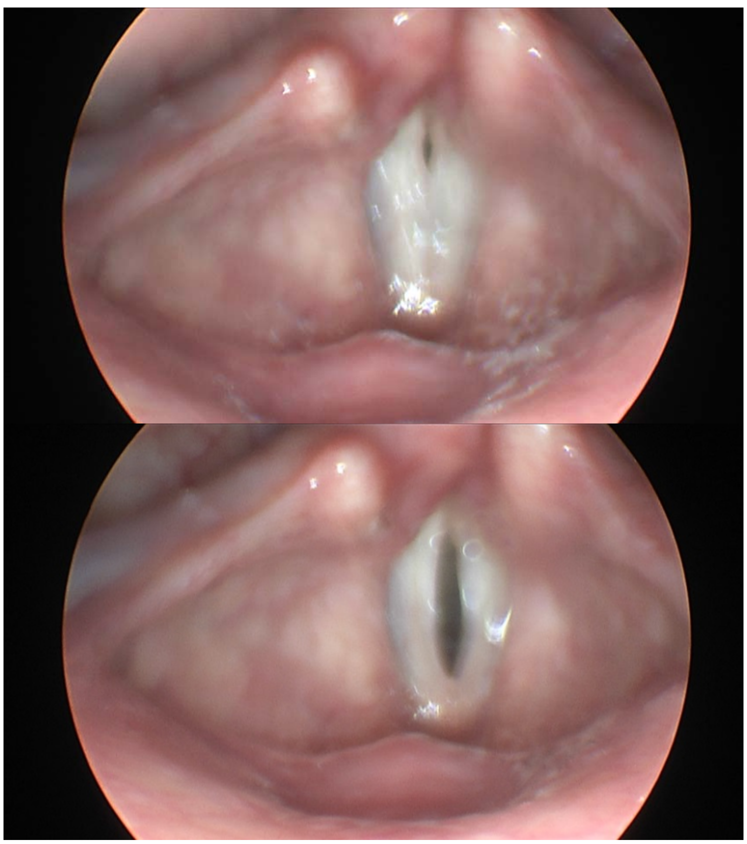

Figure 3.1 AAM #1 Age 27 Top: adducted view; bottom: abducted view

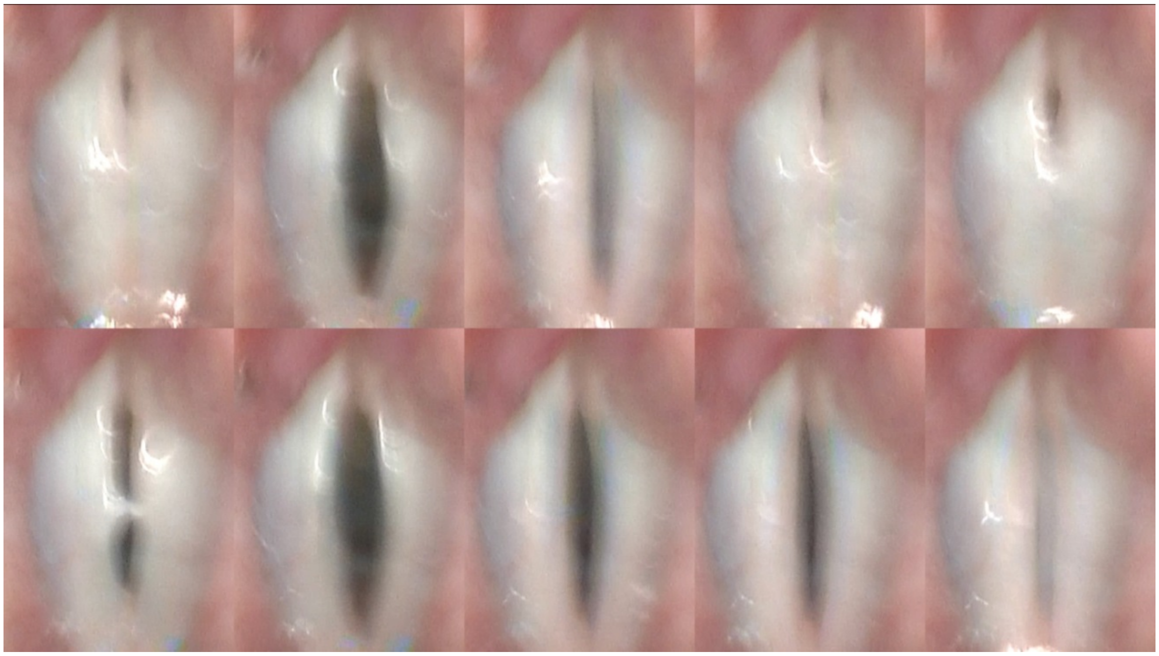

Figure 3.2 AAM #1 Montage View

In Figure 3.1 the AAM #1 had normal color and appearance of the VFs. Adequate VF margins and symmetry were present. A normal mucosal wave was also evident under video. The glottal closure was rated as incomplete with a prominent posterior chink as well as tightening of the interarytenoid space with a slight asymmetry of the arytenoids. In Figure 3.2 the montage view depicts an incomplete glottal closure pattern with a prominent posterior chink.

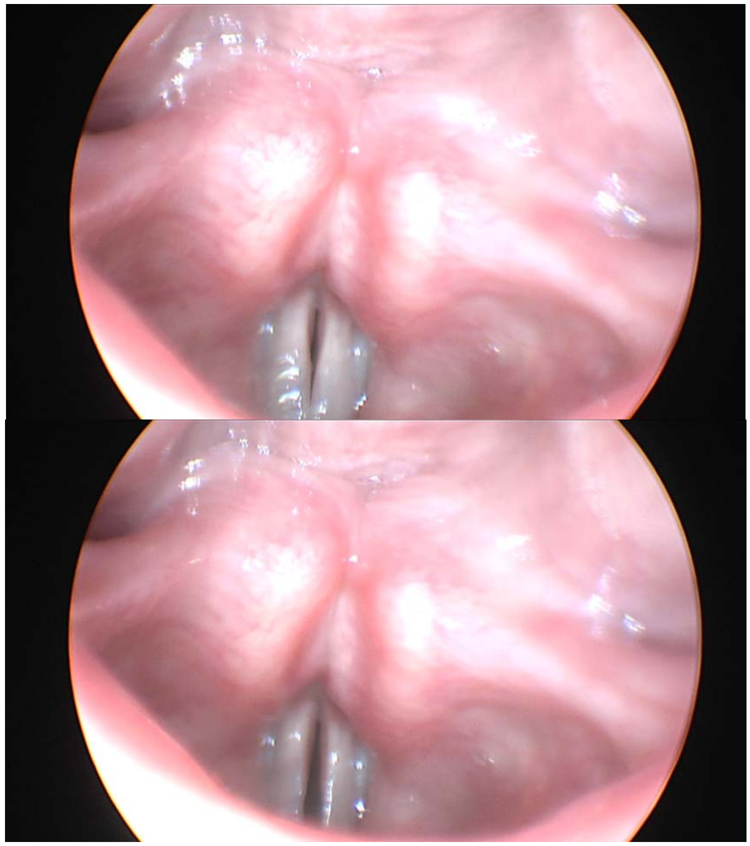

Figure 3.3 AAF #1 age 25 Top: adducted view; bottom: abducted view

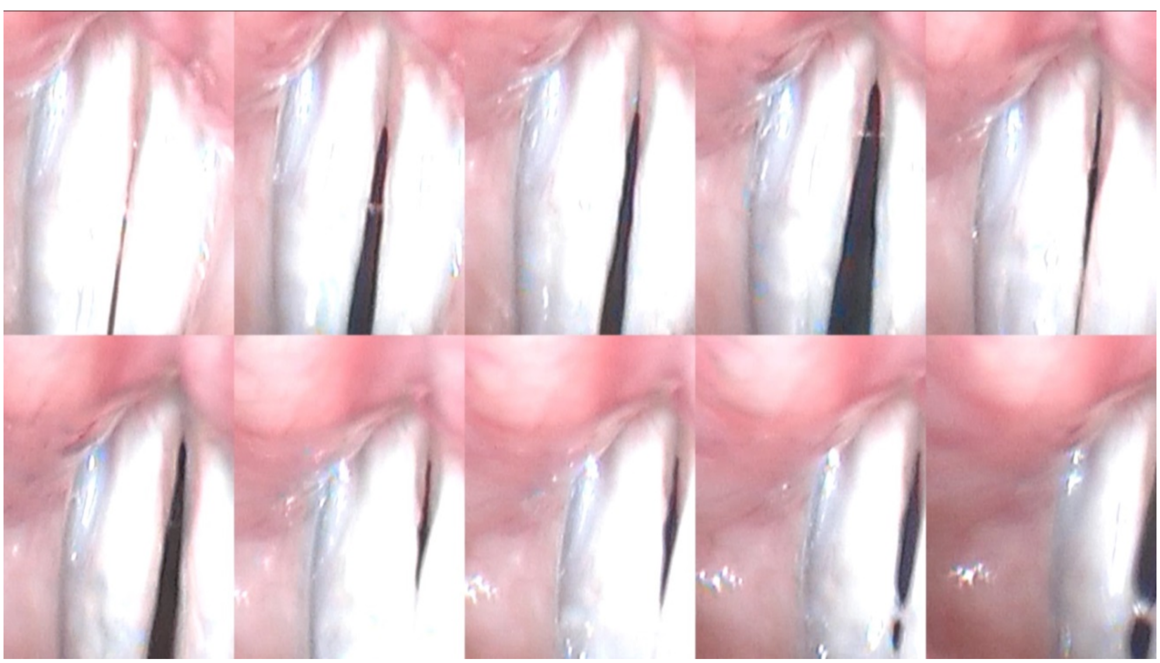

Figure 3.4: AAF #1 Montage view

In Figure 3.3 AAF #1 had adequate VF margin and symmetry. A normal mucosal wave was present. The glottal closure pattern was incomplete with a prominent posterior chink (top view). Also noted was the thickness of the VFs compared to CF #1 and #2 counterparts. The upper montage view depicts incomplete glottal closure with a more prominent posterior chink.

