2018-10-01

# Acoustic Correlates of Aging and Familial Relationship

Samantha Michelle Taylor
*Brigham Young University*

Acoustic Correlates of Aging and Familial Relationship

Samantha Michelle Taylor

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Christopher Dromey, Chair
Kristine Tanner
Shawn Nissen

Department of Communication Disorders

Brigham Young University

# ABSTRACT

Acoustic Correlates of Aging and Familial Relationship

Samantha Michelle Taylor
Department of Communication Disorders, BYU
Master of Science

The purpose of this study was to examine the potential differences in selected acoustic measures of speech as a function of age, across sexes, and between families. The data used in this study were previously collected for a larger project on voice production at the University of Utah. Participants included 169 individuals, 79 men and 90 women, from 18 Utah families, ranging in age from 17 to 87 years. All participants had no history of articulation disorders, stroke or active neurologic disease, or severe-profound hearing loss. Participants were recorded reading two passages aloud in a sound booth. These two passages were selected as connected speech tasks from which to extract the following acoustic metrics: fricative spectral measures (center of gravity, standard deviation, skewness, and kurtosis), mean fundamental frequency ($F_0$), semitone standard deviation (STSD), speaking time ratio, and cepstral peak prominence smoothed (CPPS). Results indicated significant aging effects on spectral center of gravity and skewness, mean $F_0$, and STSD. There was a significant sex effect for spectral center of gravity and kurtosis, mean $F_0$, speaking time ratio, and CPPS. Familial relationship had a significant effect for spectral skewness, STSD, and CPPS. Findings from the current study indicate that certain speech and voicing features point to a decline with age and that aging affects the speech of men and women differently. Additionally, these data suggest that related speakers may demonstrate similar patterns for prosody, voicing, and articulation behavior, although the statistical testing did not allow us to draw specific inferences about such similarities. These findings describe some normal variations in the speech production of persons of differing age, sex, and familial background. An understanding of these normal speech differences in healthy individuals is valuable for differentiating between typical and pathological speech patterns in a clinical setting.

Keywords: acoustic speech measures, age, familial relationship

## ACKNOWLEDGMENTS

The completion of this project would have been impossible without the innumerable contributions of my advisor Dr. Christopher Dromey. Thank you for your seemingly endless patience, positivity, and wisdom as I waded through this process. I would also like to thank my other committee members for lending their expertise as many times as I asked for it, as well as Dr. Dennis Eggett for performing the statistical analysis and Dr. Kim Corbin-Lewis at Utah State University for providing access to the data. The guidance of my professors made a research project of this magnitude seem less intimidating than I once feared it would be.

I also owe a great debt of gratitude to the other students in my cohort. The bond we forged in the early mornings, long days, and late nights (and sometimes more) of graduate school is one I will not soon forget. Our shared learning and friendship have made me a better clinician, researcher, and person. I would also like to extend gratitude to the friends from USGA (Understanding Sexuality, Gender, and Allyship) at BYU who became family as I navigated a difficult chapter of my life. Thank you for providing me with motivation and distractions in appropriate doses.

Finally, I am eternally grateful to my actual family for their constant, unconditional love and support throughout my education and all of life's journeys. Thank you for extending me the patience, encouragement, trust, and confidence I needed during this difficult process. Any accomplishment of mine belongs to all of you as well.

# TABLE OF CONTENTS

# LIST OF TABLES

## LIST OF FIGURES

# DESCRIPTION OF THESIS STRUCTURE AND CONTENT

This thesis, *Acoustic Correlates of Aging and Familial Relationship*, is written in a hybrid format which combines traditional thesis requirements with communication disorders journal publication formats. The body of this thesis is written as a manuscript suitable for submission to the university. An annotated bibliography is found in Appendix A, reading passages used in this study are included in Appendix B, and descriptive statistics are reported in Appendix C.

## Introduction

When designing effective, evidence-based speech therapy, clinicians rely on findings from recent clinical research to inform treatment planning. Such research is conducted to evaluate the efficacy of various treatment designs by assessing an individual's speech performance pre and post treatment. There are a variety of approaches available to researchers for evaluating treatment-related change, including perceptual, kinematic, and acoustic methods of measurement. Each approach has unique benefits as well as inherent drawbacks that require consideration.

Perceptual evaluations of articulation, phonation, and resonance have great clinical value in that they are noninvasive and reflect a speaker's speech performance in a naturalistic environment. Additionally, clients and caregivers generally rely on perception to determine whether treatment has been successful, rather than inquiring about changes in acoustic/kinematic precision. Many clinicians rely on both formal and informal perceptual evaluations to track treatment-related progress. Some studies have found listener perception to correlate well with more objective measures for certain aspects of speech, including jitter, signal-to-noise ratio, and fricative spectral measures (Eadie & Doyle, 2005; McFarland, Baum, & Chabot, 1996). Others, however, suggest that auditory-perceptual ratings are susceptible to biases across multiple dimensions and often do not reliably describe speech performance, especially in disordered populations (Kent, 1996). For this reason, more objective measures are preferred for tracking speech performance in a clinical research setting.

Kinematic measures of speech production provide quantitative, objective information about articulation by directly tracking movement of structures in the oral cavity during speech. This can be particularly valuable for differentiating between normal and disordered speech in a clinical research setting. However, kinematic measures involve attaching transducers directly to articulatory structures (e.g., tongue and lips), which introduces a slight oral-articulatory perturbation. Such an invasive approach can affect natural articulation patterns, thus providing information that may not closely match everyday speech performance. Research has also found kinematic measures to affect sensation and tactile feedback during speech, which play an important role in speech motor control (Lametti, Nasir, & Ostry, 2012). Although typical speakers have been found to adapt over time to the presence of transducers in the mouth, some studies suggest that speech patterns may never fully return to normal in the presence of perturbation, particularly in speakers with a history of articulation disorders (McFarland et al., 1996).

Acoustic metrics allow for measurement that is objective, quantitative, and noninvasive. Analyzing acoustic properties of speech output provides valuable information about the behaviors of the speech mechanism without affecting speech patterns. For example, formants are associated with the shape of the vocal tract during sonorant production (Simpson, 2002); fundamental frequency measures change according to the vibratory rate of the vocal folds (Nishio & Niimi, 2008); and voice onset times reflect the coordination of the larynx with the articulators (Westrop, 2000). These and other acoustic metrics are valuable for describing both normal and disordered speech, as well as quantitatively tracking treatment-related progress. While instrumental measures such as these may be impractical for routine use in a clinic, they are valuable for evaluating treatment efficacy in clinical research.

Conducting clinical research is of particular importance in elderly populations, which comprise a large portion of those who receive speech services. As of 2017, persons aged 65 years or older number 65 million in the United States: approximately one in every five Americans. By

2020, that number is expected to climb to 77 million (U.S. Department of Health and Human Services, Administration for Community Living, Administration on Aging, 2017). Because speech pathologists play an important role in serving such a large population, it is necessary to establish patterns of typically aging speech in order to distinguish healthy elderly speakers from those exhibiting pathological speech processes. This first requires an understanding of how typical aging affects the human body in general. Research suggests that, after the third decade of life, bodily function generally undergoes a linear decline which is reflected in changes in the cardiovascular, respiratory, skeletal, and muscular systems (Ramig, et al., 2001). Such changes include (a) increased arterial rigidity, (b) reduced respiratory efficiency due to calcification of rib cartilage and weakening of rib muscles, (c) reduced lingual and pharyngeal elasticity, (d) changes in the dimensions of the oral cavity, and (e) calcification of the larynx and atrophy of the laryngeal muscles (Ramig et al., 2001; Slawinski, 1994). Histological studies of human vocal folds have revealed changes in the lamina propria with age, including reduced thickness and increased elastin content (Hammond, Gray, Butler, Zhou, & Hammond, 1998).

The typical aging process also involves neurological changes, both structural and functional. Many elderly individuals exhibit a diminished number of cerebellar cells as well as atrophy of the temporal lobes and the precentral gyrus of the cerebral cortex (Slawinski, 1994). Additionally, aged individuals are generally found to have decreased synaptic activity characterized by a reduction in the release of neurotransmitters (Ramig et al., 2001). These neurological changes may cause a slowing of neural processes which, in turn, may lead to decreased speed and accuracy of motor control as well as diminished cognitive-linguistic functioning (Torre & Barlow, 2009).

These neuromuscular changes to the speech mechanism are reflected in concomitant changes to the acoustic properties of speech with age. Much research has been devoted to studying the effects of aging on the voice, particularly the study of fundamental frequency ($F_0$), the acoustic correlate of perceived pitch. It is generally agreed that, with increasing age, female speakers demonstrate a decrease in $F_0$, while male speakers demonstrate an increase in $F_0$ beyond the fourth decade (Nishio & Niimi, 2008; Ramig et al., 2001). While consensus has been reached with respect to the direction of $F_0$ change, researchers have reached widely different conclusions about the magnitude of that change (Nishio & Niimi, 2008). There exists considerable inconsistency in the literature regarding the timing and degree of $F_0$ changes with age, and further studies are needed to better establish typical $F_0$ patterns in geriatric populations.

Elderly phonation is characterized by instability, which is often reflected in perceived tremor, voice breaks, hoarseness, and/or breathiness (Gorham-Rowan & Laures-Gore, 2006). These perceptual qualities are likely related to acoustic measures of perturbation, such as amplitude perturbation quotient (APQ) and fundamental frequency standard deviation ($F_0$SD). Both APQ and $F_0$SD have been found to increase with age, indicating a reduction in amplitude and $F_0$ stability, which listeners may rely on to estimate a speaker's age (Xue & Deliyski, 2001). The perceived hoarseness/breathiness in elderly voices is also likely associated with noise in the voice signal, which can be measured as a harmonics-to-noise ratio (HNR). Research indicates that HNR is significantly lower in elderly speakers than in younger adults (Ferrand, 2002). This has been attributed to less periodic vocal fold vibration and/or incomplete vocal fold closure, which have been observed in geriatric speakers (Gorham-Rowan & Laures-Gore, 2006).

Because HNR measures are only valid for sustained phonation, some researchers have used spectral/cepstral measures to describe voice quality in connected speech, which more accurately reflects a speaker's phonation in natural speaking contexts. Measures of voice quality in connected speech are also preferred in the presence of voice disorders such as spasmodic dysphonia, in which phonation is less symptomatic in sustained phonation than in continuous speech (Awan, Roy, & Dromey, 2009). One such measure is Cepstral Peak Prominence (CPP), which has been found to closely correlate with vocal breathiness (Awan et al., 2009). In a study of typical adults, Garrett (2013) found older speakers to have significantly lower CPP than younger adults in connected speech, indicating more severe breathiness. Increases in amplitude fluctuation, $F_0$ variation and noise in the voice signal contribute to instability in elderly voices.

The process of typical human aging also has an impact on several aspects of speech motor control. Age-related neuromuscular changes affect the regulation of articulator movement amplitude, which results in generally decreased accuracy in speech production (Ramig et al., 2001; Torre & Barlow, 2009). Articulatory precision has been measured through acoustic analysis of fricative spectral moments, which consider spectra as a distribution of frequencies and describe the mean (centroid frequency), standard deviation (spread), skewness (spectral tilt), and kurtosis (peakedness) of that distribution (McFarland et al., 1996). Spectral moment analysis can reflect the configuration of the articulators during speech production. For example, the shorter, wider, and more posterior constriction /ʃ/ is associated with relatively low spectral centroid frequencies compared with /s/, which is characterized by more anterior placement and a longer, deeper groove (Tjaden & Turner, 1997). Individuals with neurodegenerative diseases have been found to produce atypical fricative spectra associated with reduced articulatory precision (Tjaden & Turner, 1997). It can thus be anticipated that, as their neuromotor control declines, aged speakers will produce fricative spectra that differ from younger adults.

In addition to changes in these characteristics of phoneme segments, older adults exhibit changes in suprasegmental features of speech production as well. Suprasegmental features – often collectively referred to as prosody – are properties of syllables and larger units of speech which serve linguistic functions, including intonation, stress, and rhythm. Research suggests that, when reading aloud, older speakers produce larger intonational ranges and more frequent inflections than younger adults, perhaps to demarcate syntactic and semantic units in continuous speech in the presence of decreased articulatory precision (Barnes, 2013). Listeners also rely on pauses to understand syntactically complex sentences. In healthy adults, pausing patterns have been found to correlate closely with syntactic boundaries (Huber, Darling, Francis, & Zhang, 2012). However, research indicates that elderly speakers exhibit different pausing structures than younger adults. Older adults demonstrate more frequent breaths, fewer syllables per breath, increased frequency and duration of pauses, and greater intraspeaker variability in temporal parameters (Ballard, Robin, Woodworth, & Zimba, 2001).

Researchers have suggested several possible factors which may contribute to these changes in speech motor control with age, including (a) declining neuromotor control, (b) impaired sensory feedback due to age-related hearing loss, (c) changes to respiratory physiology and (d) slower cognitive processing (Huber et al., 2012; Torre & Barlow, 2009). Although these and likely many other factors cause a general decline in voice quality and speech motor control among older adults, they are part of the typical human aging process and must therefore be distinguished from pathological speech behaviors. An understanding of this distinction is crucial in carrying out clinical research and tracking treatment-related progress of geriatric speakers.

Another important consideration for conducting clinical research is the need to evaluate the speech features of males and females separately. Research has clearly established the presence of biomechanical differences in the speech mechanism across sexes. Perhaps the most apparent differences lie in the characteristics of phonation. The average speaking fundamental frequency ($SF_0$) for typical adult males is generally agreed to lie between 100-146 Hz, while typical adult females average between 188-221 Hz (Gelfer & Mikos, 2005). As noted earlier, $SF_0$ changes with age for speakers of both sexes. However, aging seems to affect male and female voices differently and disproportionately. For example, Nishio and Niimi (2008) found female voices to demonstrate a clear decrease in $SF_0$ in each decade after the second decade of life, while male voices showed only a slight increase in $SF_0$ after age 70.

Although research suggests that listeners rely primarily on $SF_0$ to determine speaker sex, there are other basic differences in the phonatory mechanisms of male and female speakers that require consideration. For example, male laryngeal structure differs from that of females in several key ways. Males have larynges whose cartilage is 20% larger than females' in the anterior-posterior dimension, and vocal folds whose membranous portion is 60% longer than those of females (Titze, 1989). Also, male vocal folds are approximately 20% to 30% thicker than female vocal folds, with a greater percentage of collagenous fibers (Titze, 1989). These structural laryngeal differences result in vocal fold vibration that is not only slower, but qualitatively different in males compared with females.

Male-female anatomical differences are also present in supraglottic features of the speech mechanism. On average, male vocal tracts – consisting of the structures between the lips and the vocal folds – are approximately 15% longer than female vocal tracts, indicating that male speakers will generally resonate lower formant frequencies than female speakers (Titze, 1989). This has been corroborated by findings that vowel formant frequencies are generally higher in female speakers, particularly F1 and F2 (Simpson, 2002). Although secondary to $SF_0$, research indicates that vowel formant frequencies play an important role in differential sex identification (Titze, 1989). Simpson (2002) has suggested, however, that sex differences in acoustic metrics such as this are nonuniform because many anatomical features of the speech mechanism are not linearly proportional across sexes. For example, males and females present with different average ratios of oral-to-pharyngeal cavity length, which in turn produces nonuniform differences in resonance properties. More research is needed to understand the intricate ways in which male and female speech production differs and establish patterns of typical speech production for men and women.

Several studies have examined data which suggest the presence of familial patterns in the speech production of related individuals. Untrained listeners have been found to reliably differentiate between related and unrelated speakers (Feiser & Kleber, 2012; Vanderydt, 1998). However, there is a paucity of research which adequately describes the speech features responsible for perceived similarity in the speech of family members. One study reported that accurate listener identification of related speakers was correlated with more pronounced dialectal features (Feiser & Kleber, 2012). Another study found that family members who were perceived as sounding related had similar $F_0$, intonation patterns, and voice onset and offset timing (Vanderydt, 1998). Westrop (2000) also reported familial relationship to have a significant effect on segment durational differences, including stop-gap durations and voice onset time. However, other studies have indicated that some acoustic similarities between pairs of unrelated speakers are often stronger than similarities between siblings, even those who are rated by listeners as sounding similar (Feiser, 2009). While the data suggest that acoustic trends in familial speech production exist, it remains unclear which speech features are affected by familial relationships and what mechanisms are responsible for such trends.

Research has established that acoustic measures reflect important changes across the lifespan, differences between sexes, and even similarities within family relationships. The present study was conducted to examine selected acoustic aspects of the speech of healthy individuals within families and across the adult life span. The availability of a data set with 169 individuals aged 17 to 87 from 18 different families allowed for an acoustic examination of speech properties in a large and heterogenous cross section of the population. The data set examined in the present study was also used by Westrop (2000) to measure segment durational changes associated with age, gender, and familial relationship. She evaluated speech samples from 93 subjects across 10 families, aged 21 to 82. Each participant repeated the sentence, *The blue spot is on the key again* three times. Westrop measured vowel duration for /u/, /ɑ/, /ɪ/, and /i/, stop-gap duration, and voice onset time (VOT) with the following results: a significant age effect on /ɪ/ duration and VOT (both increased with age); a gender effect on /ɑ/ duration (longer in males) and VOT (shorter in males); and a family effect on stop-gap duration. The author suggested several possible explanations for these findings, including (a) slower speech production with increasing age, (b) differences in oral cavity shape and size across genders, and (c) genetic patterning and/or environmental influences within families.

