Acoustic and Perceptual Analyses of the Fundamental Frequencies of African American and Caucasian Males and Females

Rose A. Vargo
Cleveland State University

ACOUSTIC AND PERCEPTUAL ANALYSES OF THE FUNDAMENTAL FREQUENCIES OF AFRICAN AMERICAN AND CAUCASIAN MALES AND FEMALES

ROSE A. VARGO

Bachelor of Arts in Speech and Hearing

Cleveland State University

August 2010

Submitted in partial fulfillment of requirements for the degree

MASTER OF ARTS IN SPEECH-LANGUAGE PATHOLOGY AND AUDIOLOGY

at the

CLEVELAND STATE UNIVERSITY

December 2012

This thesis has been approved

for the department of SPEECH PATHOLOGY AND AUDIOLOGY

and the College of Graduate Studies by

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

Department & Date

Monica Gordon-Pershey, Ed.D., CCC-SLP

Department & Date

Andrew Lammers, Ph.D.

Department & Date

ACKNOWLEDGEMENTS

I would like to take advantage of this unique opportunity to formally thank the people that have changed my life in many ways. Many people have contributed to the successful completion of this project in a variety of contexts. Contributions came in forms of advice, late night editing, kind words of encouragement, but, most importantly, unconditional love and support.

My deepest thanks go to Dr. Violet Cox, without whom, none of this would have been possible. You guided me every step of the way, coaching and offering advice, yet encouraging me to take ownership of this special project and run with it. Even when I was burned out and tired, you inspired me to keep going. I could never thank you enough for being my unfaltering leader throughout this adventure. Your words and guidance will stay with me for a lifetime. I would also like to thank the members of my committee, Dr. Monica Gordon-Pershey and Dr. Andrew Lammers, whose skills, knowledge, and insight helped to make this project a success.

Thank you to the faculty of the Speech and Hearing Program, all of whom supported and cheered me on for the past five years. The strong support and reassurance

gave me confidence to achieve things I never thought possible. Dr. Myrita Wilhite, you empowered me with the strength to go on. Thank you for believing in me.

Thank you to my family. I love every one of you! When I felt like I was falling, you have always been there to pick me up. To my amazing mother, you have been there through it all. You have provided me with the motivation and confidence to achieve my dreams. You have been my biggest cheerleader and believed in me all along, even when I doubted myself. Simple words cannot express my gratitude for everything you have done for me. You are my role model, my inspiration, and my angel on earth. Like my mother, I will continue to dance!

Finally, a special thank you to my angel in heaven. This is for you. I did it, Grandma! I have not forgotten your emphasis on education; I hear your words in my head every day. I know you would be proud. Continue to watch over me and be my guardian angel.

"Gone yet not forgotten, although we are apart, your spirit lives within me, forever in my heart." –Author Unknown

ACOUSTIC AND PERCEPTUAL ANALYSES OF THE FUNDAMENTAL FREQUENCIES OF AFRICAN AMERICAN AND CAUCASIAN MALES AND FEMALES

ROSE A. VARGO

ABSTRACT

Normative data compiled by Hixon and Abbs (1980) continue to serve as guidelines for fundamental frequency (F_0). These normative data were collected solely from Caucasian participants and may not accurately reflect norms for other ethnicities. According to published literature regarding pitch differences among racial groups, African Americans are believed to have a lower F_0 than their Caucasian counterparts.

This study investigated differences in F_0 between African American and Caucasian young adults. Twenty adults between the ages of 18 and 30 were examined along three vocal parameters: sustained vowel phonation, a speaking task, and a reading task. Three experienced speech-language pathologists and three students without training in voice listened to one second samples of the sustained vowels and made judgments of the speaker's race. Acoustic results revealed no significant differences between the African American and Caucasian male participants. However, F_0 values for the sustained vowel task were higher for African American females compared to their Caucasian counterparts while the same trend was not observed in the other two tasks. This suggests that F_0 differences across the two races may be task specific rather than racially influenced.

TABLE OF CONTENTS

APPENDICES

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

The human voice has been examined both acoustically and perceptually for a variety of purposes over many decades. Medical researchers have focused on the voice as an indicator of a patient's health status as well as an indicator of the severity and possible progression of a disease. From an intellectual standpoint, the voice has been examined as a means for communicating linguistic information. Culturally, the voice has been analyzed to determine if linguistic markers are present to signify communities or dialects associated with the speaker (Andrianopoulos, Darrow, & Chen, 2001).

Boone, McFarlane, Von Berg, and Zraick (2011) note five aspects of a normal voice: (1) loudness, the voice must be able to be heard over environmental noise; (2) hygienic, the voice must not be produced in a way that may damage the laryngeal musculature or framework; (3) pleasantness, the voice should, to some degree, be pleasant to listen to and not make listeners uneasy; (4) flexibility, the ability to express emotional tones through the voice; (5) representative, the voice should represent the speaker in terms of age and gender. When one of these aspects is violated, the listener perceives an aberration of the voice. For example, if a speaker's pitch is not representative of his age or gender, then the listeners' expectations of normality in voicing are violated.

Vocal pitch is the perceptual correlate of the rate of vocal fold vibration and is a component of all aspects of the voice. The acoustic correlate of this perception is known as fundamental frequency (F_0). In other words, as F_0 changes, the listener perceives it as a change in pitch. As the speaker varies his or her intensity or loudness, F_0 will also vary. F_0 is acoustically represented as the number of vocal fold phases, which represents openings and closings, that occur in one second. This is measured in Hertz (Hz) or cycles per second.

Producing a hygienic voice is related to speaking at an appropriate pitch so that the vocal folds are not subject to abuse or misuse. Pitch is regularly examined, both informally during conversations with others and formally when a vocal pathology is suspected. When expressing emotional tones in connected speech, it is F_0 that varies to alter the prosody and inflection to reflect the emotion of the speaker. Perceptually, pitch is directly related to the pleasantness of a voice. For example, a voice that is extremely high or low pitched violates the listener's expectations and is perceived as abnormal. A person's pitch is similar to a fingerprint, in that each person's is unique. Often times listeners are able to instantly identify someone they know just by hearing his or her voice.

The speaker's pitch is dependent on several factors. The primary biomechanical determinants of F_0 are: (1) length of the vocal folds, (2) tension of the vocal folds, and (3) mass of the vocal folds (Boone et al., 2011). Vocal folds that are elongated with more tension and that have less mass will produce a high F_0; vocal folds that are shortened with less tension and have more mass will produce a low F_0.

According to Hixon and Abbs (1980), there is a predictable number of cycles per second identifiable across gender and age. A person's habitual pitch is dependent on gender, age, and race (Boone et al., 2011). In the currently established literature, the normal ranges indicate that an adult female's F_0 should be approximately 225 Hz and an adult male's F_0 is approximately 128 Hz (Williamson, 2006). Given the biomechanical determinants of F_0, it is apparent that males are expected to have a lower F_0 than females. Males have denser, or heavier, and longer vocal folds (17-21 mm) compared to females (11-15 mm), resulting in a lower pitch. The current norms state that the F_0 of a child's voice should be approximately 265 Hz (Williamson, 2006). Boone et al. (2011) explain that differences heard in the voice over the lifespan can be attributed to changes in laryngeal anatomy and physiology. Infants have a vocal fold length of only 2.5-3.0 mm. Furthermore, the vocal fold mucosa of infants are thinner and the transitional zone of the folds is a single layer. In adults, the transitional zone is comprised of the intermediate and deep layers of the lamina propria, but these two distinct layers are not evident until the end of puberty. The combination of shorter, thinner, and less dense vocal folds contributes to the elevated pitch heard in infants.

On the latter end of the lifespan is the senescent voice. Liss, Weismer, and Rosenbek (1990) discuss several physical changes that result in the senescent voice heard in the elderly. Age-related thinning of muscle tissue causes the tissue to become stiff and rigid. These factors, combined with a build-up of connective tissue, contribute to a loss of vibratory mass that may result in an elevated pitch. Conversely, increased edema will add to the mass of the vocal folds, contributing to a lower pitch. As an individual ages there is a decrease in the amount of secretions from the mucous glands, which causes an increase in the viscosity of the coating found on vocal fold tissue, therefore increasing the mass and lowering the vocal pitch. The ossification of cartilages will decrease the mobility of the laryngeal framework, making it difficult or painful to vary the pitch. These are all contributing factors to presbyphonia, an age related vocal change characterized by recognizable perceptual abnormalities in the pitch, pitch range, loudness, and quality of the older speaker's voice (Boone et al., 2011).

