Establishing the DementiaBank Protocol: Using Big Data to Understand Language Changes in Dementia

Anna K. Saylor, Matthew L. Cohen, Davida Fromm, Brian MacWhinney & Alyssa M. Lanzi

University of Delaware & Carnegie Mellon University

Introduction

- By 2050, it is expected that 12.7 million Americans over 65 years will be living with Alzheimer’s disease and related dementias (ADRDs).¹
- To support this public health crisis, research has prioritized prevention and early detection at the mild cognitive impairment (MCI) stage.²
- Although AD is primarily characterized by impairment in episodic memory,³ language abilities are often impaired and may precede the decline of other cognitive abilities.⁴
- Spoken language measurement may help researchers better understand the progression of AD and to detect early decline.⁵,⁶
- To study the progression of language, open access databases are needed. TalkBank is an open access database for transcribed multimedia data from spoken language interactions.⁷
- Our goal is to expand DementiaBank a clinical bank within TalkBank.

The overall goal of this work is to:
(1) Describe the new DementiaBank protocol
(2) Describe the Delaware Corpus data
(3) Illustrate types of analyses using CLAN and additional resources in DementiaBank/TalkBank

Delaware Corpus

- To date, participants (n=53) were recruited from previous studies at the University of Delaware.
- 29 neurotypical participants
- 33 MCI participants
- Participants completed one session lasting ~90 minutes via Zoom.
- Participants were classified based on the National Institute of Aging-Alzheimer’s Association criteria⁸ as determined by a neuropsychologist.
- To be classified as MCI, participants had to meet all four criteria specified below:
  1. Concern regarding change in cognition
  2. Impairment in 1+ cognitive domains through objective assessment
  3. Preservation of independence in functional abilities
  4. Not demented
- To be classified as neurotypical, participants had to meet criteria 3 & 4, and produce scores within normal limits on cognitive assessments.

Table 1. Delaware Corpus Demographics

<table>
<thead>
<tr>
<th>Sex</th>
<th>Neurotypical (n=20)</th>
<th>MCI (n=33)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(SD; Range)</td>
<td>(SD; Range)</td>
</tr>
<tr>
<td>Male</td>
<td>5, 16</td>
<td>7, 12</td>
</tr>
<tr>
<td>Female</td>
<td>15, 17</td>
<td>8, 11</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black/African American</td>
<td>0, 3</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>20, 30</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High School/Equivalent</td>
<td>1, 1</td>
<td></td>
</tr>
<tr>
<td>Some College</td>
<td>2, 5</td>
<td></td>
</tr>
<tr>
<td>Associates/Technical</td>
<td>1, 3</td>
<td></td>
</tr>
<tr>
<td>Bachelors or Higher</td>
<td>16, 27</td>
<td></td>
</tr>
</tbody>
</table>

Example Analyses

- Cinderella Story:
  Core Lexicon (CoreLex)
  - CoreLex is a measure used to determine specific lexical items used to tell a story and can also be compared to norms.¹⁷
  - Below, we use the CoreLex analysis to:
    - Compare MCI and neurotypical group performance from the Delaware corpus
    - Compare Delaware corpus groups to a larger group of neurotypical adults established by Dalton et al., 2020¹⁸

DementiaBank Protocol

Discourse Protocol
- Picture Description: Cookie Theft¹⁰
- Story Narrative: Cat Rescue,⁴ “Going & Coming,”¹⁰ Cinderella¹¹
- Procedural Discourse: P&BJ
  - Personal Narrative: Hometown
  - Cognitive-Linguistic Battery
    - Boston Naming Test—Short Form¹²
    - Hopkins Verbal Learning Test - Revised¹³
    - Wechsler Memory Scale—Revised: Logical Memory Subtests¹⁴
    - Montreal Cognitive Assessment¹⁵

  “All discourse protocol scripts and materials can be accessed from the DementiaBank website”

Transcription
- Audio files are transcribed into CHAT format to be analyzed using various CLAN commands.
- CHAT transcription can be completed using one of two methods: (1) manual transcription or (2) Automatic speech recognition (ASR) transcription

Figure 1. Norman Rockwell “Going and Coming” print¹⁰

Table 1. Delaware Corpus Demographics

<table>
<thead>
<tr>
<th></th>
<th>Neurotypical (n=20)</th>
<th>MCI (n=33)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Age</td>
<td>69.6 (5.5, 62-82)</td>
<td>74.8 (8.8, 61-91)</td>
</tr>
<tr>
<td>Sex</td>
<td>Male 5, 16</td>
<td>Female 15, 17</td>
</tr>
<tr>
<td>Race</td>
<td>Black/African American 0, 3</td>
<td>White 20, 30</td>
</tr>
<tr>
<td>Education</td>
<td>High School/Equivalent 1, 1</td>
<td>Some College 2, 5</td>
</tr>
<tr>
<td>CoreLex</td>
<td>Mean=74, 3</td>
<td>Mean=90.8</td>
</tr>
<tr>
<td>MCI Group</td>
<td>Mean=95.2</td>
<td>Mean=74</td>
</tr>
</tbody>
</table>

Figure 2. Illustration of CoreLex analysis between Delaware corpus groups and using Dalton et al., 2020¹⁸ norms

Figure 3. Images of DementiaBank/TalkBank educational resources

Educational Resources
- Manuals, tutorial screencasts, browsable database, collaborative commentary, and more!

Next Steps
- Expand the database to include diverse populations
- Develop more educational resources
  - Develop “Grand Rounds” tutorial to provide in-depth illustrative examples of communication challenges
  - Prepare “Classroom Activities” for students to practice language sample analysis, plan treatment goals, and compare discourse across disorders
- Refine analyses techniques to help better understand the progression of spoken language across of continuum (neurotypical, MCI, and dementia)

How Can YOU Get Involved?
- Join the TalkBank consortium (scan the QR code)
- Learn how to use the CLAN and CHAT tools for computer-based transcription, coding, and automated analyses
- Analyze data from the Delaware corpus and other DementiaBank corpora
- Contribute your own data
  - Using the DementiaBank protocol
  - Non-protocol language samples from adults across ADRDs

NIH National Institute on Aging

This work is funded by NIA/NIDCD Grant 3R01DC008524-13S1.

We would also like to thank Michelle Bourgeois & Houjun Liu for their contributions to this work.
References: