

Tutorial

DementiaBank: Theoretical Rationale, Protocol, and Illustrative Analyses

Alyssa M. Lanzi,^{a,b}  Anna K. Saylor,^a  Davida Fromm,^c  Houjun Liu,^d  Brian MacWhinney,^c 
and Matthew L. Cohen^{a,b,e} 

^aDepartment of Communication Sciences and Disorders, University of Delaware, Newark ^bDelaware Center for Cognitive Aging Research, University of Delaware, Newark ^cDepartment of Psychology, Carnegie Mellon University, Pittsburgh, PA ^dThe Nueva School, Hillsborough, CA ^eCenter for Health Assessment Research and Translation, University of Delaware, Newark

ARTICLE INFO

Article History:

Received September 5, 2022

Revision received November 1, 2022

Accepted November 25, 2022

Editor-in-Chief: Katherine C. Hustad

Editor: Kimberly Diggle Mueller

https://doi.org/10.1044/2022_AJSLP-22-00281

ABSTRACT

Purpose: Dementia from Alzheimer's disease (AD) is characterized primarily by a significant decline in memory abilities; however, language abilities are also commonly affected and may precede the decline of other cognitive abilities. To study the progression of language, there is a need for open-access databases that can be used to build algorithms to produce translational models sensitive enough to detect early declines in language abilities. DementiaBank is an open-access repository of transcribed video/audio data from communicative interactions from people with dementia, mild cognitive impairment (MCI), and controls. The aims of this tutorial are to (a) describe the newly established standardized DementiaBank discourse protocol, (b) describe the Delaware corpus data, and (c) provide examples of automated linguistic analyses that can be conducted with the Delaware corpus data and describe additional DementiaBank resources.

Method: The DementiaBank discourse protocol elicits four types of discourse: picture description, story narrative, procedural, and personal narrative. The Delaware corpus currently includes data from 20 neurotypical adults and 33 adults with MCI from possible AD who completed the DementiaBank discourse protocol and a cognitive-linguistic battery. Language samples were video- and audio-recorded, transcribed, coded, and uploaded to DementiaBank. The protocol materials and transcription programs can be accessed for free via the DementiaBank website.

Results: Illustrative analyses show the potential of the Delaware corpus data to help understand discourse metrics at the individual and group levels. In addition, they highlight analyses that could be used across TalkBank's other clinical banks (e.g., AphasiaBank). Information is also included on manual and automatic speech recognition transcription methods.

Conclusions: DementiaBank is a shared online database that can facilitate research efforts to address the gaps in knowledge about language changes associated with MCI and dementia from AD. Identifying early language markers could lead to improved assessment and treatment approaches for adults at risk for dementia.

By definition, individuals with dementia experience a cognitive decline from a previous level of performance in one or more cognitive domains that restricts their ability to function independently. These cognitive declines

may occur in one or more of the following cognitive domains: complex attention, executive function, learning and memory, language, perceptual-motor, or social cognition (American Psychiatric Association, 2013). Dementia due to Alzheimer's disease (AD) is primarily characterized by a significant decline in episodic memory abilities (Bäckman et al., 2001); however, there is strong evidence to suggest that speech and language abilities are also often affected and may in fact precede the decline of other cognitive skills (Bayles et al., 1993; Croisile et al., 1996; Hier

Correspondence to Alyssa M. Lanzi: lanzia@udel.edu. **Disclosure:** Brian MacWhinney created TalkBank which houses DementiaBank. Alyssa M. Lanzi is a member of the TalkBank governing board. The other authors have declared that no competing financial or nonfinancial interests existed at the time of publication.

et al., 1985; Murdoch et al., 1987; M. Nicholas et al., 1985; Price et al., 1993; Vuorinen et al., 2000). Therefore, there is growing interest in developing new and more precise digital tools to investigate the speech and language skills of older adults at risk for AD or with a possible early clinical presentation of AD, for example, at a stage known as mild cognitive impairment (MCI; Au et al., 2019; Kourtis et al., 2019; Tavabi et al., 2022).

MCI describes a clinical syndrome marked by modest cognitive decline but generally persevered functional independence. The most common cause of MCI and dementia in older adults is AD (Alzheimer's Association, 2022; Jicha et al., 2006; Schneider et al., 2007, 2009). However, it is not the only cause. Although MCI is the earliest symptomatic stage of AD, it is also a period of clinical uncertainty. Not all individuals with MCI will progress to dementia or have AD pathology, or necessarily any known pathology (Smith & Bondi, 2013). However, at least for research purposes, MCI is widely accepted as the transitional stage between age-related cognitive decline and early-stage dementia (Petersen, 2016). In the absence of biomarker evidence or evidence of non-AD causes, people with MCI are said to have met the core clinical criteria of MCI from AD (Albert et al., 2011).

MCI provides a unique opportunity for researchers and clinicians to identify those older adults at risk for dementia early in the AD continuum to support the delivery of interventions that may delay the onset of dementia or reduce the functional impact. Unfortunately, the most common assessment tools used to assess speech and language skills are generally coarse and not optimally able to make fine distinctions between older adults with normal cognition, MCI, and dementia. For example, measures of confrontational naming and word-finding (e.g., Boston Naming Test [Kaplan et al., 1983] or a verbal fluency test) have been criticized for their poor sensitivity to subtle errors and lack of ecological validity (Cuetos et al., 2007; Mueller et al., 2016). This imprecision may stem from the highly decontextualized nature of these assessments, which makes them different from communication in everyday contexts (Kavé & Goral, 2016).

More precise digital measurement of the speech and language skills of older adults with age-typical cognition, MCI, and dementia may help advance research and clinical practice by permitting more fine-grained distinctions between these classifications and perhaps subtyping MCI and dementia based on language profiles (Au et al., 2019; Gold et al., 2018; Tavabi et al., 2022). Analyses that rely on spoken language data may provide this needed precision while also providing an economically, accessible, and ecologically valid means of assessing the variability of cognitive and language abilities over time. Such data can be elicited by having the individual produce spoken or written responses to open-ended prompts such as picture

description tasks (e.g., the Cookie Theft picture description task from the Boston Diagnostic Aphasia Examination; Goodglass & Kaplan, 1983). Connected speech samples can then be transcribed, and various linguistic markers can be quantified (e.g., semantic, syntactical, fluency, and/or lexical structures). These discourse markers may be more sensitive to language changes than the traditional decontextualized neuropsychological assessments of naming or fluency, perhaps because they are sensitive to the multiple cognitive-communication processes that are required for the production of connected speech in functional tasks (Cuetos et al., 2007; Mueller et al., 2016).