It was hypothesized that these factors may also impact other acoustic characteristics of speech production. Therefore, the present study was conducted to examine the potential differences in segmental and suprasegmental speech features with aging, across sexes, and between families. Fricative spectral measures and spectral/cepstral analysis were selected as indices of articulatory precision and vocal quality, respectively. It was anticipated that both measures would indicate a decline in speech performance with age and reflect physiologic differences between sexes and, potentially, unrelated individuals.  In order to characterize prosodic differences among and between these participant groups, the present study relied on $F_0$ variability and a ratio of speaking/pausing time in connected speech. Taken together, these segmental and suprasegmental measures may describe normal variations in the speech production of persons of differing age, sex, and familial background. An understanding of normal speech differences in healthy individuals is crucial for distinguishing between typical and pathological speech patterns in the clinical setting.

## Methods

### Participants

The data used in this study were previously collected for a larger project on voice production as a part of the Utah Genetic Reference Project (UGRP) at the University of Utah, funded by the Keck Foundation. The present study included speech samples from 169 individuals in 18 Utah families. The availability of data from related speakers allowed for the analysis of speech acoustics without the complicating influence of diverse dialects. Participants included 79 men and 90 women ranging in age from 17 to 87 years. All participants were native speakers of American English with no history of articulation disorders, stroke or active neurological disease, or severe-profound hearing loss. Participants were not excluded for mild-moderate health problems typical of their age group, such as high blood pressure, thyroid problems, diabetes, or mild-moderate hearing loss.

### Procedures

Participants were recorded in a sound booth with a digital audio tape (DAT) recorder and an AKG Acoustics C420 head-mounted condenser microphone with a constant 3 cm mic-to-mouth distance. Each sample was perceptually judged to be normal (i.e., without overt articulation deficits) with 100% agreement by two certified speech-language pathologists with at least five years clinical experience. Recordings were transferred from DAT to 48 kHz audio files using a Dell computer with the Kay Elemetrics (Tokyo, Japan) DAT interface system.

The data collection protocol for this study included two reading passages: (a) *Goldilocks and the Three Bears*, and (2) the *Rainbow Passage*. Participants were given large-print copies of each passage to read aloud. Prior to reading the Goldilocks passage, each speaker was given the following instruction, “Read this story as if you were reading to a small child, changing your voice for each bear and Goldilocks.” For the *Rainbow Passage*, participants were instructed to read in their normal, conversational voice. These two passages were selected as connected speech tasks from which to extract fricative spectral measures, mean fundamental frequency, semitone standard deviation (STSD), speaking time ratio, and the cepstral peak prominence smoothed (CPPS).

**Acoustic Analysis**

For each participant, fricative center of gravity and standard deviation were measured from /s/ in each of five occurrences of the word *said* in the Goldilocks passage and then averaged. This context was chosen to control for coarticulation effects and ensure consistency of fricative production. Using the Praat software program (version 5.4; Boersma & Weenink, 2016), each /s/ was segmented and saved as a separate .wav file. The boundaries of each /s/ were identified using the combined audio waveform and spectrographic display in Praat, confirmed by listening. Next, the .wav files were read into a custom Matlab application to further segment each fricative to save only the middle 50% of the total phoneme duration. Praat’s spectral slice function was then used to compute the center of gravity, standard deviation, skewness, and kurtosis in Hz for each segmented fricative. The five sets of spectral values for each participant were then averaged prior to statistical analysis.

Measures of prosody, including $F_0$ variability and the proportional amount of time participants spoke and paused, were computed from readings of the *Rainbow Passage*. Recordings of this passage were imported into Praat and trimmed to exclude experimenter speech, pauses at the beginning and end of each reading, and nonspeech behaviors (coughing, laughing, etc.) prior to analysis. $F_0$ mean and standard deviation were measured from these trimmed recordings using Praat's voicing report, adjusting the $F_0$ range as needed to avoid tracking errors. Because males and females usually have different $F_0$ ranges, *SD* values were converted into semitones using a spreadsheet equation, allowing for a direct comparison of $F_0$ variability for all speakers.

Trimmed recordings of the *Rainbow Passage* were also used to compute each speaker's speaking time ratio, expressed as a proportion of time spent speaking relative to the total passage duration. For example, a speaking time ratio of 1.0 would indicate all speaking and no pausing, while a ratio of 0.75 would indicate 75% speaking and 25% pausing. These ratios were computed using a custom Matlab application, which defined pauses as lasting longer than 200 ms. The application normalized the intensity of the audio signal to 100. Ten percent of the normalized maximum was selected as the threshold, above which was operationally defined as speaking and below which was defined as pausing.

Finally, the *Rainbow Passage* files were also used to perform cepstral/spectral measures as an index of vocal quality in connected speech for each participant. Trimmed recordings of the Rainbow Passage were imported into Praat. Praat used to compute cepstral peak prominence smoothed (CPPS), which is a measure reflecting the energy of the periodic signal above the linear regression line of other noise in the signal. Higher values indicate a less noisy signal and vice versa.

### Statistical Analysis

The acoustic dependent variables were tested with Statistical Analysis System (SAS) software. Dr. Dennis Eggett of the Brigham Young University Department of Statistics performed the statistical analysis using a mixed model ANCOVA. The independent variables were age, sex, and family membership of the speakers. Dependent variables included fricative spectral measures (center of gravity, spread, skewness, and kurtosis), $F_0$ mean, STSD, speaking time ratio, and CPPS.

## Results

### Fricative Spectral Measures

**Sex.** Testing revealed a significant difference in the spectral center of gravity for /s/ between male and female speakers, $F(1, 17) = 8.74$, $p = .009$. As seen in Figure 1, females were found to have a higher spectral center of gravity. There was no significant effect of sex on spectral spread or spectral skewness for /s/. However, ANCOVA testing did reveal that females had a significantly higher spectral kurtosis measure for /s/ than males, $F(1, 17) = 12.09$, $p = .003$, as shown in Figure 2.

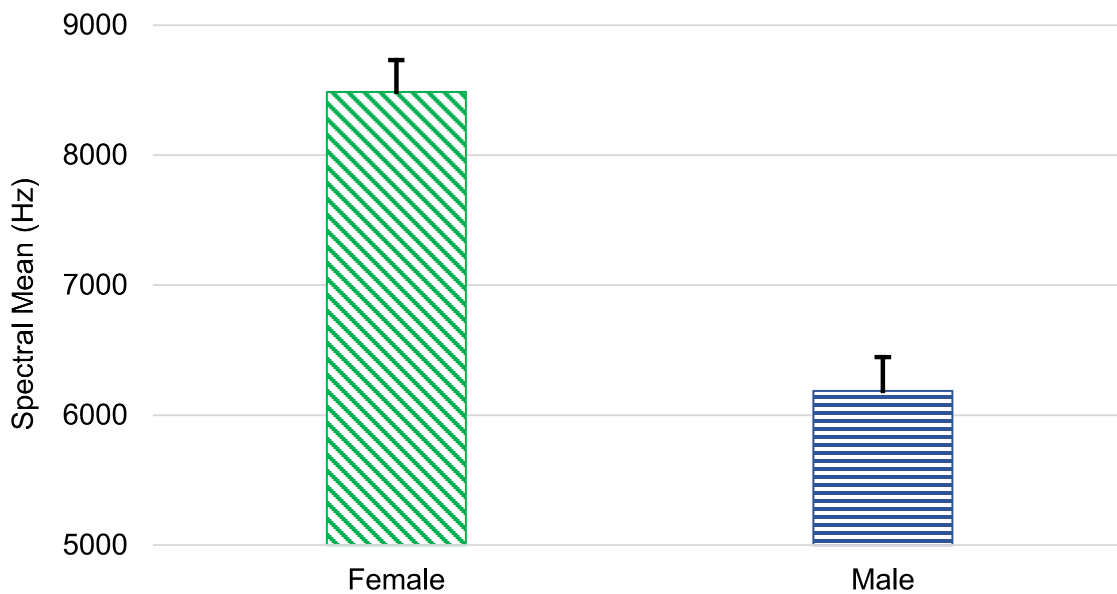


*Figure 1.* Mean (and 95% confidence interval) of the spectral mean for /s/ in Hz for women and men.

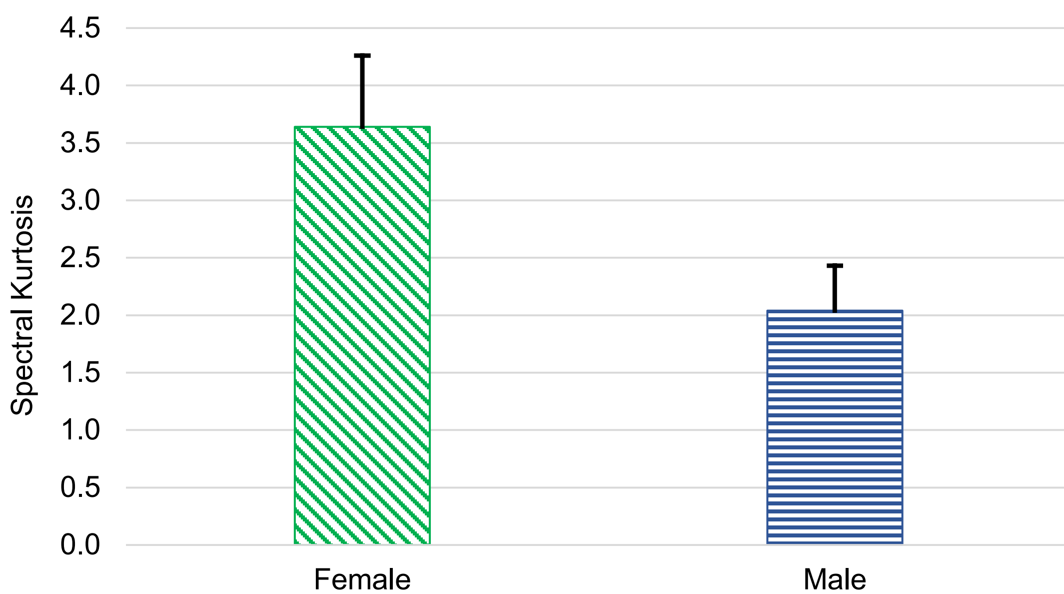


*Figure 2.* Mean (and 95% confidence interval) of the spectral kurtosis for /s/ for women and men.

**Age.** Speaker age was found to have a significant effect on spectral center of gravity for /s/, $F(1, 131) = 24.19$, $p < .001$. As shown in Figure 3, an age-by-sex interaction was found, indicating that males show a greater decline in spectral center of gravity with age than females, $F(1, 131) = 24.19$, $p < .001$. Age had no significant effect on spectral spread for /s/. There was, however, a significant age effect for spectral skewness, $F(1, 131) = 9.78$, $p = .002$. As seen in Figure 4, an age-by-sex interaction revealed that, while there was no age effect for females, male speakers' spectral skewness increased with age. There was no significant difference in spectral kurtosis for /s/ with age.

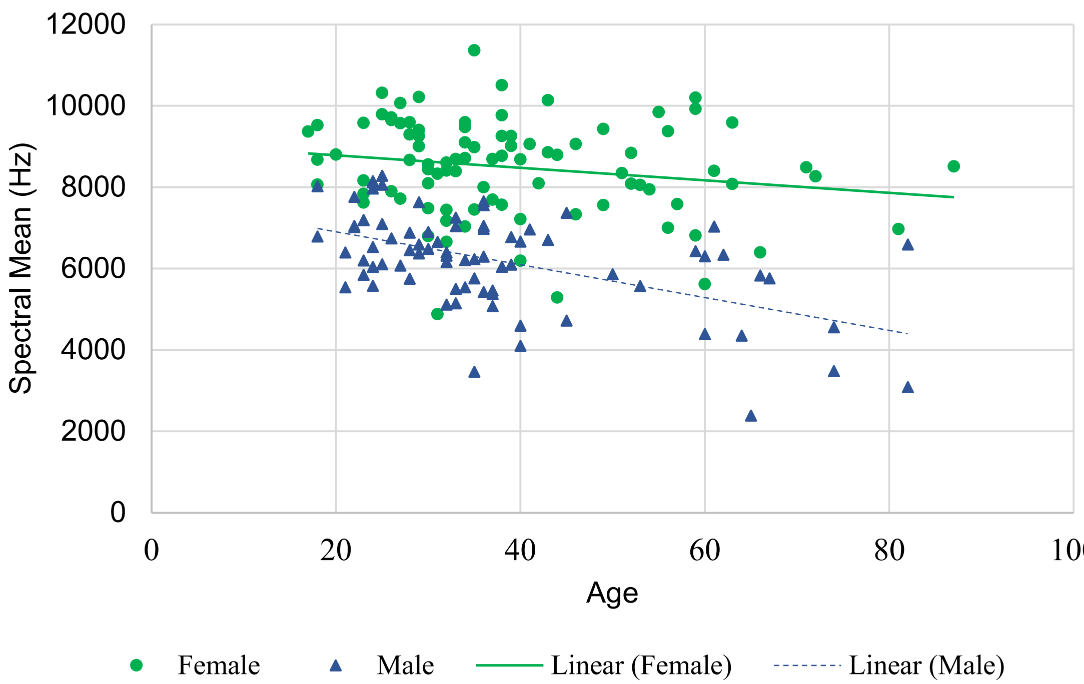


*Figure 3.* Spectral mean for /s/ in Hz for women and men as a function of age (lines represent a linear regression analysis).

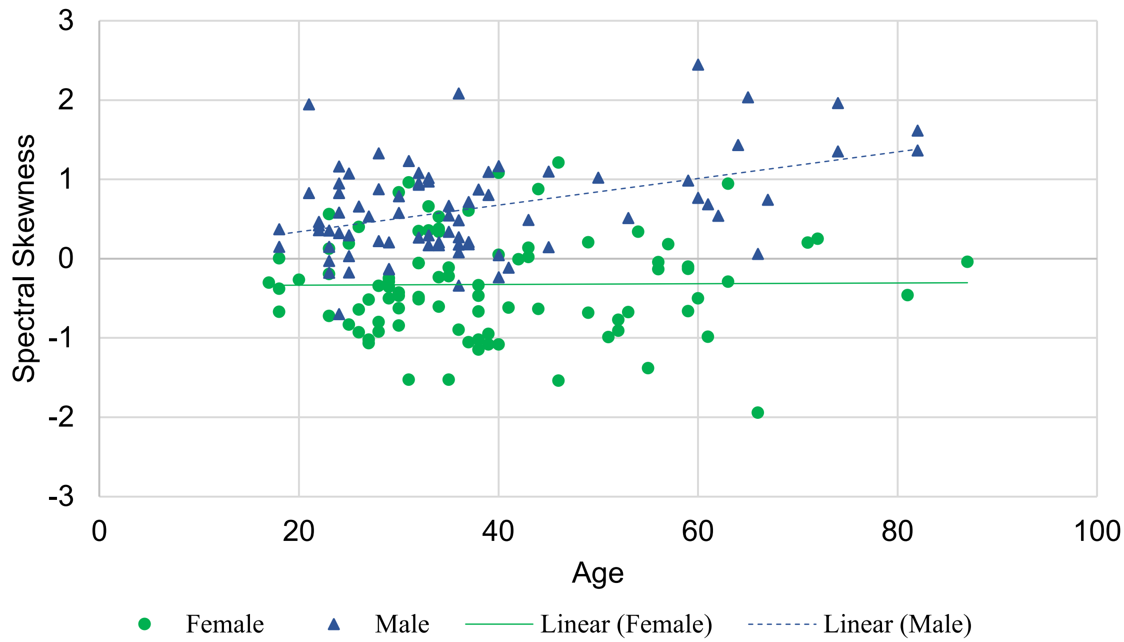


*Figure 4.* Spectral skewness for /s/ for women and men as a function of age (lines represent a linear regression analysis).

**Family.** ANCOVA testing revealed that family relationship had no significant effect on spectral center of gravity, spread, or kurtosis for /s/. However, there was significant variability among the families for spectral skewness, $F(17, 147) = 2.47$, $p = .002$. Descriptive statistics for spectral center of gravity and spread are reported in Table 1, and descriptive statistics for spectral skewness and kurtosis are reported in Table 2.

**Fundamental Frequency Measures**

**Sex.** As shown in Figure 5, female speakers were found to have a significantly higher F0 *M* than male speakers, $F(1, 17) = 195.82$, $p < .001$. Speaker sex had no significant effect on STSD.

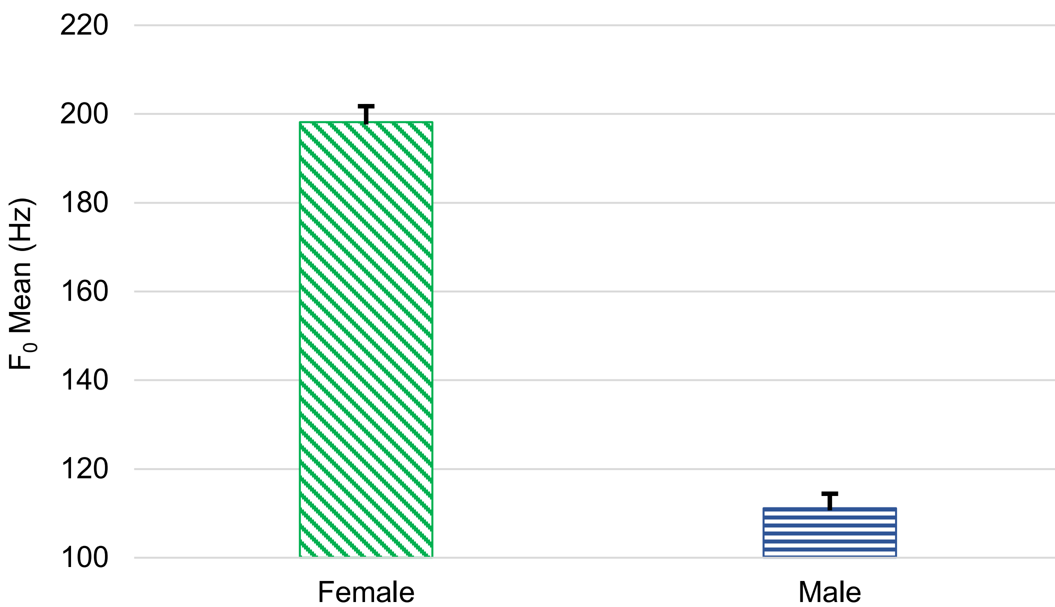


*Figure 5.* Mean (and 95% confidence interval) fundamental frequency in Hz for women and men.

