Forecasting of change in speech function using remote assessment of acoustics

Submitter Leif Simmatis

Affiliation University of Toronto

SUBMISSION DETAILS

I agree to provide poster pdf for attendee download. Yes

Methodological Issue Being Addressed This work presents a novel analytical approach to forecasting changes in speech function, which has applications to clinical trials and clinical management for neurological diseases (e.g., amyotrophic lateral sclerosis).

Introduction Background: Amyotrophic lateral sclerosis (ALS) is a debilitating neurodegenerative disease that presently has no cure. Bulbar symptoms cause accelerated disease progression, and they affect over 90% of patients over the disease course. Speech is the most well-established indicator of bulbar disease in ALS, and so there has been a considerable recent push to develop automated speech assessments to capture the early signs of bulbar disease onset, its severity and progression. A critical next step is to identify predictors of bulbar decline as well as evaluate the use of methods to reduce acoustic predictor dimensionality, all of which will be invaluable for developing tools to stratify patients for clinical trials and for optimizing clinical management.

Objective: To evaluate the ability of acoustic features of speech to predict clinical bulbar deterioration over 6 months and determine whether feature grouping approaches improve predictive performance.

Methods Methods: We retrospectively analyzed data from 33 individuals with varying severities of bulbar ALS. Participants were assessed twice, with an average inter-assessment interval of 192.4 ± 41.2 days. The average change in ALS Functional Rating Scale -Revised score was -3.9 ±4.0 points. The Sentence Intelligibility Test (SIT) was used to derive overall speech intelligibility (%) and create bulbar disease severity stages.
Acoustic data consisted of recordings of participants reading a standard 99-word passage (at both assessments), collected using a custom, web-based speech assessment platform. Acoustic features were extracted using an end-to-end pipeline incorporating various validated measures of speech motor control. Here, we considered 6 categories of acoustic features encompassing 33 measures in total as predictors: (A) first and second formant statistical moments; (B) first formant derivatives; (C) second formant derivatives; (D) speech rate and articulatory rate; (E) pause timing statistics; (F) vocal tract configuration (Mel-frequency cepstral coefficients/MFCCs). The prediction target was whether intelligibility changed from >85% intelligibility at first assessment to ≤85% at the second (i.e., binary outcome).
Base prediction models were random forest classifiers trained on 80% of the data and evaluated on the remaining 20% as a held-out test set. We explored three methods of prediction: (1) using individual features with ranking, (2) data-driven feature aggregation using principal components

analysis (PCA), and (3) knowledge-driven feature aggregation via feature groups A-F. Models were fitted using nested cross-validation and were evaluated using the area under the receiver operating characteristic curve (AUROC).

Results Results: Individual-feature models forecasted bulbar change with AUROC of 0.72 ± 0.06, PCA-based aggregation achieved 0.79 ± 0.09, and knowledge-based grouping yielded 0.80 ± 0.19. Selected features in the latter case were typically in categories (B) and (C) (i.e., formant derivatives) or categories (D) and (E) (i.e., timing).

Conclusion Conclusions: Our end-to-end speech analysis system was able to predict the likelihood of bulbar deterioration with good accuracy, providing a critical insight into the mechanisms behind bulbar loss of function in patients with ALS. Knowledge-driven grouping performed as well as data-driven grouping, with increased interpretability. Future studies will leverage larger sample sizes and multimodal analysis to potentially further improve upon predictive performance.

Co-Authors

Leif Simmatis[1], Yana Yunusova[1]

[1] University of Toronto

Keywords

Keywords
Speech
Machine learning
Prediction

Guidelines I have read and understand the Poster Guidelines

Disclosures N/A

Related Tables and Supporting Materials <blank>