Several studies that have investigated connected speech in written and spoken language samples suggest that language differences in AD can be detected well before the onset of dementia (Mueller, Kosciak, Hermann, et al., 2018; Mueller et al., 2021). Indeed, the neuropathological process of AD begins years to decades before the onset of obvious clinical symptoms (Golde et al., 2011; Jack et al., 2009; Lloret et al., 2019; Snowden et al., 1996). Garrard et al. (2005) examined three written texts spanning 41 years representing the early, middle, and end of the writing career of a renowned author with neuropathology-confirmed AD. The comparison of the texts revealed more restricted vocabulary in the late-life text, although the syntactic structure was relatively spared. Le et al. (2011) also reported differences in lexical diversity and lexical characteristics (i.e., reduced vocabulary, more repetition of fixed phrases and content words, a deficit in noun tokens, and an increase in fillers) but no consistent differences in the syntax of written samples from three novelists: one who died of AD, one who was suspected to have AD, and one who was cognitively healthy. These findings provide initial evidence for specific lexical markers as linguistic indices of AD in written language samples. However, the texts analyzed were not reflective of connected speech as used in ordinary oral conversations.

Cuetos et al. (2007) provided additional evidence for language markers of AD by analyzing the Cookie Theft picture descriptions provided by 19 preclinical adults who carried the E280A mutation in the presenilin-1 gene that is deterministic of AD and 21 noncarrier family members. These carriers performed significantly lower than non-carriers on two semantic markers: the total number of semantic units and the total number of object situations present in the picture card. Similarly, Ahmed et al. (2013) analyzed connected speech samples elicited from the Cookie Theft picture of 18 adults who were diagnosed with pathology-confirmed AD and 18 pathology-free, cognitively healthy adults. The researchers reported significantly lower scores for the AD group than matched controls in the total number of semantic units and a decrease in efficiency in the number of semantic units per time. Mueller et al. (2016) reported similar findings in their

retrospective analysis of connected speech samples from 39 cognitively healthy older adults and 39 adults with amnesic MCI who were recruited from the Wisconsin Registry for Alzheimer's Prevention study. When completing the Cookie Theft picture descriptions, participants with amnesic MCI performed significantly differently from adults in the cognitively healthy group by having a lower idea density and producing fewer semantic units and fewer unique words.

Taken together, these findings provide evidence for the value of analyzing connected speech samples to understand language deficits in three early-stage populations: early stages of pathology-confirmed AD, preclinical stages of disease carriers of the E280A mutation, and preclinical MCI (Koscik et al., 2014). However, there are several methodological limitations with these studies and others that hinder advancements in research investigating connected speech before the onset of dementia. First, most of these studies lack a robust sample size (Clarke et al., 2020). Second, the lack of agreement on which discourse metrics should be used and how they are to be defined limits generalization and cross-study comparison (Kavé & Goral, 2017). For example, there is no clear agreement on how to analyze pauses and hesitations in connected speech (Pistono et al., 2016, 2019; Sluis et al., 2020). Third, there is limited cognitive testing, and the clinical syndrome of participants is not well described (Kim et al., 2019; Kim & Lee, 2021). Fourth, the elicitation methods used across studies are variable (e.g., personal narrative or picture description), limiting cross-study comparisons, and several studies only analyzed one method of elicitation, which limits cross-method analysis (Kavé & Goral, 2017).

Mueller, Koscik, Hermann, et al. (2018) were among the first to investigate connected speech in a large longitudinal study of middle-aged adults (aged 40–65 years) at risk for sporadic (i.e., not genetically deterministic) AD. Participants were recruited from the Wisconsin Registry for Alzheimer's Prevention study and completed the Cookie Theft picture description task and a comprehensive neuropsychological assessment at two time points, 2 years apart. Participants had a consensus diagnosis of either cognitively healthy or early MCI at both time points. The final sample included in the analysis consisted of 200 adults classified as cognitively healthy and 64 adults classified as having early MCI. To analyze the connected speech samples, the researchers used discourse measures that represented four latent discourse factors (i.e., semantic, syntax, lexical, and fluency) derived from previous psychometric research conducted by their group (Mueller, Koscik, Clark, et al., 2018). Furthermore, the language samples were transcribed using Codes for the Human Analysis of Transcripts (CHAT; MacWhinney, 2000) and automatically analyzed using the Computerized Language Analysis (CLAN; MacWhinney, 2000) program. The major finding

from this study was that those with MCI declined more rapidly over the 2 years in the semantic and fluency domains than those who were cognitively healthy. A surprising finding from this study was that the lexical factor did not differ between groups. The researchers hypothesized that this may be due to the relatively short length of their language sample. The strengths of this study were that it included a relatively large sample size, used a rigorous latent structure to classify and describe discourse measures, and supported replication by using standardized procedures to elicit discourse samples and to code and extract data (i.e., CHAT/CLAN). Despite these strong findings, several barriers persist and limit further advancements in this area of research. That is, speech samples are difficult to obtain from adequately large numbers of participants who represent diverse populations (Clarke et al., 2020), that they are labor-intensive to transcribe and analyze, and that the analyses are complex and require methods and insights from multiple disciplines (Fraser et al., 2019).

The TalkBank project (<https://talkbank.org/>) helps to address these common methodological barriers in connected speech research. TalkBank is the world's largest open-access integrated web-based repository for transcribed video and audio data on communicative interactions and is grounded in six basic principles: (a) maximally open data sharing, (b) CHAT transcription format, (c) CHAT-compatible software, (d) interoperability, (e) responsivity to research community needs, and (f) international standards (MacWhinney, 2019). The CHAT transcription format is an inclusive transcription standard that harmonizes data across disciplines that have historically used different methods. The TalkBank system includes a collection of programs for data analysis. The core analysis system, referred to as CLAN, combines 30 analysis commands and 25 utility commands and provides automatic analysis for parts of speech and grammatical dependent structures. Within TalkBank are 14 component clinical language banks that span unique research areas (e.g., child language, phonology, and aphasia). Each bank consists of multiple corpora that adhere to the same standardized CHAT transcription format and database organizational standards. The maximally open-access data and analysis methods of the TalkBank system have led to thousands of publications investigating spoken language (MacWhinney, 2019).

Open-access databases are essential to advance dementia science because they facilitate collaborations among researchers, support advancements in computational analyses, improve the dissemination of research to stakeholder groups, and create a richer research environment for generating new hypotheses and opportunities for innovation well beyond the capacity of a single team of researchers (Au et al., 2019; Toga et al., 2016). The Alzheimer's Disease Neuroimaging Initiative (ADNI; Mueller

et al., 2005) and the Health and Retirement Study (Juster & Suzman, 1995) are examples of databases that have already led to major advancements in AD research and provide models and frameworks for data sharing. Based on the successes of TalkBank's clinical language banks, DementiaBank was established to create an open-access data sharing platform for the study of language abilities across the progression of dementia. Currently, the Pitt corpus within DementiaBank is the most heavily relied on corpus in the research domain. It contains transcriptions and audio files from a longitudinal analysis of 104 healthy adults, 208 adults with dementia due to possible Alzheimer's disease and related dementias (ADRD), and 85 older adults with an unknown diagnosis (Becker et al., 1994). This corpus has been used for a wide variety of analyses in over 250 publications. Many of these publications focus on the task of discovering vocal and linguistic features that discriminate between dementia, MCI, and controls. For example, in the ADReSS (Alzheimer's Dementia Recognition through Spontaneous Speech) challenge (Luz et al., 2020, 2021a) at INTERSPEECH 2020 (<https://dementia.talkbank.org/ADReSS-2020/>), there were 25 groups from eight different countries competing to achieve the best F score in terms of precision and accuracy for making this distinction. In that year, competitors were allowed to use both auditory features and features derived from the morphosyntactic analysis in the transcripts. In 2021, a further INTERSPEECH ADReSS challenge was designed to provide competitors with only the sound recordings from which they would have to automatically extract features. This limitation was imposed to more closely approximate the scenario in which the software would be deployed. The results of the second competition were published in 15 articles in a special issue of *Frontiers in Aging Neuroscience* (Luz et al., 2021b). Apart from an exploration of features for classification, the different contributions explored variations in machine learning algorithms for this task. The best-performing classifier from the Baidu team reached an F score of 0.92. Other uses of the corpus have focused on its use to train speech recognizers as well as the extraction of pragmatic, semantic, and discourse features.