**Age.** ANCOVA testing revealed a trend for an age-by-sex interaction for $F_0$ *M*. As seen in Figure 6, male speakers' values increased with age while female speakers' decreased with age, $F(1, 131) = 3.58$, $p = 0.061$. As shown in Figure 7, STSD was found to increase significantly with age, $F(1, 132) = 29.09$, $p < .001$.

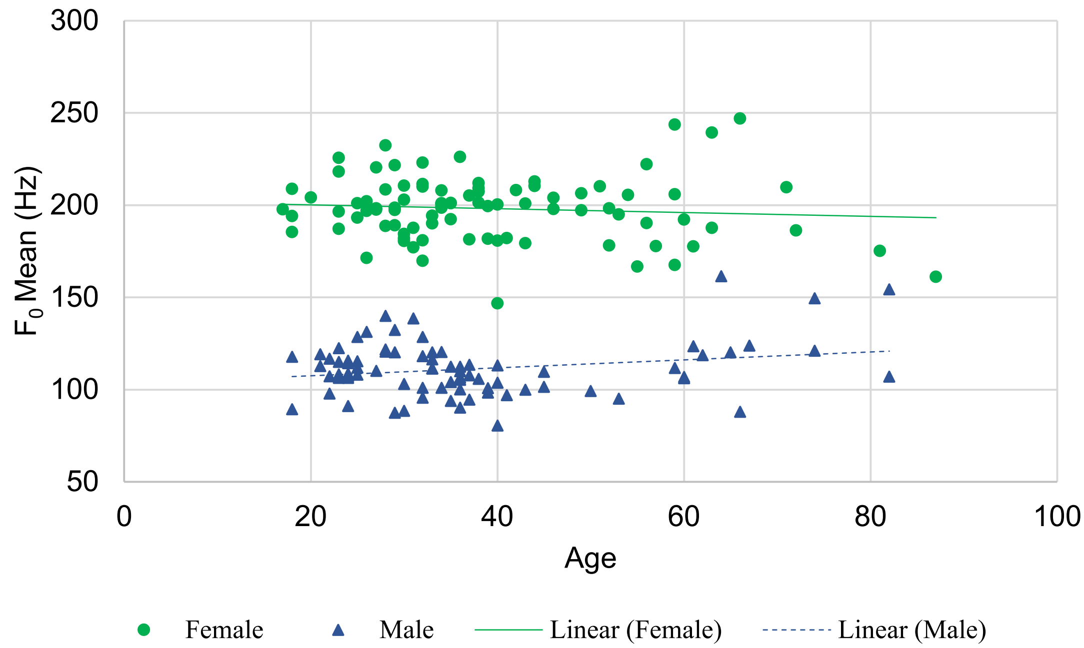


*Figure 6.* Fundamental frequency mean in Hz for women and men as a function of age (lines represent a linear regression analysis).

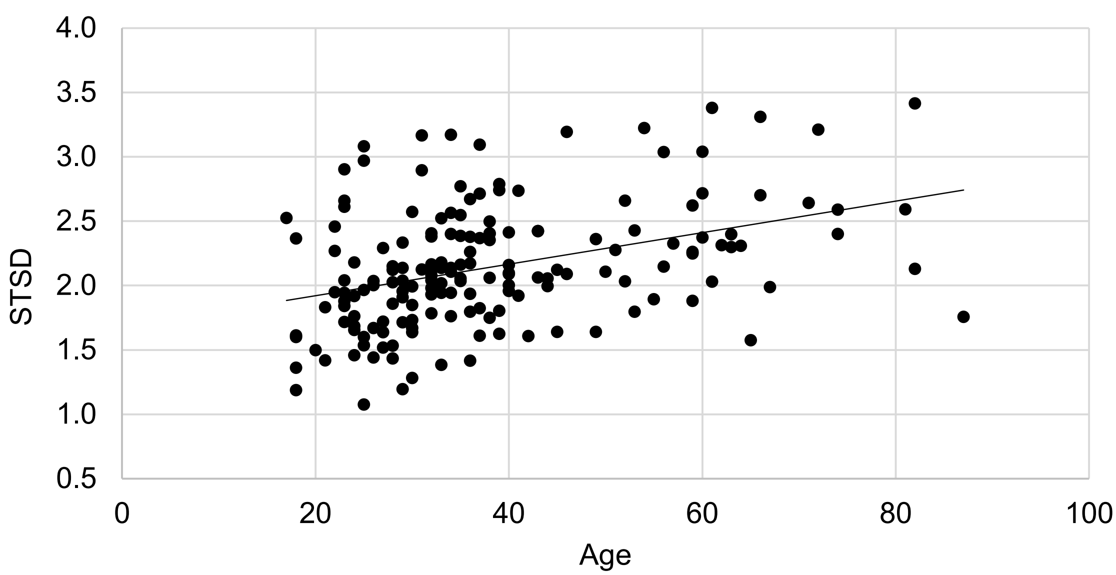


*Figure 7.* Semitone standard deviation with age for all speakers (line represents a linear regression analysis).

**Family.** Family relationship had no significant effect on speaker $F_0$ *M*. However, there was a significant difference across the families for STSD, $F(17, 147) = 2.10$, $p = 0.010$. Descriptive statistics for $F_0$ *M* and STSD are reported in Table 3.

**Speaking Time Ratio**

**Sex.** ANCOVA testing revealed a significant sex effect on speaking time ratio, $F(1, 17) = 5.63$, $p = 0.030$. As seen in Figure 8, females had a significantly higher ratio than males.

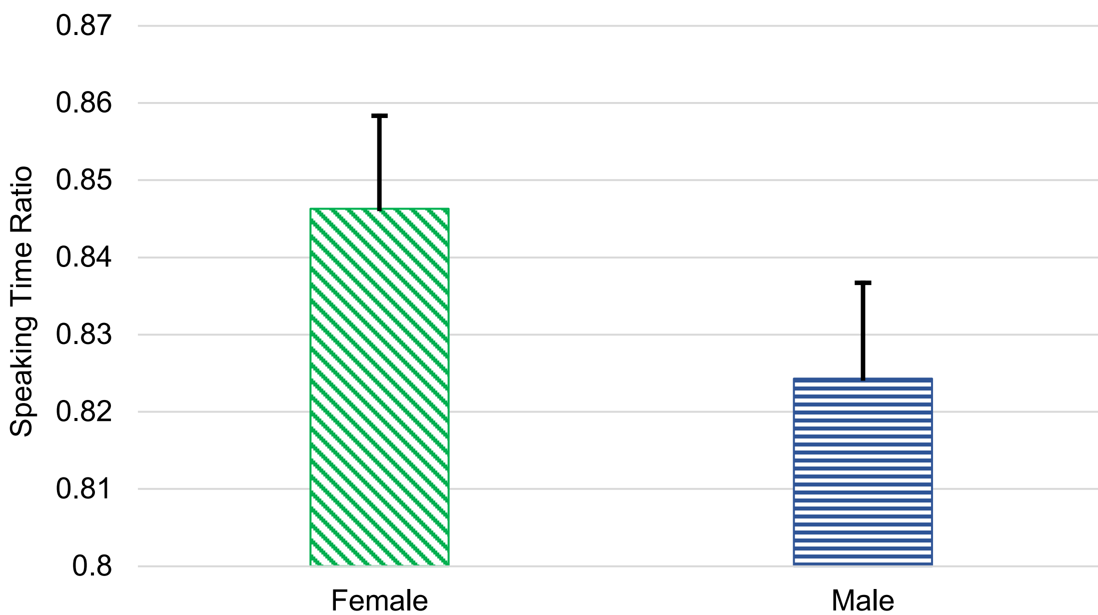


*Figure 8.* Mean (and 95% confidence interval) speaking time ratio for women and men.

**Age.** Age had no significant effect on speaking time ratio.

**Family.** Family relationship had no significant effect on speaking time ratio. Descriptive statistics for speaking time ratio are reported in Table 4.

**Cepstral Peak Prominence**

**Sex.** A significant difference was found in the CPPS between male and female speakers, $F(1, 17) = 6.37$, $p = 0.022$. As shown in Figure 9, female speakers had a higher CPPS measure than male speakers.

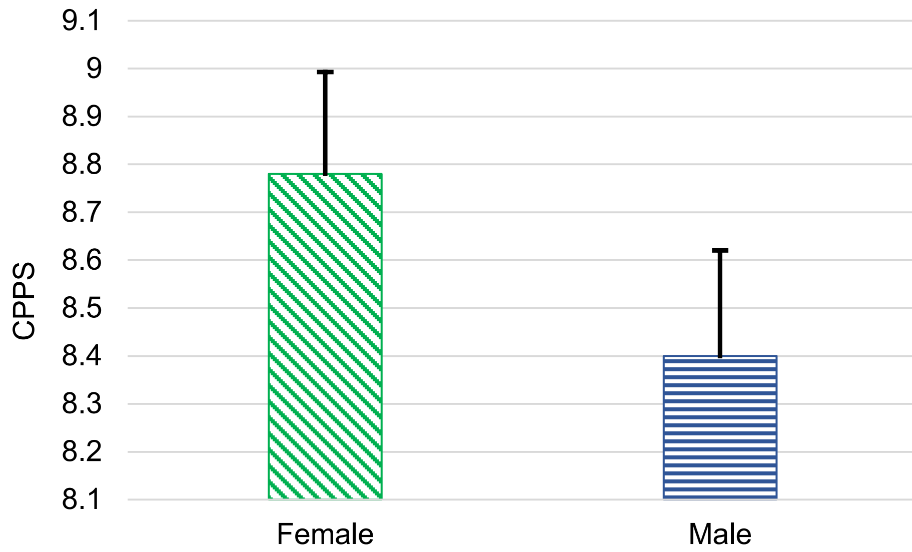


*Figure 9.* Mean (and 95% confidence interval) cepstral peak prominence smoothed for women and men.

**Age.** Speaker age had no significant effect on CPPS.

**Family.** ANCOVA testing revealed significant differences across the families for CPPS, $F(17, 147) = 3.51$, $p < .001$. Descriptive statistics for CPPS are reported in Table 5.

## Discussion

The present study was conducted to examine the potential differences in segmental and suprasegmental speech features with aging, across sexes, and between families. It was hypothesized that these measures would indicate a general decline in speech performance and vocal quality with age and reflect physiologic differences between sexes and, potentially, unrelated individuals. These trends were found on a number of measures with a few notable exceptions.

### Effects of Age

Research has well established that typical aging is accompanied by atrophy of some neural structures as well as decreased synaptic activity (Ramig et al., 2001; Slawinski, 1994).

These age-related neural changes are thought to affect the regulation of articulator movement amplitude, which may result in generally decreased accuracy in speech production (Ramig et al., 2001; Torre & Barlow, 2009). Tjaden and Turner (1997) reported that atypical fricative spectra were associated with reduced articulatory precision in individuals with neurodegenerative diseases. It was thus anticipated that, in the current study, typically aging speakers would demonstrate fricative spectra that differed from those of younger adults.

As hypothesized, we found that the spectral mean for /s/ decreased significantly with age, suggesting a decline in articulatory precision. An age-by-sex interaction revealed a greater decline in spectral mean in males than in females with age, supporting the hypothesis that aging affects the speech mechanisms of males and females differently. Although there was no age effect on spectral skewness for females, males' spectral skewness was found to increase significantly with age. While this finding may speak to age-related changes in production of /s/, the physiologic mechanisms that underlie changes in spectral skewness are unclear (Nissen & Fox, 2009) and more research on the subject is needed.

Much research has been devoted to evaluating the effects of aging on $F_0$. It has generally been found that females demonstrate a steady decrease in $F_0$ across the lifespan, while males demonstrate a decrease through adolescence into adulthood, followed by a gradual increase after middle age (Nishio & Niimi, 2008; Ramig et al., 2001). The current study's findings are consistent with this research. Testing revealed an age-by-sex interaction, where females' $F_0$ *M* decreased with age and males' increased. These trends may be attributed to changes in the thickness and elastin content of the vocal folds with age, as described by Hammond et al. (1998). The present study also found STSD to increase significantly with age, indicating greater variability in $F_0$ in older adults than in younger adults. This is consistent with previous findings that $F_0$ *SD* increases with age (Xue & Deliyski, 2001), and that elderly voices are characterized by instability (Gorham-Rowan & Laures-Gore, 2006). Additionally, Barnes (2013) has suggested that older speakers produce larger intonational ranges and more frequent inflections than younger adults, perhaps to demarcate syntactic and semantic units in connected speech in the presence of reduced articulatory precision.

As an additional measure of prosody in the present study, a speaking time ratio was computed for each participant, expressed as a proportion of time spent speaking relative to the total passage duration. Testing indicated that age had no significant effect on speaking time ratio. This is contrary to previous findings that older adults tend to use more frequent breaths, fewer syllables per breath, increased frequency and duration of pauses, and greater intraspeaker variability in temporal parameters (Ballard et al., 2001).

Previous studies have used CPPS as an index of vocal quality, particularly in connected speech tasks. In one such study, Garrett (2013) found older speakers to have significantly lower cepstral peak prominence values in connected speech, indicating less periodicity and more severe breathiness in the voice signal. This trend may be due to changes in laryngeal structure and function with age, which cause increased instability in elderly voices (Gorham-Rowan & Laures-Gore, 2006). It was thus hypothesized that CPPS would decrease with age, indicating a decline in vocal quality. However, testing of the present dataset revealed no significant effect of speaker age on CPPS.

**Effects of Sex**

The current study showed that females had higher spectral *M* and kurtosis values for /s/ than males. It was hypothesized that spectral moment measures would differ between men and women due to sex differences in the size, shape, and proportions of the vocal tract (Nissen &

Fox, 2009; Simpson, 2002; Titze, 1989). The physiologic mechanisms that underlie differences in spectral moments are unclear. However, Tjaden & Turner (1997) reported that the anterior, longer, deeper groove required for /s/ production is associated with higher spectral *M* values relative to other phonemes with shorter, wider, and more posterior constriction. It may then be the case that females in the present study produced /s/ with a more anterior and/or deeper grove than males. It may also be that females had a smaller anterior resonating cavity during /s/ production. Further investigation into the relationship between fricative spectral moment measures and physiologic speech patterns is needed.

There is general consensus in the literature that the average speaking $F_0$ for males lies between 100 and 146 Hz, while the average for females is between 188 and 221 Hz (Gelfer & Mikos, 2005). The results of the present study are consistent with earlier reports. The average $F_0$ for males was 111.15 Hz, and the average for females was 198.16 Hz. These measures reflect sex differences in laryngeal structure. Male laryngeal cartilage is 20% larger than that of females in the anterior-posterior dimension. Also, male vocal folds are 60% longer and 20% to 30% thicker than female vocal folds, with a greater percentage of collagenous fibers (Titze, 1989). It was hypothesized that STSD would also differ between sexes. However, no significant effect of sex on STSD was found. This indicates that males and females spoke with similar $F_0$ variability.

Speaking time ratio was used in the present study to investigate differences in suprasegmental features between sexes. Females had significantly higher speaking time ratio measures than males, meaning that females spent more time speaking and less time pausing than males. Although we did not anticipate this result, it was consistent with the findings of a recent research project which found that in certain divided attention conditions, females demonstrated higher speaking time ratios than males (Glenn, 2017). It may be the case that male and female participants in the present study responded differently to the demands of completing a read speech task in a highly controlled environment. It may also be possible that the sex effect found in the present study is due to learned or behavioral factors. Further investigation into sex differences in pausing structure, perhaps in a more naturalistic context, is needed.

CPPS was used in the current study as a measure of vocal quality. It was not anticipated that there would be a significant difference in this measure between sexes. However, females were found to have higher CPPS measures than males, indicating a greater degree of periodicity in the voice signal of female speakers during a connected speech task. This finding may reflect sex differences in the dimensions of the larynx and the makeup of vocal fold tissue structure as described by Titze (1989). More research is needed to investigate what physiologic differences, if any, contribute to changes in CPPS.

**Effects of Familial Relationship**

Several studies have found that untrained listeners are able to reliably differentiate between related and unrelated speakers, indicating that acoustic trends in familial speech production exist (Feiser & Kleber, 2012; Vanderydt, 1998; Westrop, 2000). However, very little research exists which adequately describes the speech features affected by familial relationship and what mechanisms are responsible for such trends. Vanderydt found that related speakers who were perceived as sounding similar had similar mean $F_0$ and intonation patterns. It was thus hypothesized that, in the current study, familial relationship would have an effect on mean $F_0$, STSD, and possibly speaking time ratio.