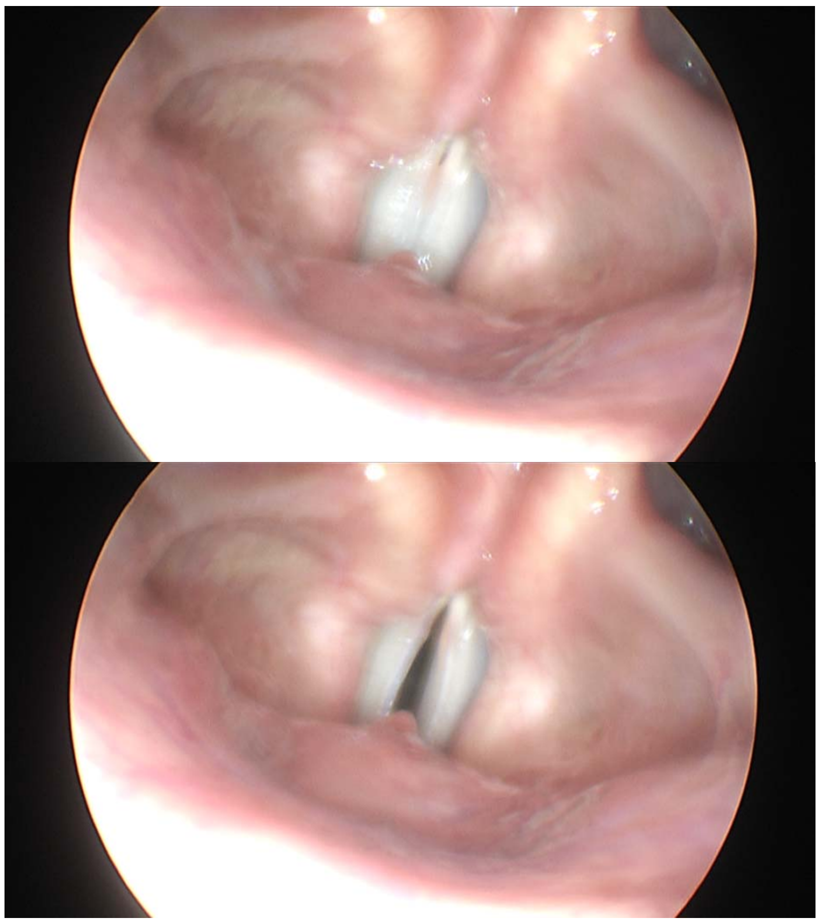

Figure 3.5 AAF #2 age 31 Top: adducted view; bottom: abducted view

In Figure 3.5 AAF #2 presented with normal VF appearance and color. Adequate VF margins and symmetry were observed throughout the examination. A normal mucosal wave occurred during phonation. There was incomplete glottal closure with a slight posterior chink (top view). Also noted was the thickness of the VFs.

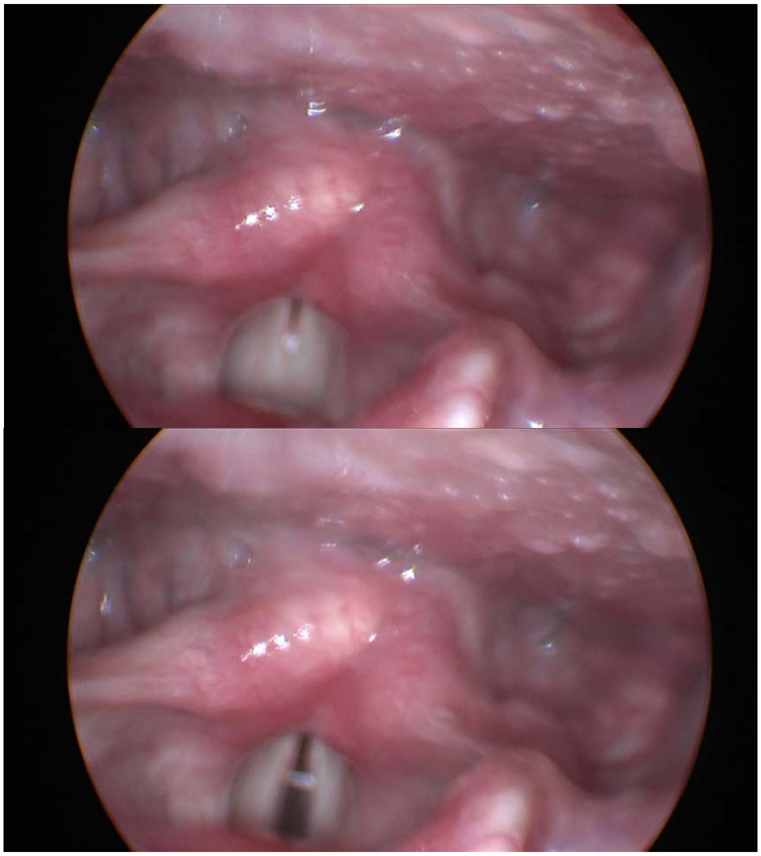

Figure 3.6 ASM #1 age 21 Top: adducted view; bottom: abducted view

In Figure 3.6 ASM #1 presented with normal VF margins and symmetry. A normal mucosal wave was observed during phonation. The glottal closure was incomplete with a prominent posterior chink (top view) was observed consistently throughout the examination. There was squeezing of the interarytenoid space during phonation and some asymmetry noted. The right arytenoid appeared larger and was further across the midline during phonation and appeared edematous.

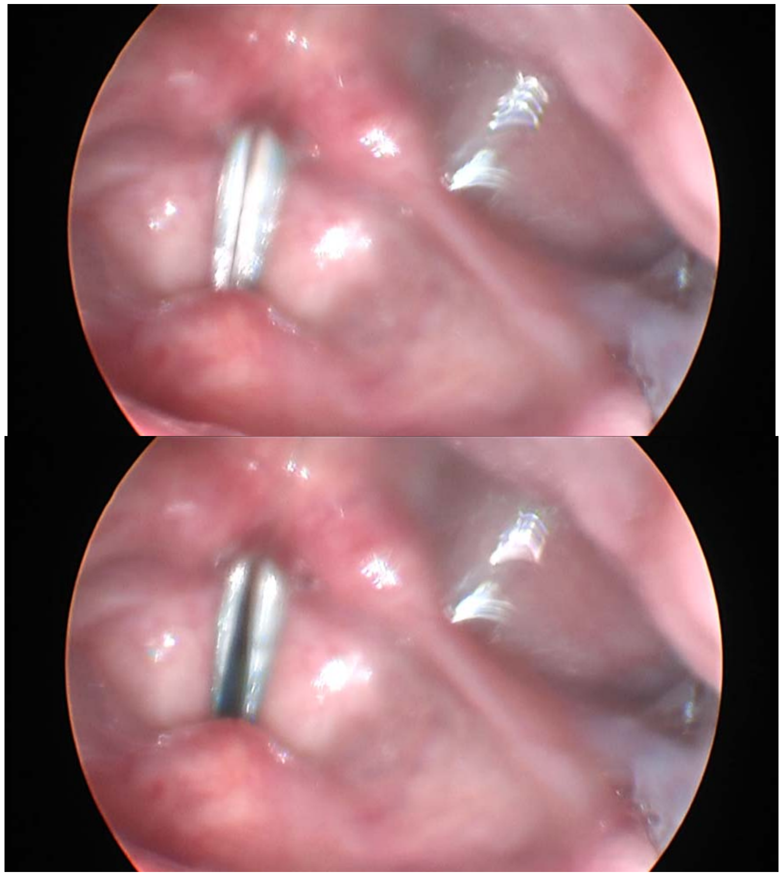

Figure 3.7 ASM #2 age 21 Top: adducted view; bottom: abducted view

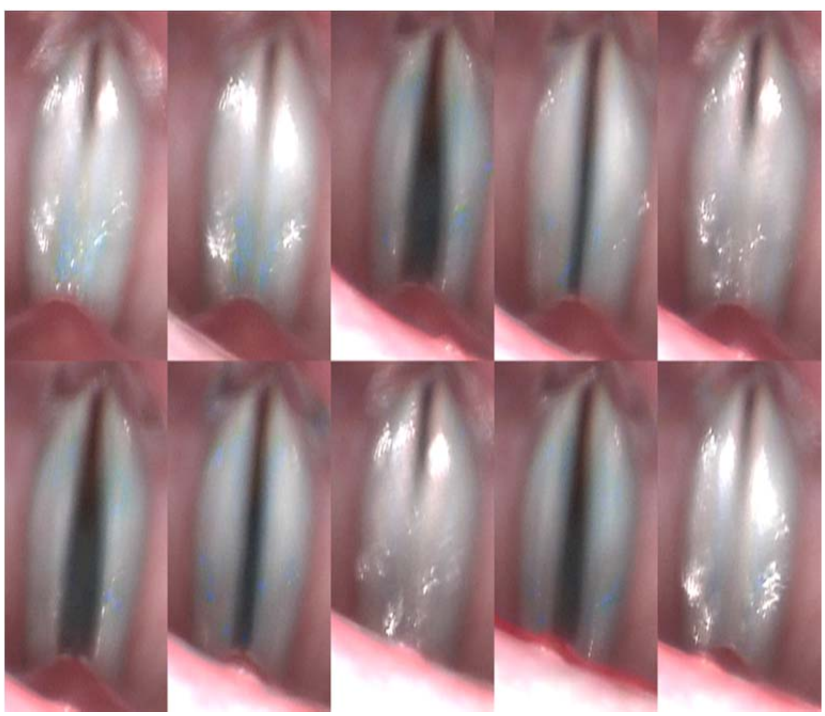

Figure 3.8: ASM #2 Montage view

In Figure 3.7 ASM #2 presented with normal VF color and appearance. Adequate VF margins and symmetry were evident during the examination. A normal mucosal wave was observed. The glottic closure pattern was incomplete with a posterior chink (top view) present. The arytenoids were symmetrical but a tightened interarytenoid space was observed during phonation. In Figure 3.8 the upper montage view demonstrates an incomplete glottal closure with a prominent posterior chink.

