Age and Speech Production: A Longitudinal Study of 50 Years

Eric J. Hunter, PhD1,2,3, Mara Kapsner-Smith, M.S.1, Patrick Pead1, Megan Zito Engar, B.S.1, and Wesley R. Brown, M.D.1

1National Center for Voice and Speech, University of Utah, Salt Lake City, UT
2Center on Aging, University of Utah, Salt Lake City, UT
3Division of Otolaryngology–Head and Neck Surgery, University of Utah School of Medicine, University of Utah, Salt Lake City, UT

To the Editor.

Speech changes with age, affecting quality of life1,2. Underlying degenerative processes include laryngeal neuromuscular degeneration through atrophy and dystrophy, and edema in the vocal fold cover3–6. Because voice production structures share physiological territory with the aerodigestive tract, age-related degeneration of the voice may coincide with degeneration of other key functions such as breathing, swallowing, and airway protection. Historically, age-related voice studies have been cross-sectional in nature, identifying age-related vocal characteristics by comparing an elderly subject group to a younger group. Although the use of subject groups provides general trends, longitudinal case studies may provide additional insights by tracking the progression of voice, swallowing and breathing characteristics with age without the effects of inter-subject statistical averaging and variability.

The current case study uses 50 years (1958–2007, 48–98 y/o) of speech recordings. The subject is a male lay leader of an international church. In addition to the unique longitudinal breadth of his speeches, this subject and his body of speeches are unique because (1) he received no training as a public speaker and used none of the traditional rhetorical characteristics of sermons; (2) he avoided smoking, coffee, and alcohol, common vocal irritants that might obfuscate age-specific changes to the voice; (3) the acoustical environment were consistent, one of two multi-purpose university arenas; and (4) all of the speeches were long enough to provide a sustained representative voice sample for analysis. Two types of analyses were employed: speech fundamental frequency to reveal the current health of the laryngeal physiology, as well as length of speech breath groups to indicate efficiency of laryngeal valving and/or lung vital capacity.

Overall, the subject’s voice changed significantly in the mid to latter part of the sixth decade (Figure 1), which could be traced to age-related physiological processes. Generally, speech...
fundamental frequency decreased until about age 68 (Figure 1a). From age 68 to 98 years, average pitch increased from 140 to 160 Hz and the range (inter-quartile range) decreased by 20 percent. Because speech fundamental frequency depends on the physiology of the vocal folds and control of the musculature of the larynx, changes in mean and range may suggest a deterioration of the state of the tissue and general motor control with age. For example, age-related loss of mass of itself would increase the average speech fundamental frequency; however, decreased mass in the vocal folds could cause the vocal folds to begin to bow\(^7\). Further, if the subject adjusted for the bowing by increasing the stretch of the vocal fold to assist with glottal closure during phonation, this would also raise average speech fundamental frequency.

Changes in speech fundamental frequency corresponded with a reduction of speech breathing length. The subject increased the number of breath groups per minute (6.3% per decade), losing about 6–6.5 percent of speech breath group length per decade (Figure 1b). This change was almost imperceptible until the sixth decade. Simultaneously, the standard deviation of words per breath group decreased nearly linearly throughout the observation period. Thus, the subject could not sustain the same number of words in a breath group and needed to breathe more frequently while speaking. This change might have been caused by (1) a less flexible rib cage and the loss of vital capacity; or (2) increased glottal chink or bowing of the vocal folds\(^8\), resulting in more air leakage during speaking and reduce the air available.

It is possible the results were affected because variations of recording environment, recording equipment and compression of the audio were not controlled. Nevertheless, the effects were likely minimal because (1) the venues and communication context were similar; (2) the metrics used are less sensitive to these variabilities; and (3) the results were similar to other reports in the literature. Further, while the longitudinal breadth of the study period makes these results valuable, they are nevertheless preliminary because only one subject was examined.

Systemic neuromuscular changes can be inferred from changes in speech fundamental frequency and speech breathing. Other changes, such as increased risk of dysphagia (the inability to swallow safely and efficiently), may also correlate with these changes. Additional studies may identify indicators of when further assessments and treatments of age-related changes (e.g., dysphagia, dysphonia) are needed, or when preventative exercise may assist in slowing age indicators\(^9,10\). Future longitudinal studies using more subjects (both genders) may further understanding of normal changes due to aging versus pathology. However, such a corpus of recordings must first be filtered based on communicative intent, venues, knowledge of vocal coaching and related information.

**Acknowledgments**

The authors would like to thank the many members of the NCVS who provided insights, feedback to the study design, and interpretation of results during normal laboratory meetings.

**Funding Sources and related paper presentations:** Work supported by National Institute on Deafness and Other Communication Disorders (grant number R01DC04224).

**Sponsor’s Role:** The sponsor had no role in the design or management of the study, or the preparation of this manuscript.

**REFERENCES**


Figure 1.
(a) Speaking fundamental frequency changes over a lifetime: mode, mean, and median. (b) Average (diamonds) number of words per breath group and standard deviation (squares) of words per breath group, as counted by a reviewer. Solid filled symbols represent a second reviewer for rater reliability testing.