Validity of Automatic Acoustic Workflow

Justine Heider, supervised by Carrie Rountrey, Ph.D., CCC-SLP
University of Cincinnati
Department of Communication Sciences and Disorders
Cincinnati Intelligibility in Neurogenic Communication Impairments-lab (CINCI-lab)

CINCI-lab

Abstract

Speech intelligibility is a perceptual measure, and although it is currently the gold standard in clinical practice, it is subject to bias and reliability issues. Acoustic measures, by comparison, are quantifiable and correlated with traditional perceptual intelligibility outcomes. Speech intelligibility is also correlated with quality of life and disease progression in Parkinson's Disease (PD). This project aims to compare calculated acoustic values via automatic acoustic workflow to hand calculated values for validity. The acoustic workflow is the result of an interdisciplinary collaboration between computer science students and the CINCI-lab, targeting four acoustic measurements from elicited and spontaneous speech samples taken from healthy volunteers and persons with PD. The analyzed measurements include: F0 interquartile range, F2 slope, vowel duration, and vowel articulation index. Each measurement was chosen based on its significant correlation to elicited speech intelligibility in previous research. From a larger dataset, a random sample was selected for hand analysis via PRAAT and compared to the automatic acoustic workflow. Assuming coding was run correctly and considering human error, we predict hand calculations will be comparable to the automatic measurement, lending validity to the workflow. This is a crucial first step for translation to clinical practice. If validated, the workflow could be used to automatically measure acoustic values in a variety of speech samples, including both elicited and spontaneous speech. A valid workflow that measures acoustic correlates to functional intelligibility may provide quantitative outcome measures related to quality of life and ascertain reliable disease progression benchmarks for PD patients.

Introduction

Intelligibility has been shown to be correlated to several quantifiable acoustic markers in elicited speech, such as **vowel articulation index** [1], **F2 slope** [2], and **interquartile range of fundamental frequency** [3]. However, using elicited speech to determine intelligibility based on these variables presents limited ecological validity. Using spontaneous speech samples would increase the ecological validity of functional intelligibility in Parkinson's Disease (PD) and correspond more closely to the way that speech is naturally used patient's lives. The current study aims to validate an automatic computed calculation of VAI, F2 slope, and F0-IQR against hand calculations of these variables, in spontaneous and elicited speech samples of PD patients and healthy controls. A sample proportion of the spontaneous speech samples was run through the PRAAT for acoustic analysis and the three variables were calculated by hand. These results were then compared to the results given by the automatic computed calculation to determine if computed calculation is valid and reliable.

Spontaneous speech samples are necessary for determining functional intelligibility. Acoustic analysis of spontaneous speech lends to detection of health-related quality of life, reliable disease progression benchmarks, and quantifiable outcomes measures that relate more closely to speech in natural contexts. However, lack of automaticity limits clinical usage of acoustic measures for spontaneous speech. This project seeks to bridge the gap between clinical calculations by hand and automatic analysis by computers. If validated, this will be an important first step to efficient clinical application of acoustic measures for PD.

Methods and Materials

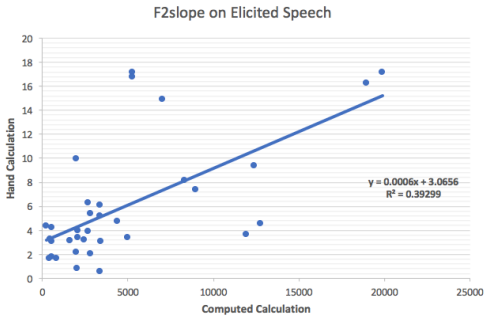
Random Number Generator: Patient speech samples were randomly selected from a previous dataset.
Dataset: Dataset was previously collected and identified as a part of a dissertation project [4]. Each patient had about 10 sentences in each task that were recorded on a LENA device, which meets Nyquist Theorem Standards .
Inclusion/Exclusion Criteria: R-colored vowels were excluded from any analysis. If alignment was more than 10ms off, the data was excluded (n=1).
VAI: The averages of F1 and F2 were calculated in excel for each vowel. Then, using the formula: VAI = (F2/i/+F1/a/)/(F1/i/ +F1/u/ +F2/u/ +F2/a/), VAI was calculated for each task. [1]
F2 slope: The second formant of diphthong vowels (/aɪ/, /eɪ/, /əʊ/, /aʊ/, /ɔɪ/) were measured as well as time elapsed. The formula ΔHz/ms was used to find the F2 slope of each diphthong vowel in each sentence. [2]
F0IQR: Interquartile range of each sentence was calculated via the range in Hz from the 25th percentile to the 75th percentile. [3]
Computed values: A workflow was written using a virtual machine, python, R, and calling on PRAAT to replicate the above acoustic calculations
Comparisons: Individual correlations were run with regression equations fitted to the data. R^2 values were considered significant if >0.35, in accordance with common practice in social social sciences. [5]

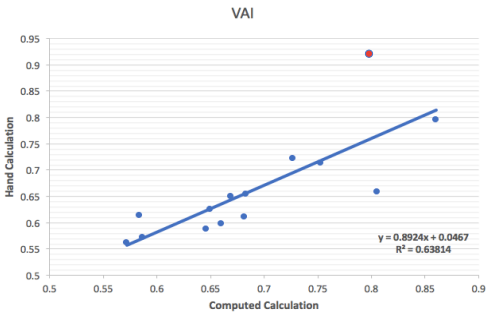
Results

Individual correlations on hand versus computer calculations were as follows: VAI* (elicited and spontaneous together), F2 slope (narrow, wide, and elicited* and spontaneous separately), and F0IQR (elicited and spontaneous separately, healthy controls, and PD patients).

*The following R^2 values were significant at the R^2 >0.35 level: VAI (R^2= 0.6381) and F2 slope (narrow) elicited (R^2= 0.393).

These significant R^2 values signify that there is a meaningful linear correlation between the hand and computer calculations for these two variables.
The remaining individual correlations had R^2 values <0.35.





is considered an outlier. Upon inspection, this was the least impaired participant, with near-normal intelligibility, so it is to be expected that this participant's VAI would be the highest. When the outlier is removed, R^2=0.7552, indicating increased correlation between hand and computer calculations with more typical parkinsonian speech patterns.

Discussion

Acoustic measures that rely on the midpoint or nucleus of a vowel, such as VAI, may be more highly correlated between hand and computed calculations. Because of the strength of the vowel at the midpoint of measurement, there is little discrepancy between the human and computer calculations. Whereas, when acoustic measures rely on the range of the vowel (F2 slope and F0-IQR), there is greater variation in alignment for the computed output and the and the human eye/ear. This may cause greater discrepancies between the two calculations. Therefore, measurements that require the length of a vowel may be less predictive of human measurements than measurements that rely only on the midpoint of a vowel.

Future Directions

Because of the strong linear correlation of VAI hand and computer calculations, CINCI-lab can confidently use this automatic acoustic analysis for future VAI studies, which is useful for predicting disease progression in PD patients and for tracking outcomes measures in a truly quantifiable fashion .

The computed acoustic workflow needs to be investigated further for alignment issues for F2 slope and F0IQR.

CINCI-lab also wishes to use this automatic acoustic analysis for vowel duration and is currently writing code for this purpose.

Contact

Justine Heider
University of Cincinnati, Undergraduate Scholarly Showcase
Email: heiderjv@mail.uc.edu

References

1. Skodda S, Grönheit W, Schlegel U (2012) Impairment of Vowel Articulation as a Possible Marker of Disease Progression in Parkinson's Disease. PLoS ONE 7(2): e32132. https://doi.org/10.1371/journal.pone.0032132
2. Kim, Yunjung, Raymond D. Kent, and Gary Weismer. "An acoustic study of the relationships among neurologic disease, dysarthria type, and severity of dysarthria." *Journal of speech, language, and hearing research* (2011).
3. Bowen, Leah K., et al. "Effects of Parkinson's disease on fundamental frequency variability in running speech." *Journal of medical speech-language pathology* 21.3 (2013): 235.
4. Rountrey, C. E. (2015). Speech Function in Persons with Parkinson's Disease: Effects of Environment, Task and Treatment.