Specific documented differences in physiology and pitch also exist in the second factor in determining F_0, which is gender. For women, there is a decrease in pitch after menopause due to the secretion of excess androgenic hormones and the thickening of the glottal membrane, both contributing to vocal fold mass. After this post-menopausal decrease, a female's voice tends to stay relatively constant. However, elderly males demonstrate an increase in pitch, most likely due to the loss of muscle mass.

Few studies have explored race or ethnicity as a determining factor of pitch, however, Boone et al. (2011) suggest it is possible that such a connection exists. It is important to distinguish between the terms race, culture, and ethnicity. Bauman-Waengler (2012) offers clear definitions for these concepts. Race is considered a biological 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 purposes of this study, the term race will be used to describe individuals as either African American or Caucasian.

There is a body of literature that currently documents vocal acoustic and perceptual differences among cultural and ethnic groups, but not many studies have attempted to quantify these differences across *racial* boundaries. Consequently, this study will attempt to examine possible F_0 differences across two racial groups. This type of investigation becomes of paramount interest to the field of speech-language pathology in terms of remediation of voice disorders from a very significant standpoint. The initial studies by Hixon and Abbs (1980) establishing F_0 norms for males and females were performed on solely Caucasian individuals. If there is a difference in F_0 across racial boundaries, then these established norms may indeed be inappropriately applied to patients of all races. A further implication is that since a vocal pathology is signaled by a change in F_0, a possibility exists that an inaccurate norm could lead to an over- or under-diagnosis of a vocal pathology in speakers of various racial groups.

If race is a determining factor in pitch, the implication is that physiological differences may be the cause. A 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 F_0 of the voice. Contraction and relaxation of these muscles allows 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.

Conversely, some researchers specify that it is ethnicity that has the impact on F_0, which suggests a cultural or environmental influence on vocal production. In another physiological study, Wise (1933) found no differences in laryngeal dimensions or resonance cavities, but found that African Americans demonstrated a more forward tongue position in the mouth, which he attributes to a learned behavior characteristic of one's demographic or cultural group. The altered tongue position could affect the quality of the voice and possibly the F_0 by altering resonance in the oral cavity.

Another view of pitch differences is that they are caused by 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. A study by Deutsch, Jinghong, Shen, and Henthorn (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 F_0 values were clustered within each village but differed by approximately three semitones. These data support the claim that F_0 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.

Available studies investigating F_0 differences can be divided into three categories: acoustic, perceptual, and combination. Acoustic studies are those that take samples from participants and analyze them with software to determine specific parameters of the voice. The majority of the published studies concerning racial differences in voice characteristics rely solely on acoustic methods. Acoustic methods are very reliable but do not take into account the perspective of the listener, who is an integral part of communication. Perceptual studies in this area of research aim to determine if it is possible for listeners to identify the race of the speaker based on certain characteristics of their voice. Perceptual studies collect data on a listener's interpretation of the voice sample. The most compelling studies are the combination studies. These studies employ both methods and are rare, although they provide the most valuable information. Cox and Mueller (2004) suggest that, "perceptual studies that are carried out in tandem with acoustic studies may offer more compelling data in establishing norms for vocal characteristics of different ethnic groups" (p. 49).

Acoustic Studies

The relevant acoustic studies compared F_0 of African Americans to Caucasians either by studying both groups simultaneously or by studying one group of African Americans then comparing the results to published norms or to data compiled from other studies. Comparative acoustic studies have been performed on individuals across the lifespan. A difference in F_0 was observed in children of different ethnicities as young as five years old. Awan and Mueller (1996) studied a significant number of Caucasian, African American, and Hispanic male and female kindergarteners. The children were compared on measures of mean speaking fundamental frequency (SFF), maximum and minimum SFF, pitch sigma, and speaking range. The results indicated significant differences across ethnic groups. These researchers found that the African American children had lower F_0 values than the Caucasian and Hispanic children.

Another study by Wheat and Hudson (1988) examined the F_0 of spontaneous speech in six-year-old African American children and compared them to the published norms. These normative data state that children should have a F_0 of approximately 265 Hz (Williamson, 2006). Wheat and Hudson (1988) found that the mean F_0 of African American boys and girls was 211.3 Hz and 219.5 Hz, respectively. Cox and Mueller (2004) in their review of the literature reported that the “comparison of the SFF values of the Black speakers with established norms of matched White boys and girls revealed that the White speakers had a higher SFF, thus supporting previous studies that claimed racial differences in the SFF of African American and White speakers” (p. 50).

Studies done by Hollien and Malcik (1962; 1967) compared voice characteristics of African American and Caucasian adolescents. Their first study (1962) determined the median F_0 of 10-year-old, 14-year-old, and 18-year-old southern African American boys and compared the results with results from a study by Curry (1940) that investigated the F_0 of northern Caucasian boys in the same age groups. The differences were most pronounced in the 14-year-olds, where it was found that the African American boys had a median F_0 of 162.7 Hz and the Caucasian boys had a median F_0 of 241.5 Hz. The differences were the least pronounced at 18 years old, with median F_0’s of 137.1 Hz and 124.4 Hz for Caucasians and African Americans, respectively. Hollien and Malcik (1962) found that African American 10-year-old boys had a median F_0 of 223.4 Hz, while Curry (1940) found that 10-year-old Caucasian boys had a median F_0 of 269.7 Hz. This led Hollien and Malcik to the conclusion that African American boys were maturing physiologically at an earlier age than Caucasian boys. However, validity may be questioned, as the geographic region was not controlled: one study looked at southern boys, while the other looked at northern. Those who believe that culture and dialect may influence pitch may question the methodology of this study.

Hollien and Malcik (1967) completed a subsequent study. In this investigation, they employed controls for geographic region, age, physical size, intellectual ability, speaking ability, and reading ability. In this study, the reading fundamental frequency (RFF) was determined. This study compared data gathered on 10-year-old, 14-year-old, and 18-year-old southern Caucasian boys to data obtained from their earlier study (1962) of southern African American boys. These results showed that the African American boys had a lower RFF at each age group. The 10-year-old Caucasian boys had a mean RFF of 235.4 Hz, while the African Americans had a RFF of 210.2 Hz. The biggest difference in RFF was seen in the 14-year-olds, where it was determined that the African American boys had a RFF of 158.2 Hz and the Caucasian boys had a RFF of 232.7 Hz. The smallest disparity was seen at 18 years old, where African American boys had a RFF of 124.4 Hz and the Caucasian boys had a RFF of 137.1 Hz.

Hudson and Holbrook (1982) investigated F_0 values of African American college students. The researchers collected data on prompted spontaneous speech and reading tasks. The authors compiled data for the African American students and compared them to data gathered by previous researches on Caucasians of the same age group. They found lower mean F_0's in African Americans, however, these differences were more pronounced in the female participants.

Few acoustic studies investigating racial or ethnic influences on F_0 have been completed on elderly participants to fully examine pitch changes throughout the lifespan. Xue, Neeley, Hagstrom, and Hao (2000) examined the mean SFF of 70- to 80-year-old African American and Caucasian males and females. No significant difference was found between the male groups. However, mean SFF values for Caucasian women were found to be 182.94 Hz, while mean SFF values for African American women were 163.78 Hz, a difference of approximately 19 Hz. The majority of acoustic studies that have investigated cultural and linguistic factors in voice show that these do, in fact, influence F_0. However, one cannot rely on acoustic data alone.

