Development of Clinician-Friendly Discourse Analysis Tools: Main Concept Analysis

Sarah Grace Hudspeth, Sarah Campbell, Nicole Williams, Emily Dillow, Jessica D. Richardson
Neuroscience of Rehabilitation Laboratory, Department of Communication Sciences and Disorders
University of South Carolina, USA

Background
• Discourse is a highly individualized and complex speech act essential for effective communication in daily life.
• Individuals with language disorders demonstrate impaired narrative ability and a resultant decline in functional communication.1
• While current discourse measures have standardized instructions, norm-referenced discourse measures are lacking, making it difficult for clinicians to quantify how differently a client is performing compared to his/her peers.
• Main Concept Analysis has been established as a reliable and valid method of assessing narrative adequacy.2
• In previous main concept research, investigators determined which main concepts they considered to be central to successful retell of a story or picture sequence. No studies have to date generated lists based on narratives of control speakers.
• The authors’ experience while completing the study, as well as peer-reviewed research, has suggested potential age-related differences in narrative quality,3 warranting further investigation. In addition to age-related differences, we also examined differences in sex and years of education.
• Standardized main concept lists for three narrative tasks involving different types of discourse may allow clinicians to quickly gain insight into patients’ functional communication.

Methods
• Relevant concepts are defined as correct utterances about the story that contained a subject, one main verb, and object (if appropriate).
• May contain subordinate clauses, but must contain ONLY ONE MAIN VERB.
• A list of relevant concepts was developed for each age group for the three stories.
• All concepts were simplified to S-V-O form for ease of comparison across participants.
• Each participant’s transcript was analyzed. Every RC identified resulted in an additional point for the participant and for the concept. As new concepts were identified, they were added to the list.
• Age-related differences, a median test was conducted by age group for each story on the number of RCs produced by each participant.

Main Concepts (MC)
• Relevant concepts spoken by 50% or more of individuals in an age group were considered as a Main Concept.
• After identifying Main Concepts, all normal language transcripts were scored according to the guidelines published by Nicholas and Brooks.2 The possible scores are:
  - Absent (A): The participant did not produce any portion of the MC.
  - Inaccurate/Incomplete (I): The participant attempted to produce a portion of the MC, but it was missing at least one essential element and another essential element was incorrect.
  - Inaccurate/Complete (IC): The participant produced a complete MC, but at least one essential element was incorrect.
  - Accurate/Incomplete (AI): The participant produced an accurate MC, but at least one essential element was incorrect.
  - Accurate/Complete (AC): The participant produced all essential elements, and all essential elements were correct.
• After scoring was completed, scores were converted to a numeric composite score with AB = 0, BI = 1, IC = 2, AI = 2, AC = 3
• To examine potential differences by sex and years of education, proportion scores were calculated using the composite scores for each age group.

Results

Relevant Concepts:
• The total number of RCs spoken for each age group for each story are below:
  - Cinderella: 309, 387, 585, and 237 from Young to Oldest.
  - Breaking Window: 32, 33, 45, and 32 from Young to Oldest.
  - PB&J: 35, 58, 199, and 59 from Young to Oldest.
• A median test was performed to investigate differences in age groups.
• Results revealed a significant difference between the Middle and Old and Young and Oldest age groups for Cinderella story only (p = .001 and p = .8414, 1 - .004, respectively)(See Fig. 1).
• Main Concepts:
  - Cinderella and PB&J stories had different MC lists across all ages.
  - Breaking Window story had a different MC list for the Oldest group only.
  - A median test was completed using the proportion scores for sex and education.
  - No significant differences by sex were found only for Cinderella (p = .4802, p = .028).
  - No differences were found for any story based on years of education (p = .764, .259, and .234, respectively).

Table 1: Proportion of Nicholas and Brookshire score by age for all stories

<table>
<thead>
<tr>
<th>Age Groups</th>
<th>PB&amp;J</th>
<th>CINDERELLA</th>
<th>BREAKING WINDOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>0</td>
<td>0.20</td>
<td>0.23</td>
</tr>
<tr>
<td>BI</td>
<td>0</td>
<td>0.39</td>
<td>0.67</td>
</tr>
<tr>
<td>AI</td>
<td>1</td>
<td>0.25</td>
<td>0.37</td>
</tr>
<tr>
<td>AC</td>
<td>2</td>
<td>0.42</td>
<td>0.61</td>
</tr>
<tr>
<td>O</td>
<td>3</td>
<td>0.56</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Table 2: Proportion of each Nicholas and Brookshire score by age for all stories

<table>
<thead>
<tr>
<th>Age Groups</th>
<th>PB&amp;J</th>
<th>CINDERELLA</th>
<th>BREAKING WINDOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>0</td>
<td>0.20</td>
<td>0.23</td>
</tr>
<tr>
<td>BI</td>
<td>0</td>
<td>0.39</td>
<td>0.67</td>
</tr>
<tr>
<td>AI</td>
<td>1</td>
<td>0.25</td>
<td>0.37</td>
</tr>
<tr>
<td>AC</td>
<td>2</td>
<td>0.42</td>
<td>0.61</td>
</tr>
<tr>
<td>O</td>
<td>3</td>
<td>0.56</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Table 3: Proportion of Nicholas and Brookshire score by age for all stories

<table>
<thead>
<tr>
<th>Age Groups</th>
<th>PB&amp;J</th>
<th>CINDERELLA</th>
<th>BREAKING WINDOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>0</td>
<td>0.20</td>
<td>0.23</td>
</tr>
<tr>
<td>BI</td>
<td>0</td>
<td>0.39</td>
<td>0.67</td>
</tr>
<tr>
<td>AI</td>
<td>1</td>
<td>0.25</td>
<td>0.37</td>
</tr>
<tr>
<td>AC</td>
<td>2</td>
<td>0.42</td>
<td>0.61</td>
</tr>
<tr>
<td>O</td>
<td>3</td>
<td>0.56</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Results (Cont.)

- The data in Table 2 show a general trend for a proportionally more scores of AC received by the younger age groups, which steadily decreases as age increases. The reverse trend is also true, with proportionally more scores of AB received by the older age groups, and steadily decreases as age decreases.

Discussion
- Although there were only two significant differences in the median number of relevant concepts spoken across age groups, our analysis showed that there are important age-related differences in the content produced for each task.
- The difference found between males and females in retelling the Cinderella story indicates there may be an inherent gender bias, with females having greater knowledge and familiarity with the Cinderella story.
- Future studies should explore the use of a gender neutral story to eliminate bias.
- The discourse of older individuals may be less coherent and cohesive as demonstrated by a trend of more AB scores for older age groups (although it has not been determined if there is a statistically significant difference).
- We have established a set of standardized and normed results for three discourse tasks that clinicians can use to quickly evaluate a client’s speech and compare to same age peers.
- The generated MC lists could serve as a clinically useful non-transcription-based checklist for narrative assessment when Cinderella, Breaking Window, or PB&J stories are elicited according to AphasiaBank protocols.

Future Directions
- Analyze narratives of speakers with aphasia and TBI using generated MC lists.
- Correlate number of MCs produced to listener ratings of narrative adequacy to establish functional relevance.
- Develop an automated MC analysis tool within Computerized Language Analysis (CLAN).
- The potential for even more efficient discourse measurements, such as core lexicon production, will also be investigated.