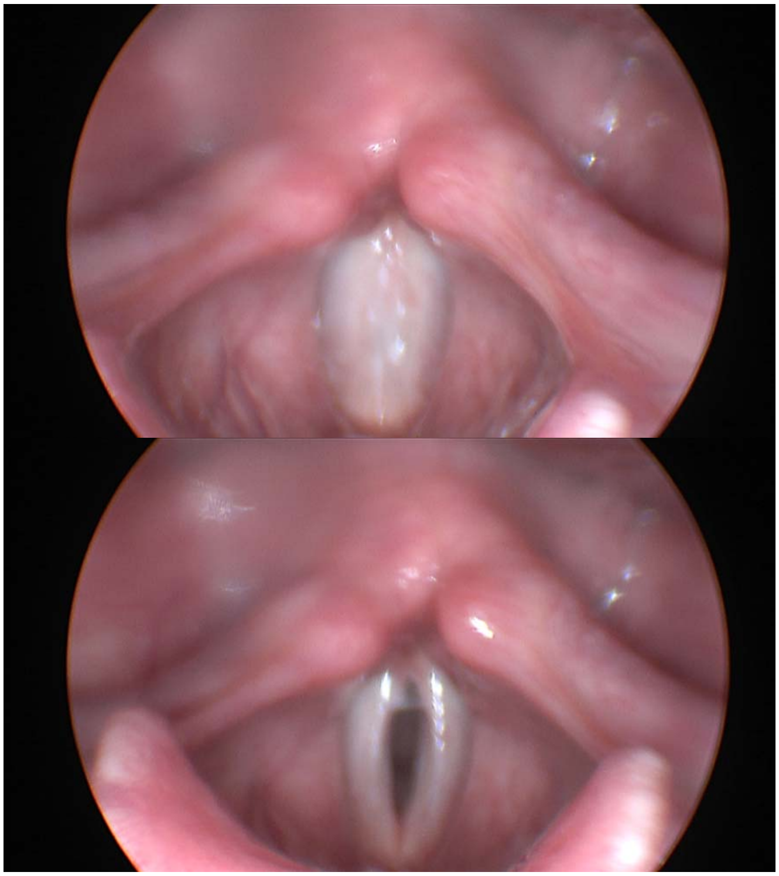

Figure 3.9 CM #1 age 21 Top: adducted view; bottom: abducted view

In Figure 3.9 Caucasian male participant age 21 presented with normal VF color and appearance. Normal VF margins and symmetry were present in the examination. Prominent amplitude of the mucosal wave was observed. The glottal closure pattern was complete no posterior chinks observed throughout the examination as seen in the montage view in Figure 3.10.

Figure 3.10: CM #1 Montage view

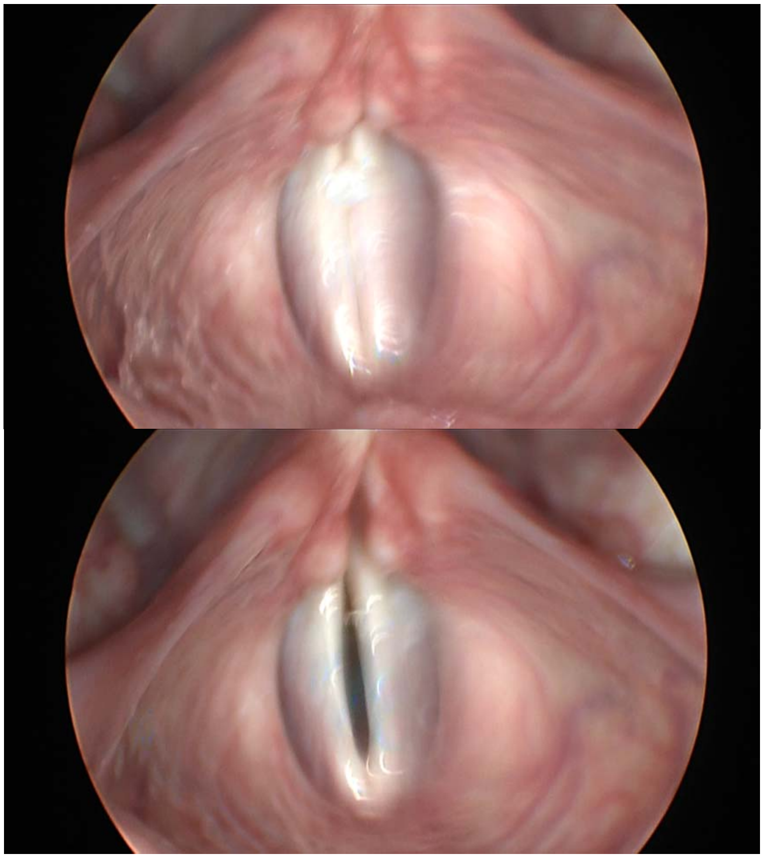

Figure 3.11 CM #2 age 25 Top: adducted view; bottom: abducted view

In Figure 3.11 CM #2 presented with normal VF margins and symmetry. A normal mucosal wave was present during the examination. The glottal closure pattern was complete.

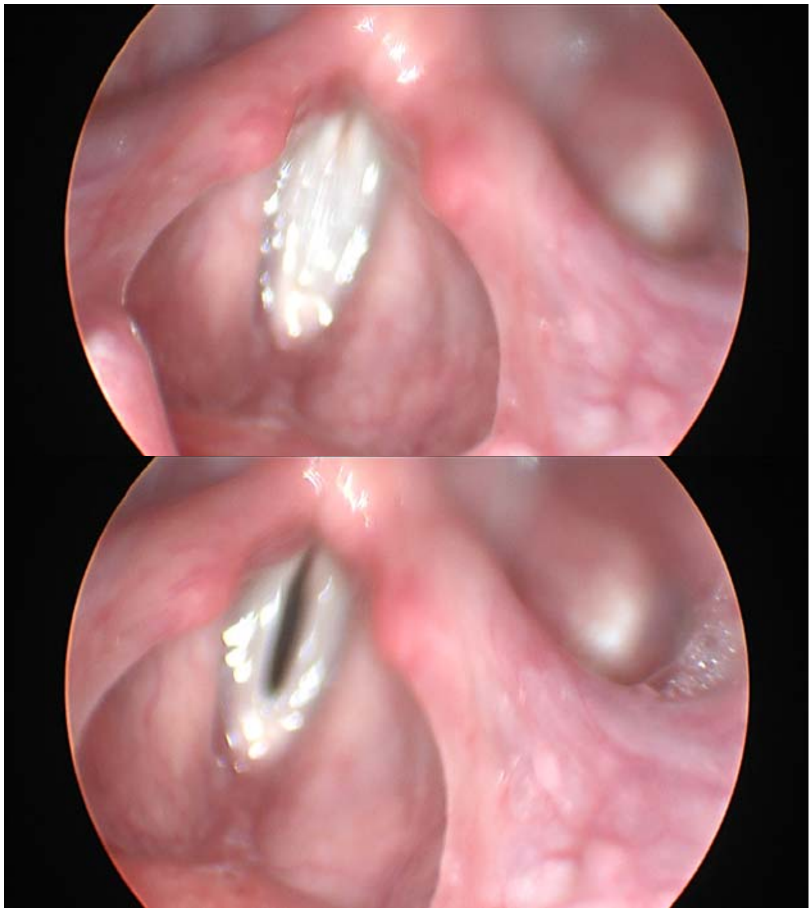

Figure 3.12 CF #1 age 19 Top: adducted view; bottom: abducted view

In Figure 3.12 CM #2 presented with normal VF appearance and symmetry. A normal mucosal wave was observed. There was incomplete glottal closure with of a posterior chink present. There was also noted marked interarytenoid squeezing. The upper video montage depicts good glottal closure and an obvious posterior chink in Figure 3.13.