Perceptual Studies

Previous perceptual studies have shown that listeners are able to identify African American and Caucasian speakers from tape recorded speech samples at all levels of phonetic complexity (Walton & Orlikoff, 1994). Perceptual studies rely on listeners, either trained or untrained in voice disorders, to identify a particular vocal quality or a characteristic of the speaker based on various types of samples. Saniga, Carlin, and Farrell (1984) investigated the use of fry register in African American and Caucasian female speakers in order to quantify the presence of fry within the African American Vernacular English [AAVE] dialect. Fry (pulse) register occupies frequencies below habitual pitch. This register generates a syncopated secondary beat perceived as a crackling sound (Boone et al., 2010). The use of fry register is associated with a lower fundamental frequency (Cox & Mueller, 2004). Thirty women (10 Caucasian speakers of Standard American English [SAE], 10 African American speakers of SAE, and 10

African American speakers of AAVE) retold stories and were recorded. Ten judges listened to the samples and perceived the speakers of AAVE to use more fry register. The authors concluded that dialect was the confounding variable causing speakers to produce fry register, not race. These findings support the notion of cultural influences in pitch determination rather than race.

The larger focus of this line of research in perceptual studies is the possibility of accurate racial identifications of speakers from recorded samples. The majority of these studies conclude that it is possible to make accurate racial identifications of the speaker. Hollien and Jackson (1973) noted that lower frequencies and more frequency variability may contribute to the listener's accuracy in identifying the race of the speaker.

Dickens and Sawyer (1962) found that 70% of judges in their study were able to make accurate racial identifications from a recorded speech sample. They also observed that the listeners were more accurate when identifying members of their own race and that males generally made more accurate judgments than females.

Irwin (1977) completed a study that was concerned with listeners' comparative judgments of the vocal quality, speech fluency, and confidence of African American and Caucasian college students. Twenty-five African Americans and 25 Caucasians were asked to familiarize themselves with a reading passage, then were recorded while reading it aloud. Thirty-six judges from the university were asked to identify the race of each speaker. Irwin (1977) found that the listeners were 90% accurate in their judgments.

The validity of any of these studies, however, could be called into question. The nature of the extended samples provided to the listeners may allow for stereotypical notions and biases to influence judgments. <u>If *pitch* is what is to be examined, then all</u>

other linguistic factors should be removed from the sample. For example, Roberts (1966) recorded African American and Caucasian speakers while prolonging vowels and diphthongs. The listeners were considered to be experts in the field of voice and were instructed to identify the race of each speaker. These listeners were 80% accurate in their judgments of the isolated vowel samples. The experts noted that differences in nasality and pitch enabled them to make their judgments of speaker race.

Combined Studies

Cox and Mueller (2004) state that "Perceptual data alone cannot reliably establish ethnic voice differences, consequently research that quantifies perceptual and acoustic data may provide more valuable information" (p. 49). Hanley (1951) completed a study that investigated frequency and duration characteristics of speakers of General English, Southern English, and Eastern English dialects. Data were compiled for mean RFF and SFF. The judges were professionally trained at Louisiana State University and Columbia University. The listeners were instructed to identify the dialect and pass judgment on the "acceptability" of the samples. The judges were 96% accurate when identifying speakers of General English, 92% accurate for speakers of Southern English, and 90% accurate for speakers of Eastern English. Acoustically, different F_0 values were calculated for each dialect. For both RFF and SFF, speakers of General English demonstrated the lowest pitch, while speakers of Southern English had the highest F_0 values. This has implications for African Americans using the AAVE dialect as well, as it is a possibility this dialect may also influence F_0 values.

Bryden (1968) examined isolated vowels in a similar manner. The vowels /i/ and /u/ were recorded in isolation from African American and Caucasian speakers. The first formant frequencies were determined and found to be significantly lower in African American speakers. The first formant of /i/ was 260 Hz in African Americans and 319 Hz in Caucasians. The first formant of /u/ was 284 Hz in African Americans and 400Hz in Caucasians. Based on the vowel samples, listeners were asked to make a perceptual judgment of the speaker's race. The judges were 85% accurate in identifying Caucasian speakers and 82% accurate in identifying African American speakers.

Walton and Orlikoff (1994) studied 100 African American and Caucasian participants between the ages of 18 and 57. Acoustic data were collected on samples of a sustained vowel (/a/). Their acoustic analyses did not reveal any significant differences between racial groups. The perceptual component of the study revealed that given an isolated vowel that is one second in duration, the judges were 60% accurate in their racial identifications, which is significantly greater than chance. Cox and Mueller (2004) suggested that listeners used perceived differences in vocal quality as a cue for race identification, rather than the vowel itself.

Rationale for the Study and Research Questions

According to the published literature regarding pitch differences between racial groups, in general, African Americans are believed to have lower F_0's than their Caucasian counterparts. More studies concerning African American speech should be conducted in order to establish normative data and thereby provide a basis for the accurate evaluation and treatment of vocal pathologies. According to Cox and Mueller

(2004), one clinical imperative that justifies establishing normative data for African American voice production characteristics is that African American speakers with voice disorders are frequently evaluated based on data derived from Caucasian speakers. If a difference in F_0 characteristics across racial groups truly exists, as the literature seems to suggest, then more empirical data comparisons should be available, so that accurate assessment of phonatory function for African American speakers can be provided.

While the majority of the above published data show that African Americans have a lower F_0, more research is needed to clearly establish data for racial groups other than Caucasian. The previous studies did not examine sustained isolated vowels both acoustically and perceptually, nor did the research incorporate other measures of F_0 to compare across racial groups. Many of the previous studies have used published norms or previous studies by which to compare results, rather than testing both racial groups simultaneously with the same equipment and under the same conditions. According to Xue and Mueller (1996), more studies are needed with both African American and Caucasian participants under similar testing conditions in order to ensure valid and reliable cross-ethnic comparisons. Awan and Mueller (1996) agree, also citing the need for studies to directly compare racial groups in order to establish the presence of possible differences in voice characteristics among racial groups.

The purpose of the present study is to investigate the differences in F_0 between African American and Caucasian adults simultaneously by employing a combination of acoustic and perceptual measures. The investigators of this study suspect that the F_0 of African American participants will be lower than that of Caucasian participants. If the suspected results are found to be true, this study has the potential to affect speech-language-pathology practice and research. New normative data may need to be collected on the F_0's of other racial groups, which would prompt many new areas of research. Additionally, the knowledge that individuals of different racial groups speak at different pitches will affect all levels of voice disorder diagnosis and treatment, from evaluation procedures to treatment techniques as well as discharge criteria.

The knowledge that there is an inherent difference in fundamental frequency as a function of race would be of significance for many disciplines. The results of this study could impact the ever-growing world of technology. Voice recognition software, in addition to speech-to-text software, are becoming increasingly popular in mainstream society. These types of software come standard on most cell phones and are available on almost all devices used in daily life. Knowledge of perceived pitch differences may influence further refinement of the accuracy of these types of software. Many users cite the inaccuracy of these systems as their main complaint. The data collected from this study may point out specific perceptual characteristics of the voice that the software systems are currently neglecting. Devices that speak *for* the user are also becoming more popular. Many individuals use Alternative Augmentative Communication (AAC) devices. Patients who are unable to speak for themselves can use specialized equipment to “regain” their voice and independence. Data from this study may be useful in programming voices in order to personalize the device for the user. Implications may also me made in the field of criminology. Knowledge of distinct vocal characteristics in racial groups can aid in identification of individuals when voice recordings are available.

The present study will attempt to answer this specific question: Is there a difference in F_0 in African Americans when compared to Caucasians? This will be accomplished through the investigation of: 1) possible F_0 differences between Caucasian and African American men and women on three tasks; 2) perceptual racial identification through pitch, using sustained vowels; and 3) the reliability and accuracy of these perceptual judgments.

CHAPTER II

METHODOLOGY

Participants

Twenty participants between the ages of 18 and 30 from Cleveland State University (CSU) in Cleveland, Ohio served as speakers for this study. Participants were recruited through announcements on campus bulletin boards and were enlisted by the investigators themselves. Gift cards worth $5 to a popular restaurant close to the CSU campus were provided as incentives for participation. The participants consisted of four groups of five members each based on race and gender. The groups were as follows: five African American men, five African American women, five Caucasian men, and five Caucasian women.

As a part of the screening process, prospective participants were asked to report their ethnicity. For purposes of the investigation, African American is defined as having both parents of African American descent. The operational definition of Caucasian is having both parents of Caucasian descent. As a result, prospective participants that consider themselves biracial, or having parents of different races, were excluded from the study.

Based on their own personal 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 a history of or current smoking were excluded from the study. Those who currently had a head cold or upper respiratory infection or chronic sinusitis were excluded. Additionally, participants judged by the investigator to have a foreign accent were excluded. These precautions were taken as a measure to ensure the validity and reliability of the results obtained from this study. The questionnaire used to screen participants can be found in Appendix A.

Three certified and experienced speech-language pathologists (SLPs) who all hold Certificates of Clinical Competence from the American Speech-Language-Hearing Association and three students who were inexperienced in judging vocal quality served as the listening participants for this study. The SLPs were volunteers selected from the CSU Speech and Hearing faculty and staff. The inexperienced students were volunteers selected from CSU as well. In an attempt to determine how evident the perceptual differences in pitch really are, both experienced and inexperienced listeners were used. Accuracy and agreement were calculated for the listeners, both individually and in groups. This data were further analyzed to observe any patterns in perceptual judgments.

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 investigator conducted all testing. The Institutional Review Board (IRB) of CSU approved the methodology for this study. All participants received an informed consent form and checked a box that indicated he or she was aware of their rights and was willing to participate in the study. This form also stated that all information provided would be anonymous and confidential. See Appendix B for a copy of this form.

The speaking participants completed three audio recorded tasks: sustaining a vowel (/a/), reading a printed sentence, and describing a picture. Participants were offered a brief rehearsal period prior to each task to ensure comprehension of the directions and to attempt to increase their general comfort prior to the actual recording period.

Speaking participants were first asked to sustain the vowel /a/ for five 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 F_0 (Fairbanks, 1940). The participants were asked to perform this task twice to ensure proper recording of the sustained vowel.

Participants were then instructed to read a printed sentence aloud in a natural voice. The sentence was printed in large black 36-point font on white paper. The sentence "Kick the ball straight and follow through" was taken from Harvard Sentences, List 2 (Institute of Electrical and Electronics Engineers, 1969). This sentence is phonetically balanced, meaning that the phonemes in the sentence are among the most frequently used in the English language. Phonetically balanced sentences are widely used in research where standardized and repeatable sequences of speech are needed.

In the final task, speaking participants were asked to describe a line drawing in one simple sentence. The line drawing can be found in Appendix D. The line drawing depicted an androgynous stick figure riding a bicycle. This rather bland picture was chosen to reduce the possibility that any emotional expression in the voice of the speakers might impact F_0.

Data Acquisition

All data were collected in a voice lab with minimal distractions and environmental noise. There are no windows in this room. Speaking participants spoke into a microphone placed four to six inches away from their mouths. The Visi-Pitch IV (Model 3950)/Sona-Speech II (Model 3650) uses a sampling rate of 50 kHz, which works best for pitch extraction accuracy (Kay Elemetrics Corp., 2004). Acoustic data were collected and analyzed using the Real-Time Pitch module, which allows the user to capture a speech signal and instantly perform a variety of acoustic analyses.

Perceptual Procedures

One-second samples were extracted from the middle portion of each of the 20 speakers' vowel prolongations. The middle portion of the sample is used 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 one sample from each race (African American and Caucasian). 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 judgment of the race of each speaker, with the understanding that an African American speaker and a Caucasian speaker are both present in the pair. This forced-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 speaker race after each vowel sample, they would be correct approximately 50% of the time and that a forced-choice method of presentation helps to eliminate this variable. This method of pair comparison was chosen because this study attempts to focus on the listener's 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 African American and Caucasian speakers.

Specific directions provided to the listening participants can be found in Appendix E. Listeners were provided with a form and instructed to write the order of presentation of the ten paired voice samples with African American denoted by "A" and Caucasian denoted by "C." A copy of the form provided to participants can be found in Appendix F. A pre-planned order of presentation was given to all listening participants in order to make comparisons between the judgments of the listeners on the same sample pairs. This order of presentation can be found in Appendix G. African American females were presented first in two out of the five samples. African American males were presented first in three out of the five samples. Repetitions of all samples were allowed upon request, with both voices in the pair played.

CHAPTER III

RESULTS

Acoustic Analyses

The data were analyzed separately for males and females from the two racial groups with respect to F_0 in an attempt to determine if the values are significantly different. The raw data for all participants can be found in Appendix H. Descriptive statistics were calculated for these data as part of the acoustic analysis. The results are as follows.

Table I: Mean Ages of Speaking Participants

Group	Mean Age (in years)	Standard Deviation	Range (in years)
AAF	24	5.15	12
CF	25.2	3.90	10
AAM	21.8	3.63	8
CM	22.8	3.56	9

AAF= African American Female; CF= Caucasian Female; AAM= African American Male; CM= Caucasian Male

Table I displays the mean age and the standard deviation for each group of participants. Both groups of Caucasian participants were approximately one year older than their African American counterparts. This miniscule age difference is not believed to affect the validity of results. The greatest variability in age is seen in the African American female group. The ages of all participants can be found in Appendix H.

Sustained Vowel Task

In the sustained vowel task, the entire five second sample was analyzed to determine the F_0. The mean F_0 values and standard deviations for each group are provided in Table II. Caucasian males demonstrate slightly higher (5.23 Hz) F_0 values than African American males. The opposite trend is seen with females. African American females demonstrate a F_0 of 256.58 Hz, while Caucasian females average a F_0 of 216.78 Hz. This is a difference of almost 40 Hz, much greater than the discrepancy seen between the male groups.

Table II: Mean Fundamental Frequencies and Standard Deviations for the Sustained Vowel Task

Group	Mean F_0 Values (in Hertz)	Standard Deviation
AAM	123.22	13.62
CM	128.45	47.37
AAF	256.58	39.20
CF	216.78	28.68

AAM= African American Male; CM= Caucasian Male; AAF= African American Female; CF= Caucasian Female

Reading Task

In the reading task, RFF was calculated by using the entire sentence read into the microphone. Mean F_0 values and standard deviations are provided in Table III. The data for RFF show that African American males have a slightly higher F_0 than Caucasian males on this task (by 5.07 Hz). In regards to females, African American females display a lower F_0 than their Caucasian counterparts (by 1.85 Hz). While the comparisons between the means of the female groups may seem insignificant, a noteworthy trend is observed in the variability of the values calculated for the African American female group. The mean F_0 calculated for African American females in the sustained vowel task is 256.58 Hz, while the RFF calculated for the same group is 201.64 Hz. This is a difference of approximately 55 Hz between the two tasks. The Caucasian females do not display this trend.

Table III: Mean Fundamental Frequencies and Standard Deviations for the Reading Task

Group	RFF (Hz)	Standard Deviation
AAM	129.42	15.39
CM	124.35	34.11
AAF	201.64	31.54
CF	203.49	27.17

AAF= African American Female; CF= Caucasian Female; AAM= African American Male; CM= Caucasian Male

Spontaneous Speech Task

In the final task of the acoustic portion of the present study, the participant's entire spontaneous sentence was used for analysis. Mean F_0 values and standard deviations for this task can be found in Table IV. African American men continue to show slightly elevated F_0 values when compared to their Caucasian counterparts; a discrepancy of 4.64 Hz was observed in this task. However, the trend observed in the female groups on the reading task continues into spontaneous speech. African American females demonstrated a SFF approximately 10 Hz lower than Caucasian females for this task. The calculated SFF for African American females is roughly 7 Hz lower than the RFF calculated for the same group.

Table IV: Mean Fundamental Frequencies and Standard Deviations for the Spontaneous Speech Task

Group	Fundamental Frequency (Hz)	Standard Deviation
AAM	122.64	14.73
CM	118.00	25.22
AAF	194.71	36.12
CF	205.19	28.13

AAF= African American Female; CF= Caucasian Female; AAM= African American Male; CM= Caucasian Male

Intra-Participant Variability

In order to further investigate the phenomena seen in the F_0 values of the African American female group, data on intra-participant variability were gathered. The data collected on the variability of all female participants can be seen in Table V. This variability was measured by calculating the difference between the F_0 values for the sustained vowel and spontaneous speech task. In every participant tested, the frequencies for spontaneous speech were lower than those calculated for the sustained vowel task, however, the greatest discrepancy is seen in the African American female group. On average, there is more than a 60 Hz difference between the F_0's calculated for the sustained vowel task and the spontaneous speech task. The Caucasian females, however, stay relatively constant between all speaking tasks with only minimal variability, as do the male participants. These data were further analyzed via significance tests as reported in the following section.

Table V: Intra-Participant Variability: Differences Between Sustained Vowel and Spontaneous Speech Tasks

Race and Gender	Age	Difference in Samples (in Hz)
AAF	30	102.40
AAF	29	90.23
AAF	18	8.03
AAF	22	34.18
AAF	20	74.49
CF	24	2.59
CF	30	0.13
CF	24	39.50
CF	20	12.88
CF	28	2.86

African American Female= AAF; Caucasian Female= CF

ANOVA and Tukey Post Hoc Testing

A one-way analysis of variance (ANOVA) was performed on this data set using Systat software. This study has one independent variable (F_0) indicating that a one-way ANOVA was warranted in this case. By nature, an ANOVA takes means from all groups into account for analysis. In this case, it meant that the means from the female participants and the means from the male participants were both used. Not surprisingly, the ANOVA determined significant interactions between the means across genders. A more rigorous statistical analysis was needed in order to parcel out the groups of participants and analyze specific data interactions. A Tukey Post Hoc test was used to determine if any of the findings are significant ($p \leq .05$). The *p*-values calculated from the Tukey pairwise comparisons can be found in Table VI. In the sustained vowel task, the *p*-value calculated for the females was roughly 0.3 and 0.99 for males. For the reading task, the *p*-values were found to be approximately 0.99 in the case of females and males. Lastly, in the spontaneous speech task, similar results were yielded: females have a *p*-value of approximately 0.93, while males have a *p*-value of 0.99. The Tukey test revealed that these results are not statistically significant, as the *p*-values are all greater than 0.05.

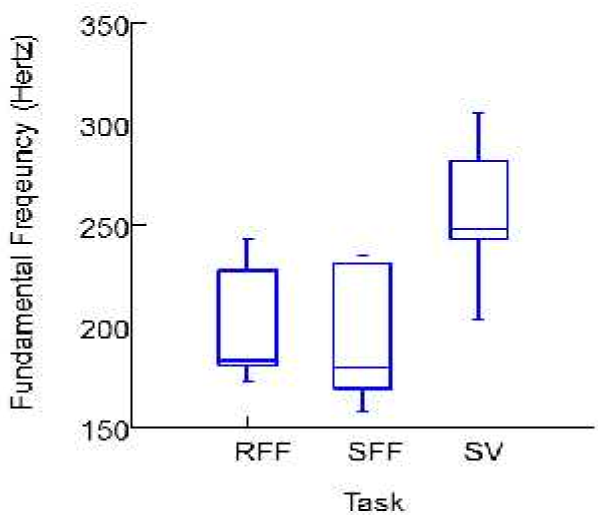
Table VI: Pairwise Comparisons from Tukey Post Hoc Test

Task	AAM to CM (p=)	AAF to CF (p=)
SV	.994995	.300875
RFF	.991529	.999961
SFF	.992809	.927404

AAM= African American Male; CM= Caucasian Male; AAF= African American Female; CF= Caucasian Female
SV= Sustained Vowel; RFF= Reading Fundamental Frequency; SFF= Speaking Fundamental Frequency; $p \leq .05$

An additional Tukey Post Hoc test was performed on the acoustic data provided by the African American female participants. A box plot of this data can be found in Figure 1. The African American females show statistically significant differences between the values obtained for the sustained vowel and values obtained for the spontaneous speech task (p = 0.044383). No significant differences were found between the sustained vowel and reading task (p = 0.075713) or the reading task and the spontaneous speech (p = 0.949834). Figure 1 shows that the medians calculated in each task reside in the lower values of the data set. The greatest range in F_0 for African American females is seen in the sustained vowel task, with values ranging almost 100 Hz. Less variability is seen in the other two tasks.

Figure 1: Within-Group Analysis Box Plot for African American Females



RFF= Reading Fundamental Frequency; SFF= Speaking Fundamental Frequency; SV= Sustained Vowel

Perceptual Analyses

A total of six listeners participated in this portion of the study. The three inexperienced listeners consisted of one male and two females. The three experienced listeners were all females. A percentage was calculated for the accuracy of their perceptual judgments, both individually and collectively as a group. Data collected on accuracy can be found in Table VII. Agreement between group members was also calculated and is found in Table VIII. Agreement is operationally defined as all three group members agreeing on the order of presentation.

Table VII: Accuracy for Individual Listening Participants

Listener	Female Samples	Male Samples	Total Percent Correct
INEX #1	4/5	4/5	80%
INEX #2	3/5	5/5	80%
INEX #3	3/5	4/5	70%
EXP #1	3/5	5/5	80%
EXP #2	2/5	5/5	70%
EXP #3	3/5	5/5	80%

INEX= Inexperienced Listener; EXP= Experienced Listener

Table VIII: Presence of Agreement Within Listening Groups for Each Sample

Paired Sample	INEX	EXP
1	+	+
2	–	–
3	–	–
4	+	–
5	+	+
6	–	+
7	+	+
8	+	+
9	–	+
10	+	+

INEX= Inexperienced Listener; EXP= Experienced Listener
"+" signifies all members were in agreement; "-" signifies all members were not in agreement

Inexperienced Listeners

Specific data collected on accuracy are found in Table VII. Inexperienced listeners were approximately 77% accurate overall in their judgments of the speaker's race. This level is reasonably greater than chance. These listeners were more accurate when judging the male voice samples than the female. As shown in Table VIII, these participants were in agreement for six out of 10 samples, however, all of the three participants were wrong on the first sample in the set. They were in agreement, but, in this case, were not accurate.

Experienced Listeners

Specific data collected on accuracy can be found in Table VII. The experienced listeners were, as a group, approximately 77% accurate overall in their perceptual judgments. All three experienced listeners were 100% accurate in their judgments of the male voice samples. The experienced listeners were in agreement for seven of the 10 samples, which is one more than the inexperienced listeners, as seen in Table VIII. Additionally, as seen in the inexperienced group, all participants were in agreement on the first sample, but were incorrect in this identification. This is the exact paired sample that was guessed incorrectly by all members of the inexperienced group. Both groups were in agreement, meaning all six listeners made identical perceptual judgments for five out of 10 paired samples, two of which were female voices and three of which were male samples.

To further investigate the commonly missed pairs, two samples were analyzed in direct comparison to each other. Table IX shows these data. In the sample that all six listeners judged incorrectly (Pair 1), the African American voice is 63.79 Hz higher than the Caucasian voice. In the sample in which two experienced listeners and one inexperienced listener judged incorrectly (Pair 3), the African American voice is 76.48 Hz higher than the Caucasian voice that followed. These are notable increases in pitch. In both cases, the African American voice was presented first, at a higher pitch, and the lower pitched Caucasian voice followed.

Table IX: Further Analysis of Commonly Missed Paired Samples

Pair	First Sample	Second Sample
1	AAF 243.3 Hz	CF 179.51 Hz
3	AAF 305.89 Hz	CF 229.41 Hz

AAF= African American Female; CF= Caucasian Female

CHAPTER IV

DISCUSSION

Acoustic Data

Between-group comparisons did not yield any statistically significant data. There are, however, interesting interactions between the means of each of the groups. In the male participants, the Caucasian males demonstrate slightly higher F_0 values for the sustained vowel than the African American males. In the following two tasks, where connected speech is involved, the African American males demonstrate slightly higher F_0 values that their Caucasian counterparts. However, the F_0 values obtained for both groups of males are within 10 Hz of each other for all three tasks. This information can be found in Tables II, III, and IV in Chapter III.