Due to the advancements in language analysis as a tool to detect and quantify early cognitive changes, our team is updating and expanding DementiaBank. We have established the Delaware corpus, a new corpus that consists of multimedia language samples of older adults at risk for dementia due to possible AD. The Delaware corpus will expand upon previous connected speech research and databases in several ways. First, there is a large amount of variation in the elicitation methods used in previous research, and many studies lacked standardized elicitation procedures (Drummond et al., 2015; Kavé & Goral, 2017). Most studies have included only the Cookie Theft picture description task (Ahmed et al., 2013; Cuetos et al., 2007; Mueller et al., 2016; Mueller, Koscik, Hermann, et al.,

2018), which has several limitations. Language samples elicited from the Cookie Theft picture description are usually short, and findings from these studies may be specific to the Cookie Theft task and may not generalize to other forms of speech (e.g., conversation; Yeung et al., 2021). Therefore, the new corpus includes a standardized protocol to elicit multiple types of discourse to support comparative analyses across methods and samples. Second, the participants in most AD studies lacked ethnic and cultural diversity. For example, although the ADNI study has led to major contributions to the field's understanding of pathological changes, African Americans represented less than 5% of ADNI's sample. The new Delaware corpus includes procedures for both in-person and remote (video-conference) delivery to support the recruitment and retention of diverse participants. Future planned data collection studies will target specific minoritized populations through community outreach. Third, previous studies have struggled to recruit large sample sizes because clinical data are expensive and time consuming to collect and analyze (Ahmed et al., 2013; Cuetos et al., 2007; Le et al., 2011; Mueller et al., 2016; Mueller, Koscik, Hermann, et al., 2018). To build a large open-access dataset, the new corpus is housed in the already well-established TalkBank system that facilitates open data sharing. The DementiaBank team is also multidisciplinary, consisting of clinicians, researchers, and programmers who will support streamlined data collection and analysis (Fraser et al., 2019). Last, previous methods used for language data collection and analysis lacked standardization, the quality of the audio recordings was variable, and transcription/coding work was labor intensive (Mueller, Hermann, et al., 2018). Therefore, construction of the new corpus involves standardized data collection and analysis procedures using state-of-the-art methods within the TalkBank system including an automatic speech recognition (ASR) pipeline.

To summarize, the overall goal of this work is to expand DementiaBank to include a large dataset of multimedia spoken language samples of neurotypical older adults and older adults at risk for dementia (i.e., by virtue of having MCI due to possible AD) to help better understand speech and language abilities before the onset of dementia. This database could lead to improved analysis algorithms, improved early detection metrics, improved diagnostic criteria for MCI subtypes, treatment approaches that promote brain health for those at risk for dementia, and educational resources for clinicians and students. The aims of this tutorial are to (a) describe the new DementiaBank discourse protocol, a standardized discourse protocol and battery of cognitive-linguistic assessments; (b) describe the Delaware corpus data; and (c) illustrate the types of analyses that can be conducted using the CLAN program and additional resources in TalkBank. More detailed and

rigorous findings will be published after more data are collected.

Method

Development of the DementiaBank Protocol

A key aspect of the clinical banks within TalkBank is the establishment of a standardized protocol that can be used to collect data across sites and investigators. To create the DementiaBank protocol, a multidisciplinary research/clinical team (a) reviewed protocols used in other clinical banks (i.e., AphasiaBank, TBIBank, and RHDBank); (b) reviewed the connected speech literature for AD, aphasia, traumatic brain injury, and right hemisphere brain damage; (c) reviewed the Pitt corpus protocol and literature that has made use of the corpus data; and (d) engaged in discussions with experts in the field of connected speech and AD. The types of discourse stimuli selected were based on several considerations. First, because previous research primarily focused on one discourse type (picture description: Cookie Theft), we wanted to establish a protocol that would include other types (i.e., descriptive, story narrative, procedural, and personal) to depict a wider array of functional contexts. This would support the comparison across speech-sampling methods and across studies using the same protocol. Second, we wanted to select discourse tasks and methods that would overlap with those used in the AphasiaBank (MacWhinney et al., 2011), TBIBank (Togher et al., 2014), and RHDBank (Minga et al., 2021) protocols to allow for comparisons across patient populations. Third, we wanted to be mindful of participants' time and effort constraints, so we limited the administration time for the discourse protocol to approximately 20 min. Fourth, we wanted to select tasks that could be flexibly administered in-person or remotely via videoconference. The final DementiaBank protocol consists of a discourse protocol and a cognitive–linguistic battery.

Discourse Protocol

The discourse protocol includes four task types: picture description, story narrative, procedural discourse, and personal narrative. To support standardized administration across investigators, scripts and materials were developed for three testing scenarios (i.e., remote assessment using the DementiaBank website, remote assessment using PowerPoint, and in-person administration). The scripts were modeled after the AphasiaBank (MacWhinney et al., 2011) and RHDBank (Minga et al., 2021) protocols. The script includes initial prompts and second- and third-level prompts that may be used if the participant does not respond to the initial prompt within 10 s and/or if the

response is fewer than two utterances. For each task, the script includes the approximate times (in minutes) of each task. If the participant continues talking once the time has elapsed, the administrator is instructed to end the task by saying, “I am going to stop you there so we can move on to the next item.” The examiner must adhere to the script and keep additional verbal encouragement to a minimum. Instead of ad-lib verbal encouragement, the protocol asks examiners to use nonverbal encouragement (e.g., head nods or facial expressions) to support transcription. The scripts and materials can be accessed from the DementiaBank website.

Picture description. The “Cookie Theft” picture (Goodglass & Kaplan, 1983) is presented with the prompt, “Please tell me everything you see going on in this picture.”

Narrative discourse. Three story narratives are elicited. First, the “Cat Rescue” picture (L. E. Nicholas & Brookshire, 1993) is presented with the prompt, “Look at everything that’s happening and then tell me a story about what you see. Tell me the story with a beginning, a middle, and an end.” Then, participants are presented with the Norman Rockwell print “Coming and Going” (Rockwell, 1947), and the examiner uses the same prompt as for the Cat Rescue task. For the third task, participants are asked to look through a picture book of the “Cinderella” story (Grimes, 2005). After reviewing the pictures, the participants are then prompted to tell the story in their own words. If the participant is unfamiliar with the Cinderella story, another fairytale or a similar story may be substituted.

