Variation in speech and language variables based on demographic factors

Jessica Robin¹, Mengdan Xu¹, William Simpson^{1,2}

Winterlight Labs, Toronto, ON, Canada, (2) Department of Psychiatry and Behavioural Neuroscience, McMaster University, Hamilton, ON, Canada

Background

Speech is a promising modality for developing digital biomarkers for psychiatric and neurological disorders.¹ Variations in speech and language have been shown to occur in a broad spectrum of neurological and psychiatric indications, making these measures potentially useful for detecting and monitoring disease.^{2,3,4} To use speech assessments to accurately measure disease, however, variations in speech and language based on demographic factors including age, sex and education must be understood and accounted for. In this study, we determine the relationship between demographic factors and speech variables based on a normative dataset of older adults, in order to understand variations in these measures that occur independently from disease-related changes.



Figure 2: Example speech variables that differ by sex and age



Figure 3: Example Speech variables that differ by education level



Results

- A small number of the >500 acoustic and linguistic variables showed significant associations with sex.
- Variables that differed by sex included acoustic variables such as fundamental frequency, and the use of possessive pronouns and use of the filled pause "uh".
- Similarly, a small number of variables had significant associations with the age of the speaker.
- All variables with significant associations with age were acoustic, including variance and skewness of Mel-frequency cepstral coefficients (MFCCs) and the variance in intensity of the recording, suggesting that older participants have higher variance in their vocal characteristics.
- A higher proportion of speech variables had significant correlations with the years of education of the speaker.
- Variables associated with years of education included the duration of speech, length of pauses, length of utterances and the coherence and graph organization of language.

Conclusions

These results demonstrate that speech and language patterns largely have minimal associations with the age and sex of the speaker in this normative sample, with a few exceptions. There is acoustic variation in speech based on the sex of the speaker reflecting different vocal pitches, and usage of certain word categories (i.e. possessive pronouns, filled pauses) differs by sex. Similarly, there is some variation in acoustic features, including the variance in intensity of the voice, that is associated with the age of the speaker. In contrast, a number of acoustic and linguistic properties of language were associated with the number of years of formal education of the speaker. For example, our results indicate that speakers with more years of education tend to speak longer, with shorter pauses and their speech is more coherent.

Together, these findings indicate that age, sex and especially vears of education should be controlled for when analyzing speech and language patterns. These findings have implications for the statistical analysis of novel speech-based biomarkers as exploratory endpoints in clinical trials and explain the variation that might occur in such measures based on the demographic variability of a population.

References

- Robin, J. et al. Evaluation of Speech-Based Digital Biomarkers: Review and Recommendations. Digit Biomark 99–108 (2020) doi:10.1159/000510820. Szatloczki, G., Hoffmann, I., Vincze, V., Kalman, J. & Pakaski, M. Speaking in Alzheimer's (2)
- Satelloca, G., Boundardi, L., Vintze, V., Animahi, J., & Padaski, M., Speaking in Arzleitine's Disease Front. Ading Neurosci, 7 (2015). Disease Front. Ading Neurosci, 7 (2015). Poole, M. L. Brodmann, A., Darby, D. & Vogel, A. P. Motor Speech Phenotypes of Frontotemporal Dementia, Primary Progressive Aphasia, and Progressive Apraxia of Speech. J Speech Lang Hear Res 60, 897-911 (2017). Low, D. M., Bentely, K. H. & Ghosh, S. S. Automated assessment of psychiatric disorders using (3)
- speech: A systematic review. Laryngoscope Investigative Otolaryngology 5, 96-116 (2020).



Methods

- Speech recordings were collected from 164 community-dwelling study volunteers (57 M, 107 F, mean age = 70, range = 50-95, mean years of education = 15, range = 6-26) who were enrolled in a normative data collection study.
- Speech samples were elicited by an open-ended picture description task.
- Picture descriptions were recorded and analyzed using natural language processing tools, generating >500 variables per recording measuring the different acoustic and linguistic characteristics of speech.
- To determine the relationship between demographic factors and speech variables, non-parametric group comparisons based on sex were made using Mann-Whitney U tests, and correlations between age and years of education and speech variables were tested using Spearman rank-order correlations.

Winterlight Speech Analysis Pipeline



WINTERLIGHT



valu

ch