Much more variability is seen between the groups of female participants. In the sustained vowel task, African American females demonstrate F_0 values notably above their Caucasian counterparts, with a difference of 40 Hz. However, in the two subsequent tasks, F_0 values for African American females are observed to have fallen for the reading task and are even lower in the speaking task. In the case of the Caucasian female group, the calculated mean F_0 values are within 13 Hz for each of the three tasks. These data can be viewed in Tables II, III, and IV in Chapter III. The F_0 values obtained from the African American female group for the sustained vowel task are 55 Hz higher than those obtained for the same group in the reading task and 60 Hz higher than those obtained in the speaking task. These data are somewhat inconsistent, as similar performances would be expected for each task presented. Further statistical analysis of this phenomenon reveals significant interaction between the F_0 values calculated for the sustained vowel and spontaneous speech task, meaning the values are significantly different (p = .044383). Some researchers suggest that measures utilizing connected speech are more representative of an individual's habitual pitch. This is supported by data collected in this study from African American females that show significant differences in F_0 between tasks.

These acoustic data do not support the hypothesis proposing differences in F_0 between racial groups. However, it is not thought by the investigator that these results disprove the work done by previous researchers in this area. Several possible contributing factors must be taken into consideration. First and foremost, the sample size is extremely small. With only five members in each group, distinctive patterns in data cannot be observed. Additionally, the order of presentation of the speaking tasks may have evoked some psychological effects that impacted the performance of some of the participants. The first task presented was the sustained vowel, which is an unnatural context. Some participants may have felt uneasy or as if they had to “sing” rather than simply phonate. This may be the cause of the elevated pitch observed in the African American female group. The subsequent two tasks were more of a natural speaking context for the participants. Reading and spontaneous speech are parts of daily life, whereas sustaining a vowel sound is not. If the order of presentation of the tasks were reversed or randomized, it may yield different results.

Also, there were no controls put in place for the participant’s body size (i.e., height and weight). It is known that body size and mass have a direct impact on F_0 (Boone et al., 2011). Furthermore, there were no controls in place for the influence of the participant’s linguistic community. The Speech Accommodation Theory proposed by Howard Giles suggests that individuals adjust their voice, mannerisms, and gestures according to their interactions (Miller, 2005). Based on this theory, it is possible that the participants, considering the formality of the situation, adjusted their vocal patterns to accommodate this interaction. Another explanation based on the Speech Accommodation Theory could be that depending on the participant’s linguistic community, his or her voice may already be accommodated. In the case of the first paired sample presented in the perceptual portion of the study in which all listeners guessed incorrectly, the Caucasian female’s voice is 179.51Hz, well below the norm of 225 Hz (See Table IX). The Speech Accommodation Theory suggests that linguistic communities impact F_0. Potentially, in the first sample presented, one of the females may be from a linguistic community,\ or speak a dialect that alters the expected F_0.

Perceptual Data

The perceptual data show that inexperienced and experienced listeners yield the same performance. Both groups were approximately 77% accurate on all 10 samples. These data lend the conclusion that a listener does not need to have formal training in voice disorders to be able to consistently and accurately identify a speaker based on a one-second vowel sample. There must be a particular feature or clue in the voice that one's innate human ability is able to quickly identify and use to pass an accurate judgment. Collectively, the listeners were 60% accurate in judging the female samples and 93% accurate on the male samples. Information on listener accuracy can be viewed in Table VII in Chapter III.

Upon conclusion of the listening tasks, each participant commented that the male samples were easier to distinguish than the female samples. This is evidenced by the increased accuracy seen in the judgments of male voices, although acoustically speaking, a greater difference was observed in females than in male participants. This lends itself to the conclusion that other parameters of voice production may be employed to make racial identifications, but these parameters cannot be determined by the present study.

Table X: Information on the Samples Found to be in Agreement by All Listeners

Pair	First Voice	Second Voice	Difference in F_0 (in Hz)
1	AAF 243.3 Hz	CF 179.51 Hz	63.79
5	CF 251.23 Hz	AAF 248.1 Hz	3.13
7	CM 211.9 Hz	AAM 116.38 Hz	95.52
8	CM 120.9 Hz	AAM 113.46 Hz	7.44
10	AAM 110.38 Hz	CM 106.84 Hz	3.54

AAF= African American Female; CF= Caucasian Female; AAM= African American Male; CM= Caucasian Male

As Table X shows, the six listeners were in agreement for five out of the 10 paired samples, meaning all listeners made identical judgments for the same samples. Two of the samples were female, and three were male. However, in Pair 1 all six listeners made incorrect judgments. In four out of the five samples in which all listeners were in agreement, the listeners identified the voice with a higher F_0 as Caucasian. This information shows that listeners instinctively associate African Americans as having a lower F_0, and, in most cases, their judgments were accurate.

These perceptual data, however, are contradicted by the acoustic data. Perceptual data hold merit because communication involves listening. Everyday analysis of voice is done instantaneously using the listener's own anatomy. By adulthood, the ear and brain are fine-tuned to speech. The human ear and brain may be better suited to differentiate acoustic subtleties and various vocal parameters than acoustical analysis software can, as humans subconsciously make perceptual judgments all day long. For purposes of this study, F_0 was the only vocal parameter examined acoustically. It must not be discounted that another parameter of voice was used by listeners as a basis by which to make a perceptual judgment of speaker race.

Limitations of this Study

Several limitations in the study design must be mentioned. The biggest limitation of this study is the small sample size. With a total of 20 speakers and six listeners, it is nearly impossible to yield statistically significant data. Secondly, without controls for body size or linguistic community (dialect), it can be assumed that F_0 was possibly affected by these factors. Additionally, the order of presentation of tasks may have induced psychological effects that further influenced F_0. By starting with the least natural speaking context, participants may have felt performance anxiety or uneasiness about the task at hand. If the order of tasks were reversed, participants would be "warmed up" to the testing situation and results may have been different. Another limiting factor of this study is that there is no way to confirm the answers to the participant questionnaire. Participants read the form themselves and silently checked boxes. Perhaps if the questionnaire were read aloud by the investigator and the answers were recorded for the participants, false answers may be deterred. Finally, the pitch analysis equipment used and the pitch analyses performed may be considered to be antiquated. Equipment such as videostroboscopy that uses a microphone placed flush to the speaker's neck may yield more accurate data.

Future Research

Future studies should employ a larger sample size in order to increase the power of the results. Additional studies in this area should take body size and linguistic community into account in an attempt to control for any extraneous influences on F_0. If more than one speaking task is involved, the order of presentation should begin with the most natural context in order to make the participant comfortable and able to produce the most genuine results possible. If at all possible, the investigator may wish to do a physical screening in order to rule out any vocal pathology or daily living habits that may have affected the integrity of the vocal folds. Using newer equipment such as videostroboscopy would assist in not only ensuring the integrity of the vocal folds but in yielding more accurate data. Future investigators may also wish to use lengthier reading and speaking tasks to analyze more linguistic information. Additionally, the future studies may benefit from further examination of other vocal parameters as grounds for racial identification.

Conclusion

In conclusion, the results of this present study support the notion that listeners, both trained and untrained in voice disorders, are able to make accurate racial identifications, even with a minimal amount of linguistic information. This evidence implies that the vocal parameter used in perceptual judgments, it is evident even to the naïve listener. However, the acoustic findings do not support the hypothesis suggesting differences in F_0 between racial groups. Acoustic data do demonstrate some small differences in F_0 that may be task specific, rather than racially influenced.

REFERENCES

Abrams, A. (1973). *Minimal auditory cues for distinguishing black from white talkers* (Unpublished doctoral dissertation). City University of New York, New York, NY.

Aguilar, L. (2011). *Gender and identity threat in same & mixed-gender negotiations: Speech accomodation and relational outcomes* (Unpublished doctoral dissertation). Columbia University, New York, NY.

Alvarenga, J. (1971). An investigation of the ability of listeners to differentiate race on the basis of tape recorded evidence (Unpublished master's degree thesis). Herbert H. Lehman College, New York, NY.

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.

Awan, S., & Mueller, P. (1996). Speaking fundamental frequency characteristics of White, African-American, and Hispanic kindergarteners. *Journal of Speech and Hearing Research, 39*, 573-577.