Procedural discourse. Participants are prompted to tell the examiner how to make a “peanut butter and jelly sandwich.” Another simple meal may be substituted if not appropriate culturally or otherwise.

Personal narrative. Finally, participants are asked to tell the examiner about their “hometown” or where they grew up. Participants may also be invited to tell a different personal narrative, if necessary.

Cognitive–Linguistic Battery

The cognitive–linguistic battery consists of four standardized measures spanning cognitive domains to provide contextual and complementary information for the discourse metrics collected from the discourse protocol.

Boston Naming Test–Short Form. The Boston Naming Test–Short Form (Mack et al., 1992) is a 15-item confrontational naming assessment. Participants are presented with line drawings and asked to name each image spontaneously. If the word is not recalled spontaneously, the examiner may prompt participants with phonemic or multiple-choice cues; however, only spontaneous naming is scored as accurate (maximum score = 15). This assessment is included in the cognitive–linguistic battery to

support comparative analyses across clinical banks within TalkBank.

Hopkins Verbal Learning Test–Revised. The Hopkins Verbal Learning Test–Revised (HVLT-R; Benedict et al., 1998) is a measure of episodic memory. During administration, the participant is presented with a list of 12 words that comprise three semantic categories. The participant recalls the list immediately after each of the three learning trials and again following a 20- to 25-min delay. In the recall tasks, participants are scored on each accurate word remembered. Raw scores are converted to *T* scores ($M = 50$, $SD = 10$) that are based on age-matched norms for each of these tasks (i.e., immediate recall, delayed recall).

Wechsler Memory Scale–Revised Logical Memory I and II. The Wechsler Memory Scale–Revised (WMS-R; Wechsler, 1987) Logical Memory subtest is an additional measure of episodic memory. In this subtest, the examiner reads two short vignettes (i.e., “Anna Thompson” and “Robert Miller”) and asks the participant to retell the details of each paragraph immediately after hearing them (Logical Memory I) and also following a 20- to 30-min delay (Logical Memory II). The total raw score (sum of the total score for both stories) is converted to a scaled score ($M = 10$, $SD = 3$) for Logical Memory I and II.

Montreal Cognitive Assessment (v7.1). The Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005) is a cognitive screening test designed to detect MCI by sampling multiple cognitive domains (e.g., language, memory, executive functions). A single total raw score is interpreted (range: 0–30) against a cutoff score to classify cognitive impairment. Although there may be overlap in score ranges based on impairment classification, a cutoff score of 25/26 is typically used to separate normal cognition and MCI, and a cutoff score of 17/18 is often used to distinguish MCI from dementia (Nasreddine et al., 2005; Trzepacz et al., 2015).

Delaware Corpus Data Collection

Participants

Data collection for the Delaware corpus was approved by the University of Delaware Institutional Review Board. Participant recruitment is ongoing, but at the time of this writing, there are 53 participants in the Delaware corpus. Recruitment efforts target older adults who either are neurotypical or meet the clinical criteria for MCI due to possible AD. Participants have been recruited primarily from a registry of community-dwelling older adults through the Delaware Center for Cognitive Aging at the University of Delaware. These adults have previously participated either in memory research in the lab of authors A. M. L. and M. L. C. or through promotional materials distributed to targeted groups of community-dwelling older adults. Participants were also recruited who had previously completed memory research studies in the lab of authors A. M. L. and M. L. C. Table 1 provides a summary of participant demographic information, and Table 2 provides a summary of participant assessment scores.

To rule out other systemic or brain diseases that could cause cognitive decline and increase the likelihood that the underlying disease might be AD, we collected self-reported demographic/medical data using a questionnaire that was developed by the geriatric psychiatrist and clinical liaison for the Delaware Center for Cognitive Aging Research, in line with the guidance from the National Institute on Aging-Alzheimer’s Association criteria (Albert et al., 2011). Participants were included in the study if they were 60 years of age or older and met the following inclusion/exclusion criteria: speak and understand spoken English; have adequate hearing and vision; are stable on (or no) nootropic medications; and have negative self-reported history of major psychiatric disorder, untreated major depression, or other systemic medical

Table 1. Self-reported participant demographic information.

Demographics	Classification	
	Neurotypical ($n = 20$)	MCI ($n = 33$)
Mean (<i>SD</i> ; range) age at testing (years)	69.6 (5.9; 62–82)	74.8 (8.8; 61–91)
Sex (n)		
Male	5	16
Female	15	17
Race (n)		
Black or African American	0	3
White	20	30
Education (n)		
High school or equivalent	1	1
Some college	2	2
Associate degree/technical degree	1	3
Bachelor’s degree or higher	16	27

Note. MCI = mild cognitive impairment.

Table 2. Mean participant assessment scores.

Assessment (SD; range)	Classification	
	Neurotypical (<i>n</i> = 20)	MCI (<i>n</i> = 33)
BNT-SF	14.6 (0.8; 13–15)	14.0 (1.3; 10–15)
MoCA	27.3 (2.0; 22–30)	24.2 (2.3; 20–28)
WMS-R		
Logical Memory I	11.6 (3.0; 5–17)	10.2 (3.4; 5–17)
Logical Memory II	11.7 (2.5; 7–17)	9.1 (3.5; 2–15)
HVLT-R		
Immediate Memory	52.1 (8.1; 31–64)	41.6 (10.3; < 20–61)
Delayed Memory	50.5 (10.3; 23–63)	39.6 (14.7; < 20–63)

Note. BNT-SF and MoCA scores are reported as total raw score. WMS-R scores are reported as scaled scores. HVLT-R scores are reported as *t* scores. MCI = mild cognitive impairment; BNT-SF = Boston Naming Test–Short Form; MoCA = Montreal Cognitive Assessment; WMS-R = Wechsler Memory Scale–Revised; HVLT-R = Hopkins Verbal Learning Test–Revised.

illness that could cause cognitive decline (e.g., brain tumor, acquired brain injury, or repeated or recent concussion). Participants included in the corpora were classified as either neurotypical (*n* = 20) or MCI (*n* = 33). MCI classification was based on the National Institute on Aging–Alzheimer’s Association criteria (Albert et al., 2011) and determined by a neuropsychologist, author M. L. C.

Specific criteria for the MCI classification were the following.

1. A decline in cognition. This criterion was met if the participant endorsed at least one item (i.e., a score of 2–4) on the full 39-item Everyday Cognition (ECog; Farias et al., 2021) questionnaire. The ECog consists of 39 items that measure an individual’s ability to perform daily tasks relative to 10 years ago (e.g., “Compared to 10 years ago, has there been any change in recalling conversations a few days later?”). Participants rate items on the following scale: 1 = *no change or actually performs better than 10 years ago*, 2 = *occasionally performs the task worse than 10 years ago but not all of the time*, 3 = *consistently performs the task a little worse than 10 years ago*, 4 = *performs the task much worse than 10 years ago*, and 5 = *do not know* (these responses were treated as missing values). Participants were also recruited by promotional materials targeting those concerned about their memory.
2. Impairment in one or more cognitive domains through objective assessment. Participants met this criterion if they produced a score greater than or equal to -1.5 SD below age-matched norms on the HVLT-R or WMS-R. This criterion could also be met if the participant produced a total score between 18 and 25 on the MoCA (Nasreddine et al., 2005), which is a range most characteristic of MCI (Nasreddine et al., 2005; Trzepacz et al., 2015).