Figure 3.13: Montage view of CF#1

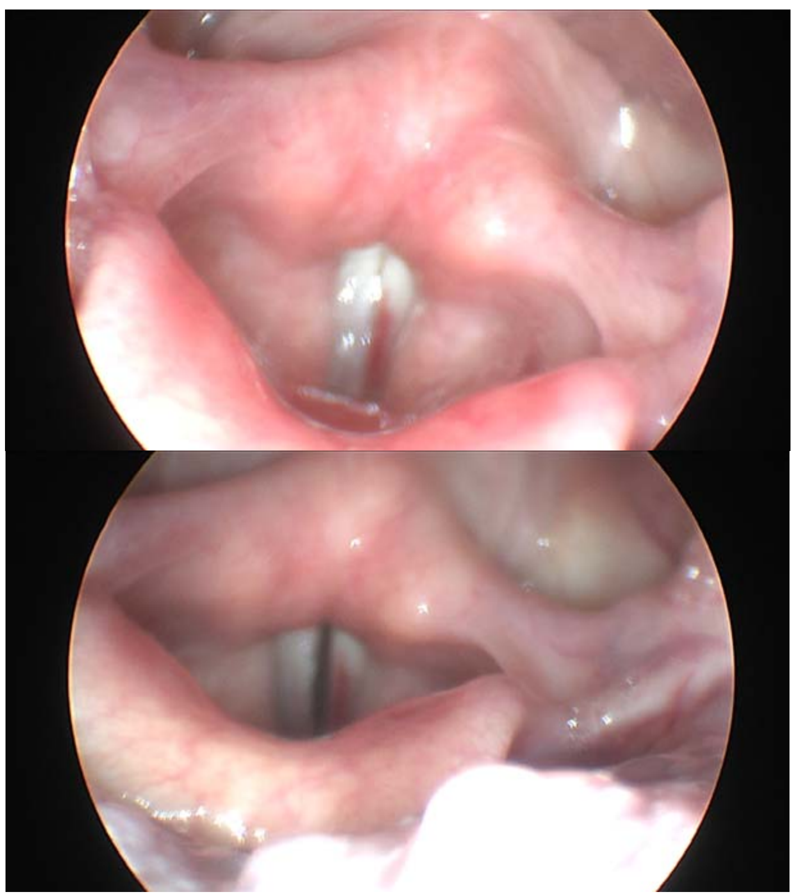

Figure 3.14 CF #2 age 24 Top: adducted view; bottom: abducted view

In Figure 3.14 CF #2 presented with redness on the left VF. The participant had been coughing recently. The VF margins appeared normal in appearance. A normal mucosal wave was observed in the examination. There was a small posterior chink present. Also noted was a prominent interarytenoid squeeze.

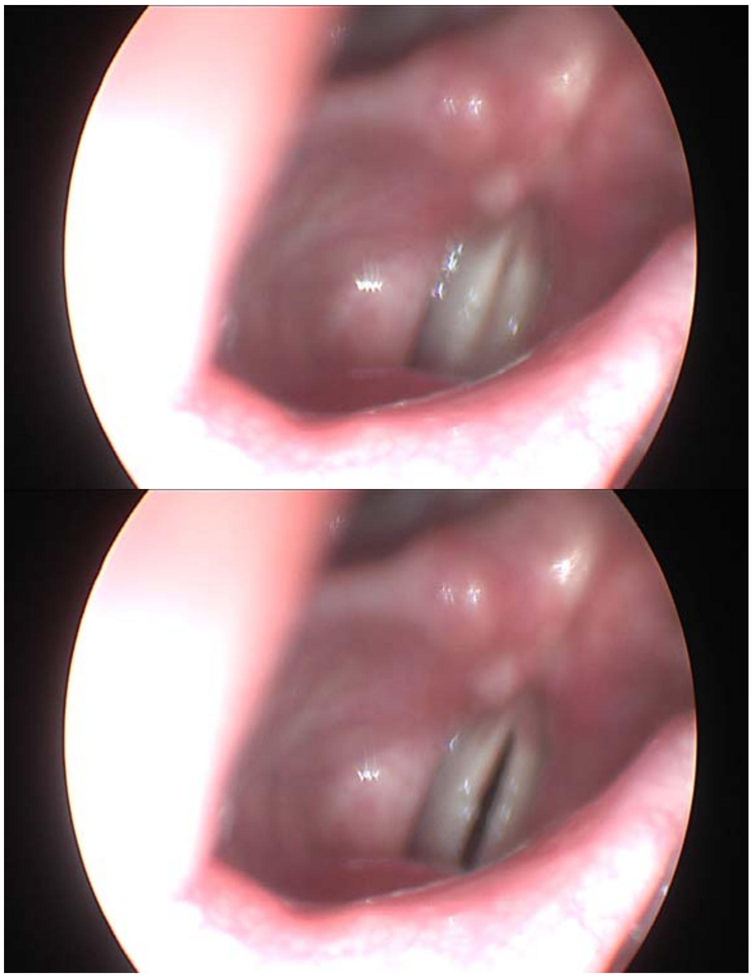

Figure 3.15 HM #1 age 19 Top: adducted view; bottom: abducted view

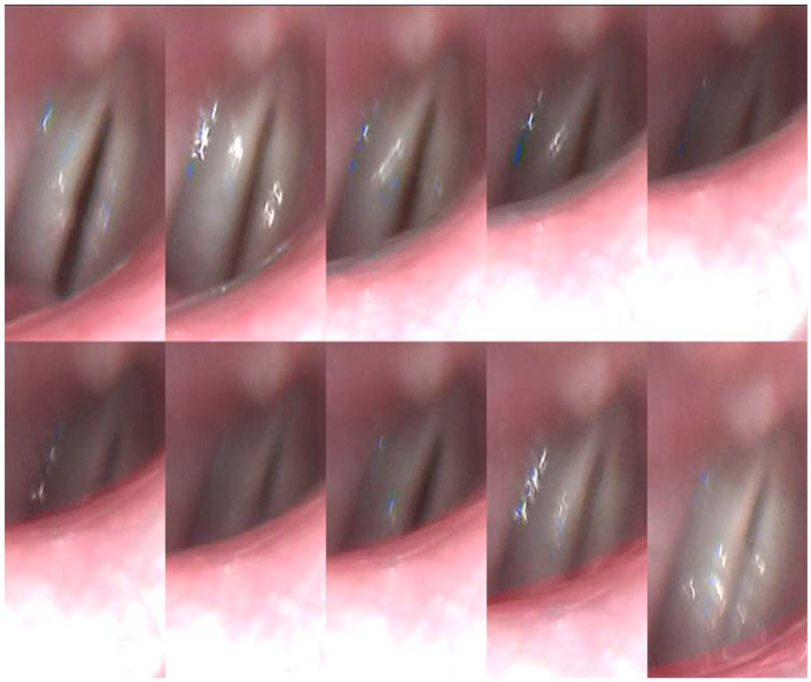

Figure 3.16: HM #1 Montage view

In Figure 3.15 HM #1 presented with normal VF margins and symmetry. A normal mucosal wave was observed. The glottal closure was incomplete; there was a prominent posterior chink that was observed consistently throughout phonation. This is well seen in the montage view. The VFs appear thinner in length within the montage view in Figure 3.16.

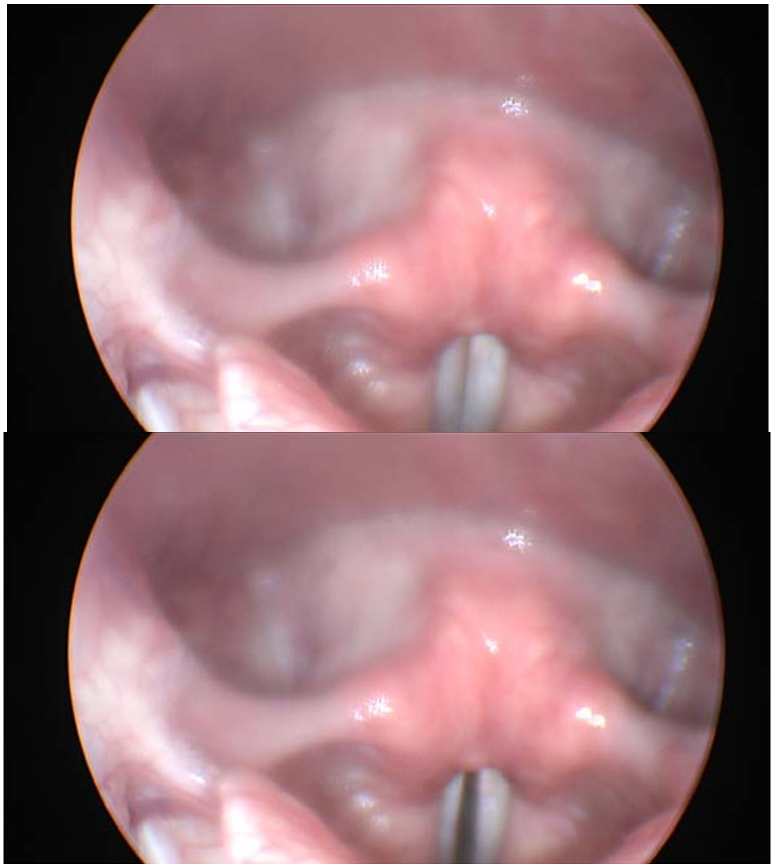

Figure 3.17 HF #1 age 24 Top: adducted view; bottom: abducted view

In Figure 3.17 HF #1 age 24 presented with normal VF color and appearance. Adequate VF margins and symmetry was present. A normal mucosal wave was observed during the examination. The VF closure was complete during phonation. The arytenoids were symmetrical in appearance but some interarytenoidal squeezing was present. The montage view depicts a complete glottal closure pattern in Figure 3.18.