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. (2011). *The voice and voice therapy* (8th ed.). Boston, MA: Allyn & Bacon.

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

Bryden, J. (1968). *An acoustic and social dialect analysis of perceptual variables in listener identification and rating of Negro speakers* (Unpublished doctoral dissertation). University of Virginia, Charlottesville, VA.

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

Curry, E. (1940). The pitch characteristics of the adolescent male voice. *Speech Monographs, 7*, 48-62.

Deutsch, D., Jinghong, L., Shen, S., & Henthorn, T. (2009). The pitch levels of female speech in two Chinese villages. *Journal of the Acoustical Society of America*, *125*(5), 71-76.

Dickens, M., & Sawyer, G.M. (1962). An experimental comparison of vocal quality among mixed groups of Whites and Negroes. *Southern Speech Journal, 18,* 178-185.

Ducote, C. (1983). *A study of the reading and speaking fundamental frequency of aging Black adults* (Unpublished doctoral dissertation). Louisiana State University, Baton Rouge, LA.

Fairbanks, G. (1940). *Voice and articulation drill book.* New York, NY: Harper and Brothers.

Fasold, R. (1981). Relation between Black and White speech in the south. *American Speech*, *56*(3), 163-189.

Felippe, A., Grillo, M., & Grecht, T. (2006). Standardization of acoustic measures for normal voice patterns. *Brazilian Journal of Otorhinolaryngology*, *72*(5), 659-664.

Hanley, T. (1951). An analysis of vocal frequency and duration characteristics of selected samples of speech from three American dialect regions. *Speech Monographs, 12,* 78-93.

Hasek,C., Singh, S., & Murry, T. (1980). Acoustic attributes of preadolescent voices. *Journal of the Acoustical Society of America*, *68,* 1262-1265.

Hecker, M. (1971). Speaker recognition: An interpretive survey of the literature. *ASHA Monographs, 16.*

Hibler, M. (1960). *A comparative study of speech patterns of selected Negro and White kindergarten children* (Unpublished doctoral dissertation). University of Southern California, Los Angeles, CA.

Hixon, T.J., & Abbs, J.H. (1980). Normal speech production. In T.J. Hixon, J.D. Shriberg, & J.H. Saxman (Eds.), *Introduction to communication disorders.* Upper Saddle River, NJ: Prentice Hall.

Hollien, H., Green, R., & Massey, K. (1994). Longitudinal research on adolescent voice change in males. *Journal of the Acoustical Society of America*, *96,* 2646-2654.

Hollien, H., & Jackson, B. (1973). Normative data on the speaking fundamental frequency characteristics of young adult males. *Journal of Phonetics*, *1*, 117-120.

Hollien, H., & Malcik, E. (1962). Adolescent voice change in southern Negro males. *Speech Monographs*, *29*, 53-58.

Hollien, H., & Malcik, E. (1967). Evaluation of cross-sectional studies of adolescent voice change in males. *Speech Monographs, 34,* 80-84.

Hollien, H., Malcik, E., & Hollien, B. (1965). Adolescent voice change in southern White males. *Speech Monographs, 32,* 87-90.

Horli, Y. (1975). Some statistical characteristics of voice fundamental frequency. *Journal of Speech and Hearing Research, 18,* 192-201.

Horli, Y. (1980). Vocal shimmer in sustained phonation. *Journal of Speech and Hearing Research, 23,* 202-209.

Hudson, A. & Holbrook, A. (1981). A study of the reading fundamental vocal frequency of young Black adults. *Journal of Speech and Hearing Research, 24*, 197-201.

Hudson, A. & Holbrook, A. (1982). Fundamental frequency characteristics of young black adults: Spontaneous speaking and oral reading. *Journal of Speech and Hearing Research, 25,* 25-28.

Institute of Electrical and Electronics Engineers Subcommittee on Subjective Measurements [IEEE]. (1969). IEEE recommended practices for speech quality measurements. *IEEE Transactions on Audio and Electroacoustics, 17,* 227-246.

Irwin, R. (1977). Judgments of vocal quality, speech fluency, and confidence of southern Black and White speakers. *Language and Speech, 20,* 261-266.

Kay Elemetrics Corp. (2004). *Instruction manual for Visi-Pitch IV, Model 3950/Sona-Speech II, Model 3650.* Lincoln Park, NJ: Kay Elemetrics Corp.

Lass, N., Mertz, P., Kimmel, K. (1978). The effect of temporal speech alterations on speaker race and sex identifications. *Language and Speech, 20,* 279-290.

Lass, N., Almerino, C., Jordan, J., & Walsh, J. (1980). The effect of filtered speech on speaker race and sex identifications. *Journal of Phonetics, 8,* 101-112.

Lass, N., Tecca, J., Mancuso, R., & Black, W. (1979). The effect of phonetic complexity on speaker race and sex identifications. *Journal of Phonetics, 7,* 105-118.

Liss, J., Weismer, G., & Rosenbek, J. (1990). Selected acoustic characteristics of speech production in very old males. *Journal of Gerontology, 45,* 35-45.

Mayo, R. (1990*). Fundamental frequency and vowel formant frequency characteristics of normal African-American and European American adults* (Unpublished doctoral dissertation). Memphis State University, Memphis, TN.

Mayo, R., & Manning, W. (1994). Vocal tract characteristics of African-American and European-American adult males. *Texas Journal of Audiology and Speech-Pathology, 20,* 33-36.

McGione, R., & McGione, J. (1972). Speaking fundamental frequency of 8 year old girls. *Folia Phoniatrica, 24,* 313-317.

Miller, K. (2005). *Communication theories: Perspective, processes and contexts*. New York: McGraw-Hill.

Morris, R. (1997). Speaking fundamental frequency characteristics of 8- through 10-year old White- and African-American boys. *Journal of Communication Disorders, 30*, 101-116.

Murry, T. (1988). Vocal tract parameters associated with voice quality preference. *Journal of Voice, 2,* 111-117.

Murry, T., Brown, W., & Morris, R. (1995). Patterns of fundamental frequency for three types of voice samples. *Journal of Voice, 9,* 282-289.

Peterson, G., & Barney, H. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America, 24,* 175-184.

Pronovost, W. (1942). An experimental study of methods for determining natural and habitual pitch. *Speech Monographs*, *9*(1), 111-123.

Roberts, M. (1966). *The pronunciation of vowels in Negro speech* (Unpublished doctoral dissertation). The Ohio State University, Columbus, OH.

Saniga, R., Carlin, F., & Farrell, S. (1984). Perception of fry register in Black Dialect and Standard English. *Perceptual and Motor Skills, 59,* 885-886.

Saxman, J., & Burk, K. (1967). Speaking fundamental frequency characteristics of middle aged women. *Folia Phoniatrica, 19,* 167-172.

Schiwitz, C. (2011). *The effect of task type on speaking fundamental frequency in women* (Unpublished master's thesis). Auburn University, Auburn, AL.

Stoicheff, M. (1981). Speaking fundamental frequency characteristics of nonsmoking female adults. *Journal of Speech and Hearing Research, 24,* 437-441.

Van Den Berg, J. (1958). Myo-elastic-aerodynamic theory of voice production. *Journal of Speech and Hearing Research, 1,* 227-244.

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

Wheat, M., & Hudson, A. (1988). Spontaneous speaking fundamental of 6 year old Black children. *Journal of Speech and Hearing Research, 31*, 723-725.

Williamson, G. (2006). *Human communication: A linguistic introduction* (2nd edition). Billingham, UK: Speech Language Services.

Wise, M. (1933). Negro dialect. *Quarterly Journal of Speech, 19*(4), 522-528.

Xue, A. & Mueller, P. (1996). Speaking fundamental frequency of elderly African-American nursing home residents: Preliminary data. *Clinical Linguistics & Phonetics, 10,* 65-70.

Xue, A., Neely, R., Hagstrom, F., & Hao, J. (2000). Speaking F_0 characteristics of Euro-American and African-American elderly speakers: Building a clinical comparative platform. *Clinical Linguistics & Phonetics, 15*(3), 245-252.

Zraick, R., Gregg, B., & Whitehouse, E. (2006). Speech and voice characteristics of geriatric speakers: A review of the literature and a call for research and training. *Journal of Medical Speech-Language Pathology*, *14*(3), 133-142.

APPENDICES

APPENDIX A

SCREENING QUESTIONNAIRE FOR SPEAKING PARTICIPANTS

Age: __________ Gender: ___________ Ethnicity: _____________________

Please check YES or NO for each of the following.

	YES	NO
1. Do you currently have a history of asthma?		
2. Do you currently have a history of acid reflux (GERD)?		
3. Have you had a diagnosis of any voice disorders (i.e., cancer, polyps, nodules, etc)?		
4. Do you currently have a history of smoking?		
6. Do you currently have a head cold or upper respiratory infection?		
7. Do you have emphysema or any other respiratory diagnosis?		
8. Do you have chronic sinusitis?		
9. Do you currently have any medical diagnoses?		

Your participation is greatly appreciated! Thank you!

APPENDIX B

INFORMED CONSENT FORM

I am a graduate student working with Dr. Violet Cox at **Cleveland State University**. I am comparing fundamental frequency (pitch) across ethnic groups. This will help to provide me with information regarding the treatment and diagnosis of voice disorders in minority populations.

I would like to analyze the pitch of your voice in several different contexts in in the Speech Lab of the Cleveland State University Speech and Hearing Clinic in MC 429. This will occur over only one session lasting 30 minutes. Your participation is voluntary. No identifying information will be asked of you in any way. Your participation in this research will be anonymous and all data collected will be confidential.

If you want to know more about this project, please contact me at (440) 269-1052 or my Advisor, Dr. Violet Cox, at (216) 687-6909. This project has been approved by the Institutional Review Board of Cleveland State University. If you have further questions you may contact the Institutional Review Board.

“I understand that if I have any questions about my rights as a research subject I can contact the CSU Institutional Review Board at (216) 687-3630.”

By checking the box below and signing your name, you are stating that you understand the above information and agree to participate in this research.

I understand the above information and consent to participate in this study. ☐

Signature: ____ Date:__________

APPENDIX C

SPECIFIC INSTRUCTIONS PROVIDED TO SPEAKING PARTICIPANTS FOR SUSTAINED VOWEL TASK

"Hold the sound 'ah' at a pitch which seems most natural and easy for you. Don't try to strike any particular pitch, simply relax and speak. Hold this tone for five seconds, I will hold my hand up for you to stop."

APPENDIX D

LINE DRAWING FOR PICTURE DESCRIPTION TASK

APPENDIX E

SPECIFIC INSTRUCTIONS TO LISTENING PARTICIPANTS

"You will hear ten paired samples that are one second in length. One speaker in the pair will be African American and one speaker will be Caucasian. What I would like for you to do is indicate the order of presentation. For example, if you heard a Caucasian voice first and an African American voice second, indicate C/A on the form provided, and vice versa. If you would like to hear a sample again, please ask. The first five samples will be women, while the second five will be men."

APPENDIX F

FORM PROVIDED TO LISTENING PARTICIPANTS

Pair 1: ____________________ ____________________

Pair 2: ____________________ ____________________

Pair 3: ____________________ ____________________

Pair 4: ____________________ ____________________

Pair 5: ____________________ ____________________

Pair 6: ____________________ ____________________

Pair 7: ____________________ ____________________

Pair 8: ____________________ ____________________

Pair 9: ____________________ ____________________

Pair 10: ____________________ ____________________

APPENDIX G

ORDER OF PRESENTATION FOR PAIRED VOWEL SAMPLES

Pair	Voice 1	Voice 2
1	AAF 3	CF 1
2	CF 2	AAF 4
3	AAF 5	CF 4
4	CF 6	CF 6
5	CF 7	CF 7
6	AAM 1	CM 1
7	CM 2	AAM 2
8	CM 4	AAM 4
9	AAM 5	CM 5
10	AAM 6	CM 6

AAF= African American Female; CF= Caucasian Female; AAM= African American Male; CM= Caucasian Male

Numbers correspond to raw data

APPENDIX H

RAW DATA FOR ALL FEMALE PARTICIPANTS

		Mean Frequency	Mean F0	Standard Deviation
BF6	SV	282.14	282.09	3.96
30 yo	RFF	191.36	183.56	34.32
	SFF	185.7	179.69	31.2
BF7	SV	248.84	248.1	9.33
29 yo	RFF	186.75	180.66	32.86
	SFF	159.9	157.87	14.93
BF3	SV	272.88	243.3	66.87
18 yo	RFF	252.7	243.28	43.25
	SFF	237.88	235.27	25.95
BF4	SV	203.69	203.52	4.52
22 yo	RFF	182.32	173.19	35.21
	SFF	175.45	169.34	26.48
BF5	SV	318.87	305.89	41.57
20 yo	RFF	236.71	227.51	38.64
	SFF	232.63	231.4	16.94

WF1	SV	179.71	179.51	7.69
24 yo	RFF	177.2	175.87	15.77
	SFF	179.39	176.92	25.72
WF2	SV	197.83	195.97	13.99
30 yo	RFF	189.53	182.22	35.89
	SFF	199.1	195.84	24.26
WF6	SV	227.85	227.79	4.3
24 yo	RFF	206	192.43	50.69
	SFF	194.23	188.29	33.15
WF4	SV	231.23	229.41	15.08
20 yo	RFF	229.68	223.48	38.16
	SFF	238.84	216.53	77.09
WF7	SV	251.31	251.23	4.51
28 yo	RFF	247.52	238.43	38.81
	SFF	271.85	248.37	71.32

APPENDIX H

RAW DATA FOR ALL MALE PARTICIPANTS

		Mean Frequency	Mean F0	Standard Deviation
BM1	SV	137.11	136.96	5.8
24 yo	RFF	137.5	135.59	21.8
	SFF	132.87	132.55	6.54
BM2	SV	116.48	116.38	4.33
18 yo	RFF	115.09	114.49	8.72
	SFF	108.02	107.76	5.64
BM6	SV	111.47	110.38	18.85
26 yo	RFF	128.02	121.75	38.52
	SFF	113.91	113.56	6.4
BM4	SV	113.48	113.46	1.51
23 yo	RFF	122.55	121.93	9.19
	SFF	116.59	116.12	7.37
BM5	SV	140.01	138.92	9.32
18 yo	RFF	170.23	153.32	63.28
	SFF	145.12	143.22	17.12

WM1	SV	98.7	98.68	1.23
23 yo	RFF	97.46	96.42	10.57
	SFF	106.98	101.69	36.48
WM2	SV	211.92	211.9	2.08
20 yo	RFF	187	183.48	26.32
	SFF	161.9	158.42	22.68
WM6	SV	106.9	106.84	2.63
19 yo	RFF	125.78	119.26	34.3
	SFF	131.27	124.59	36.94
WM4	SV	130.11	120.9	49.72
24 yo	RFF	114.88	113.72	11.67
	SFF	111.33	110.72	8.18
WM5	SV	104.95	103.94	18.33
28 yo	RFF	110.12	108.89	12.77
	SFF	94.61	94.57	2.17

APPENDIX H

RAW DATA FOR AVERAGES FOR EACH GROUP OF PARTICIPANTS

			Mean frequency	Mean F0	Standard Deviation	Range
Black Females						
	Avg Age: 24	SV	265.28	256.58	25.25	102.37
	19, 20, 22, 29, 30	RFF	209.97	201.64	36.86	54.32
		SFF	198.31	194.71	23.1	77.4
White Females						
	Avg Age: 25.2	SV	217.59	216.78	9.11	71.72
	20, 24, 24, 28, 30	RFF	209.98	203.49	35.86	62.56
		SFF	216.68	205.19	46.31	71.45
Black Males						
	Avg Age: 21.8	SV	123.71	123.22	7.96	28.54
	18, 18, 23, 24, 26	RFF	134.68	129.42	28.3	38.83
		SFF	123.3	122.64	8.61	35.46
White Males						
	Avg Age: 22.8	SV	130.52	128.45	15	113.22
	19, 20, 23, 24, 28	RFF	127.05	124.35	19.12	87.06
		SFF	121.22	118	21.29	63.85