Background & Objective

- Amyotrophic lateral sclerosis (ALS) is a progressive neurodegenerative disorder that affects motor neurons
- People with ALS (pALS) experience speech impairment
- Vowel features can characterize speech motor control [1]
- Vowel space area (VSA) is a commonly utilized measure [2,3]
- Formant centralization ratio (FCR) is less sensitive to inter-speaker variability compared to VSA [4]
- Previous studies are based on recordings in clean lab conditions using specific recording hardware
- To our knowledge, this is the first analysis of vowel characteristics for identifying dysarthric speech that is based on remotely recorded speech with diverse consumer-grade hardware

Objective: Assess the feasibility and usefulness of automatically measured vowel characteristics – based on remotely recorded speech samples in varying conditions – for progress monitoring in ALS

Methods and Materials

- 72 pALS (33 females, mean age 62.7, SD: 8.6 years) participated in a self-administered speech assessment using a web-based multimodal dialong system [5] (561 recordings sessions total)
- A virtual guide, Tina, guided participants through a set of speech exercises
- Stimuli used for this study: consonant-vowel-consonant (CVC) words 'peep' (pi:p), 'peep' (ps:p), 'peep' (pu:p), and 'pop' (po:p)
- The bilabial /p/ was used at onset and coda to reduce co-articulation effects
- A virtual guide, Tina, guided participants through a set of speech exercises
- Stimuli used for this study: 72 pALS

Statistics

- Stratification: following [6], samples were grouped into two cohorts based on self-reported ALSFRS-R [7] bulbar sub score (ranges from 0 to 12)
  - Bulbar symptomatic (BUL): 273 sessions
  - Bulbar pre-symptomatic (PRE): 288 sessions
- At baseline visit, 31 participants in PRE, 41 participants in BUL cohort
- We ran non-parametric Kruskal-Wallis tests to identify statistically significant differences in vowel characteristics between the cohorts
- Receiver operating characteristics (ROC) analysis was performed, both with static feature values and with features' rate of change over time, to investigate usefulness for monitoring progression in ALS

Findings

- F2/F1 ratio on average higher in BUL group
- Formant centralization ratio (FCR) on average higher in BUL group
- Stratifiction: following [6], Bulbar symptomatic (BUL) vs. Bulbar pre-symptomatic (PRE)
- ROC analysis (discriminate between the two cohorts): best results in terms of area under the curve (AUC) for FCR and F2/F2u ratio (AUC=0.68 for both)
- In the longitudinal analysis, the change rate of the VSA yielded the best result (AUC=0.70)

Conclusions

- We investigated the clinical utility of automatically extracted vowel features for progress monitoring in ALS
- Findings indicate that vowel features reflect differences in speech motor control between bulbar symptomatic and pre-symptomatic people with ALS
- In this dataset, differences are more pronounced in female participants
- Observations underscore the utility of remotely collected speech samples recorded with consumer-grade hardware, using an interactive dialog system for naturalistic speech elicitation

References


Table 1: Vowel features based on first two formants F1 and F2

<table>
<thead>
<tr>
<th>Feature</th>
<th>Formula</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vowel space area (VSA)</td>
<td>(F2u - F1u) / (F2i - F1i)</td>
<td>Area of quadrilateral formed by four corner vowels on F1-F2 space, in Hz; calculated with the shoelace formula</td>
</tr>
<tr>
<td>Formant centralization ratio (FCR) F2i/F2u ratio</td>
<td>F2i / F2u</td>
<td>Vowel centralization measure that reduces sensitivity to inter-speaker variability</td>
</tr>
<tr>
<td>F2i/F2u</td>
<td></td>
<td>Sensitive to changes in articulatory movements (anterior-posterior tongue movement and lip rounding)</td>
</tr>
</tbody>
</table>

Table 2: Feature values based on first two formants F1 and F2

<table>
<thead>
<tr>
<th>Feature</th>
<th>Formula</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>VSA (Hz²)</td>
<td>(F2u - F1u)² / (F2i - F1i)²</td>
<td>Area of quadrilateral formed by four corner vowels on F1-F2 space, in Hz²; calculated with the shoelace formula</td>
</tr>
</tbody>
</table>

Fig. 1: Schematic diagram of the Modality.AI dialog platform.

Fig. 2: Distribution of the three vowel features in the BUL and PRE cohorts.

Fig. 3: Bulbar symptomatic pALS exhibit on average a smaller vowel space area (red lines) than pre-symptomatic pALS (blue lines). Quadrilaterals are drawn based on the centroids of each corner vowel for each cohort.

This work was funded by the National Institutes of Health grant R42DC019877. We thank all study participants for their time and we gratefully acknowledge the contribution of the Peter Cohen Foundation and EverythingALS towards participant recruitment and data collection.