As predicted, significant variability among the families was found for STSD, suggesting that related speakers demonstrate similarities in $F_0$ variability. However, average $F_0$ and speaking time ratio were not significantly affected by familial relationship. One possible explanation for this finding is that, in Vanderydt's study (1998), similarities in $F_0$ and intonation patterns were only found in related speakers who were judged as sounding similar. In contrast, the present study did not involve perceptual measurement of any kind.

Testing revealed some additional, unanticipated results. There was found to be significant differences across the families for spectral skewness for /s/. Although the physiologic mechanisms responsible for changes in spectral skewness are unclear, this finding may suggest similarities in the articulation of /s/ for related speakers. Testing also revealed a significant effect of familial relationship on CPPS. This indicates that some families demonstrated a higher degree of periodicity in the voice signal than others. These findings could potentially be attributed to genetically-based, anatomical similarities in the speech mechanism within families, learned articulation and voicing behaviors, or a combination of these factors. Further investigation into the speech and voice patterns affected by familial relationship is needed.

**Limitations of the Present Study and Directions for Future Research**

Although spectral moment analysis has been used as an index of articulatory precision in previous studies (McFarland et al., 1996), research has yet to establish a clear relationship between fricative spectral measures and physiologic speech patterns. It is therefore difficult to draw conclusions about trends in articulatory behavior from the present study's findings that some fricative spectral moments vary with aging, between sexes, and across families. Future research could address this gap in the literature by pairing spectral moment analysis with kinematic measures to investigate possible links between spectral measures and changes in physiologic speech patterns.

An additional constraint of the present study is that all of the speech samples used were recordings of reading passages. While read speech still provides valuable information about speech patterns, it may not be entirely representative of natural, everyday speech, particularly for suprasegmental measures such as STSD and speaking time ratio, which were used in the current study. Future projects could collect these and similar measures during a spontaneous speech task to explore prosody patterns in a more naturalistic context.

Lastly, future studies could include a perceptual component to investigate what links, if any, exist between the acoustic trends we identified and perceived differences in speech and voicing behaviors. This could provide more information about the salience of differences in acoustic measures to listeners. The inclusion of a perceptual component may be especially valuable when investigating possible effects of familial relationship on acoustic metrics, since previous research suggests that these measures correlate more in related speakers who are judged as sounding similar by unfamiliar listeners (Vanderydt, 1998).

**Conclusion**

The availability of a large dataset with such a wide age range allowed for an in-depth investigation of acoustic changes across the lifespan. These data not only confirmed previous findings that certain speech and voicing features indicate a decline with age, they also provide insight into how aging affects the speech of men and women differently. Additionally, few studies have considered these acoustic changes in the context of related individuals across multiple generations. Findings from the current study suggest that related speakers may demonstrate similar patterns for prosody, voicing, and articulation behavior, although the statistical testing did not allow us to draw specific inferences about such similarities.

These findings have valuable clinical implications for speech and language treatment. The data from this study describe some normal variations in the speech production of persons of differing age, sex, and familial background. This is particularly important for the treatment of elderly individuals, for whom it is normal to demonstrate some decline in speech and voicing performance compared to younger adults. An understanding of these normal speech differences in healthy individuals is vital for differentiating between typical and pathological speech patterns in a clinical setting.

# References


Awan, S. N., Roy, N., & Dromey, C. (2009). Estimating dysphonia severity in continuous speech: Application of a multi-parameter spectral/cepstral model. *Clinical Linguistics & Phonetics, 23*, 825-841. doi: 10.3109/02699200903242988

Ballard, K. J., Robin, D. A., Woodworth, G., & Zimba, L. D. (2001). Age-related changes in motor control during articulator visuomotor tracking. *Journal of Speech, Language, and Hearing Research: JSLHR, 44*, 763-777. doi: 10.1044/1092-4388(2001/060)

Barnes, D. R. (2013). *Age-related changes in the production of linguistic prosody* (Master's thesis). Purdue University, West Lafayette, Indiana. Retrieved from https://docs.lib.purdue.edu/open_access_theses/17/

Boersma, P., & Weenink, D. (2016). Praat: Doing phonetics by computer [Computer program]. Version 5.4, retrieved 30 March 2017 from http://www.praat.org/

Eadie, T. L., & Doyle, P. C. (2005). Classification of dysphonic voice: Acoustic and auditory-perceptual measures. *Journal of Voice, 19*, 1-14.

Feiser, H. S. (2009). Acoustic similarities and differences in the voices of same-sex siblings. In Proc. 18th Annual Conference of the International Association for Forensic Phonetics and Acoustics (IAFPA) 2009, Cambridge, UK.

Feiser, H. S., & Kleber, F. (2012). Voice similarity among brothers: Evidence from a perception experiment. *Proceedings of the 21st Annual Conference of the International Association for Forensic Phonetics and Acoustics (IAFPA) 2012*. Santander, Spain.

Ferrand, C. T. (2002). Harmonics-to-noise ratio: An index of vocal aging. *Journal of Voice, 16*, 480-487. doi: 10.1016/S0892-1997(02)00123-6

Garrett, R. (2013). *Cepstral- and spectral-based acoustic measures of normal voices* (Master's thesis). University of Wisconsin, Milwaukee, Wisconsin. Retrieved from https://dc.uwm.edu/etd/217/

Gelfer, M. P., & Mikos, V. A. (2005). The relative contributions of speaking fundamental frequency and formant frequencies to gender identification based on isolated vowels. *Journal of Voice, 19*, 544-554. doi: 10.1016/j.jvoice.2004.10.006

Glenn, K. (2017). *Effects of conversational modalities on driving and speaking performance* (Master's thesis). Brigham Young University, Provo, Utah. Retrieved from https://scholarsarchive.byu.edu/etd/6294/

Gorham-Rowan, M., & Laures-Gore, J. (2006). Acoustic-perceptual correlates of voice quality in elderly men and women. *Journal of Communication Disorders, 39*, 171-184. doi: 10.1016/j.jcomdis.2005.11.005

Hammond, T. H., Gray, S. D., Butler, J., Zhou, R., & Hammond, E. (1998). Age- and gender-related elastin distribution changes in human vocal folds. *Otolaryngology and Head and Neck Surgery, 119*, 314-322. doi: 10.1016/S0194-5998(98)70071-3

Huber, J. E., Darling, M., Francis, E. J., & Zhang, D. (2012). Impact of typical aging and Parkinson's disease on the relationship among breath pausing, syntax, and punctuation. *American Journal of Speech-Language Pathology, 21*, 368–379. doi.org/10.1044/1058-0360(2012/11-0059)

Kent, R. D. (1996). Hearing and believing: Some limits to the auditory-perceptual assessment of speech and voice disorders. *American Journal of Speech-Language Pathology, 5*, 7-23. doi: 10.1044/1058-0360.0503.07

Lametti, D. R., Nasir, S. M., & Ostry, D. J. (2012). Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *Journal of Neuroscience, 32*, 9351-9358. doi: 10.1523/JNEUROSCI.0404-12.2012

McFarland, D. H., Baum, S. R., & Chabot, C. (1996). Speech compensation to structural modifications of the oral cavity. *Journal of the Acoustical Society of America, 100*, 1093-1104. doi: 10.1121/1.416286

Nissen, S. L., & Fox, R. A. (2009). Acoustic and spectral patterns in young children's stop consonant productions. *Acoustical Society of America, 126*, 1369-1378. doi:10.1121/1.3192350

Nishio, M., & Niimi, S. (2008). Changes in speaking fundamental frequency characteristics with aging. *Folia Phoniatrica et Logopaedica, 60*, 120-127. doi: 10.1159/000118510

Ramig, L. O., Gray, S., Baker, K., Corbin-Lewis, K., Buder, E., Luschei, E., . . . Smith, M. (2001). The aging voice: A review, treatment data and familial and genetic perspectives. *Folia Phoniatrica et Logopaedica, 53*, 252-265. doi: 10.1159/000052680

Simpson, A. P. (2002). Gender-specific articulatory-acoustic relations in vowel sequences. *Journal of Phonetics, 30*, 417-435. doi: 10.1006/jpho.2002.0171

Slawinski, E. (1994). Acoustic correlates of [b] and [w] produced by normal young to elderly adults. *Journal of the Acoustical Society of America, 95*, 2221-2230. doi: 10.1121/1.408682

Titze, I. R. (1989). Physiologic and acoustic differences between male and female voices. *Journal of the Acoustical Society of America, 85*, 1699-1707. doi: 10.1121/1.397959

Tjaden, K., & Turner, G. S. (1997). Spectral properties of fricatives in amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research, 40*, 1358-1372. doi: 10.1044/jslhr.4006.1358

Torre, P., & Barlow, J. A. (2009). Age-related changes in acoustic characteristics of adult speech. *Journal of Communication Disorders, 42*, 324-333. doi: 10.1016/j.jcomdis.2009.03.001

U.S. Department of Health and Human Services, Administration for Community Living, Administration on Aging. (2017). *Projected future growth of older population*. Retrieved from https://www.acl.gov/node/579.

Vanderydt, R. E. (1998). *A study of voice similarities among family members* (Master's thesis). Massachusetts General Institute of Health Professions, Boston, MA. Retrieved from https://search-proquest-com.erl.lib.byu.edu/comdisdome/docview/85147153

Westrop, R. M. (2000). *Segment durational changes from young adults to elderly: A familial study* (Master's thesis). Utah State University, Logan, Utah. Retrieved from https://search-proquest-com.erl.lib.byu.edu/comdisdome/docview/85146865/

Xue, S. A., & Deliyski, D. (2001). Effects of aging on selected acoustic voice parameters: Preliminary normative data and educational implications. *Educational Gerontology, 27*, 159-168. doi:10.1080/03601270151075561

APPENDIX A

## Annotated Bibliography

Awan, S. N., Roy, N., & Dromey, C. (2009). Estimating dysphonia severity in continuous speech: Application of a multi-parameter spectral/cepstral model. *Clinical Linguistics & Phonetics, 23*, 825-841. doi: 10.3109/02699200903242988

*Purpose:* The purpose of the study was to identify methods of cepstral/spectral-based analysis that accurately predict severity of dysphonia in connected speech, as measured by auditory-perceptual analysis. *Method:* Participants in this study included 104 women with primary muscle tension dysphonia (MTD). All participants were diagnosed with primary MTD after an otolaryngologist and speech-language pathologist completed a full medical and voice history as well as a laryngeal examination. Pre- and post-therapy voice recordings of each participant were collected during routine clinical practice. Treatment consisted of manual circumlaryngeal maneuvering completed in a single extended session. Each participant was instructed to read 'The Rainbow Passage' at a comfortable pitch and loudness. All recordings were trimmed to include only the 2nd and 3rd sentences of the reading passage and were analyzed using a Windows-based computer program for implementation of spectral and cepstral analysis methods. Voice samples were also perceptually judged by five master's degree students for voice quality pre- and post-treatment. Recordings were presented auditorily via headphones and each judge was asked to rate the severity of voice quality using a 100-point visual analogue scale (VAS) with one end of the scale labelled as 'normal' and the other as 'profoundly abnormal'. *Results:* The researchers computed a predicted speech dysphonia severity value for each participant's pre-treatment voice sample using the following cepstral/spectral-based measures: cepstral peak prominence (CPP), direct Fourier transform ratio (DFTR), and DFTR SD. Of these measures, CPP was found to be the strongest individual correlate of perceived dysphonia severity. The researchers computed a mean predicted dysphonia severity rating for all pre- and post-treatment samples (41.77) versus the mean perceived dysphonia severity rating (41.78). Statistical analysis revealed no significant difference between mean predicted and mean perceived dysphonia severity ratings. However, a significant difference was found between predicted and perceived pre-treatment ratings, as well as between predicted and perceived post-treatment ratings. Results did indicate significant decreases in post-treatment dysphonia severity according to both predictive and perceived measures. *Conclusion:* Based on these results, the researchers concluded that strong predictions of dysphonia severity can be achieved from continuous speech samples, although there are a number of reasons why acoustic predictions and auditory-perceptual ratings of dysphonia may not always coincide. The automatic voice analysis methods described in this study can provide objective and accurate measurement of dysphonia severity. *Relevance to the current work:* This study found cepstral/spectral methods of analysis to be accurate predictors of vocal quality.

Ballard, K. J., Robin, D. A., Woodworth, G., & Zimba, L. D. (2001). Age-related changes in motor control during articulator visuomotor tracking. *Journal of Speech, Language, and Hearing Research: JSLHR, 44*, 763-777. doi: 10.1044/1092-4388(2001/060)

*Purpose:* This study was conducted to examine changes in visuomotor control of oral-facial structures during nonspeech tasks across the lifespan. *Method:* Participants in this study included 52 females and 35 males ranging in age from 8 to 84 years. All participants reported a negative history of speech, language, or neurological disorder and/or substance abuse, as well as visual and auditory ability within normal limits (corrected or uncorrected). Participants were asked to complete visuomotor tracking tasks using a target signal displayed on an oscilloscope screen, which was placed 1 meter from a dental chair where participants were seated with their heads immobilized. Visuomotor tracking performance was measured for the lower lip and jaw using a strain gauge. To track the larynx ($F_0$), the researchers transduced participants' sustained phonation of /a/. The transduced signals from the lower lip, jaw, and $F_0$ were represented with a dot on the oscilloscope screen, centered horizontally on the bar of the target signal. Participants were asked to keep the dot on the bar throughout the bar's vertical movement for all tracking conditions (sinusoids of 0.3, 0.6, and 0.9 Hz, and an unpredictable complex signal). All participants were given 40 seconds of practice in each condition. Both the target signal and articulatory movements were compared with measures of cross correlation, gain ratio, phase shift, and average target-tracker amplitude difference (TTD) for each articulator. *Results:* Results of ANOVA indicated that target and tracker correlation improved from 8 years, leveled around 15-20 years, and began to decline at about 40-45 years. Results indicated significant main effects of target frequency and age group for all articulators. Similar trends were found in analysis of gain ratio. There was a significant main effect of target frequency for lip and significant main effect of age group for all articulators. Although younger adults demonstrated the greatest gains, all age-group comparisons were significant. There was a significant main effect of target frequency for the lip and $F_0$, of age group for $F_0$ only, and of gender for the jaw and $F_0$. Finally, results of ANOVA indicated significant main effects of target frequency and age group and a significant interaction effect of target frequency by age group for all articulators. This interaction effect was a result of older adults performing more poorly than children when tracking the predictable signals, but generally better in the unpredictable condition. *Conclusion:* Results of this study indicate that increasing predictable target frequency decreased tracking accuracy. Generally, tracking performance was lowest in the unpredictable condition. These results are consistent with previous research. This study provides a normative database of motor control of the articulators during nonspeech tasks across the lifespan. *Relevance to the current work:* This study evaluated changes to neuromotor control of the articulators as a function of age.

Barnes, Daniel R. (2013). Age-related changes in the production of linguistic prosody (Unpublished master's thesis). Purdue University, West Lafayette, Indiana. Retrieved from https://docs.lib.purdue.edu/open_access_theses/17/

*Purpose:* This study was conducted to investigate the effects of age-related changes to physiology and cognition on the production of two types of linguistic prosody: lexical stress and the disambiguation of syntactically ambiguous utterances. *Method:* Ten young adults (age 18-30) and ten older adults (age 65 or older) participated in this study. Each age group include 5 female and 5 male speakers, all of whom were native English speakers with no history of speech-language, voice, respiratory, or neurological problems. All participants completed The Gray Oral Reading Test (GORT), Cognitive-Linguistic Quick Test (CLQT), and Test of Adolescent and Adult Language (TOAL), to determine age-appropriate language, reading, and cognitive ability. In the lexical stress paradigm, participants were asked to produce noun-verb word pairs that were differentiated with the use of strong-weak or weak-strong stress patterns. In the ambiguous sentences paradigm, participants read sentences with ambiguous prepositional phrase attachments. Researchers measured acoustic correlates of stress, including mean intensity, mean F0, segment duration, and duration of pauses and words. *Results:* ANOVA completed on data from the lexical task indicated that there was a significant effect of stress pattern on intensity, as measured by sound pressure level (SPL). SPL was significantly higher for the strong-weak stress pattern compared with weak-strong. Age had no significant effect on SPL. There was also found to be a significant effect of stress pattern on F0. F0 was found to be significantly different across stress patterns for older adults, but not for younger adults. Next, the data revealed a significant effect of stress pattern and age on syllable duration. Young adults had more negative duration values compared with older adults. Older adults were found to produce target syllables with anticipated prosodic features with greater accuracy than young adults. In the ambiguous sentences paradigm, older adults were found to have significantly greater durations of direct objects and prepositional phrases than younger adults. *Conclusions:* All participants were found to use SPL and F0 to differentiate between strong-weak and weak-strong lexical stress patterns, although older adults used wider F0 ranges than younger adults. Also, older adults were generally found to have higher mean segment durations than younger adults, which may be due to cognitive slowing with age. *Relevance to the current work:* The researchers established significant differences in the use of prosody during reading tasks between younger adults and older adults.

Eadie, T. L., & Doyle, P. C. (2005). Classification of dysphonic voice: Acoustic and auditory-perceptual measures. *Journal of Voice, 19*, 1-14.

*Purpose:* The purpose of this study was to evaluate the correlation between acoustic and auditory-perceptual measures of overall voice quality and to assess the ability of such measures to discriminate between normal and dysphonic voices. *Method:* Participants in this study included 24 adult speakers (12 male, 12 female) ranging in age from 20 to 65 years. All speakers had diagnosed vocal pathology across a wide range of voice quality severities, including those without visible laryngeal pathology and those with benign mass lesions of the vocal folds. Additionally, normal voice samples from 3 males and 3 females were included, spanning the same age range as participants with voice pathology. The voice samples used in the study consisted of the second sentence of the Rainbow Passage. Twelve graduate students with no history of hearing, speech, voice, or language difficulties participated as listeners. They evaluated each voice sample for overall severity and pleasantness using direct magnitude estimation (DME) scaling procedures. Listeners were first familiarized with a modulus sample which was assigned an arbitrary value of 100 on the DME scale. They were instructed to rate all subsequent samples relative to the modulus. Listener ratings of pleasantness and severity were averaged for each speaker. In addition to the auditory-perceptual ratings, the researchers performed six acoustic measurements from the same voice sample: LTAS spectral tilt, voiced segments of spectral tilt, harmonics-to-noise ratio, linear prediction signal-to-noise ratio, pitch amplitude, and spectral flatness ratio. A stepwise linear regression analysis was then used to determine how well acoustic parameters predicted perceptual dimensions of overall severity and pleasantness. *Results:* Data analysis revealed that most acoustic dimensions intercorrelated with the auditor-perceptual ratings of overall severity and pleasantness. Harmonics-to-noise ratio did not significantly correlate with any dimension. The linear predicted dimensions (signal-to-noise ratio, pitch amplitude, and spectral flatness ratio) were found to intercorrelate significantly with auditory-perceptual ratings of vocal quality. Stepwise linear regression analysis revealed that, when all speakers were included, only pitch amplitude was a significant predictor of vocal severity. When normal speakers were excluded, results indicated that both pitch amplitude and harmonics-to-noise ratio were significant predictors, accounting for 48% of variance in overall severity for dysphonic speakers. Also, pitch amplitude was found to be the only significant predictor of vocal pleasantness for both dysphonic and normal speakers accounting for 40% of the variance. *Conclusion:* This study found auditory-perceptual ratings of overall voice pleasantness and severity to be highly correlated. Measures based on linear predictive analysis techniques were particularly intercorrelated. The acoustic metric with the most significant auditory-perceptual intercorrelation was pitch amplitude. The data indicate that ratings of both voice severity and pleasantness were accurate in the classification of dysphonic voice, with slightly more difficulty observed in ratings of normal voice. The researchers concluded that both acoustic and auditory-perceptual measures are valuable in rating voice quality. *Relevance to the current work:* This study found significant correlations between certain acoustic metrics and auditory-perceptual dimensions.

Feiser, H.S. (2009). Acoustic similarities and differences in the voices of same-sex siblings. In Proc. 18th Annual Conference of the International Association for Forensic Phonetics and Acoustics (IAFPA) 2009, Cambridge, UK.

*Purpose:* This study was designed to evaluate the voices of same-sex siblings who are not twins to determine which acoustic factors, if any, may contribute to similar sounding voices in related speakers. *Methods:* Participants included 10 male siblings and 10 female siblings, all of whom spoke German. All siblings fell in the age range between 19 and 27 years. Participants were recorded reading a passage from a text written in German. Preliminary auditory-perceptual analysis of telephone-filtered versions of the speech recordings indicated that the sibling pairs in this study were similar sounding. Acoustic-phonetic analysis consisted of measuring average fundamental frequency, the frequencies of formant F1-F4 in selected vowels, and Long-Term Average Spectra (LTAS). For $F_0$ and LTAS, degrees of similarity were calculated between sibling-internal comparisons and sibling-external comparisons. For formants F1-F4, a related method was used to contrast sibling-internal against external similarities. *Results:* Acoustic analysis revealed that similarities between the $F_0$ of unrelated speakers (sibling-external comparisons) were significantly stronger than similarities between siblings (sibling-internal comparisons). Sibling-external comparisons of formants F1-F4 also showed stronger similarities than sibling-internal comparisons, though less often than $F_0$. Lastly, researchers reported that sibling-internal comparisons of LTAS were more closely related than sibling-external comparisons. *Conclusion:* The results of this study show that acoustic analysis can reveal differences between siblings who are similar-sounding, indicating that acoustic measurement is more sensitive to sibling-internal speaker differences than auditory-perceptual analysis alone. This study also shows that there is some acoustic basis for perceived similarities in the voices of siblings. *Relevance to the current work:* This study explored acoustic features of related speakers rated as sounding similar.

Feiser, H.S. & Kleber, F. (2012). Voice similarity among brothers: Evidence from a perception experiment. In Proc. 21st Annual Conference of the International Association for Forensic Phonetics and Acoustics (IAFPA) 2012. Santander, ES.

*Purpose:* The purpose of this study was to evaluate whether unfamiliar listeners are able to perceptually identify brothers and thus distinguish them from non-related male speakers. *Methods:* Participants included five pairs of brothers (ten total speakers) between ages 22 and 29, with an age difference of one to three years between brothers. All speaking participants were speakers of Standard German. Each speaker read the German version of the North Wind and the Sun text. This study also included twenty listeners (nine female, eleven male) between the age of 22 and 45 who spoke Standard German. Recordings were presented to listeners via Praat in an ABX discrimination test. The A and B stimuli were productions by non-related speakers and X was always a production from the brother of either A or B. A total of 160 ABX triplets were presented with no repetitions. Sessions lasted approximately 30 minutes for each listener. Most of the listeners were phoneticians; however, all listeners were unfamiliar with the speakers and the task. *Results:* Researchers performed a one-sample proportions test with continuity correction. Results of the analysis indicated that pairs of brothers were correctly identified by their voice in 78.5% of opportunities, which is significantly above chance. Then, a two-sample test for equality of proportions with continuity corrections was performed. It revealed no significant difference between the accuracy of male and female listeners' identification of related speakers. Finally, researchers used a general linear mixed model with correct identification as the dependent variable, brother as an independent variable, and listener as a random factor. Analysis showed that brother pair M4 was significantly better identified than the other four pairs and that the number of false identifications was greatest for brother pairs M2 and M5. *Conclusion:* The results of this study clearly show that unfamiliar listeners are able to distinguish between related and non-related speakers in connected speech. The researchers suggest that listeners may rely, in part, on dialect features for identification, as the brothers who showed more pronounced dialect features (M4) were better identified than the other pairs. *Relevance to the current work:* This study suggests that there is some significant, perceptible similarity in the voices of related speakers.

Ferrand, C. T. (2002). Harmonics-to-noise ratio: An index of vocal aging. *Journal of Voice, 16*, 480-487. doi: 10.1016/S0892-1997(02)00123-6

*Purpose:* This study was designed to accomplish three goals: (1) compare harmonics-to-noise ratios (HNRs) in three age groups of adult women, (2) compare HNR and jitter measures in terms of sensitivity to differences in vocal function, and (3) to determine $F_0$ differences between the three age groups. *Methods:* Participants in this study included a total of 42 adult women with normal voices, with 14 speakers in each of the following groups: young adults (age 21-34), middle-aged adults (age 40-63), and older adults (age 70-90). All participants reported a negative history of neurologic disease, dementia, hearing impairment, or voice therapy. All elderly speakers passed a hearing screening. Participants were recorded vocalizing the vowel /a/ five times each, using a comfortable pitch and loudness level, with each vocalization lasting 3 seconds and a 2-3 second pause between each trial. Recordings were input directly from a tape recorder to a Kay Elemetrics Computerized Speech Lab. The voice analysis function of the CSL was used to obtain HNR, as well as jitter and average $F_0$. *Results:* Values for HNR, jitter, and $F_0$ were averaged over the five vowel prolongations for each participant. Three separate one-way ANOVAs were performed, one for each dependent variable, with age as the grouping variable. Age was found to have an effect on HNR, with significant differences between the young and elderly values, and between the middle-aged and elderly values. Age was also found to have a significant effect on $F_0$, again with significant differences between the young and elderly values, and between the middle-aged and elderly values. No significant differences were found between the young and middle-aged groups for any variables. Finally, no significant differences in jitter were found between any age groups. *Conclusion:* The results of this study suggest that HNR appears to be stable in the young and middle years, then decreases during elderly years. This is likely an indication of instability of vocal fold vibration in older adults. This study is consistent with previous research which suggests that jitter is an ineffective measure for distinguishing between vocal function in different groups. Also consistent with previous studies, these data indicate a significant lowering of $F_0$ in elderly female speakers, which has often been attributed to postmenopausal vocal fold edema. *Relevance to the current work:* This investigation yielded results that describe changes to vocal function with age in female speakers.

Garrett, R. (2013). Cepstral- and spectral-based acoustic measures of normal voices (Unpublished master's thesis). University of Wisconsin-Milwaukee, Milwaukee, Wisconsin. Retrieved from https://dc.uwm.edu/etd/217/

*Purpose:* The purpose of this study was to establish normative data for Long-Term Average spectral- and cepstral-based measurements for male and female speakers in two age groups. *Methods:* Participants in this study consisted of 15 males and 15 females, ages 20-30 years, and 15 males and 15 females, ages 40-50, totaling 60 participants. All were native speakers of American English with a negative history of voice disorders, neurological disorders, hearing impairment, or speech-language therapy. Participants were recorded in a sound booth with a constant 6-inch mouth-to-mic distance. Speakers were asked to sustain the vowels /a/ and /i/ for approximately 3 seconds at a 75 dB (+/- 2 dB) intensity level. Additionally, participants read out loud stimuli from the CAPE-V assessment as well as two sentences from the Rainbow Passage at the same intensity level to provide connected speech samples for voice analysis. Prior to statistical analysis, the most stable one-second portion of each sustained vowel sample was isolated. Data were obtained using the Analysis of Dysphonia in Speech and Voice (ADSV) subprogram of the Kay-Pentax Multi-Speech software. Dependent variables included Cepstral Peak Prominence (CPP), Low-to-High Spectral Ratio (L/H spectral ratio), and the fundamental frequency of the Cepstral Peak Prominence (CPP $F_0$). These acoustic measurements were performed in seven different contexts for each speaker: two sustained vowels, and five connected speech segments. Independent variables in this study included gender (2 levels: male and female) and ag (2 levels: 20-30 years and 40-50 years). *Results:* Statistical analysis revealed noticeably higher CPP and L/H spectral ratio values for both vowels /a/ and /i/ for males compared with females. For CPP $F_0$, values were in the expected frequency ranges for male and female speakers, with male $F_0$ approximately one octave below females. Generally, age did not appear to have a significant effect on the dependent variables. However, there seemed to be a drop in CPP $F_0$ with age for females, while for males, a slight increase in CPP $F_0$ was observed for the vowel /a/. Two significant interactions occurred between age and gender: one for CPP $F_0$ in productions of /a/, and one for CPP in the production of /i/. Results showed that older female speakers had a lower CPP $F_0$ for /a/ than the younger female group. Older males had a slightly higher CPP $F_0$ than younger males. In productions of the vowel /i/, CPP was found to decrease with age in male speakers and slightly increase with age in female speakers. Results of the MANOVA for the five different connected speech segments revealed a significant main effect for age in CPP for all segments, and CPP $F_0$ for segment 5, generally showing a decrease in voice quality with age. *Conclusions:* The results of this study showed significant differences by gender in all dependent variables for both vowels and connected speech segments. Age did not appear to have a significant effect on vocal quality for sustained vowels /a/ and /i/; however, age appeared to have a significant effect on CPP for all connected speech segments tested. *Relevance to the current work:* The results of this study indicate a significant decline in certain measures of vocal quality with age, which may be tied to perceived breathiness in elderly speakers.

Gelfer, M. P., & Mikos, V. A. (2005). The relative contributions of speaking fundamental frequency and formant frequencies to gender identification based on isolated vowels. *Journal of Voice, 19*, 544-554. doi: 10.1016/j.jvoice.2004.10.006

*Purpose:* This study was designed to evaluate the accuracy with which listeners could identify the gender of a speaker from a synthesized isolated vowel based on the natural production of that speaker under three conditions: (1) the $F_0$ was consistent with the speaker's gender, (2) the $F_0$ was inconsistent with the speaker's gender, and (3) the speaker was transgender. *Methods:* Participants in this study included 30 total individuals: 10 cisgender men, 10 cisgender women, and 10 male-to-female transgendered persons. In order to qualify for this study, transgender participants had to have begun social and/or physical transitioning, report that she had been at least "somewhat successful" in developing a feminine voice, and produce a sustained vowel with a $F_0$ of 165 Hz. Transgender participants ranged in age from 23 to 57 years. Cisgender participants were selected and matched to the transgender participants by age. All speakers were instructed to sustain the vowels /i/, /u/, and / ɝ/ for 5 seconds each. Using the Dr. Speech program, two synthesized vowel files were created for each sustained vowel, one with a $F_0$ of 120 Hz, and one at 240 Hz. These vowel files were presented to 30 normal-hearing young adult listeners between 18 and 35 years of age. They were seated in an IAC booth and instructed to identify each speaker as a man or a woman, and rate how confident they were on a five-point scale from 'guessing' to 'very confident.' *Results:* For all three groups, results indicated significantly higher rates of accurate gender identification when synthetic vowels were presented with both gender-appropriate formant and fundamental frequencies, compared to when formant frequencies and pitch were mismatched. ANOVA revealed that there was a significant interaction between fundamental frequency and formant frequencies. Statistical analysis also determined that there was no significant difference in the level of identification accuracy in samples of female compared with transgendered speakers when using a female-appropriate $F_0$. *Conclusion:* The results of this study indicate that $F_0$ cues were unequivocally more salient to listeners than formant frequency cues. The data suggest that this is also true for male-to-female transgender speakers, who were accurately identified more often than biological females when formant and $F_0$ cues were gender appropriate. *Relevance to the current work:* This study evaluated the differences between male and female voices and which cues contribute most to listener differentiation of voices by gender.

Glenn, K. (2017). Effects of Conversational Modalities on Driving and Speaking Performance (Master's thesis). Brigham Young University, Provo, Utah. Retrieved from https://scholarsarchive.byu.edu/etd/6294/

*Purpose:* This study was conducted to investigate the bidirectional effects of a simulated driving exercise and speaking behaviors across conditions using different speaking modalities. *Methods:* Participants included 60 individuals from the local community, 30 men and 30 women, divided evenly into three age groups: young adults (ages 20-30 years), middle-aged adults (ages 40-50 years), and older adults (ages 60-71 years). All participants self-reported normal or corrected-to-normal vision, a negative history for speech, language or hearing disorders, and a valid driver's license. Participants sat in a sound booth and were fitted with a headset microphone which recorded their speech to a computer. They first completed five practice trials on a driving simulator software using a steering wheel and a gas/brake pedal unit, given instructions to maintain a constant speed, remain in the center of the lane, and take the first exit. Then, participants completed each of the following seven tasks in random order: a) driving without speaking; b) conversing on a hand-held cell phone without driving; c) conversing on a hands-free cell phone without driving; d) conversing with a passenger in the "car" (sound booth) without driving; e) driving while having a conversation on a hand-held phone; f) driving while having a conversation on a hands-free phone; g) driving while having a conversation with a passenger seated in the sound booth. They were instructed to merge onto the freeway, drive in the center of the right lane at a constant speed, and take the first exit. Researchers used Praat to calculate selected speech measures, including speaking time ratio, *M* and *SD* of intensity, and *M* and *SD* of fundamental frequency in semitones. The driving measures collected included *SD* of lane position, *M* and *SD* of speed, *SD* of steering wheel position, and the average number of steering wheel turns. *Results:* Repeated measures ANOVA testing revealed significant effects of driving while speaking for mean intensity, speaking time ratio, *SD* of steering wheel position, and number of steering wheel turns. Significant gender effects were found for speaking time ratio, and *M* and *SD* of intensity. Females demonstrated higher speaking time ratios while men had higher values for *M* and *SD* of intensity. Lastly, the older adults were found to have significantly lower *M* fundamental frequency, and higher *SD* of lane position and steering wheel position, than younger adults. *Conclusions:* These findings indicate that conversation can negatively impact driving performance, but also that driving has an impact on speaking performance. This has important implications for speech-language pathologists providing therapy in a clinical setting. In order to more closely replicate everyday speaking conditions, clinicians should introduce divided attention tasks where appropriate. *Relevance to the current work:* this project evaluated acoustic properties of speech between genders and with age across multiple conditions and modalities.

Gorham-Rowan, M., & Laures-Gore, J. (2006). Acoustic-perceptual correlates of voice quality in elderly men and women. *Journal of Communication Disorders, 39*, 171-184. doi: 10.1016/j.jcomdis.2005.11.005

*Purpose:* The purpose of this study was to examine the relationship between scaled perception of breathiness and hoarseness and selected acoustic variables ($F_0$SD, amplitude perturbation quotient, harmonic-to-noise ratio, and H1-A1) in young and elderly individuals. *Methods:* Participants in this study included 112 healthy volunteers. They were split evenly into four groups: 28 young women (avg. age = 24.7), 28 young men (age = 25.4), 28 elderly women (age = 70.7), and 28 elderly men (age = 69.6). All participants were nonsmokers in good physical health with normal hearing and voice quality within normal limits as judged by an experienced voice therapist. Speakers were recorded producing the vowel /a/ three times for approximately 3-5 seconds at a comfortable pitch and effort level in a sound attenuated booth. The middle one second of phonation was isolated and used to obtain measures of noise-to-harmonic ratio (NHR), amplitude perturbation quotient (APQ), $F_0$SD, and H1-A1. Ten untrained listeners (eight women, two men) were recruited to participate in the study. Listeners were asked to rate each voice sample, first to reflect the perceived amount of breathiness, then to estimate hoarseness (larger numbers = very breathy or hoarse voice quality). *Results:* ANCOVA of the acoustic measurements indicated that the elderly speakers demonstrated greater vocal instability. This was evidenced by higher values for $F_0$SD, APQ, and NHR. A paired samples *t*-test was performed to evaluate differences in perceived breathiness or hoarseness between age groups. There was no significant difference in perceived breathiness between young and elderly women nor between young and elderly men. However, elderly women were rated as sounding significantly more hoarse than young women. No difference was observed in the perceived hoarseness of male speakers. Breathiness was moderately correlated with H1-A1 for young women and with $F_0$SD for elderly men. Hoarseness was moderately correlated with APQ for both young men and elderly women. Perceived hoarseness was also moderately correlated with NHR for elderly speakers of both genders. *Conclusion:* As a group, elderly speakers demonstrated significantly higher $F_0$SD, NHR, and APQ values than their younger counterparts. These data suggest increased noise and instability in the acoustic signal with age. Additionally, the results of this study indicate significant correlations between selected acoustic properties and auditory-perceptual ratings of voice quality. *Relevance to the current work:* This study indicates that there are changes to vocal output with age, and it establishes a connection between acoustic measurements and perceived breathiness and hoarseness.

Hammond, T. H., Gray, S. D., Butler, J., Zhou, R., & Hammond, E. (1998). Age- and gender-related elastin distribution changes in human vocal folds. *Otolaryngology and Head and Neck Surgery, 119*, 314-322. doi: 10.1016/S0194-5998(98)70071-3

*Purpose:* This study was designed to evaluate possible relationships between age and gender in elastin content and distribution in the vocal folds. *Methods:* Larynges were obtained from the state medical examiner within 24 hours of death of the donor. Cases with suspected laryngeal injury were excluded by the medical examiner. Larynges were obtained from 8 female and 11 male subjects. Donors were divided into the following age groups: infant (1 to 5 months), adult (30 to 39 years), and geriatric (65 to 82 years). The vocal folds from these donors were processed with alcohol dehydration and paraffin embedment to obtain midmembranous coronal sections of the folds. Sections were stained with Verhoeff's elastic issue stain (EVG) which stains elastin fibers black. Image analysis software was used to obtain a grayscale measurement of the amount of staining. Strips were cut from the epithelial surface to the vocal fold muscle in order to capture the layered distribution pattern of elastin throughout the lamina propria. *Results:* Researchers found statistically significant differences in the amount of elastin between all age groups, increasing chronologically with age. Infant larynges had almost no measurable elastin in the lamina propria, with no significant differences between male and female infants. Adult larynges were found to have substantial elastin throughout the lamina propria. Although female larynges were found to have slightly more, both male and female adults had less elastin in the superficial region, increasing toward deeper regions. Geriatric larynges had more elastin staining in the lamina propria than infant or adult subjects. Female geriatric larynges had a higher average elastin content and a smaller superficial layer than their male counterparts. *Conclusion:* The results of this study indicate a significant increase in the amount of elastin distribution in the lamina propria with age. Also, the data indicate changes in the distribution patterns of elastin between age groups. No significant differences were found in the elastin content between gender groups. The authors suggested that this may have been due to a small sample size. *Relevance to the current work:* The results of this investigation demonstrate physiologic changes to the vocal folds with aging, which may have an impact on vocal quality in aged populations.

Huber, J. E., Darling, M., Francis, E. J., & Zhang, D. (2012). Impact of typical aging and Parkinson's disease on the relationship among breath pausing, syntax, and punctuation. *American Journal of Speech-Language Pathology, 21*, 368–379. doi.org/10.1044/1058-0360(2012/11-0059)

*Purpose:* This study was conducted to evaluate the impact of typical aging and Parkinson's disease (PD) on the relationship between breath pausing, syntax, and punctuation. *Methods:* Participants included 30 young adults (15 men, 15 women), 25 older adults (10 men, 15 women), and 15 individuals with PD (9 men, 6 women), totaling 71 participants. All were native English speakers. The young adult speakers ranged in age from 20 to 35 years, the older adults were between 66-82 years, and those with PD were between 51 and 83 years old. All 15 individuals with PD were age- and gender-matched with healthy speakers. Each participant was instructed to read a short passage twice at a comfortable pitch and loudness. From these recordings, researchers measured the speakers' utterance length, location of breath pauses relative to punctuation and syntax, and number of disfluencies and mazes. A breath group was operationally defined as all of the words produced within one breath. Utterance length was defined as the number of syllables produced within one breath group. The location of a breath was signified by a sharp upward deflection in the sum signal from a Respitrace placed around the speakers' rib cage. The occurrence of a breath was corroborated by perceptual analysis of the microphone signal as needed. Breaths taken at periods or commas were counted as breaths at punctuation. Breaths were also analyzed according to syntax using the following categories: major syntactic boundary, minor syntactic boundary, and locations unrelated to a syntactic boundary. *Results:* For typically aging adults, older adults were found to produce shorter utterances than young adults. Age also had a significant effect on breaths at punctuation. Older adults took a smaller percent of breaths at periods and a larger percent of breaths at commas than young adults. Next, older adults took a significantly smaller percent of breaths at major boundaries and a larger percent of breaths at minor boundaries than young adults. There was no significant effect of age on the number of mazes or disfluencies for typically aging adults. For individuals with PD, utterance length was not found to be significantly different than that of age-matched speakers. There were also no significant group effects on breaths at punctuation. Results did indicate a significant effect of group on percent of breaths at locations unrelated to syntax, with individuals with PD taking a larger percentage of breaths at locations unrelated to syntax. Finally, speakers with PD were found to produce significantly more mazes than control subjects. Disfluencies did not differ significantly between groups. *Conclusion:* This study provided a normative data set for typically aging adults, indicating a relationship between syntax and breath pausing. Typically aging adults appear to breathe in syntactically-appropriate places, but they also often breathe at minor boundaries. This does not appear to affect intelligibility. The researchers suggest that changes in breath pausing patterns with age are likely due to changes in respiratory physiology. They also suggest that, in individuals with PD, these changes are likely the result of a combination of changes to respiratory physiology and cognition. *Relevance to the current work:* This study demonstrates the changes to pausing patterns that occur with typical aging.

Kent, R. D. (1996). Hearing and believing: Some limits to the auditory-perceptual assessment of speech and voice disorders. *American Journal of Speech-Language Pathology, 5*, 7-23. doi: 10.1044/1058-0360.0503.07

*Purpose:* the purpose of this article was to highlight some of the limitations of auditory-perceptual assessment of voice disorders, stuttering, dysarthria, aphasia, and apraxia of speech. *Content:* Auditory-perceptual judgment of speech and voice output plays an important role in clinical decision making; it is often the primary – and sometimes only – method used to evaluate the efficacy of an intervention program. While such measurements are convenient, economic, and robust, they are also susceptible to a variety of error and biases from a number of sources. For instance, listeners have been found to hear speech sounds that are not really present, as demonstrated in auditory illusion studies. Phonemic false evaluation – or the mistaken recognition of phonemes that were not produced – has been observed in several studies. It is believed that this phenomenon takes place in natural communication as a result of listener normalization that may override detection of subtle errors in the speech signal. Auditory-perceptual processing has also been found to be influenced by the lexical status of sound stimuli as well as simultaneous visual input (i.e. the McGurk Effect). Many research studies have used auditory-perceptual judgment to describe disordered voices, but Kent suggests that these studies lack convergence on the basic descriptors used. When judging stuttering, listeners have demonstrated very low agreement for consistently differentiating between stuttering instances and normal disfluencies. Auditory-perceptual judgments of dysarthria have been found to vary a great deal with the predictability of its syntactic and semantic content as well as the judges' familiarity with the dysarthric patterns. Even phonetically trained listeners demonstrate a lack of agreement in the transcription of errors in aphasia and apraxia of speech. Auditory-perceptual judgments are subject to biases based on speaker non-speech characteristics, auditory salience of the speech characteristics in question, listener characteristics, and even by the mode of stimulus presentation. In order to improve the reliability of auditory-perceptual assessment, Kent suggests providing listeners with reference samples, standardizing procedures of measurement, and controlling the qualifications and characteristics of judges. *Relevance to the current work:* This paper speaks to the need for more objective, instrumental measurement in the assessment of speech production to supplement auditory-perceptual ratings.

Lametti, D. R., Nasir, S. M., & Ostry, D. J. (2012). Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *Journal of Neuroscience, 32*, 9351-9358. doi: 10.1523/JNEUROSCI.0404-12.2012

*Purpose:* This study was designed to investigate the relative roles of somatosensory and auditory feedback during speech production of healthy adults. *Methods:* Participants in this study included 75 native English speakers (23 males, 52 females). Custom acrylic and metal dental appliances were made for each participant to fit his or her upper and lower teeth. The lower appliance was attached to a small robotic device which tracked jaw movement and could also apply force. Participants were seated wearing headphones with their heads stabilized facing a computer monitor during testing. They were asked to read the word displayed on the monitor (either "had" or "head") repeatedly until the word was removed. On average, each participant repeated the word 11 times. Somatosensory feedback perturbation was introduced by altering the movement path of the lower jaw. The small robot attachment applied force that pulled the jaw outward, perpendicular to the path of its normal movement. In order to perturb auditory feedback, researchers altered the first formant (F1) in real time. An acoustical effects processor was used to shift F1 downward and introduce 70 dB speech-shaped masking noise. The altered signal was played back to the participants through the headphones. Participants were divided into 5 groups: (1) auditory perturbation only, (2) somatosensory perturbation only, (3) both auditory and somatosensory perturbation, (4) first auditory only, then both, and (5) first somatosensory only, then both. The researchers evaluated increases in F1 frequency as a measure of compensation for auditory perturbation, and a decrease in robot-induced movement deviation to measure somatosensory compensation. *Results:* The results of this study indicated, first, that the effects of each perturbation were independent of each other (i.e. somatosensory perturbation did not affect formant frequencies and auditory perturbation did not alter jaw movement). The data also demonstrate that the participants who simultaneously experienced both somatosensory and auditory perturbation showed a preference for sensory feedback. They were found to compensate for auditory perturbation significantly less or not at all. Lastly, the results suggested a negative correlation between the amount of compensation for each perturbation. In conditions 3, 4, and 5, participants who adapted more to somatosensory perturbation adapted less to auditory perturbation and vice versa. *Conclusion:* By introducing auditory and somatosensory perturbations alone and then in different combinations, the authors of this study found individuals to show preferential reliance for either acoustic or somatosensory feedback during speech production. This resulted in a negative correlation between adaptation to one perturbation and adaptation to the other. *Relevance to the current work:* The results of this study highlight the important role that sensory feedback plays in speech production, which carries implications for the use of invasive kinematic measures.

McFarland, D. H., Baum, S. R., & Chabot, C. (1996). Speech compensation to structural modifications of the oral cavity. *The Journal of the Acoustical Society of America, 100*, 1093-1104. doi: 10.1121/1.416286

*Purpose:* The purpose of this study was to investigate the effects of oral-articulatory perturbation on the production of vowels, stop consonants, and fricatives, measured both acoustically and auditory-perceptually. *Methods:* Participants in this study included 15 adult females between ages 20 and 23 years old. All were native speakers of (Quebec) French with normal hearing and a negative history of speech and/or language disorders. For each participant, two acrylic alveolar-palatal acrylic appliances were custom made. One appliance was 6mm thick at the midline of the cuspid-to-cuspid plane, while the other was 3mm thick. Stimuli included the vowels /i/, /a/, and /u/ in isolation, prevocalic voiceless stop consonants /p/, /t/, and /k/, and prevocalic voiceless fricatives /s/ and / ʃ/. Participants were presented with each stimulus in random order and instructed to produce five repetitions of each. These stimuli were elicited immediately after introduction of the perturbation and then again after a conversation, to allow for adaptation. In each of these conditions, stimuli were elicited with first with no palate, then with the thin palate, and then with the thick palate. The following acoustic measures were obtained: vowel duration, stop consonant duration, fricative duration, vowel formants F1 and F2, and spectral moments one (centroid), three (skewness), and four (kurtosis) for stop consonants and fricatives. Additionally, ten adult native speakers of (Quebec) French were recruited to participate. These listeners were presented with three random productions from each speaker in each condition. Listeners were instructed to first identify the sound presented from a limited field of options, and then to rate the quality of the sound on a five-point scale from *intelligible* to *perfect*. *Results:* In the immediate condition, duration of /t/ was significantly longer with the thick palate than with the thin palate or no palate. Duration of /k/ was found to be significantly longer in the thick and thin palate conditions than in the no palate condition. Palate condition did not have a significant effect on vowel formants in the immediate condition. Palate condition did have a significant effect on mean centroid frequencies, skewness, and kurtosis for phonemes /t/, /k/, and /s/. Significant differences between thick and thin palate sizes were only observed for /t/ and /k/ centroid frequency. No significant differences were found for /p/ and / ʃ/ spectral measures in the immediate condition. In the post-conversation condition, the artificial palate had no significant effect on segment durations. Results indicated very little difference in post-conversation average F1 or F2 across palate conditions. Finally, artificial palate had a significant effect on post-conversation spectral measures for /k/ and /s/. No other consonantal spectral measures were affected by the artificial palate in the post-conversation condition. ANOVA of data from the perceptual study revealed that quality ratings were significantly higher for /i/ with no palate than with the thick or thin palate in the immediate condition. Identification accuracy tended to be lower for stop consonants with artificial palates than with no palate. Although identification accuracy was quite high across vowels and palate conditions, ANOVA indicated a significant vowel by condition interaction for vowel quality ratings. Statistical analysis also revealed significant main effects of condition on percent correct identification and quality ratings for fricatives. *Conclusion:* The results of this study suggest that different speech sounds may be differentially impaired by the presence of perturbation. This study aligned with previous research in its finding of adaptation to oral-articulatory perturbation over time. Both statistical and perceptual data suggest that fricative spectra are particularly susceptible to alterations under conditions of perturbation. *Relevance to the current work:* This study demonstrates the role of sensory feedback in speech production and it suggests that fricative spectra may be a sensitive measure of speech performance. Also, it demonstrates an alignment of perceptual and acoustic data.

Nissen, S. L., & Fox, R. A. (2009). Acoustic and spectral patterns in young children's stop consonant productions. *Acoustical Society of America, 126*, 1369-1378. doi:10.1121/1.3192350

*Purpose:* This study was conducted to investigate acoustic and spectral characteristics of stop consonants in the speech of prepubescent children. *Methods:* Participants in this study included 10 adults between 18 and 40 years old and 90 children evenly divided into three age groups: 3-year-olds, 4-year-olds, and 5-year-olds. Each group contained an equal number of male and female participants, all with no diagnosed history of speech, language, or hearing problems. Researchers used a custom MATLAB software to present age-appropriate pictures representing the target words in a random order. All target words consisted of a voiceless obstruent (i.e. /p/,

/t/, or /k/) in word-initial position followed by a monophthong vowel (i.e. /i/, /ɑ/, or /u/). Participants produced each target word three times embedded in a carrier phrase. Productions were recorded with a head-mounted microphone. Adobe Audition was used to obtain values for normalized amplitude and spectral moment measures (slope, mean, variance, skewness, and kurtosis). Repeated measures ANOVAs were used to examine potential acoustic differences in articulation of stop consonants as a function of place of articulation, vowel context, speaker sex, and age group. *Results:* Statistical analysis revealed that normalized amplitude and spectral slope were significantly different across all three places of articulation and all three vowel contexts. Also, a main effect was found for both the sex and age group of the speaker, with sex-specific differences beginning with 5-year-olds and extending to the adults. Spectral mean was significantly affected by place of articulation and vowel context. *Post hoc* testing revealed significant differences between all three places of articulation, while only the /i/ vowel context was significantly different than /a/ and /u/. The /i/ vowel context effect was increased in /k/ and relatively reduced in /p/ and /b/. Spectral variance was significantly affected by place of articulation and vowel context. The spectral variance of /p/ was significantly higher than /t/ and /k/, and the variance was significantly lower when followed by an /i/ vowel. For spectral skewness, ANOVA testing revealed significant main effects of place of articulation, vowel context, sex, and age. Finally, spectral kurtosis was only significantly affected by place of articulation. *Conclusions:* These findings indicate significant variation across all three places of articulation for all measures except spectral variance. Vowel context also has a clear effect on the spectral and acoustic properties of stop consonants. The sex-related differences in the spectral characteristics of stop consonants in children found in this study are likely due, at least in part, to learned or behavioral factors. *Relevance to the current work:* This study suggests links between the place of articulation and changes in consonant spectral measures.

Nishio, M., & Niimi, S. (2008). Changes in speaking fundamental frequency characteristics with aging. *Folia Phoniatrica et Logopaedica, 60*, 120-127. doi: 10.1159/000118510

*Purpose:* This study was designed to evaluate changes to speaking fundamental frequency patterns in typically aging speakers, both male and female. *Methods:* Participants in this study included 374 native speakers of Japanese (187 females, 187 males). Exclusion criteria included laryngeal pathology, neurological disease, professional voice training, and vision and/or hearing impairment. History of smoking was not among the exclusion criteria in order to provide data representative of the general population in Japan. Participants were split into the following age groupings: young speakers (77 males and 77 females 19-34 years old), middle-aged speakers (55 males and 55 females 35-59 years old), and elderly speakers (55 males and 55 females 60 years and older). All participants were instructed to read *The North Wind and the Sun* passage in Japanese at a comfortable pitch, loudness, and rate. Speaking fundamental frequency (SFF) measures were obtained for each sentence in the passage, and then averaged for each speaker. Participants were recorded in a sound-treated studio with a constant mouth-to-mic distance of 15 cm. SFF measures were obtained using the Computerized Speech Lab (CSL) system by Kay Elemetrics. Mean and standard deviation of the SFF values were calculated for males and females of each age group. *Results:* For male speakers, significant differences were observed in the mean and standard deviation SFF of elderly speakers compared with both the young and middle-aged groups, which were not significantly different. When male participants were grouped according to age by decade, results indicated small differences from the 20s through the

60s, but the mean SFF did increase in speakers in their 70s and older. No statistically significant differences were observed between male speakers grouped by decade. For female speakers, a one-way ANOVA revealed significant differences in mean SFF between the young group, middle-aged group, and elderly group. The highest values were observed in the young group, followed by the middle-aged group and then the elderly group. When grouped by decade, female speakers demonstrated significantly lower values in their 30s and 40s than in their 20s, with an age-related decrease observed across all decade groups. *Conclusion:* Male speakers exhibited no significant trend for SFF changes with age, although a slight increase was found in participants aged 70 years or older. Female speakers in their 30s and 40s exhibited markedly lower SFFs than those in their 20s. A moderate negative correlation was observed between age and SFF in female speakers. The degree of SFF change with aging was much larger in females than in males. *Relevance to the current work:* These results provide normative data for SFF in typically aging males and females.

Ramig, L. O., Gray, S., Baker, K., Corbin-Lewis, K., Buder, E., Luschei, E., . . . Smith, M. (2001). The aging voice: A review, treatment data and familial and genetic perspectives. *Folia Phoniatrica et Logopaedica, 53*, 252-265. doi: 10.1159/000052680

*Purpose:* This article was written to review aspects of the typically aging voice in reference to general physiologic aging. *Content:* It is generally reported that we reach peak bodily function at 30 years age and thereafter experience a steady decline. Ramig et al. highlight the well-documented physiologic changes that accompany typical aging, including changes to immunological, respiratory, gastrointestinal, skeletal, muscular, and dermatological systems. Also of note, they report structural and functional changes to the neurological system, including loss of neurons, decrease in synaptic activity, and reduced excitatory and inhibitory neurotransmitters. Research indicates that the effects of the aging process may not be uniform across all groups of people. Several studies have documented differential effects of aging between males and females. For example, females tend to have a longer life-expectancy and fewer age-related degenerative changes than their male counterparts. Because aging affects every individual at a different rate, physiologic age has been suggested as a preferred descriptor of the aging process compared with chronological age. The authors discuss the effects of vocal aging on perceptual, acoustic, aerodynamic, kinematic, electromyographic, and histological levels. For example, they cite studies which associate elderly voices with perceived hoarseness, harshness, lower pitch, vocal tremor, increased breathiness, increased strain, and unsteadiness. Also, aging male voices often present with higher $F_0$, higher $F_0$ and intensity variability, increased perturbation, and greater spectral noise. Elderly female voices are often characterized by lower $F_0$, increased $F_0$ variation and increased jitter. Ramig et al. cite previous studies which show that perceived vocal age correlates closely with physiologic age. Research also indicates that familial and genetic characteristics may play a role in determining the differential effects of aging on different body systems. *Relevance to the current work:* This report provides a very detailed description of the effects of physiologic aging on the voice and how the aging process differentially affects male and female voices.

Simpson, A. P. (2002). Gender-specific articulatory-acoustic relations in vowel sequences. *Journal of Phonetics, 30*, 417-435. doi: 10.1006/jpho.2002.0171

*Purpose:* The purpose of this study was to evaluate how gender differences in vocal tract dimensions may affect features of articulation. *Methods:* Data used in this study were obtained from an existing database containing acoustic and articulatory records from 22 male and 26 female speakers, between ages 18 and 37. Participants completed speech tasks with four gold pellets placed equidistantly along the midline of the tongue. These sensors tracked lingual movement during production of the sentence, *They all know what I said.* This sentence was selected because it involves a large amount of tongue body movement from a front and close vocalic stricture to a back and open configuration. Each participant was asked to repeat this sentence five times at a comfortable pitch and loudness. The ESPS program *formant* was used to obtain measures of F1, F2, and F3 during the vocalic stretch /eɪ ɑː/. Both visual and auditory evaluation of spectra were used to segment the vocalic stretch of interest. The dependent variables of this study were vowel stretch duration, formant tracks, and magnitude and speed of lingual movement. *Results:* A one-tailed *t*-test revealed that females had significantly longer durations for the vowel stretch in question. However, no significant gender differences were found in durations of the entire utterance. Analysis of formant tracks revealed that female acoustic excursion during the vocal stretch from the maximum to minimum was greater for both F1 and F2 compared with male speakers. Particularly, F1 excursion for female speakers was 59% greater than for male speakers. The F1 data suggest a perceptually more closed vowel quality for males and a more open female quality. The mean F2 and F3 tracks for males and females were found to run parallel with little change in the tonotopic distance between them throughout the vocalic stretch. Finally, movement of the lingual pellets during the vocalic stretch was analyzed to evaluate the magnitude and rate of tongue movement in male and female speakers. General movement patterns were found to be strikingly similar for all participants. However, male speakers exhibited significantly greater overall lingual displacement and faster tongue movement than female speakers. *Conclusion:* Gender-specific differences were found in both the acoustic and articulatory space traversed. Female speakers made greater acoustic excursions for shorter articulatory distances. The authors suggest that nonuniform differences in palate shape leads to nonuniform dynamic tongue movement patterns for males and females. *Relevance to the current work:* This study makes a number of important observations about physiologic and articulatory differences between genders.

Slawinski, E. (1994). Acoustic correlates of [b] and [w] produced by normal young to elderly adults. *Journal of the Acoustical Society of America, 95*, 2221-2230. doi: 10.1121/1.408682

*Purpose:* This study was designed to investigate age-related acoustic changes in the production of a phonetic contrast: stop consonant /b/ compared with the semivowel /w/. *Methods:* Participants in this study consisted of 39 total individuals who were divided equally into the following age groups: 20-30 years, 50-60 years, and 70-80 years. Each group was comprised of 13 individuals, 6 male and 7 female. All participants were native, monolinguals speakers of Canadian English with normal hearing and a negative history of articulation or hearing problems. Participants were seated in an anechoic chamber and asked to produce the CV syllables /wɑ/ and /bɑ/ at a comfortable rate and loudness level. Syllable stimuli were presented in random order, totaling five times each. Participants were instructed to pause in between the production of consecutive syllables to ensure initial-position productions. For each CV syllable produced, a waveform editing program was used to obtain the following acoustic measures: (1) onset of pre-voicing, (2) the release point, (3) the transition onset, (4) the onset of the steady state of the vowel, and (5) the change in amplitude energy from the onset of voicing to the release point. *Results:* MANOVA was performed with age grouping as the between-subjects variable and consonant as the within-subjects variable. Analysis revealed a significant effect of age on syllable duration for /bɑ/, with older adults producing significantly longer syllables than either younger or middle-aged subjects. For /wɑ/, older adults produced significantly longer syllables than younger and middle-aged adults, and younger adults produced significantly shorter utterances than middle-aged adults. Next, older adults were observed to produce significantly longer pre-voicing segments than the other age groups for both /bɑ/ and /wɑ/. The same pattern was observed for release segments, with the duration of this segment increasing progressively with age. Statistical analysis did not indicate any significant change in the duration of vowel transitions with age for /bɑ/. However, older adults did produce significantly shorter vowel transitions than either the young or middle-aged group. No significant age effect was found in the amplitude ratio for /bɑ/ or /wɑ/. *Conclusion:* This study revealed significant changes to temporal features of subphonemic segments with aging. These changes suggest a general decline in speech performance which may be due to a number of physiologic and neurologic changes that accompany the typical aging process. *Relevance to the current work:* The authors of this study documented a general slowing of speech production in multiple consonants and even identified physiologic processes which may contribute to that slowing.

Titze, I. R. (1989). Physiologic and acoustic differences between male and female voices. *The Journal of the Acoustical Society of America, 85*, 1699-1707. doi: 10.1121/1.397959

*Purpose:* The purpose of this article was to draw comparisons between male and female larynges by overall larynx size, vocal fold membranous length, elastic properties of vocal tissue, and pre-phonatory glottal shape. *Content:* The authors of this report highlight several size differences between male and female larynges. In the anterior-posterior dimension, male laryngeal cartilage is 20% larger than female cartilage, indicating a linear scale factor of 1.2. However, differences in vocal fold length are not uniform with cartilage dimensions. The membranous vocal fold length for male speakers is 60% longer than that of females, suggesting a linear scale factor of 1.6. This indicates that male vocal folds grow in the anterior-posterior direction disproportionately to other larynx dimensions. The authors of this article also cite previous research to discuss the role of variations in vocal fold thickness in the control of $F_0$. While there has been speculation that gender differences in $F_0$ may be accounted for by differences in vocal fold thickness, Titze writes that this is unlikely. This is because $F_0$ depends on the ratio of stiffness to mass, and simply adding tissue mass adds stiffness at the same time. Instead, Titze suggests that gender differences in $F_0$ may be due to the depth of vibration and membranous length rather than vocal fold thickness. Additionally, male vocal folds have been found to have a greater percentage of collagenous fibers than female vocal folds. This may account for differences in tissue linearity which, in turn, affects vibration patterns. *Conclusion:* Titze suggests that the scaling factor 1.6 based on vocal fold membranous length is the primary reason for gender differences in mean $F_0$. This factor, taken together with gender variations in vocal fold depth and histology, likely account for male-female differences in pitch, mean airflow, and aerodynamic power produced. *Relevance to the current work:* This article details anatomic and physiologic differences in male and female larynges as well as the acoustic implications of those differences.

Tjaden, K., & Turner, G. S. (1997). Spectral properties of fricatives in amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research, 40*, 1358-72. doi: 10.1044/jslhr.4006.1358

*Purpose:* The purpose of this study was to use fricative spectral moments as a measure of articulatory precision in speakers with ALS and healthy controls. This study also sought to evaluate the relationship between spectral moments and potential perceptual correlates of articulation quality. *Methods:* Participants in this study included 7 individuals with ALS and 7 age- and gender-matched controls (4 male and 3 female speakers in each group). Participants with ALS exhibited mild to moderate speech intelligibility deficits. These individuals were part of a larger study investigating speech rate change in ALS. All participants were instructed to read the Farm Passage at a habitual speaking rate. For the present study, 14 words containing initial /s/ and 2 words with initial /ʃ/ in the reading passage were analyzed. Only word-initial fricatives were used to control for possible position effects. Fricatives were segmented visually using a combined waveform and spectrographic display. Coefficients for the first through fourth spectral moments were calculated for the initial 40 ms of frication of each token. Coefficients for each moment were grouped for each 10 ms time slice across all repetitions of each target phoneme for each speaker. Additionally, 4 graduate-level students in communication disorders were recruited to serve as judges. Listeners were presented with speech samples in random order and asked to rate precision of particular consonants on a 7-point interval scale. Ratings for each sample were averaged across the four listeners and plotted against first moment difference scores. *Results:* Female speakers with ALS demonstrated significantly lower first moment coefficients than healthy females at all time slices for /s/. Two of the three female speakers with ALS demonstrated significantly higher mean values for /ʃ/. Males with ALS also exhibited a tendency for higher first moment coefficients for /ʃ/ than their healthy counterparts. However, no significant differences were found between mean first moment coefficients of healthy speakers and those with ALS. Analysis of the second and fourth spectral moments revealed no significant between group differences. For female speakers with ALS, third moment coefficients tended to fall toward the top of the distribution values for /s/ and toward the bottom for /ʃ/. For males with ALS, third moment coefficients tended to fall toward the lower end of the distributions for both /s/ and /ʃ/, although no significant between-group differences were observed. Finally, the data indicate a significant relationship between spectral moment coefficients and listener-rated consonant precision. *Conclusion:* The results of this study demonstrate that ALS can have a significant effect on articulation as measured by spectral moment coefficients. Also, the relationship between spectral moment coefficients and listener ratings of consonant precision suggest that fricative spectral analysis may be a salient measure of speech performance. *Relevance to the current work:* These data show that reduced neuromotor control (which often occurs as a part of typical aging) can impact articulatory precision. Also, this study makes a case for using fricative spectral moment analysis as a measure of speech production.

Torre, P., & Barlow, J. A. (2009). Age-related changes in acoustic characteristics of adult speech. *Journal of Communication Disorders, 42*, 324-333. doi: 10.1016/j.jcomdis.2009.03.001

*Purpose:* This study was conducted to investigate the effects of age and sex on fundamental frequency, formant frequencies, and voice onset time. *Methods:* Participants in this study were divided into two groups by age. The younger adult group consisted of 27 individuals between ages 20 and 35 years (15 women and 12 men). The older adult group was comprised of 59 individuals between ages 60 and 89 (32 women and 27 men). All individuals in the older adult group had some degree of hearing loss but were allowed to wear their hearing aids during the study. All participants were seated in a sound booth and given five written lists of 22 CVC words each. They were instructed to read each word with the carrier phrase, *Say the word* _____ at a comfortable pitch and loudness. Speech recordings were imported into Praat speech analysis software in order to obtain the following acoustic measures: $F_0$, F1, F2, F3, and VOT. These measures (excluding VOT) were obtained at the steady-state portion of each monophthong vowel for each CVC word. VOT was measured for all words with initial oral stops. *Results:* Repeated measures ANOVA revealed significant sex-by-age group interactions for $F_0$ and F1 across all six vowels, with the exception of F1 for /i/, /u/, and /ʌ/. For all vowels, $F_0$ decreased with age for women and increased with age for men. F1 values for male and female speakers were found to decrease with age for /ɪ/, /ɛ/, /æ/, and /ʌ/. No significant age group differences were observed for F2 and F3 values. Generally, the older adult group demonstrated larger standard deviations for these frequency measures than the younger adult group. Next, statistical analysis revealed significant sex-by-age group interactions for VOT. Older speakers demonstrated shorter VOTs for oral stops than their younger counterparts. A significant main effect of sex on VOT was only observed for /b/, with men showing longer VOTs than women. Finally, older adults showed larger VOT standard deviations than younger adults. *Conclusion:* The results of this study highlight significant changes to the acoustic properties of speech production that accompany typical aging. Additionally, these data show that older adults exhibit greater intraspeaker variability in selected acoustic features. *Relevance to the current work:* This study evaluated acoustic properties of typically aging speech production for both male and female speakers.

U.S. Department of Health and Human Services, Administration for Community Living, Administration on Aging. (2017). *Projected future growth of older population*. Retrieved from https://www.acl.gov/node/579.

*Purpose:* The purpose of this report was to describe the aging and disabled populations as they currently exist in the United States as well as to explain projections for future growth of the older population in the country. *Content:* According to this report, there are currently 65 million individuals in the U.S. who are elderly (defined as those who are 60 years and older). The vast majority of this group reside in non-institutional settings. Likewise, approximately 57 million people with disabilities also live in non-institutional settings. This report states that elderly adults are one of the most rapidly-growing populations in the nation. The Administration for Community Living estimates that, by 2020, older adults in America may number more than 77 million. *Relevance to the current work:* This report highlights that the elderly make up a large portion of the population in the U.S. and, by extension, a large portion of those who require speech and language services. This speaks to the need to investigate how speech production patterns vary for this population.

Vanderydt, R.E. (1998). A study of voice similarities among family members (Unpublished master's thesis). Massachusetts General Institute of Health Professions, Boston, Massachusetts. Retrieved from https://search-proquest-com.erl.lib.byu.edu/comdisdome/docview/85147153

*Purpose:* The purpose of this study was to examine the voices of similar-sounding family members to determine which, if any, acoustic properties support perceptual ratings of voice similarity. *Methods:* Speech sample recordings were obtained from 9 speakers: 2 mother/daughter pairs, 1 sister pair, 1 brother pair and 1 unrelated male. Participants were all native speakers of American English with no history of voice or hearing problems. They were asked to read two sentences, *the debate hurt Bob* and *Steve eats candy canes*. Next, they read the following reiterant syllable imitations: *ma mama ma ma, ha haha ha ha*, *ma ma mama ma*, and *ha ha haha ha*. These sentences were selected because of their differing stress patterns. Stimulus sentences were paired and presented to 26 listeners in 3 subtests. The first consisted of 42 recordings of full sentence and syllable imitations. The second test contained 42 trimmed sentences and syllables (e.g. "Candy canes" and "haha ha"). The third test was comprised of 23 pairs of stimuli from the previous test. In each test, listeners were asked to rate the likeness of the two voices on a scale from 1 to 5. The stimulus pairs were presented 2 to 3 times in each test and the ratings given to each set was averaged. *Results:* Perceptual evaluation resulted in pairs of related voices being rated as sounding significantly more similar than pairs of unrelated voices across all three subtests. This trend held up regardless of differences in speaker volume, prosody, and in some cases, age. Only one related pair (mother/daughter) was not judged as sounding similar in subtests 2 and 3. Acoustic analysis of vowel spectra was performed to obtain measures of $F_0$ and timing of voice onset and offset. Pairs of related female voices that were rated as sounding similar were found to have visually similar $F_0$ contours. Related female voices (except the dissimilar-sounding mother/daughter pair) were found to produce voiced syllables of very similar lengths. On average, the voiced region of syllables produced by each speaker varied by only 50 ms for the first syllable, 10 ms for the second, and 30 ms for the third. Unrelated female speakers varied by approximately 130 ms for the first syllable and 170 ms for the second. Fundamental frequency contours did not appear to play as much of a role in the similarity of male voices, as both related and unrelated males presented similar $F_0$ contour shapes. Related males produced voiced syllables that differed only by 40 ms for the first syllable, 50 ms for the second, and 20 ms for the third. Unrelated male speakers were only significantly different for the second syllable, varying by approximately 100 ms. The mother/daughter pair that was rated as sounding dissimilar were observed to have very different fundamental frequency contours. Also, they displayed voice duration differences of 210 ms for syllable 2 and 200 ms for syllable 3. *Conclusion:* The results of this study show that the voices of related speakers were consistently rated as sounding similar. Acoustic analysis revealed that the speakers who sound similar also share similar $F_0$ contours and voice timing. These acoustic features were not shared by pairs of unrelated speakers or related speakers who were not rated as sounding similar. *Relevance to the current work:* This study indicates that related speakers are often perceived as sounding similar and it suggests that pitch and timing characteristics support perceptual ratings of similarity.

Westrop, R. M. (2000). Segment durational changes from young adults to elderly: A familial study (Unpublished master's thesis). Utah State University, Logan, Utah. Retrieved from https://search-proquest-com.erl.lib.byu.edu/comdisdome/docview/85146865/

*Purpose*: This study was designed to investigate differences in speech segment durations as a function of age, gender, and family membership. *Method:* This study used preexisting data which were previously collected for an investigation of voice production at the University of Utah. Speech samples from 10 families were used, totaling 93 individuals spanning three generations, ranging from 21 to 82 years of age, all of whom had no history of articulation disorder, stroke, or active neurological disease, or profound hearing loss. All the speech samples used in this study were from Utah families to control for differences in dialect. Each participant was instructed to repeat the sentence *The blue spot is on the key again* three times, from which were measured vowel, stop-gap, and voice onset time (VOT) durations. This sentence was selected because it contained vowels representing widely separated locations on the vowel quadrilateral (i.e. /u/, /ɑ/, /ɪ/, and /i/). Speech samples were collected using a head-mounted microphone positioned 3 cm from each speaker's mouth. Total utterance duration, vowel duration, stop-gap duration, and VOT were hand measured in milliseconds (ms) from wide-band (300 Hz) spectrograms. For each speaker, duration ratios were computed by dividing each individual durational measure (vowel, stop-gap, and VOT) by the total utterance duration to compensate for speaking rate and pauses. *Results:* The results revealed that increased age had a statistically significant effect on vowel durations for /ɪ/, and /i/, but not for /u/ and /ɑ/. A gender effect on vowel duration was only observed with the vowel /ɑ/, with male subjects' duration longer than females'. There were no statistically significant effects of family membership on vowel duration. Although no significant interactions between age or gender and stop-gap duration were found, family membership did have a significant effect on stop-gap duration. VOT was observed to decrease as age increased. Also, male speakers' VOTs were significantly shorter than those of female speakers. Family membership did not have an effect on VOT. C*onclusions:* Based on these results, the researchers concluded that age, gender, and family membership affect individual speech parameters differently. As a general rule, elderly speech is characterized by slowness, but durational measures may also be influenced by environment and speaking context. *Relevance to the current work:* The study evaluated the effects of age, gender, and family relationship on acoustic characteristics of speech.

Xue, S. A., & Deliyski, D. (2001). Effects of aging on selected acoustic voice parameters: Preliminary normative data and educational implications. *Educational Gerontology, 27*, 159-168. doi:10.1080/03601270151075561

*Purpose:* This study was conducted to obtain normative acoustic data for elderly male and female voices and to compare these acoustic norms and thresholds to those of younger and middle-aged adults. *Methods:* Participants in this study included 21 elderly men (mean age = 75.4 years) and 23 elderly women (mean age = 74.8). All were free of neurological disease as well as speech or voice disorders and passed a hearing screening at 35 dB HL in at least one ear. After a 5-minute warm-up exercise, participants were asked to sustain phonation of the vowel /ɑ/ three times at a comfortable pitch and loudness for at least 3 seconds per production. The middle 1-s portion of the second production was used for acoustic analysis. The Multi-Dimensional Voice Program (MDVP) of the Computerized Speech Lab by Kay Elemetrics was used to obtain measurements of 15 acoustic parameters: average $F_0$, absolute jitter, jitter percent, phonatory fundamental frequency range (PFR), pitch period perturbation quotient (PPQ), relative average perturbation (RAP), smoothed pitch period perturbation quotient (sPPQ), $F_0$ standard deviation ($F_0$SD), $F_0$ variation (v$F_0$), amplitude perturbation quotient (APQ), shimmer in decibels (ShDB), shimmer percent, noise-to-harmonic ratio (NHR), soft phonation index (SPI), and voice turbulence index (VTI). Values for these acoustic measures were obtained for elderly speakers and compared to published norms for younger and middle-aged adults. *Results:* MANOVA was performed using age group as the independent variable and the selected acoustic measures as dependent variables. Results indicated that elderly male and female participants had significantly lower $F_0$ than young and middle-aged speakers. For the other 14 acoustic measures obtained, the elderly group demonstrated significantly higher measurements than the other two age groups. This suggests that individuals over age 70 had significantly greater vocal frequency variation (absolute jitter, jitter percent, PFR, PPQ, RAP, sPPQ, $F_0$SD, and v$F_0$), higher intensity variations (APQ, ShdB, and shimmer percent), and more noise in the voice signal (NHR, SPI, and VTI) compared with younger and middle-aged speakers. *Conclusion:* The 15 acoustic parameters obtained in this study indicate that elderly individuals exhibit significantly different voicing features that are generally consistent with poorer vocal output. *Relevance to the current work:* The results of this study reveal characteristics of the typically aging voice and support the notion of a general decline in speech performance with age.

# APPENDIX B

## Reading Passages

Goldilocks and the Three Bears passage:

The three bears went into the living room. “Someone has been sitting in my chair”, said the father bear. “Someone has been sitting in my chair”, said the mother bear. “And someone has been sitting in my chair and has broken it all to pieces,” cried the baby bear. Then the three bears went upstairs. “Someone has been sleeping in my bed”, said the father. “Someone has been sleeping in my bed”, said the mother bear. “Someone has been sleeping in my bed”, cried the baby bear. Goldilocks opened her eyes and saw the three bears standing over her. “Oh dear”, Goldilocks said.

Rainbow passage:

When the sunlight strikes rain drops in the air, they act like a prism and form a rainbow. The rainbow is a division of white light into many beautiful colors. These take the shape of a long, round arch, with its path high above and its two ends apparently beyond the horizon. There is, according to legend, a boiling pot of gold at one end. People look, but no one ever finds it. When a man looks for something beyond his reach, his friends say he is looking for the pot of gold at the end of the rainbow.

APPENDIX C

**Descriptive Statistics by Family and Sex**

Table 1

*Descriptive Statistics for Average Spectral Mean and SD for /s/ in Hz*

| | Spectral *M* | | | | Spectral *SD* | | | |
|---|---|---|---|---|---|---|---|---|
| | Male | | Female | | Male | | Female | |
| Family | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| 17 | 5244.50 | 1160.42 | 8307.68 | 1643.88 | 2910.47 | 497.82 | 2410.52 | 523.79 |
| 18 | N/A | N/A | 8080.40 | 797.98 | N/A | N/A | 3484.49 | 590.92 |
| 19 | 6761.62 | 1017.39 | 8458.04 | 2078.13 | 3227.29 | 580.67 | 2745.87 | 850.56 |
| 20 | 5612.52 | 1129.45 | 8296.53 | 618.05 | 3123.93 | 529.50 | 2886.99 | 687.35 |
| 21 | 6275.92 | 1042.06 | 7740.58 | 1336.01 | 3020.97 | 635.14 | 3615.92 | 80.73 |
| 22 | 6723.05 | 711.03 | 8474.94 | 965.25 | 3038.12 | 474.78 | 3061.20 | 1002.25 |
| 23 | 6754.48 | 999.17 | 8715.07 | 1081.23 | 2805.81 | 572.26 | 3288.65 | 1350.68 |
| 24 | 6460.15 | 454.79 | 7832.77 | 1081.60 | 3485.37 | 404.58 | 3148.78 | 863.34 |
| 27 | 5586.79 | 1562.06 | 8698.12 | 1030.10 | 3590.94 | 542.96 | 2661.05 | 548.95 |
| 28 | 6099.85 | 1660.35 | 8884.62 | 858.51 | 3106.64 | 411.35 | 2517.06 | 780.04 |
| 29 | 5977.93 | 884.56 | 8049.63 | 1327.85 | 3250.13 | 376.88 | 3001.14 | 469.39 |
| 30 | 6270.17 | 542.64 | 9005.98 | 586.83 | 2412.32 | 269.23 | 2953.12 | 1050.62 |
| 31 | 9281.61 | 1273.62 | 9038.94 | 1355.52 | 2488.33 | 248.66 | 2268.31 | 612.92 |
| 32 | 6698.32 | 2237.15 | 8355.89 | 1555.61 | 2340.71 | 161.22 | 2941.41 | 356.38 |
| 33 | 6088.02 | 2175.41 | 8731.69 | 2242.36 | 3197.13 | 428.20 | 3111.20 | 972.44 |
| 35 | 6668.06 | 667.65 | 8421.01 | 933.99 | 2826.16 | 176.97 | 2330.95 | 818.43 |
| 37 | 6402.08 | 1021.07 | 8600.23 | 1109.73 | 2923.47 | 414.71 | 2554.37 | 803.88 |

*Note:* N/A indicates a family with only one participant of the designated sex.

Table 2

*Descriptive Statistics for Average Spectral Skewness and Kurtosis for /s/ in Hz*

| | Spectral Skewness | | | | Spectral Kurtosis | | | |
|---|---|---|---|---|---|---|---|---|
| | Male | | Female | | Male | | Female | |
| Family | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| 17 | 0.03 | 0.42 | -1.04 | 0.63 | 0.81 | 1.19 | 4.09 | 2.30 |
| 18 | N/A | N/A | -0.35 | 0.55 | N/A | N/A | 0.92 | 1.33 |
| 19 | 0.29 | 0.43 | -0.24 | 0.84 | 2.46 | 2.51 | 3.70 | 1.87 |
| 20 | 0.67 | 0.85 | -0.35 | 0.57 | 2.59 | 2.60 | 3.38 | 1.96 |
| 21 | 0.29 | 0.18 | -0.76 | 0.41 | 2.08 | 2.04 | 1.50 | 0.34 |
| 22 | 0.39 | 0.52 | -0.39 | 0.52 | 1.34 | 1.22 | 2.01 | 1.28 |
| 23 | 0.92 | 0.59 | 0.22 | 0.40 | 2.46 | 1.53 | 2.88 | 3.58 |
| 24 | 0.49 | 0.48 | -0.25 | 0.53 | 0.48 | 0.52 | 2.97 | 2.04 |
| 27 | 0.93 | 0.43 | -0.04 | 0.72 | 1.83 | 1.52 | 4.64 | 3.62 |
| 28 | 0.45 | 0.65 | -0.83 | 0.51 | 1.91 | 1.95 | 5.73 | 3.66 |
| 29 | 0.49 | 0.35 | -0.02 | 0.73 | 1.48 | 1.05 | 2.45 | 1.47 |
| 30 | 0.67 | 0.48 | -0.43 | 0.44 | 3.82 | 1.03 | 3.34 | 2.23 |
| 31 | 1.04 | 1.05 | 0.00 | 0.34 | 2.68 | 3.88 | 3.69 | 2.24 |
| 32 | 0.48 | 0.63 | -0.38 | 0.64 | 3.20 | 0.75 | 2.42 | 1.89 |
| 33 | 1.21 | 0.76 | -0.29 | 0.95 | 2.96 | 2.28 | 3.25 | 2.97 |
| 35 | 1.11 | 0.25 | -0.78 | 0.59 | 2.30 | 0.63 | 6.96 | 4.84 |
| 37 | 0.75 | 0.37 | 0.43 | 0.09 | 2.08 | 0.76 | 5.79 | 5.53 |

*Note:* N/A indicates a family with only one participant of the designated sex.

Table 3

*Descriptive Statistics for $F_0$ and STSD in Hz*

| | $F_0$ | | | | STSD | | | |
|---|---|---|---|---|---|---|---|---|
| | Male | | Female | | Male | | Female | |
| Family | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| 17 | 105.71 | 8.57 | 212.20 | 25.86 | 2.02 | 0.41 | 2.40 | 0.67 |
| 18 | N/A | N/A | 177.58 | 7.77 | N/A | N/A | 2.17 | 0.17 |
| 19 | 105.09 | 15.66 | 183.42 | 3.90 | 2.47 | 0.39 | 2.35 | 0.30 |
| 20 | 111.07 | 8.31 | 202.77 | 9.07 | 2.03 | 0.35 | 2.08 | 0.61 |
| 21 | 104.12 | 4.48 | 180.96 | 0.71 | 2.30 | 0.23 | 2.18 | 0.26 |
| 22 | 111.69 | 6.87 | 187.99 | 15.09 | 2.20 | 0.26 | 2.11 | 0.28 |
| 23 | 95.37 | 8.43 | 206.06 | 11.92 | 1.63 | 0.37 | 2.07 | 0.21 |
| 24 | 106.19 | 12.88 | 220.40 | 13.63 | 2.17 | 0.58 | 2.14 | 0.03 |
| 27 | 120.34 | 21.29 | 200.38 | 9.31 | 2.31 | 0.20 | 2.25 | 0.36 |
| 28 | 106.12 | 9.03 | 192.70 | 10.72 | 2.27 | 0.37 | 2.27 | 0.72 |
| 29 | 113.57 | 15.48 | 215.23 | 13.59 | 1.90 | 0.40 | 2.00 | 0.28 |
| 30 | 115.41 | 8.35 | 195.92 | 8.82 | 2.22 | 0.80 | 2.28 | 0.69 |
| 31 | 115.57 | 15.54 | 190.92 | 8.75 | 2.28 | 0.82 | 2.16 | 0.33 |
| 32 | 121.98 | 9.33 | 199.99 | 22.24 | 1.73 | 0.92 | 1.77 | 0.51 |
| 33 | 115.65 | 11.95 | 190.53 | 34.79 | 1.93 | 0.51 | 2.20 | 0.25 |
| 35 | 123.27 | 26.99 | 191.77 | 15.23 | 2.53 | 0.76 | 2.62 | 0.44 |
| 37 | 114.98 | 17.53 | 205.35 | 7.63 | 2.12 | 0.74 | 1.94 | 0.48 |

*Note:* N/A indicates a family with only one participant of the designated sex.

Table 4

*Descriptive Statistics for Speaking/Pausing Ratio*

| | Male | | Female | |
|---|---|---|---|---|
| Family | *M* | *SD* | *M* | *SD* |
| 17 | 0.85 | 0.03 | 0.78 | 0.08 |
| 18 | N/A | N/A | 0.83 | 0.05 |
| 19 | 0.87 | 0.04 | 0.86 | 0.05 |
| 20 | 0.85 | 0.06 | 0.89 | 0.05 |
| 21 | 0.80 | 0.02 | 0.84 | 0.02 |
| 22 | 0.81 | 0.03 | 0.89 | 0.02 |
| 23 | 0.81 | 0.05 | 0.85 | 0.03 |
| 24 | 0.83 | 0.02 | 0.88 | 0.06 |
| 27 | 0.83 | 0.06 | 0.85 | 0.06 |
| 28 | 0.80 | 0.13 | 0.87 | 0.05 |
| 29 | 0.82 | 0.05 | 0.84 | 0.07 |
| 30 | 0.80 | 0.05 | 0.82 | 0.06 |
| 31 | 0.82 | 0.06 | 0.86 | 0.05 |
| 32 | 0.82 | 0.02 | 0.82 | 0.05 |
| 33 | 0.83 | 0.04 | 0.87 | 0.05 |
| 35 | 0.77 | 0.02 | 0.80 | 0.06 |
| 37 | 0.87 | 0.05 | 0.85 | 0.03 |

*Note:* N/A indicates a family with only one participant of the designated sex.

Table 5

*Descriptive Statistics for CPPS*

| | Male | | Female | |
|---|---|---|---|---|
| Family | *M* | *SD* | *M* | *SD* |
| 17 | 8.89 | 0.54 | 9.03 | 0.63 |
| 18 | N/A | N/A | 8.57 | 0.41 |
| 19 | 7.93 | 1.01 | 8.01 | 0.67 |
| 20 | 8.37 | 1.29 | 8.87 | 0.59 |
| 21 | 8.74 | 0.55 | 9.15 | 0.05 |
| 22 | 8.39 | 0.75 | 8.97 | 0.91 |
| 23 | 8.03 | 1.02 | 8.16 | 0.57 |
| 24 | 9.56 | 0.75 | 9.82 | 0.42 |
| 27 | 8.28 | 1.05 | 9.88 | 1.24 |
| 28 | 8.19 | 1.46 | 8.90 | 1.05 |
| 29 | 7.66 | 0.98 | 7.90 | 0.86 |
| 30 | 8.03 | 0.92 | 8.24 | 0.85 |
| 31 | 9.22 | 0.59 | 9.26 | 0.72 |
| 32 | 9.40 | 0.17 | 8.32 | 0.96 |
| 33 | 8.81 | 0.55 | 8.98 | 1.14 |
| 35 | 7.86 | 0.75 | 8.46 | 1.11 |
| 37 | 8.04 | 0.44 | 8.98 | 1.31 |

*Note:* N/A indicates a family with only one participant of the designated sex.