3. Relative preservation of independence in functional abilities. This criterion was met by obtaining a Clinical Dementia Rating (CDR) scale (Morris, 1993) global score of less than or equal to 0.5. The CDR is a semistructured interview with a knowledgeable informant to provide information about cognitive and functional performance across six domains. Participants also self-reported completing activities of daily living independently.
4. Not demented. This was supported by a CDR global score < 1.0 (interview with loved one) and a MoCA score above 17, and/or self-report of independent living.

Procedure

After providing informed consent, participants completed the DementiaBank discourse protocol and cognitive–linguistic assessment battery. To administer the discourse protocol, the administrator followed the script and used the materials for remote assessment using a PowerPoint located on the DementiaBank website. The testing was completed in a single session via videoconference that lasted approximately 90 min. Administration of the discourse protocol was audio-recorded following the guidelines for high-quality audio recording on the TalkBank website. All other study data were collected and managed directly using Research Electronic Data Capture (Harris et al., 2009, 2019) tools hosted by the University of Delaware.

Transcription and Coding

To analyze the language samples using CLAN, the audio files need to be transcribed into CHAT format. The personal narrative discourse samples were manually transcribed by a trained researcher who listened to the language samples and orthographically transcribed the audio into utterances including appropriate codes. The transcript was

then reviewed by a secondary researcher for reliability. To learn more about the manual transcription method, readers are directed to AphasiaBank (MacWhinney et al., 2011) and RHDBank (Minga et al., 2021) publications. Furthermore, the TalkBank website has manuals and tutorial screencasts for CHAT transcription (<https://talkbank.org/>), which then allows for a large variety of analyses using CLAN. Historically, all discourse samples in TalkBank have been transcribed manually by trained researchers or research assistants, which can be time consuming and tedious. Recently, we have begun to use ASR as an additional transcription method to streamline the transcription process.

The TalkBank project is working on establishing an automated pipeline to process the raw audio files into CHAT transcription format, complete with utterance- and word-level alignments. The automated pipeline is composed of six stages: (a) ASR, (b) automatic utterance segmentation, (c) automatic transcript coding, (d) audio-transcript forced alignment, (e) optional human-assisted transcript correction, and (f) automatic morphology and fluency analysis. The Delaware corpus has been used to establish and pilot the automated pipeline using the audio samples from the picture description, narrative discourse, and procedural discourse tasks. For Stage 5, a trained researcher reviewed the transcript while simultaneously listening to the audio to confirm utterance splitting, correct errors, and include additional codes. To confirm transcript accuracy, 20% of audio files were transcribed using both manual and ASR transcription methods, and a mean of 97% point-to-point reliability was obtained (range: 96%–100%). The ASR pipeline code and procedures are further described on the TalkBank website (<https://github.org/talkbank> in the “batchalign” repository). Following transcription, CHAT files were linked to the corresponding media file and uploaded to the DementiaBank shared database (<https://dementia.talkbank.org/access/English/Protocol/Delaware-protect/Delaware.html>).

Results and Discussion

Illustrative Analyses

The combination of standard language protocols, shared databases, consistent transcription formats, and the automated analyses available in the CLAN program allows for the development of new analysis tools and more efficient analysis options for measuring language metrics before the onset of dementia. In the section below, we provide examples of some of the tools and analyses that can be used to further the study of discourse in this area, as well as information on where to find educational resources to learn more about analyses.

EVAL-D

The EVAL command (Forbes et al., 2012) was originally developed for AphasiaBank to facilitate the “evaluation” of language samples from the AphasiaBank discourse protocol. EVAL-D was developed to do the same thing for the specific elements of the standard DementiaBank discourse protocol described above (e.g., Cookie Theft, Rockwell’s Coming and Going, Hometown) and relevant comparison databases for the ADRD population (e.g., based on age, diagnosis). Like EVAL, the EVAL-D command creates a composite profile with 34 outcome measures (e.g., total words, total utterances, mean length of utterance, type–token ratio, words per minute, propositional idea density, open/closed class word ratio) relevant to adult language analysis. The command can be used in these three ways: (a) to analyze the performance of an individual or group of individuals on any discourse task(s), (b) to compare an individual’s performance on any of the standard discourse protocol tasks to database norms from controls or others with the same diagnosis, and (c) to compare an individual’s discourse performance at various time points. We provide examples of the EVAL-D function to analyze discourse samples at the individual and group levels.

Figure 1 shows how to compare an individual’s performance to others in the DementiaBank database. Users can select a specific comparison group or a set of parameters for comparison (e.g., neurotypical vs. MCI, age range, sex, and any or all gems). In this example, we analyzed the Cookie Theft picture description from an 84-year-old woman classified as having MCI and compared it to the Cookie Theft picture descriptions from all female control

Figure 1. EVAL-D dialog box.

PLEASE SELECT ONLY ONE SPEAKER CODE:

Speaker: *INV *PAR

Database types:

Control MCI Memory
 PossibleAD ProbableAD Vascular

Age range: Male only Female only

Gem choices:

Cookie Cinderella Hometown
 Cat Sandwich Rockwell

participants in DementiaBank. Figure 2 shows a small section of the spreadsheet results including this participant's results (Row 2); the standard deviation between her results and the mean from the comparison database (Row 3); the standard deviation between the means of the participant and control group (Row 4); the mean, minimum, maximum, and standard deviation of the database (Rows 5–8); and information about the comparison database (Rows 11–14). These few columns of data show the degree to which this participant's Cookie Theft picture description had a shorter duration, fewer total words, and fewer total utterances than those from the comparison group. In addition, her speech rate was significantly slower, and she had significantly fewer verbs per utterance.

At the group level, EVAL-D was used to compare Cinderella language samples from the MCI ($n = 33$) and neurotypical ($n = 20$) groups in the Delaware corpus. Two-tailed t -test results indicate that the groups differed significantly (all $ps < .05$) on several measures. For example, the MCI group had significantly shorter samples in total duration (128 vs. 205 s), fewer total words (294 vs. 505), fewer unique words (113 vs. 163), shorter mean lengths of utterance (9.9 vs. 11.9 words), fewer verbs per utterance (1.8 vs. 2.0), a smaller percentage of nouns (16.7 vs. 18.2), and a larger percentage of pronouns (14.2 vs. 11.9).

Core Lexicon

S. G. H. Dalton, Kim, et al. (2020) published a compendium of core lexicon (CoreLex) checklists for a number of discourse tasks, several of which are included in the DementiaBank discourse protocol. Further publications included norms (S. G. H. Dalton, Hubbard, & Richardson, 2020) for CoreLex results and validation of an automated CLAN command for computing CoreLex (S. G. Dalton et al., 2022). The checklists include the complete list of words used by control participants for each discourse task. Norms for CoreLex indicate how

many of those words were used at least once by 50% of control participants, thus providing an assessment of the typicality of lexical usage.

The automated CoreLex command was used to examine the Delaware corpus groups' performance on the Cinderella storytelling task and the Cat Rescue picture description. For Cinderella, the MCI group used a mean of 61.2 ($SD = 13.7$) CoreLex checklist words compared with 74 ($SD = 10.1$) used by the neurotypical group. This difference was significant ($t = -3.62, p = .0007$). Thus, along with the EVAL-D results, this shows that in addition to fewer unique words and fewer words overall in the Cinderella narratives, the MCI group also produced significantly fewer words from the normed CoreLex list for this task. Interestingly, however, the mean score of the MCI group was within 1 SD of the mean score of 69.8 ($SD = 15.5$) for the much larger group of control participants ($n = 133$) from S. G. H. Dalton, Hubbard, and Richardson (2020). For the Cat Rescue task, the MCI group used a mean of 26.4 ($SD = 3.0$) CoreLex words, and the neurotypical group used 25.6 ($SD = 3.6$). The difference between these means was not significant. Moreover, these means were very similar to those reported by S. G. H. Dalton, Hubbard, and Richardson. It should be noted that the mean ages for the Delaware controls and the Dalton et al. controls were younger than that of the Delaware corpus MCI group. Again, these analyses are being presented mostly for illustrative purposes and should be investigated more rigorously as the DementiaBank database grows to include more participants.

Fillers

A final example was inspired by the recent work of Farzana et al. (2022) who demonstrated the importance of automatic disfluency analysis in predicting dementia. Using the `FREQ` command in CLAN, we computed the total number of verbal fillers (e.g., "uh" and "um") and

Figure 2. Segment of EVAL-D spreadsheet output.

	A	B	C	D	E	F	G	H	I
1	File_DB	Age	Sex	Group	Duration_(sec)	Total_Utts	MLU_Words	Words_Min	Verbs_Utt
2	01.cookie.cex	84;00.	female	MCI	21	5	13.4	208.571	3
3	+/-SD				-00:00:01.047	-1.058	1.334	2.308	2.761
4					*	*	*	**	**
5	Mean Databas				00:00:46	10.165	9.438	131.959	1.527
6	Min Database				00:00:09	2	2.455	49.091	0.125
7	Max Database				00:02:24	29	20.571	201.429	3.5
8	SD Database				00:00:24.057	4.882	2.97	33.189	0.533
9									
10									
11	+/- SD * = 1 SD, ** = 2 SD								
12	Database Gems:Cookie								
13	Database keywords:Control,female								
14	#_files_in_database:170								

the total number of words (including fillers and words used in repetitions and revisions) in both the MCI and neurotypical participants from the Delaware corpus for the Cookie Theft picture description and the Cinderella story. The mean percentage of fillers per total words for the picture description was 3.1% for the MCI group ($SD = 2.8$) and 4.5% for the neurotypical group ($SD = 4.4$). For the storytelling task, results were 4.9% ($SD = 4.2$) for the MCI group and 4.2% ($SD = 3.9$) for the neurotypical group. Although the two groups did not differ on percentage of fillers (per total words) for either task, within-group comparisons indicated that, as one might predict based on task demands, the MCI group used a significantly higher percentage of fillers in the storytelling task compared with the simple picture description task ($t = 2.09$, $p = .02$, one-tailed). Interestingly, the increased cognitive demands of the Cinderella storytelling task did not significantly affect the percentage of fillers used by the neurotypical group. This would be an interesting avenue to pursue with other associated indices of disfluency as well as other groups in the database and other discourse tasks, particularly when the number of participants (and thus statistical power) increases.

Educational Resources

There are several resources available on the TalkBank website to support users in conducting analyses like the ones described here. These include manuals and screencasts (available at <https://talkbank.org>) that explain and demonstrate the various CLAN commands and a Discourse Analysis page (<https://aphasia.talkbank.org/discourse/>) that provides more information on CoreLex and other measures. These resources are constantly updated with information and resources regarding new developments in discourse analysis programming and software.

Additional Resources

Nonprotocol Corpora

The DementiaBank website also includes other corpora consisting of language samples that do not use the standard discourse protocol described here (<https://dementia.talkbank.org/>). For example, the Lanzi corpus has language samples obtained from semistructured interviews with six individuals with MCI approximately 1.5 years after they completed the Structured External Memory Aid Treatment (Lanzi et al., 2021). There are also several corpora that have language samples in different languages (i.e., German, Mandarin, Spanish, and Taiwanese).

Presentations/Publications

The DementiaBank website has a bibliography listing over 350 publications, presentations, and theses that have made use of DementiaBank data. Links are available

for several conference poster presentations. Public dissemination of these articles and publications may help fuel future research ideas and collaborations.

Requesting Access

In accordance with TalkBank policies, all participant data in the Delaware corpus are password protected and only available to DementiaBank consortium members. Interested users should request membership as per instructions on the main TalkBank webpage. Interested users should read the ground rules and then e-mail macw@cmu.edu with their contact information, affiliation, and a brief statement about how they intend to use the data. Students who are interested in becoming members must ask their faculty advisor to join as DementiaBank members.

Conclusions and Future Directions

Using TalkBank's established open science frameworks, DementiaBank will help researchers build and sharpen predictive algorithms that are well beyond the capacity of any single researcher or research team. Analyses conducted using the Delaware corpus data may help with early detection of subtle changes in language and cognition and provide insight into MCI and dementia subtypes based on discourse profiles. Findings from these analyses may also help identify candidates for clinical trials and improve therapeutic approaches before the onset of dementia. The next steps for this project involve recruiting more researchers, clinicians, and students to use the new DementiaBank protocol to contribute data to grow the database. We hope to include data from participants who have MCI from various geriatric neurodegenerative conditions and from non-English speakers. Furthermore, we intend to increase our recruitment efforts to target community-dwelling older adults who represent minoritized populations to expand the demographics of participants in the Delaware corpus. Finally, we also intend to include longitudinal data to help examine variability and progression of language and cognitive abilities over time. With more data, researchers can conduct more powerful and robust analyses to support a holistic understanding of the language before the onset of dementia, thus improving assessment, and treatment practices at a much earlier point than is typical in our current healthcare practices.

Acknowledgments

This work was supported by the National Institute of Aging of the National Institutes of Health under Grant

K23AG070185-01 awarded to Alyssa M. Lanzi, and the National Institute on Deafness and Other Communication Disorders under Grant R01DC008524-13S1 awarded to Brian MacWhinney. The authors would like to thank Michelle Bourgeois for her strong contributions to DementiaBank that supported the foundational work for the Delaware corpus.

References

- Ahmed, S., de Jager, C. A., Haigh, A.-M., & Garrard, P. (2013). Semantic processing in connected speech at a uniformly early stage of autopsy-confirmed Alzheimer's disease. *Neuropsychology*, 27(1), 79–85. <https://doi.org/10.1037/a0031288>
- Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Gamst, A., Holtzman, D. M., Jagust, W. J., Petersen, R. C., Snyder, P. J., Carrillo, M. C., Thies, B., & Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 270–279. <https://doi.org/10.1016/j.jalz.2011.03.008>
- Alzheimer's Association. (2022). *2022 Alzheimer's disease facts and figures*. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, 18(4), 700–789. <https://doi.org/10.1002/alz.12638>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.).
- Au, R., Ritchie, M., Hardy, S., Ang, T. F. A., & Honghuan, L. (2019). Aging well: Using precision to drive down costs and increase health quality. *Advances in Geriatric Medicine and Research*, 1, e190003. <https://doi.org/10.20900/agmr20190003>
- Bayles, K. A., Tomoeda, C. K., & Trosset, M. W. (1993). Alzheimer's disease: Effects on language. *Developmental Neuropsychology*, 9(2), 131–160. <https://doi.org/10.1080/87565649109540549>
- Bäckman, L., Small, B. J., & Fratiglioni, L. (2001). Stability of the preclinical episodic memory deficit in Alzheimer's disease. *Brain: A Journal of Neurology*, 124(1), 96–102. <https://doi.org/10.1093/brain/124.1.96>
- Becker, J. T., Boller, F., Lopez, O. L., Saxton, J., & McGonigle, K. L. (1994). The natural history of Alzheimer's disease. *Archives of Neurology*, 51(6), 585–594. <https://doi.org/10.1001/archneur.1994.00540180063015>
- Benedict, R. H. B., Schretlen, D., Groninger, L., & Brandt, J. (1998). Hopkins Verbal Learning Test–Revised: Normative data and analysis of inter-form and test–retest reliability. *The Clinical Neuropsychologist*, 12(1), 43–55. <https://doi.org/10.1076/clin.12.1.43.1726>
- Clarke, N., Foltz, P., & Garrard, P. (2020). How to do things with (thousands of) words: Computational approaches to discourse analysis in Alzheimer's disease. *Cortex*, 129, 446–463. <https://doi.org/10.1016/j.cortex.2020.05.001>
- Croisile, B., Ska, B., Brabant, M.-J., Duchene, A., Lepage, Y., Aimard, G., & Trillet, M. (1996). Comparative study of oral and written picture description in patients with Alzheimer's disease. *Brain and Language*, 53(1), 1–19. <https://doi.org/10.1006/brln.1996.0033>
- Cuetos, F., Arango-Lasprilla, J. C., Uribe, C., Valencia, C., & Lopera, F. (2007). Linguistic changes in verbal expression: A preclinical marker of Alzheimer's disease. *Journal of the International Neuropsychological Society*, 13(03), 433–439. <https://doi.org/10.1017/S1355617707070609>
- Dalton, S. G., Stark, B. C., Fromm, D., Apple, K., MacWhinney, B., Rensch, A., & Rowedder, M. (2022). Validation of an automated procedure for calculating core lexicon from transcripts. *Journal of Speech, Language, and Hearing Research*, 65(8), 2996–3003. https://doi.org/10.1044/2022_JSLHR-21-00473
- Dalton, S. G. H., Hubbard, H. I., & Richardson, J. D. (2020). Moving toward non-transcription based discourse analysis in stable and progressive aphasia. *Seminars in Speech and Language*, 41(1), 32–44. <https://doi.org/10.1055/s-0039-3400990>
- Dalton, S. G. H., Kim, H., Richardson, J. D., & Wright, H. H. (2020). A compendium of core lexicon checklists. *Seminars in Speech and Language*, 41(1), 45–60. <https://doi.org/10.1055/s-0039-3400972>
- Drummond, C., Coutinho, G., Fonseca, R. P., Assunção, N., Teldeschi, A., de Oliveira-Souza, R., Moll, J., Tovar-Moll, F., & Mattos, P. (2015). Deficits in narrative discourse elicited by visual stimuli are already present in patients with mild cognitive impairment. *Frontiers in Aging Neuroscience*, 7, 96. <https://doi.org/10.3389/fnagi.2015.00096>
- Farias, S. T., Weakley, A., Harvey, D., Chandler, J., Huss, O., & Mungas, D. (2021). The measurement of Everyday Cognition (ECog): Revisions and updates. *Alzheimer Disease & Associated Disorders*, 35(3), 258–264. <https://doi.org/10.1097/WAD.0000000000000450>
- Farzana, S., Deshpande, A., & Parde, N. (2022). How you say it matters: Measuring the impact of verbal disfluency tags on automated dementia detection. *Proceedings of the 21st Workshop on Biomedical Language Processing*, 37–48. <https://doi.org/10.18653/v1/2022.bionlp-1.4>
- Forbes, M., Fromm, D., & MacWhinney, B. (2012). Aphasia-Bank: A resource for clinicians. *Seminars in Speech and Language*, 33(3), 217–222. <https://doi.org/10.1055/s-0032-1320041>
- Fraser, K. C., Lundholm Fors, K., Eckerström, M., Öhman, F., & Kokkinakis, D. (2019). Predicting MCI status from multimodal language data using cascaded classifiers. *Frontiers in Aging Neuroscience*, 11, 205. <https://doi.org/10.3389/fnagi.2019.00205>
- Garrard, P., Maloney, L. M., Hodges, J. R., & Patterson, K. (2005). The effects of very early Alzheimer's disease on the characteristics of writing by a renowned author. *Brain*, 128(2), 250–260. <https://doi.org/10.1093/brain/awh341>
- Gold, M., Amatniek, J., Carrillo, M. C., Cedarbaum, J. M., Hendrix, J. A., Miller, B. B., Robillard, J. M., Rice, J. J., Soares, H., Tome, M. B., Tarnanas, I., Vargas, G., Bain, L. J., & Czaja, S. J. (2018). Digital technologies as biomarkers, clinical outcomes assessment, and recruitment tools in Alzheimer's disease clinical trials. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 4(1), 234–242. <https://doi.org/10.1016/j.trci.2018.04.003>
- Golde, T. E., Schneider, L. S., & Koo, E. H. (2011). Anti-A β therapeutics in Alzheimer's disease: The need for a paradigm shift. *Neuron*, 69(2), 203–213. <https://doi.org/10.1016/j.neuron.2011.01.002>
- Goodglass, H., & Kaplan, E. (1983). *Boston Diagnostic Aphasia Examination booklet*. Lea & Febiger.
- Grimes, N. (2005). *Walt Disney's Cinderella*. Random House.
- Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., McLeod, L., Delacqua, G., Delacqua, F., Kirby, J., Duda, S. N., & REDCap Consortium. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*, 95, 103208. <https://doi.org/10.1016/j.jbi.2019.103208>
- Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for

- providing translational research informatics support. *Journal of Biomedical Informatics*, 42(2), 377–381. <https://doi.org/10.1016/j.jbi.2008.08.010>
- Hier, D. B., Hagenlocker, K., & Shindler, A. G.** (1985). Language disintegration in dementia: Effects of etiology and severity. *Brain and Language*, 25(1), 117–133. [https://doi.org/10.1016/0093-934X\(85\)90124-5](https://doi.org/10.1016/0093-934X(85)90124-5)
- Jack, C. R., Lowe, V. J., Weigand, S. D., Wiste, H. J., Senjem, M. L., Knopman, D. S., Shiung, M. M., Gunter, J. L., Boeve, B. F., Kemp, B. J., Weiner, M., & Petersen, R. C.** (2009). Serial PIB and MRI in normal, mild cognitive impairment and Alzheimer's disease: Implications for sequence of pathological events in Alzheimer's disease. *Brain*, 132(5), 1355–1365. <https://doi.org/10.1093/brain/awp062>
- Jicha, G. A., Parisi, J. E., Dickson, D. W., Johnson, K., Cha, R., Ivnik, R. J., Tangalos, E. G., Boeve, B. F., Knopman, D. S., Braak, H., & Petersen, R. C.** (2006). Neuropathologic outcome of mild cognitive impairment following progression to clinical dementia. *Archives of Neurology*, 63(5), 674–681. <https://doi.org/10.1001/archneur.63.5.674>
- Juster, F. T., & Suzman, R.** (1995). An overview of the health and retirement study. *The Journal of Human Resources*, 30, S7–S56. <https://doi.org/10.2307/146277>
- Kaplan, E., Goodglass, H., & Weintraub, S.** (1983). *The Boston Naming Test*. Lea & Febiger.
- Kavé, G., & Goral, M.** (2016). Word retrieval in picture descriptions produced by individuals with Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology*, 38(9), 958–966. <https://doi.org/10.1080/13803395.2016.1179266>
- Kavé, G., & Goral, M.** (2017). Do age-related word retrieval difficulties appear (or disappear) in connected speech? *Aging, Neuropsychology, and Cognition*, 24(5), 508–527. <https://doi.org/10.1080/13825585.2016.1226249>
- Kim, B. S., Kim, Y. B., & Kim, H.** (2019). Discourse measures to differentiate between mild cognitive impairment and healthy aging. *Frontiers in Aging Neuroscience*, 11, 221. <https://doi.org/10.3389/fnagi.2019.00221>
- Kim, B. S., & Lee, M. S.** (2021). Discourse performance and related cognitive function in mild cognitive impairment and dementia: A preliminary study. *Applied Neuropsychology: Adult*. Advance online publication. <https://doi.org/10.1080/23279095.2021.1922408>
- Koscik, R. L., La Rue, A., Jonaitis, E. M., Okonkwo, O. C., Johnson, S. C., Bendlin, B. B., Hermann, B. P., & Sager, M. A.** (2014). Emergence of mild cognitive impairment in late middle-aged adults in the Wisconsin Registry for Alzheimer's Prevention. *Dementia and Geriatric Cognitive Disorders*, 38(1–2), 16–30. <https://doi.org/10.1159/000355682>
- Kourtis, L. C., Regele, O. B., Wright, J. M., & Jones, G. B.** (2019). Digital biomarkers for Alzheimer's disease: The mobile/wearable devices opportunity. *Npj Digital Medicine*, 2(1), Article 9. <https://doi.org/10.1038/s41746-019-0084-2>
- Lanzi, A. M., Wallace, S. E., Cohen, M. L., & Bourgeois, M. S.** (2021). Structured External Memory Aid Treatment (SEMAT) for older adults with mild cognitive impairment: Long-term adherence and acceptability of treatment. *Aphasiology*, 36(2), 234–250. <https://doi.org/10.1080/02687038.2020.1868395>
- Le, X., Lancashire, I., Hirst, G., & Jokel, R.** (2011). Longitudinal detection of dementia through lexical and syntactic changes in writing: A case study of three British novelists. *Literary and Linguistic Computing*, 26(4), 435–461. <https://doi.org/10.1093/lilc/lqr013>
- Lloret, A., Esteve, D., Lloret, M. A., Cervera-Ferri, A., Lopez, B., Nepomuceno, M., & Monllor, P.** (2019). When does Alzheimer's disease really start? The role of biomarkers. *International Journal of Molecular Sciences*, 20(22), 5536. <https://doi.org/10.3390/ijms20225536>
- Luz, S., Haider, F., de la Fuente Garcia, S., Fromm, D., & MacWhinney, B.** (2020). Alzheimer's dementia recognition through spontaneous speech: The ADReSS challenge. *Proceedings of INTERSPEECH 2020*, 2172–2176. <https://doi.org/10.21437/Interspeech.2020-2571>
- Luz, S., Haider, F., de la Fuente Garcia, S., Fromm, D., & MacWhinney, B.** (2021a). Detecting cognitive decline using speech only: The ADReSSo challenge. *Proceedings of INTERSPEECH 2021*, 3780–3784. <https://doi.org/10.1101/2021.03.24.21254263>
- Luz, S., Haider, F., de la Fuente Garcia, S., Fromm, D., & MacWhinney, B.** (2021b). Editorial: Alzheimer's dementia recognition through spontaneous speech. *Frontiers in Computer Science*, 3, 780169. <https://doi.org/10.3389/fcomp.2021.780169>
- Mack, W. J., Freed, D. M., Williams, B. W., & Henderson, V. W.** (1992). Boston Naming Test: Shortened versions for use in Alzheimer's disease. *Journal of Gerontology*, 47(3), P154–P158. <https://doi.org/10.1093/geronj/47.3.P154>
- MacWhinney, B.** (2000). *The CHILDES project: The database* (Vol. 2). Psychology Press.
- MacWhinney, B.** (2019). Understanding spoken language through TalkBank. *Behavior Research Methods*, 51(4), 1919–1927. <https://doi.org/10.3758/s13428-018-1174-9>
- MacWhinney, B., Fromm, D., Forbes, M., & Holland, A.** (2011). AphasiaBank: Methods for studying discourse. *Aphasiology*, 25(11), 1286–1307. <https://doi.org/10.1080/02687038.2011.589893>
- Minga, J., Johnson, M., Blake, M. L., Fromm, D., & MacWhinney, B.** (2021). Making sense of right hemisphere discourse using RHDBank. *Topics in Language Disorders*, 41(1), 99–122. <https://doi.org/10.1097/TLD.0000000000000244>
- Morris, J. C.** (1993). The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology*, 43(11), 2412–2414. <https://doi.org/10.1212/WNL.43.11.2412-a>
- Mueller, K. D., Hermann, B., Mecollari, J., & Turkstra, L. S.** (2018). Connected speech and language in mild cognitive impairment and Alzheimer's disease: A review of picture description tasks. *Journal of Clinical and Experimental Neuropsychology*, 40(9), 917–939. <https://doi.org/10.1080/13803395.2018.1446513>
- Mueller, K. D., Koscik, R. L., Clark, L. R., Hermann, B. P., Johnson, S. C., & Turkstra, L. S.** (2018). The latent structure and test–retest stability of connected language measures in the Wisconsin Registry for Alzheimer's Prevention (WRAP). *Archives of Clinical Neuropsychology*, 33(8), 993–1005. <https://doi.org/10.1093/arclin/acx116>
- Mueller, K. D., Koscik, R. L., