Aging Effects on Acoustic Characteristics of Adult Speech
Elaine Romenesko, Houri K. Vorperian, Raymond D. Kent, and Diane Austin
Vocal Tract Development Lab, Waisman Center, University of Wisconsin - Madison

Introduction
As we age, speech and voice production are affected by anatomic and physiologic changes (1). The speech production system consists of the respiratory, laryngeal (voice box), and supralaryngeal, or vocal tract (oral and pharyngeal cavity) subsystems.

• Acoustic analysis is one method used to describe and measure these changes. Figure 1 illustrates the association between acoustic variables and anatomic structures of speech and song.
  o Fundamental frequency (F0), or pitch, is the vibrating frequency of the vocal folds in the larynx.
  o Formant frequencies (F1, F2, F3, and F4) are the resonant frequencies of the vocal tract for a particular vowel production.
  o There is an inadequate understanding of the sex-specific changes in speech that occur in healthy aging, and published reports are discrepant (2). Collecting information on age-related changes will aid in understanding:
    o normal aging of speech and song production in both sexes.
    o communication difficulties in diseases that occur with aging, including neurodegenerative diseases and their progression (e.g. Parkinson’s disease, ALS or Lou Gehrig disease).
    o the acoustic information needed for speech technologies such as speech synthesis and automatic speech recognition.

• The analyses reported here focus on vowel production and are part of a larger project on the biologic bases of speech production and speech development.

• This study reports data on the fundamental frequency (F0) and the first four formants (F1-F4) of vowels in healthy male and female speakers ages 20-92 years.

Methods
Participants: 84 healthy adult participants (35 male and 49 female) ages 20 to 92 years were grouped into one of the following three age cohorts: young (20-35 years), middle (40-55 years), and older (70-92 years).

Stimuli: Participants were instructed to repeat monosyllabic words that contained one of the corner vowels: i, u, or o. There were 7 tokens per vowel (5 distinct words per vowel, with 2 words repeated). See Figure 2.

Recording protocol: Participants’ speech was recorded in a quiet room using a Shure SM48 microphone connected to a Marantz-PMD660 digital audio recorder. The words were presented visually and aurally using a laptop. The TOCS+ Platform program (3) was used for stimulus randomization. Participants were instructed to repeat the words at a normal loudness level.

Figure 1: Midsagittal CT image with the vocal tract (resonator) traced in yellow and the larynx (sounds source) cite marked with a blue oval.

Results
• Figures 4a-4e display summary plots of the data for adult male (blue) and female (red) speakers for the fundamental frequency (F0) and formant frequencies (F1-F4). When outliers were removed, approximately 1% of measurements were excluded from modeling/analysis efforts.

• These figures show the mean of the three adult age cohorts connected. Age cohort comparisons were made separately for males and females using mixed effects models, with feast effects for sex, age cohort, and sex by age cohort interaction, and a random effect for subjects. The p-values for the age cohort comparisons were adjusted using the Tukey-Kramer method.

Formant and vowel specific observations:
F0 of all vowels continues to decrease with age in females but not males.
F1 decreases significantly for all vowels except /i/ in female adult age cohorts.
F2 only increases significantly for the vowel /u/ in female adult age cohorts.
F3 shows no changes in any adult age cohorts.
F4 reveals no changes in any adult age cohorts.

Discussion – Conclusions
The carefully specified methods of acoustic analysis improved accuracy of formant estimation and reduced extraneous sources of variability, so the conclusions listed below reflect the effects of aging on speech/vowel acoustics at the word level.
• Results show that the effects of aging on vowel acoustics (F0, F1-F4) are:
  o sex-specific
  o vowel-specific

• Present findings indicate that:
  o Male speakers had no significant changes in any of the acoustic measurements for any of the four vowels.
  o Female speakers had a significant decrease in: F0 for all vowels, F1 for all vowels except /i/, and F2 for the vowel /u/ only. The most change occurred between the young adult and middle adult age groups, and the young adult and older adult age groups.

• Changes in F0 have been attributed to the calcification of laryngeal cartilages, vocal fold shortening, and epithelium thickening (4). Furthermore, changes in F0 in female speakers have been attributed to hormonal changes due to menopause. Present findings support the latter because:
  • the decrease in F0 was only present in female speakers.
  • F0 decreased the most between young adult (20-35 years) and middle (40-55 years) age groups, but there were no significant changes between middle adult and older adult (70-92 years) age groups.

• It has been hypothesized that the decrease in F1 in female speakers could be due to LPC mistakenly tracking F0 instead of F1 (5). However, others have attributed it to age-related changes in vocal tract dimensions, or positioning of speech structures (6). Our data are more consistent with the second explanation because all vowels had significantly lower F1. The first explanation relates more to high vowels than low vowels.

• The change in F2 of /u/ in female speakers is puzzling. Changes in anatomy, lips’ thickness, or lip tissue are potential explanations.

• Research on anatomic changes of vocal tract structures during aging is currently underway in this laboratory using medical imaging studies (CT and MRI).

• Additional research with systematic use of various speech stimuli is needed to more fully examine age-related acoustic variations. It is possible that vowel specific changes might help dissociate the various characteristics of aging.

• Furthermore, additional assessment of increased variability with aging, as shown in the acoustic measurements, is warranted, including inter-subject and intra-subject variability.

• Normative data of the kind reported here are needed as reference points for the analysis of speech characteristics of persons with various diseases and disorders. For example, studies have shown significant changes in formant patterns in persons with Parkinson’s disease, amyotrophic lateral sclerosis (ALS), hearing loss, stuttering, and psychological distress.

Acoustic Analysis: Based on findings from previous work (1-3), the following protocol was developed and used to make accurate F0 and F1-F4 measurements:

1. Speech recordings were segmented into separate waveforms for each word.
2. F0 and F1-F4 measurements were estimated using a modified version of the acoustic analysis software package TF32 (4).
3. The waveform and formant spectrum with the LPC formant trace were displayed and the following vowel-specific time points were used to make measurements of all formants: (See Figure 3)

• a - point with highest F2 frequency
• I - point with lowest F2 frequency
• r - point with least separation between F1 and F2 frequencies
• r - point with most overly spaced formants (taking care to avoid measurement at a point with decreasing F2-F1 difference, which reflects backing of the vowel)

• F0 was measured with TF32′s pitch trace at the same point where the formant estimates were taken.

• Missing data occurred when it was not possible to make a measurement for a specific formant.
  o Percent of missing data by measurement is: F0: 0.09%; F1: 0.04%; F2: 0.00%; F3: 1.28%; F4: 4.85%

Acknowledgments:
We thank the following individuals for assistance with data collection and acoustic analysis (listed alphabetically): Julie Eichhorn, Katelyn Timlan, and Alyssa Wil. We also thank Ying Ji Chuang for assistance with figures.

References:
(3) Hodge, M., Daniels, J., & Gotzke, C. L. (2009). TOCS+ Intelligence Measures (Version 5.3) [computer software]. Edmonton, Canada: University of Alberta.

Spectrogram
Spectral slice obtained at approximate vowel midpoint:
Formants F1, F2, F3, F4

Figure 3: Acoustic analysis display of the word “sad” using TF32 with the formants F1-F4 labeled.

Acknowledgement:
Research supported by NIH-NINDS RO1DC006282 for the project titled MRI & CT studies of the developing vocal tract.