# DementiaBank Protocol: Using "Big Data" to Understand Language Changes Across the Progression of Alzheimer's Disease

Anna K. Saylor[1], Matthew L. Cohen[1], Davida Fromm[2], Brian MacWhinney[2], & Alyssa M. Lanzi[1]

[1]University of Delaware, [2]Carnegie Mellon University

## Introduction

- Alzheimer's disease (AD) is the leading cause of dementia in older adults, affecting more than 1 in 9 adults over 65 [1].

- For older adults with AD pathology, mild cognitive impairment (MCI) is a transitional phase between normal aging and dementia that describes a detectable decline in cognitive skills but not to a degree that limits independence in everyday activities [2].

- Memory impairment is the hallmark characteristic of AD, but language impairments can also manifest throughout the progression of the disease, restricting independence and quality of life for adults [3]. Thus, there is growing interest in studying language markers for AD [4].

- **The purpose of this study is to collect preliminary data for DementiaBank: a multi-site database of multimedia interactions to study spoken communication cross-sectionally across the progression of AD in older adults (neurotypical, MCI, and dementia).**

- This project builds off of the success of TalkBank, which has been the topic of over 8,000 publications in 40+ languages across 14 research areas [5].

## Procedures

- All participants are age 60+, understand English, and do not report other conditions that can cause cognitive impairment.

- DementiaBank protocol data is collected during a single ~45 minute session completed via Zoom.

- Language samples are transcribed, coded, and linked to the media file using CHAT format, a state-of-the-art transcription system.

- Using the CLAN program, transcripts can be analyzed for acoustic, lexical, and morphosyntactic variables [6].
  - e.g., type-token ratio, pause length, fluency

### Table 1. Demographics of Participants to Date

| | Neurotypical (n=28) | MCI (n=17) | Dementia (n=3) |
|---|---|---|---|
| **Age (years)** | 62-84 mean=70.6 | 63-91 mean=76.5 | 81-90 mean=86.0 |
| **Sex** | female=17 male=11 | female=10 male=7 | female=2 male=1 |
| **BNT-SF Total Score (max=15)** | 11-15 mean=14.5 | 12-15 mean=14.3 | 8-14 mean=10.7 |
| **MoCA Total Score (max=30)** | 21-30 mean=26.5 | 17-28 mean=23.5 | 8-15 mean=12.0 |

## DementiaBank Protocol

- **Discourse**
  - Picture Description: Cookie Theft, Cat Rescue, Norman Rockwell "Going & Coming"
    - e.g., "Please tell me everything that you see going on in this picture."
  - Story Narrative: Cinderella
    - "... tell me as much of the story of Cinderella as you can…"
  - Procedural Discourse: PB&J Sandwich
    - "Tell me how you would make a peanut butter & jelly sandwich."
  - Personal Narrative: Hometown
    - "Tell me about your hometown."

- **Cognitive/Linguistic Assessment Data**
  - Boston Naming Test-Short Form (BNT-SF) [7]
  - Montreal Cognitive Assessment (MoCA v.7.1) [8]

- **Descriptive Information**
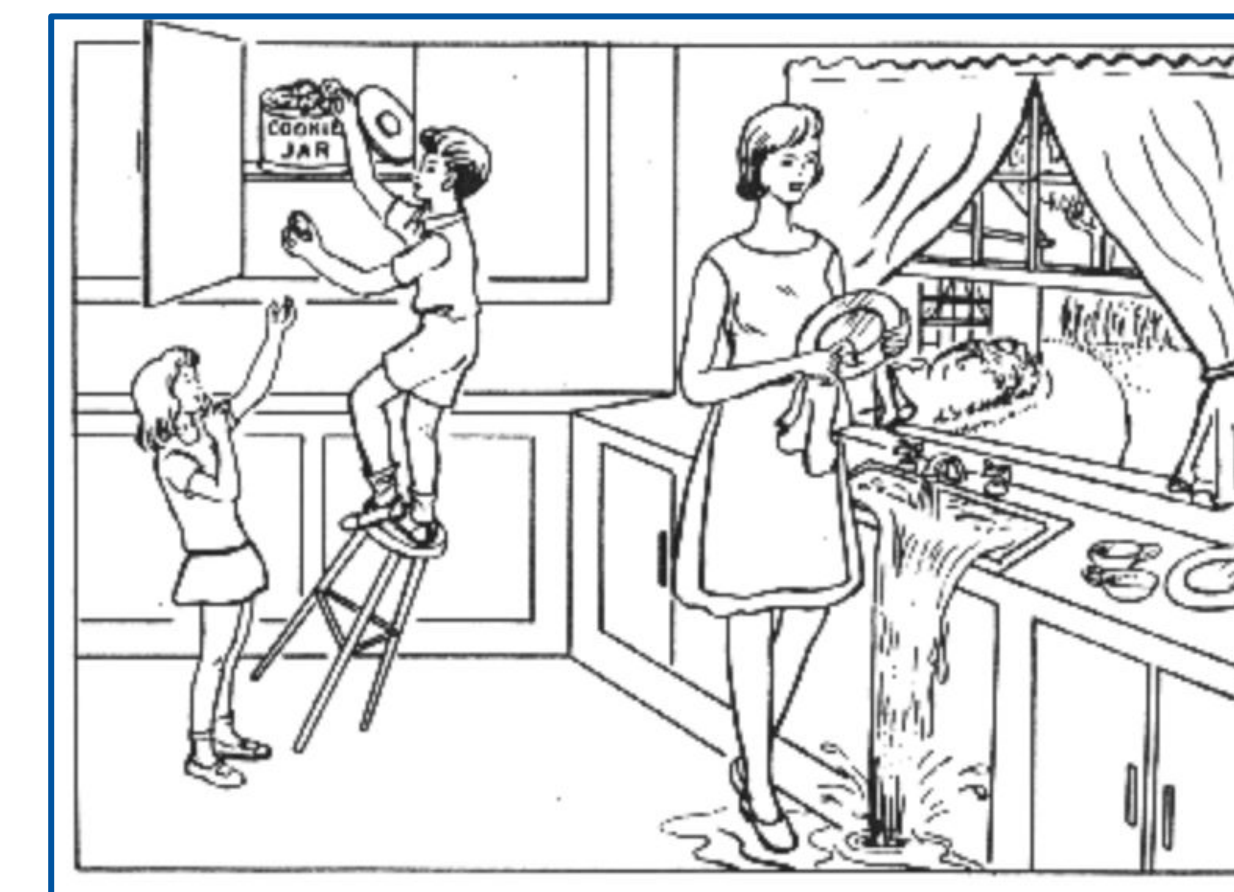  - Demographic & medical information forms



Figure 1. Cookie Theft picture used for picture description [9].



Figure 2. Cat Rescue picture used for picture description [10].

## Example Analyses: Picture Description

- Below are example analyses that can be performed using CLAN programs. We analyzed "Cat Rescue" discourse samples from two age-matched participants: a neurotypical participant and a participant with MCI.

```
*PAR:     and he can't get down out_of the tree . •
%mor:     coord|and pro:sub|he mod|can~neg|not
          v|get adv|down prep|out_of
          det:art|the n|tree .
```

Figure 3. Portion of discourse sample transcribed in CHAT format and morphosyntactically coded using CLAN.

| Neurotypical | MCI |
|---|---|
| 4 n\|cat | 9 n\|tree |
| 2 n\|dad | 6 n\|cat |
| 2 n\|department | 2 n\|department |
| 2 n\|fire | 2 n\|fire |
| 2 n\|tree | 2 n\|ladder |

Figure 4. Results from a *noun frequency analysis* of the two transcripts.
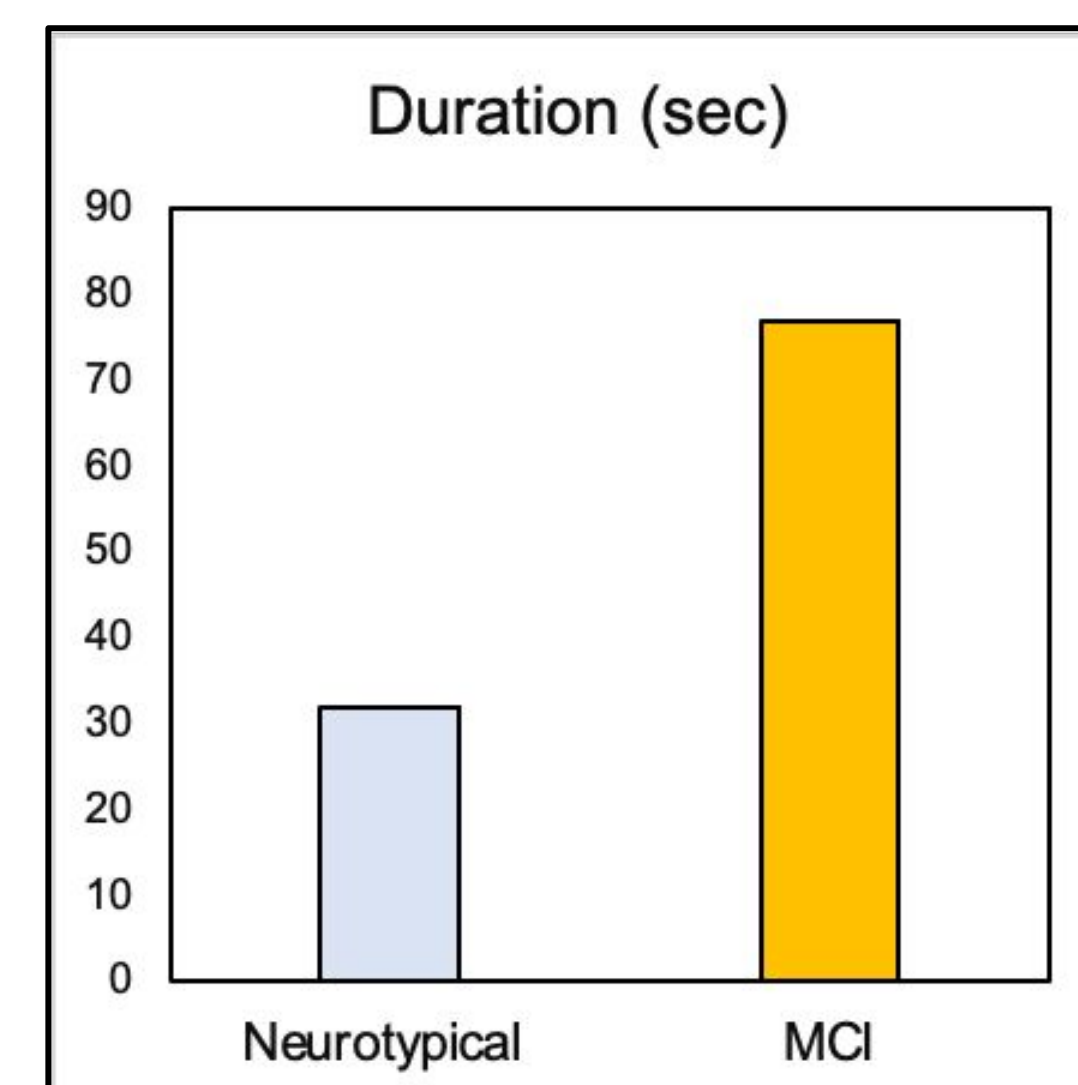


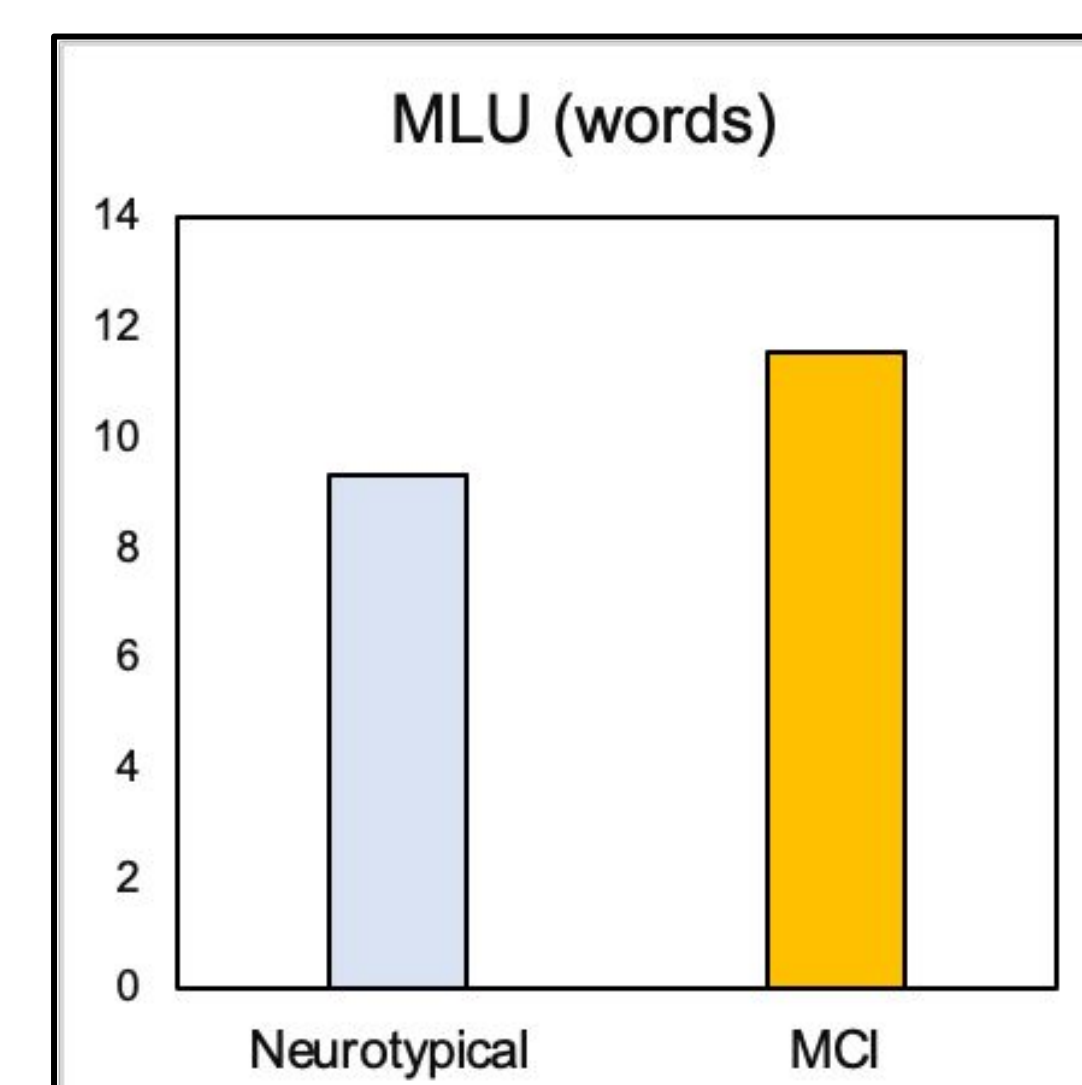Figure 5. Results of *discourse duration (sec)* of the two transcripts.



Figure 6. Results of *mean length of utterance* of the two transcripts.
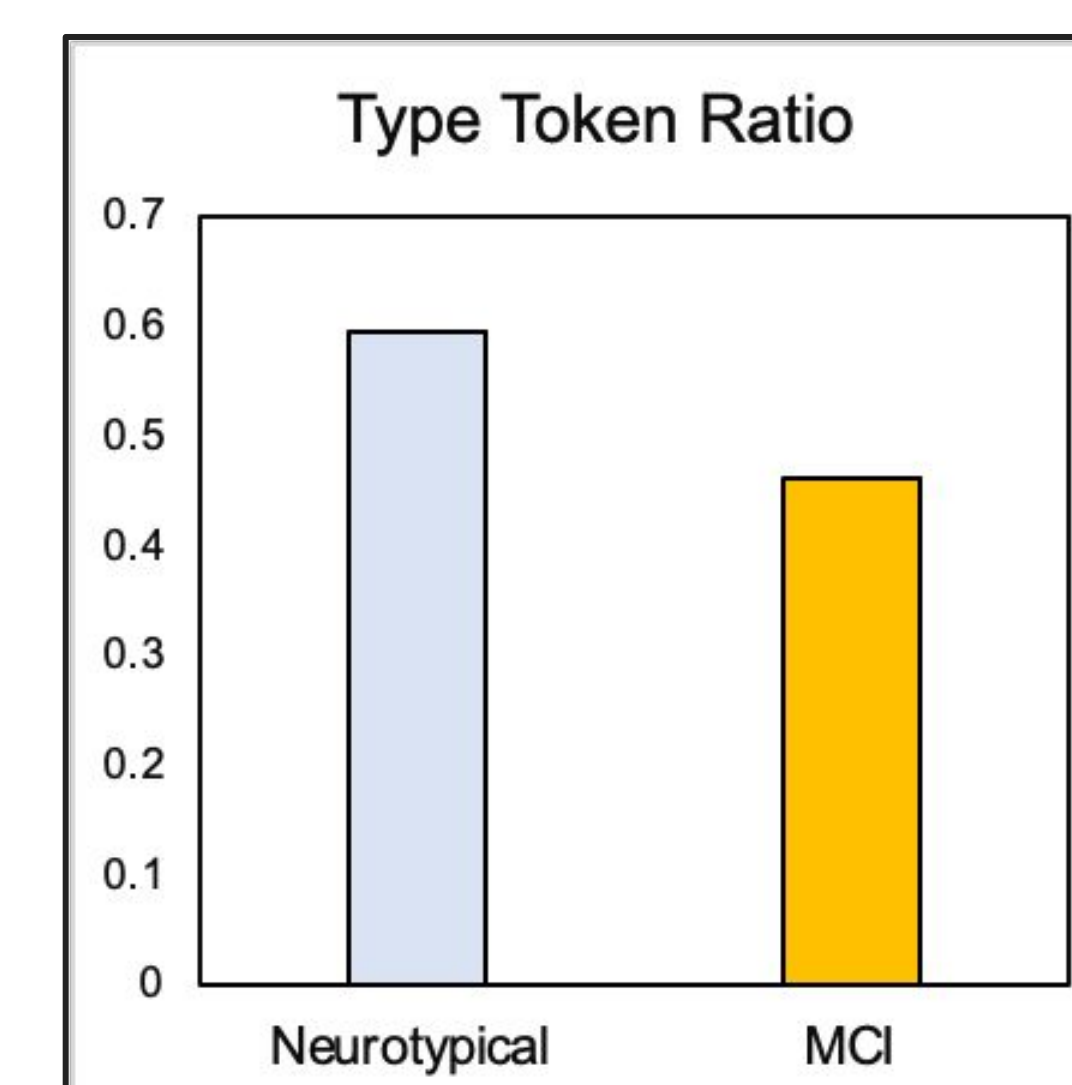


Figure 7. Results of *type token ratio* of the two transcripts.

## Planned Activities/Analyses

- **Expand database to multiple sites nationally**
  - Provide evidence for a multi-site protocol that can be used by researchers to contribute to DementiaBank.

- **Conduct large-scale discourse analyses to understand the relationships between spoken communication and AD**
  - Support early detection for MCI/dementia based on spoken discourse.
  - Help differentiate subtypes of AD (e.g., amnestic, dysexecutive, language-predominant subtypes) based on discourse.
  - Inform language treatment and support goal development based on macrostructure (e.g., main content) and microstructure (e.g., lexical diversity) level analyses.

- **Enrich teaching of AD-related communication challenges.**
  - "Grand rounds" component of the database provides in-depth illustrative examples across the progression of AD.
  - Cases allow students to practice language sample analysis and interpret clinically relevant information.

## References & Acknowledgements

[1] Alzheimer's Association. (2021). Alzheimer's disease facts and figures. *Alzheimers Dement, 17*, 327-406.

[2] American Psychiatric Association, DSM-5 Task Force. (2013). Diagnostic and Statistical Manual of Mental Disorders: DSM-5 (5th ed.). American Psychiatric Publishing. https://doi.org/10.1176/appi.books.9780890425596

[3] Woodward, M. (2013). Aspects of communication in Alzheimer's disease: clinical features and treatment options. *International psychogeriatrics, 25*(6), 877-885.

[4] Fraser, K. C., Meltzer, J. A., & Rudzicz, F. (2016). Linguistic features identify Alzheimer's disease in narrative speech. *Journal of Alzheimer's Disease, 49*(2), 407-422.

[5] MacWhinney, B. (2019). Understanding spoken language through TalkBank. *Behavior Research Methods, 51*, 1919-1927.

[6] MacWhinney, B. (2000). The CHILDES Project: Tools for Analyzing Talk. 3rd Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

[7] Kaplan, E., Goodglass, H., & Weintraub, S. (2001). Boston naming test.

[8] Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., ... & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*(4), 695-699

[9] Goodglass, H., Kaplan, E., & Barresi, B. (2001). *BDAE- 3: Boston diagnostic aphasia examination—Third edition.* Lippincott Williams & Wilkins.

[10] Nicholas, L. E., & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *Journal of Speech and Hearing Research, 36*(2), 338–350. https://doi.org/10.1044/jshr.3602.338