We selected 2174 read speech (SIT) utterances of varying lengths spoken by 40 distinct participants (comprising both PALS and healthy controls) from a large corpus of speech and video data collected remotely via an interactive dialog agent. The data was divided into four cohorts by diagnosis, responses to the Goodness of Pronunciation (GoP) question; otherwise they were considered to be pre-symptomatic for ALSFRS-R surveys, and age. BULBAR were those patients with a diagnosis who scored < 12 on the 18th Conference of the International Speech Communication Assoc.

### Methods and Materials

- We selected 2174 read speech (SIT) utterances of varying lengths spoken by 40 distinct participants (comprising both PALS and healthy controls) from a large corpus of speech and video data collected remotely via an interactive dialog agent.
- The data was divided into four cohorts by diagnosis, responses to the Goodness of Pronunciation (GoP) question; otherwise they were considered to be pre-symptomatic for ALSFRS-R surveys, and age.
- BULBAR were those patients with a diagnosis who scored < 12 on the 18th Conference of the International Speech Communication Assoc.

### Results

<table>
<thead>
<tr>
<th>Feature Set</th>
<th>GoP+CTA</th>
<th>GoP + PPT</th>
<th>CTA + PPT</th>
<th>CTA</th>
<th>PPT</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTA</td>
<td>-0.4722</td>
<td>-0.2923</td>
<td>-0.4395</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GoP</td>
<td>-0.31</td>
<td>0.455</td>
<td>0.431</td>
<td>0.17</td>
<td>0.42</td>
</tr>
<tr>
<td>PPT</td>
<td>0.66</td>
<td>0.455</td>
<td>0.431</td>
<td>0.17</td>
<td>0.42</td>
</tr>
</tbody>
</table>

Table 1. Intra-metric Correlations. Pearson’s r statistic. All correlations significant at p ≤ 0.00001. Low correlations suggest that the metrics are not redundant. CTA correlates most strongly with both GoP and PPT.

The figures above show little relationship between PPT and listener effort, though there does seem to be a relationship with respect to listener effort and GoP and CTA. With respect to GoP, though, this relationship only seems to hold and the very extreme end of the scale when the speech is unintelligible (listener effort = 100%). CTA seems a bit more robust in that listener effort scores of 75% and above correspond to lower CTA scores. Both CTA and GoP displayed moderate to high correlations with human listener effort.

### Conclusions

- Highest correlation to listener effort was achieved when combining all three metrics.
- CTA is as, if not more, informative than GoP and PPT in distinguishing controls from bulbar pre-symptomatic and bulbar symptomatic ALS patients in our cohort.
- Both CTA and GoP displayed moderate to high correlations with human listener effort and low correlation with each other. This potentially highlights the relative importance of timing over spectral information in characterizing ALS pathological speech.

### References