Lexico-semantic differences between people with PD and healthy controls observed in a story retell task
Hardik Kothare1, Andrew Exner2, Sandy Snyder2, Jessica E. Huber1 and Vikram Ramanarayanan1,3
1Modality.AI, Inc., San Francisco, CA, USA
2Purdue University, West Lafayette, IN, USA
3University of California, San Francisco, San Francisco, CA, USA

Conclusions
● People with Parkinson’s Disease (pPD), both with and without dementia, exhibit:
  ○ Word retrieval deficits
  ○ Delayed lexical activation
  ○ Verbal working memory deficits
  ○ Impaired lexico-semantic comprehension
  ○ Reduced grammatical ability and information content with progression
● Research Question: Do lexico-semantic (language) metrics extracted from a remote patient monitoring platform have the potential to capture Parkinsonian characteristics and cognitive decline relative to healthy controls?

Methods and Materials
● 39 pPD and 23 controls (see Table 1) completed 4 interactive sessions, a week apart from each other, between November 2020 and January 2022 using a cloud-based multimodal dialogue platform (Illustration in Figure 1).

● Participants were guided by Tina, a virtual agent, through a battery of tasks eliciting speech and facial behaviours: sustained vowel phonation, reading sentences and a passage, story retell, spontaneous speech on a topic of their choice, etc.

● In this work, we focus on a story retell task depicting three different stories of a boy, Bobby, and his mother (Figure 2).

● All pictures in the story were first described by Tina. Participants were then shown the pictures one more time and were asked to describe what was happening in the picture in their own words.

● Thus, the task probed the speech production network along with verbal working memory. Participants were presented with only one of the three stories during every session.

● Lexico-semantic features (Table 2) were computed from transcripts (derived through AWS Transcribe) of participant utterances using the spaCy software library (https://spacy.io/).

● Kruskal-Wallis tests were run to look for differences in the median values of these metrics between pPD and controls.

<table>
<thead>
<tr>
<th>Number of participants</th>
<th>Mean age ± standard deviation (years)</th>
<th>Median MoCA score (Q1-Q3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pPD</td>
<td>39 (19 female)</td>
<td>68.09 ± 8.83</td>
</tr>
<tr>
<td>Controls</td>
<td>23 (18 female)</td>
<td>63.82 ± 10.15</td>
</tr>
</tbody>
</table>

Table 1: Demographics; MoCA = Montreal Cognitive Assessment, Q1 = first quartile, Q3 = third quartile.

Results and Discussion
● pPD example: “They are… they have put their boots in the sun to dry.”
● Control example: “Bobby and Mom’s boots are drying on the porch.”
● pPD had a higher pronoun rate and lower noun-to-pronoun ratio, noun rate and content words percentage.
● Noun-to-pronoun ratio was also slightly positively correlated with pPD MoCA scores (Spearman’s rho = 0.21, p = 0.0158)

Conclusions
● The results indicate that as compared to controls, pPD used ‘he’ and ‘they’ more than ‘Bobby’ and ‘Bobby and his Mom’. More cognitively impaired pPD used more pronouns.
● A lower percentage of content words is in line with prior work on decreased conciseness of discourse in pPD.
● Our results suggest that lexico-semantic metrics extracted using a remote monitoring platform have the potential to capture Parkinsonian characteristics relative to healthy controls and index cognitive decline.