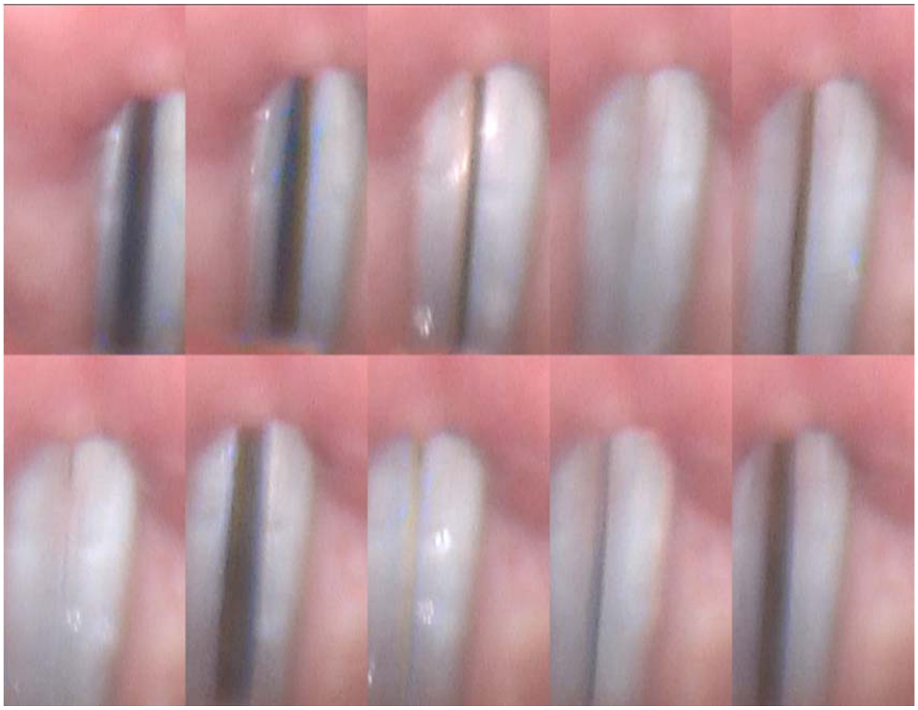

Figure 3.18: HF #1 Montage view

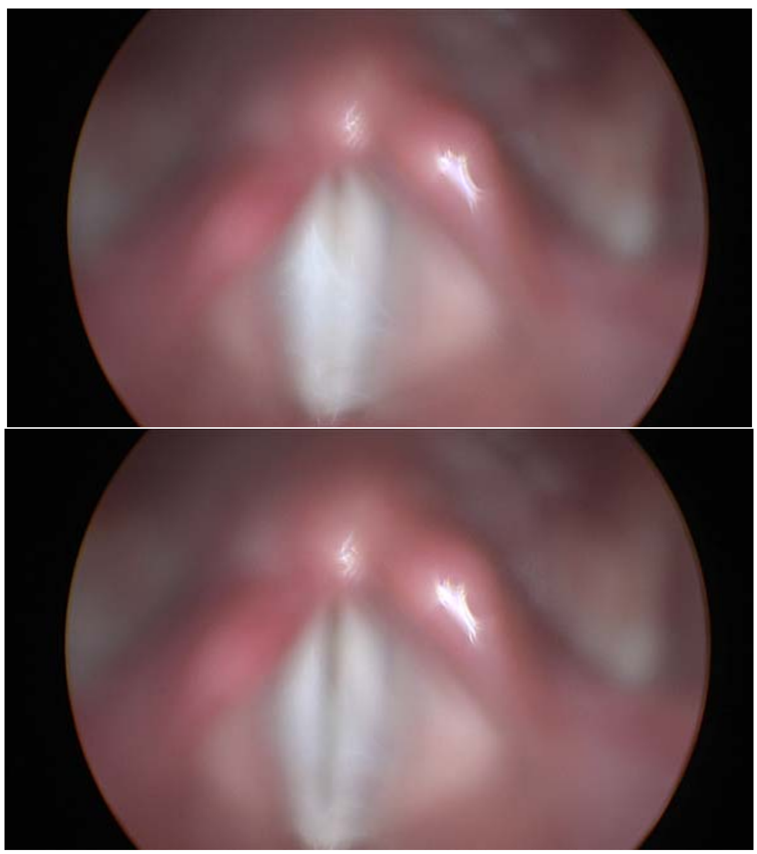

Figure 3.19 HF #2 age 23 Top: adducted view; bottom: abducted view

In Figure 3.19 HF #2 presented with normal VF color and appearance. VF margins and symmetry were normal. The glottal closure pattern presented as incomplete with an obvious posterior chink present during the examination. The arytenoids were symmetrical but some squeezing in the interarytenoid space was noted during phonation.

Acoustic Results

In the sustained vowel task the entire four second sample was analyzed to determine F0, RAP, shimmer, NHR, and VTI.

Table 3.1: Mean Vocal Parameters for Sustained Vowel Task (all participants)

Groups	Number of Participants	Mean F0 Values (in Hertz)	Mean Relative Average Perturbation (RAP)= Jitter	Mean Shimmer	Mean Noise to Harmonic Ratio (NHR)	Mean Voice Turbulence Index (VTI)
AAM	1	149.19	.764	1.850	.142	.053
AAF	3	205.54	.149	1.731	.117	.045
ASM	2	168	.286	2.723	.133	.049
ASF	1	193.9	.165	2.394	.114	.053
CM	3	133.33	.319	3.391	.127	.052
CF	2	275.5	.363	2.105	.103	.044
HM	1	222	.481	1.850	.115	.044
HF	2	260.5	.392	3.438	.117	.061

AAM= African-American Male; AAF= African-American Female; ASM= Asian Male; ASF= Asian Female; CM= Caucasian Male; CF= Caucasian Female; HM= Hispanic Male; HF= Hispanic Female.

Table 3.2: Male Mean Vocal Parameters for Sustained Vowel Task

Groups	Number of Participants	Mean F0 Values (in Hertz)	Mean Relative Average Perturbation (RAP)= Jitter	Mean Shimmer	Mean Noise to Harmonic Ratio (NHR)	Mean Voice Turbulence Index (VTI)
AAM	1	149.19	.764	1.850	.142	.053
ASM	2	168	.286	2.723	.133	.049
CM	3	133.33	.319	3.391	.127	.052
HM	1	222*	.481	1.850	.115	.044

AAM= African-American Male; AAF= ASM= Asian Male; CM= Caucasian Male; HM= Hispanic Male;. *= highest F0

Male Participants Mean Fundamental Frequency (F0) Parameters

Within the F0 measurements the HM group has the highest average F0 value of 222 Hz. The ASM group has the second highest average F0 value of 168 Hz. An average difference of 56 Hz separates the HM and ASM groups. The AAM group presents with the third highest average F0 value of 149 Hz. The average F0 difference between the HM and AAM groups is 72.81 Hz while a difference of 18.81 Hz exists between the ASM and AAM groups. The CM group have fourth highest average F0 value at 133.33 Hz. Compared to the CM group, a difference of 88.67 Hz, 34.67 Hz, and 15.86 Hz exist for the HM, ASM, and AAM groups.

Male Participants Mean Relative Average Perturbation (RAP) Parameters

Within the RAP measurements, the AAM individual has the highest average value at .764%. The HM individual has the second highest average RAP value at .481%. A difference of .283% separates the AAM and HM individuals. The CM group has the third highest average RAP value at .319%. Compared to the CM group, difference of .445% and .162% exists for the AAM and ASM groups. The ASM group has the fourth highest average RAP value at .268%The ASM group demonstrates differences of .496% .213% and .051% for the AAM, HM, and CM groups.

Male Participants Mean Shimmer Average Parameters

Within the shimmer measurements, the CM group has the highest average value at 3.391%. The HM and AAM groups have the second highest shimmer average value at 1.850%. A difference of 1.541% is demonstrated when comparing the CM group to the HM and AAM groups. The ASM group have the third highest shimmer average value at

1.731%. The ASM group demonstrates differences of 1.66 % compared to the CM group while a .119 % difference exists compared with the HM and AAM groups.

Male Participants Mean Noise to Harmonic Ratio (NHR) Parameters

Within the NHR measurements, the AAM individual has the highest average value at .142%. The CM group has the second highest average NHR value at .127%. A difference of .015% separates the AAM and CM groups. The ASM group has the third highest average NHR value at .117%. A difference of .025% and .01% separates the ASM group from the AAM and CM groups. The HM individual has the fourth highest average VTI value at .115%. When comparing the HM group, a difference of .027%, .012%, and .002% between the AAM, CM, and ASM groups.

Male Participants Mean Voice Turbulence Index (VTI) Parameters

Within the VTI measurements, The AAM individual has the highest average value at .053%. The CM group has the second highest average VTI value at .052%. When compared, a difference of .001% exists between the AAM and CM groups. The ASM group has the third highest value at .045%. The ASM group demonstrates a difference of .008% and .007% compared to the AAM and CM groups. The HM group has the fourth highest average value at .044%. The HM group demonstrates the difference of .009%, .008%, and .001% when compared to the AAM, CM, and ASM groups.

Table 3.3: Female Mean Vocal Parameters for Sustained Vowel Task

Groups	Number of Participants	Mean F0 Values (in Hertz)	Mean Relative Average Perturbation (RAP)= Jitter	Mean Shimmer	Mean Noise to Harmonic Ratio (NHR)	Mean Voice Turbulence Index (VTI)
AAF	3	205.54	.149	1.731	.117	.045
ASF	1	193.9	.165	2.394	.114	.053
CF	2	275.5*	.363	2.105	.103	.044
HF	2	260.5	.392	3.438	.117	.061

AAM= African-American Male; AAF= African-American Female; ASM= Asian Male; ASF= Asian Female; CM= Caucasian Male; CF= Caucasian Female; HM= Hispanic Male; HF= Hispanic Female. *= highest F0

Female Participants Mean Fundamental Frequency (F0) Parameters

Within the F0 measurements, the CF group has the highest average value at 275.5 Hz. The HF group has the second highest average value at 260.5 Hz. A difference of 15 Hz separates the CF and HF groups. The AAF group has the third highest average value at 205.54 Hz. A difference of 70 Hz and 54.96 Hz is demonstrated when comparing the AAF group to the CF and HF groups. The ASF group has the fourth highest average value at 193.9 Hz. The ASF group demonstrates a difference of 81.6 Hz, 66.6 Hz, and 12 Hz when compared to the CF, HF, and AAF groups.

Female Participants Mean Relative Average Perturbation (RAP) Parameters

When examining RAP measurements, the HF group has the highest average value of .392%. The CF group has the second highest average value of .363%. A difference of .029% separates the HF and CF groups. The ASF group has the third highest average value of .165%. The ASF group demonstrates a difference of .227 % and .198% compared to the HF and CF groups. The AAF group has the fourth highest average value of .149%. The AAF group demonstrates a difference of .243%, .214%, and .016% compared to the HF, CF, and ASF groups.

Female Participants Mean Shimmer Parameters

When examining shimmer measurements, the HF group has the highest average value of 3.438%. The ASF group has the second highest average value of 2.394%. A difference of 1.004 % separates the HF and ASF groups. The CF group has the third highest average value of 2.105%. The CF group demonstrates a difference of 1.333% and .289% compared to the HF and ASF groups. The AAF group has the fourth highest average value of 1.731%. The AAF group demonstrates a difference of 1.707%, .663%, and .374% when compared to the HF, ASF, and CF groups.

Female Participants Mean Noise to Harmonics Ratio (NHR) Parameters

When examining NHR measurements, the AAF and HF groups have the highest average value of .117%. The ASF group has the second highest average value of .114%. A difference of .003% exists when comparing the AAF and HF groups to the ASF group. The CF group has the third highest average value of .103%. The CF group demonstrates a difference of .014% when compared to the AAF and HF groups. A difference of .011% separates the CF and ASF groups.

Female Participants Mean Voice Turbulence Index (VTI) Parameters

When examining VTI measurements, the HF group has the highest average value of .061%. The ASF group has the second highest average value of .053%. A difference of .008 % exists between the HF and ASF groups. The AAF group has the third highest average value of .045%. The AAF group demonstrates a difference of .016% and .008% compared to the HF and ASF groups. The CF group has the fourth highest average value of .044%. The CF group demonstrates a difference of .017%, .009%, and .001% when compared to the HF, ASF, and AAF groups.

Perceptual Results

A total of two listeners participated in this portion of the study. Both are inexperienced listeners and have no knowledge of voice disorders. Both listeners are male. A percentage was calculated for the accuracy of their perceptual judgements. They listened to a one second segment of a vowel production.

Table 3.4: Accuracy for Individual Listening Participants

Listener	Male Samples	Female Samples	Percent Correct
1	4/4	4/5	88%
2	3/4	4/5	77%

The listeners were approximately 83% accurate overall in their judgment of the speakers race. The listeners were more accurate judging the female voice samples compared to the male voice samples.

CHAPTER IV

DISCUSSION

The examples of the aforementioned cases were descriptively analyzed in terms of videostroboscopic and acoustic data.

Videostrobic Data

The majority of the participant's demonstrated normal physiology of the VFs and laryngeal anatomy as expected. Glottal closure was examined on a two point scale in terms of complete closure and incomplete closure. Incomplete closure was designated by the presence of a posterior glottal chink. Complete closure was designated as complete VF closure during cycles of vibration. All African-American participants both male and female in addition to all female participants across all groups exhibited a posterior chink in VF closure patterns (Figures 3.1- 3.4). Studies have shown that females tend to have a posterior chink in their glottal closure pattern compared to males (Volkar, 2017). However posterior chink was noted in the glottal closure pattern across all males (Figures 3.1, 3.6, 3.7, 3.14) with the exception of the Caucasian male group.

AAM #1 Posterior Chink **ASM #1 Posterior Chink**

ASM #2 Posterior Chink **HM #1 Posterior Chink**

Figure 4.1: Currently there are no norm studies that have attempted to describe VF closure across ethnic groups. Consequently this data in Figure 4.1 above is unexplored.

Acoustic Data

The acoustic data measurements taken were F0, Relative Average Perturbation (RAP), shimmer, Noise-to-Harmonic Ratio (NHR), and Voice-Turbulence Index (VTI). The fundamental frequency is a measure of how high or low a person’s voice can be described in terms of pitch. The RAP (jitter) value data demonstrates that the average of three glottal cycles to a given period. This translates to the amount of hoarseness in the voice. During a sustained vibration the VFs will show a slight change in amplitude from cycle to cycle. This is called shimmer. Normal speakers will have a small amount of shimmer depending on the vowel used and the gender of the person. It is a measurement of turbulence noise in the voice signal. Colton & Casper (1996) state that the average males mean shimmer would be .33%. Females would have a mean shimmer of .25%. These fluctuations give rise to a source of sound pressure in the vocal tract at the point where the turbulence occurs. The NHR is a measurement of spectral noise which can be due to amplitude and frequency variations. Noise in the vocal signal can modify how to vocal pitch is heard. The VTI is the measurement of the “breathiness.” All of these parameters extracted from the participants are considered normal with no indication of vocal pathology.

Male Participants Acoustic Data

Table 4.1: Mean Male Vocal Parameters for Sustained Vowel Task

Group	F0	RAP	Shimmer	NHR	VTI
AAM	149.19	.764	1.850	.142	.053
ASM	168	.268	1.731	.117	.045
CM	133.33	.319	3.391	.127	.052
HM	222	.481	1.850	.115	.044

AAM=African-American Male; ASM= Asian Male; CM= Caucasian Male; HM= Hispanic Male

In the male participants the one Hispanic participant recorded the highest F0 value (222 Hz), second highest RAP value (.481 Hz), second highest shimmer value (1.850 Hz), and fourth highest NHR and VTI values (.115 Hz) and (.044 Hz). According to this data, it appears that the HM participant has the thinnest VF mass and vibrates at a faster rate. This would result in producing a higher F0 value. This could have been because the HM participant was 19 years of age. Vocal parameters differentiate depending on the age of the individual. The RAP (jitter) value in the HM group is lower than the AAM group but higher than the CM and ASM groups. According to the data it appears that the HM group has less hoarseness in relation to the AAM group but more hoarseness compared to the CM and ASM groups. It appears that the HM group demonstrates lower turbulence noise in the vocal signal than the CM group, equal noise turbulence compared to the AAM group, and higher noise turbulence compared to the ASM group. The HM group has the lowest amount of spectral noise compared to the AAM, ASM and CM groups. According to the data, it appears that the HM group demonstrates the lowest amount of breathiness compared to the AAM, ASM, and CM groups. There were two ASM participants. The ASM group yielded the second highest F0 value (168 Hz), lowest RAP value (.268%), lowest shimmer value (1.731%), third highest NHR value (.117%), and third highest VTI value (.045%). Comparing these values to the stroboscopic data, the ASM group revealed thicker VF mass that positively correlates with a slower rate of vibration; hence it was not surprising to see a much lower F0 compared to the HM. But compared to the AA male and the CM groups, the ASM exhibited much thinner VF mass and hence a higher F0. The RAP value (Table 4.1) suggests that the ASM group has the least amount of hoarseness in the vocal signal compared to the other groups. In relation to the shimmer value, it also appears that the ASM group has the least amount noise turbulence in the vocal signal. The NHR values similarly suggests that the ASM group has lower spectral noise in the vocal signal than both the AAM and CM groups but higher spectral noise compared to the HM participant. Concerning the VTI value, the data suggests that the ASM group demonstrates a lower level of breathiness in the vocal signal than both the AAM and CM groups, but higher levels of breathiness compared to the HM participant.

Of the male participants, the AAM group exhibited a higher F0 value (149.19 Hz) compared to the CM, but much lower than the HM participant and ASM group. In terms of RAP value, the AAM group had the highest jitter (.764%) value compared to all the other participants; whereas the ASM had the lowest (.268%). The AAM showed the second highest shimmer value (1.850%), highest NHR value (.142%) and highest VTI value (.053%). Observation of the VFs of the AAM showed thicker VF mass in comparison to the HM and ASM groups. It is well established that height, weight, and body index mass have an inverse effect on fundamental frequency. In this study, these variables were loosely controlled in that no formal measurement for weight, height, and body mass index were established. These variables may be contributing to a lower F0. The RAP, NHR, and VTI values appear to support the notion that the AAM participant had higher levels of hoarseness, spectral noise, and breathiness in the vocal signal compared to the other participants. The shimmer value suggests that the AAM participant had higher levels of noise turbulence than the ASM group, equal noise turbulence compared to the HM group, and lower noise turbulence in relation to the CM group. The CM group yielded the lowest F0 value (133.33 Hz), third highest RAP value (.319%), highest shimmer value (3.391%), second highest NHR value (.127%), and second highest VTI value (.052%). According to the data it appears the RAP value suggests that the CM group has lower levels of hoarseness in the vocal signal compared to the AAM and HM groups but higher levels of hoarseness in relation to the ASM group. The shimmer value appears to support the notion of the CM group having the highest level of noise turbulence in the vocal signal compared to the other groups. The NHR value seems to suggest that the CM group has a higher level of spectral noise in the vocal signal when compared to the ASM and HM groups and a lower level of spectral noise in relation to the AAM group. The VTI value appears to support idea that the CM group has a higher amount of breathiness in relation to the ASM and HM groups but a lower level of breathiness than the AAM group.

Female Participants Acoustic Data

Table 4.2: Mean Female Vocal Parameters for Sustained Vowel Task

Group	F0	RAP	Shimmer	NHR	VTI
AAF	205.54	.149	1.731	.117	.045
ASF	193.9	.165	2.394	.114	.053
CF	275.5	.363	2.105	.103	.044
HF	260.5	.392	3.438	.117	.061

AAF= African-American Female; ASF= Asian Female; CF= Caucasian Female; HF= Hispanic Female

The CF group recorded the highest F0 value (275.5 Hz), second highest RAP value (.363%), third highest shimmer (2.105%), lowest NHR value (.103%), and the lowest VTI value (.044%). Comparing the data, the RAP value suggests that the CF group has a higher level of hoarseness than the AAF and ASF groups but lower levels of hoarseness than the HF group. The shimmer value seems to support the notion that the CF group has a higher level of noise turbulence in the vocal signal than the AAF group and a lower level is demonstrated when compared to the HF and ASF groups. The NHR value seems to demonstrate that the CF group has the lowest amount of spectral noise in the vocal signal compared to the other groups. The VTI value appears to show that the CF group has the lowest level of breathiness out of all the groups. The HF group recorded the second highest F0 value (260.5 Hz), highest RAP (.392%), shimmer (3.438%), NHR (.117%), and VTI (.061%) values. The data seems to support the notion that the RAP, shimmer, NHR, and VTI values appear to suggest that the HF group have the highest levels of hoarseness, noise turbulence, spectral noise, and breathiness in the vocal signal in the female test groups. The AAF group yielded the third highest F0 value (205.54 Hz), lowest RAP (.149%) and shimmer (1.731%) values, highest NHR value (.117%), and third highest VTI value (.045%). According to the data, the RAP and shimmer data appear to support that the AAF group has the lowest level of hoarseness and noise turbulence in the vocal signal out of all the female groups. The NHR value seems to support that the AAF group has the highest level of spectral noise in the vocal signal out of the female test groups. The VTI value seems to support the notion that the AFF group has a higher level of breathiness in the vocal signal in relation to the CF group but lower levels compared to the HF and ASF groups. The ASF group yielded the lowest F0 value (193.9 Hz), third highest RAP value (.165%), second highest shimmer (2.105%) and NHR (.103%) values, and second highest VTI value (.045%). The data seems to support the notion that the RAP value of the ASF group has a higher level of hoarseness of in the vocal tract in comparison to the AAF group and a lower level of hoarseness than the HF and CF groups. The shimmer value appears to suggest that the ASF group has a higher level of noise turbulence in the vocal signal than the CF and AAF groups but a lower level than the HF group. The NHR value seems to suggest that the ASF group has a lower level of spectral noise in the vocal signal compared to the AAF and HF groups but a higher level than the CF group. The VTI value suggests the ASF group has a higher level of breathiness in the vocal signal in relation to the CF and AAF group but a lower level than the HF group.

Correlating Videostrobic and Acoustic Data

The posterior chink was a parameter that was of interest to this investigation. Acoustic data from participants with posterior chinks present in the videostrobic examination were compared to participants who did not present with a posterior chink.

Table 4.3: Male Average Acoustic Data with Posterior Chink

F0	RAP	Shimmer	NHR	VTI
176	.454	2.828	.131	.048

Table 4.4: Male Average Acoustic Data without Posterior Chink

F0	RAP	Shimmer	NHR	VTI
133	.263	3.645	.126	.044

The difference in average F0 values between males that presented with a posterior chink compared to males that did not present with a posterior chink is 43 Hz. According to this difference it appears that males with a posterior chink tend to have a higher fundamental frequency than males without a posterior chink. Four of the five acoustic parameters in males with a posterior chink were higher in value in comparison to males without a posterior chink. The posterior chink is a common feature of the female glottal closure pattern. Most adult females have a higher F0 compared to males. This difference has been attributed to the difference in laryngeal size. It is possible that in addition to the size difference in larynges glottal closure pattern may also play a role in F0 differences. This has not been reported in any known published studies and obviously warrants more in depth investigation. Additionally the males with a posterior chink had a higher average RAP value (.454%) in relation to males who did not have a posterior chink (.263%) (Refer to Tables 4.3, 4.4). This may be suggestive of the possibility that the males with a posterior chink may have more wastage of air in the vocal signal than the males without the posterior chink. The average shimmer value was higher (2.828%) in males with a posterior chink in comparison to those without a posterior chink (2.645%). This resulted in a .183% difference. This data appears to support the notion that males with a posterior chink have a higher level of noise turbulence in the vocal signal in relation to males with a posterior chink. The average NHR value for males with a posterior chink was .131% in comparison to a value of .126% for males without a posterior chink. This results in a minuscule difference of .005%. This appears to support the notion that males with and without the posterior chink have similar levels of spectral noise in the vocal signal. The males with a posterior chink recorded a slightly higher average VTI value (.048%) compared to males without a posterior chink (.044%). This difference minuscule difference may be due to individual differences.

Limitations of the Study

This study represented a small attempt to explore possible differences in vocal parameters across various ethnicities. Initially the expectation was to examine 8 groups of 6 individuals, 3 male and 3 female. However, due to the constraints of individuals meeting the inclusionary criteria for this study, some individuals had to be eliminated. The study ended with a total of 15 participants. The small sample sizes both the intended and the final size posed major limitations to any conclusions that could be drawn. With total of 13 participants who underwent the videostrobospic and acoustic examination with an additional 2 participants who only participated in the acoustic examination, it was difficult to extract data that could be statistically analyzed. Another limitation is that there were no controls for body mass index (BMI), height, weight, and dialect. Studies have shown that BMI, height, and weight significantly impact vocal characteristics. In terms of dialect, there was no control over accented speech for non-native English speakers (NNES) in this study. There was no way of knowing whether the English rendition of the speech of NNES may have affected their articulatory production of certain English sounds, thus influencing spectral or glottal variations. A further limitation is that the placement of the ridged endoscope, which is over the base of the tongue, in the collection of the videostroboscopic data limits the individuals ability to achieve maximum production to achieve optimum production of the sustained vowels /ah/ and /ee/. A flexible scope through the nasal cavity and postured towards the VFs superiorly from the velopharyngeal port, would have allowed the participant a greater degree of freedom for vowel production.

Future Research

Future studies should seek to net a larger sample size in order to yield more significant data. The studies in this area should account for BMI, height, weight and impact of dialect to as further control over extraneous influences on vocal parameters. In terms of instrumentation, a flexible scope would probably be more advantageous than the rigid scope, as the former, would provide the participant with more flexibility in speech sound productions.

Conclusion

In conclusion, even though the findings from this study have not reached statistical proportions, the evidence of this study support the notion that the slight differences of glottal closure patterns seen in the small group of participants warrant a more in depth investigation. The significance of findings as stated in the rationale for this study is great enough to impact on the evaluation and treatment protocols for individuals who may have a vocal dysfunction. In this study, the acoustic data do demonstrate small differences in F0, but there is no prominence that it is racially influenced. The videostroboscopic data reveal that all African-American participants both male and female demonstrate posterior chinks. But again, due to the small sample size this is not statistically tested.

References

Abou-Elsaad, T. (2007, April). Stroboscopic examination Stroboscopic examination of the larynx. Retrieved from http://www.alexorl.edu.eg/alexorlfiles/pptorl2007/138002.pdf

Andrianopoulos, M., Darrow, K., & Chen, J. (2001). Multimodal standardization of voice among four multicultural populations: Fundamental frequency and spectral characteristics. *Journal of Voice*, 15(2), 194-219.

Alberti, P. W. (1978). The diagnostic role of laryngeal stroboscopy. *Otolaryngology Clinics of North American*, 11 347-354

Aronson, A. Clinical Voice Disorders. Thieme, New York, NY; 1990

Awan & Mueller (1992) Speaking fundamental frequency characteristics of centenarian females. *Clinical Linguistics & Phonetics*, 6, 249-254.

Baken, R. J., & Orlikoff, R. F. (2000). Clinical measurement of speech and voice (2nd ed.). San Diego, CA: Singular Thomson Learning.

Bauman-Waengler, `J. (2012). Articulatory and phonological impairments: A clinical focus (4th ed.). Boston, MA: Allyn & Bacon, Inc.

Boone, D., McFarlane, S., Von Berg, S., & Zraick, R. (2014). *The voice and voice therapy* (9th ed.). Englewood, NJ: Pearson Education.

Bonilha HS, Dawson AE. Creating a Mastery Experience During the Voice Evaluation. *J Voice*. 2012;26(5):665-71.

Boshoff, P. (1945). The anatomy of the South African Negro larynges. *South African Journal of Medical Sciences*, 10, 35-50.

Buchtal F, Faaborg-Anderson K. Electromyography of laryngeal and respiratory muscles: correlation with phonation and respiration. *Ann Otol Rhino Layrngol*. 1964;73:118-123

Colton, R.H., & Casper, J. K. (1996). Stroboscopic signs associated with benign lesions of the vocal folds. *Journal of Voice*, 9, 312-325.

Corey, J.P., Gungor, A., Nelson, R., Liu, X., and Fredberg, J. Normative standards for nasal cross-sectional areas by race as measured by acoustic rhinometry. *Otolaryngol Head Neck Surg*. 1998; 119: 389–393

Cox, V., & Mueller, P. (2004). Ethnographic factors in voice: A review of the literature. *Journal of Speech and Hearing Research*, 47, 48-52.

D.E. Battle (Ed.) Communication Disorders in Multicultural Populations. 2nd ed. Butterworth-Heinemann, Boston, MA; 1998

Diana Deutscha, Jinghong Le, Jing Shen, Trevor Henthorn (2009). *The Journal of the Acoustical Society of American*, 125.

Fairbanks, G. (1940). Voice and articulation drillbook. New York: Harper & Brothers

Garcia MJV, Cobeta I, Martin G, Alonso-Navarro H, Jimenez-Jimenez J. Acoustic analysis of voice in Huntington's disease patients. *J Voice*. 2009;25(2):208-17.

Gay T, Strom M, Hirose H, et al. Electromyography of the intrinsic laryngeal muscles during phonation. *Ann Otol Rhinol Laryngol*. 1972;81:401-409.

Hillman RE, Mehta DD. The science of stroboscopic imaging. In: Kendall KA,

Lenoard RJ, editors. Laryngeal evaluation:indirect laryngoscopy to high-speed digital imaging. New York: Thieme *Medical Publishers*, Inc.’ 2010:101-109.

Hirano M, Koike Y, Joyner J. Style of phonation: an Electromyographic investigation of some laryngeal muscles. *Arch Otolaryngol*. 1969;89:902-907.

Hirano M. Laryngeal muscles in singing, In: Hirano M, Kirchner JA, Bless D *Neurolaryngology: Recent Advances*. Boston, Mass: Little Brown; 1987

Hollien, H., Dew, D., & Phillips, P. (1971). Normative data on the speaking fundamental frequency characteristics of young adult males. *Journal of Phonetics*, 19, 117-210

Hudson, A. & Holbrook, A. (1981). Fundamental frequency characteristics of young Black Adults: spontaneous speaking and oral reading. *Journal of Speech and Hearing Research*. 25, 25-28

"IEEE Transactions on Audio and Electroacoustics, vol 17 issue 3, 225-246, 1969"

Irwin, R. (1997). Judgements of vocal quality, speech fluency, and confidence of Southern Black and White speakers. *Language and Speech* 20, 261-266

KayPENTAX. Instruction manual: Stroboscopy systems and components 2008. Montvale, New Jersey, USA.

Kempster GB, Larson CR, Distler MK. Effects of electrical stimulation of cricothyroid and thryoarytenoid muscles on voice fundamental frequency. *J Voice*. 1988;2:221-229.

Kitizing, P. (1985). Stroboscopy- a pertinent laryngological examination. *Journal of Otolaryngology*, 14, 151-157.

Lass, N.J., Mertz, P.J., & Kimmel, K.L. (1978). The effect of temporal speech alterations on speaker race and sex identifications. *Language and Speech*. 21, 279-291

Lass, N.J., Trapp, D.S., Baldwin, M.K. Scherbick, K.A., & Wright, D.L. (1982). Effect of vocal disguise on judgements of speakers sex and race. *Perceptual and Motor Skills*. 54, 1235-1240.

Lucero, J.C. (1995). "The minimum lung pressure to sustain vocal fold oscillation". *Journal of the Acoustical Society of America*. 98: 779–784. doi:10.1121/1.414354.

Mehta, D., & Hillmana, R. (2012, December). Current role of stroboscopy in laryngeal imaging. Retrieved from https://scholar.harvard.edu/files/dmehta/files/mehtahillmancurropotohns2012_current_role_of_stroboscopy_in_laryngeal_imaging.pdf

Nawka T, Konerding U. The interrater reliability of stroboscopy evaluations. *J Voice* 2012. [Epub ahead of print]

Poburka BJ. A new stroboscopy rating form. *J Voice* 1999; 13:403–413.

Rosen CA (2005) Stroboscopy as a research instrument: development of a perceptual evaluation tool. Laryngoscope 115:423–428

Titze IE. Workshop on Acoustic Voice Analysis. Summary Statement. Iowa City, IA: *National Center for Voice and Speech*. Wendell Johnson Speech & Hearing Center, University of Iowa; 1995.

Vargo, R. (2012) Acoustic and Perceptual Analyses of the Fundamental Frequencies of African American and Caucasian Males and Females. (Electronic Thesis or Dissertation). Retrieved from https://etd.ohiolink.edu/

Volkar, C. (2017). Patterns of Vocal Fold Closure in Professional Singers. (Electronic Thesis or Dissertation). Retrieved from https://etd.ohiolink.edu/

Walton, J. H., & Orlikoff, R. F. (1994). Speaker race identification from acoustic cues in the vocal signal. *Journal of Speech & Hearing Research, 37*(4), 738-745.

Wolfe, J., Garnier, M., and Smith, J. (2009). “Vocal tract resonances in speech, singing and playing musical instruments,” *Human Frontier Sci*. Progr. J. **3**, 6–23

Xue, Hao (2006). “Normative standards for vocal tract dimensions by race as measured by acoustic pharyngometry,” *J Voice*. 2006 Sep;20(3):391-400. Epub 2005 Oct 21

Yamazawa H.a · Hollien H.b (1992) Speaking Fundamental Frequency Patterns of Japanese Women 49:128–140

Zhang Z. (2016). “Mechanics of human voice production and control” *Journal of the Acoustical Society of America*. 2614–2635. PMC5412481

Zhang, Z. (2016). Cause-effect relationship between vocal fold physiology and voice production in a three-dimensional phonation model. Retrieved April 02, 2018, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4818279/

APPENDIX A

"Are there differences in Acoustic and Perceptual Vocal Fold Parameters across Ethnic Groups?"

Dear Participant,

We are Dr. Cox and Mason Meek, a faculty member and graduate student respectively, in the Speech and Hearing Program, School of Health Sciences at Cleveland State University. We would like you to participate in a research study. This study is about understanding the features of voice across different ethnic groups. We will listen to your voice as you read a short sentence. We will look at your vocal folds during production of a vowel sound. We will use a tiny camera and light source placed over your tongue. The data collected will be confidential. Every effort will be made to maintain confidentiality of your information. Results of this study will not be traced back to you. While there are no direct benefits to you from this study, your participation will help us to understand the characteristics about voices across different ethnic groups of people. You will be tested at Cleveland State University Voice and Swallowing lab. The lab is located at 2121 Center for Innovations in Medical Professions (CIMP) Room 211. We will provide you with directions to the University Participation. This study is voluntary. You may withdraw at any time. Outside the risks associated with those daily living, there is a very slight possibility of gaggin during the oral exzamination. If you cannot suppress the gagging effect, the procedure will be stopped. Dr. Cox is a trained and licensed medical speech-language pathologist and has over twenty years of experience in this field. She will be overseeing the testing for this study. This study will take about 60 minutes to complete. If you want to know more about this project, please contact Mason Meek at (440) 590-3806 or my advisor, Dr. Violet Cox, at (216) 687-6909 or at her email: v.cox@csuohio.edu. This project has been approved by the Institutional Review Board of Cleveland State University. If you have further questions you may contact IRB at (216) 687-3606. There are two copies of this letter. Sign one copy for your records and return the other copy to Mason Meek on the day of testing. I thank you in advance for your cooperation and support. Please indicate your agreement to participate by signing below.

"I am 18 years or older and have read and understood this consent form and agree to participate."

Signature: ____

Name: ____

Date: ____________________

APPENDIX B

SCREENING QUESTIONNAIRE FOR SPEAKING PARTICIPANTS

Age: __________ Gender: __________ Ethnicity: _____________

Educational Level: High School: __________ College: _________

Please check YES or NO for each of the following.

Ethnic/Cultural Background

1. Are both parents of the same ethnicity? YES NO
2. Are you a native citizen of the USA? YES NO
3. Are you a fluent speaker of English? YES NO
4. Are you bilingual? YES NO
5. How often do you use your non-English language? (Circle most accurate)

   (Consistently) (Occasionally) (Rarely) (None/Not bilingual)

   Voice Status

   1. Do you currently have a history of asthma? YES NO
   2. Do you currently have a history of acid reflux? YES NO
   3. Do you have a diagnosis of voice disorders? YES NO
   4. Do you currently have a history of smoking? YES NO
   5. Do you currently have an upper respiratory infection? YES NO
   6. Do you have emphysema or other respiratory diagnosis? YES NO
   7. Do you have chronic sinusitius? YES NO

APPENDIX C

INSTRUCTIONS PROVIDED TO PARTICIPANTS FOR SUSTAINED VOWEL TASK

"Hold the sound 'ah' at the pitch that seems to be most natural for you. Do not try to sing or hit any particular pitch. Simply relax and say 'ah' for four seconds. I will let you know when to stop."

APPENDIX D

SPECIFIC INSTRUCTIONS TO LISTENING PARTICIPANTS

“You will listen to nine paired samples that are one second in length. I will introduce the pair of speakers for each sample. What I want you to do is designate the order of presentation. For example I will introduce the two speakers (Asian and Caucasian males). If you thought you heard the Asian male first and the Caucasian second you would circle ASM in the first column and CM in the second column of the first row. If you would like to hear a sample again, just let me know. The first four samples will be men and the second five samples will be women.”

APPENDIX E

FORM PROVIDED TO LISTENING PARTICIPANTS

ASM CM	ASM CM
AFM CM	AFM CM
HM CM	HM CM
ASM CM	ASM CM
ASF CF	ASF CF
AFF CF	AFF CF
HF CF	HF CF
AFF CF	AFF CF
AFF ASF	AFF ASF

APPENDIX G

ORDER OF PRESENTATION FOR PAIRED VOWEL SAMPLES

Pair	Voice 1	Voice 2
1	ASM 1	CM 2
2	AFM 1	CM 1
3	HM 1	CM 3
4	ASM 2	CM 2

MALE VOICES

Pair	Voice 1	Voice 2
1	ASF 1	CF 2
2	AFF 2	CF 1
3	HF 1	CF 3
4	AFF 3	HF 2
5	AFF 1	ASF 1

FEMALE VOICES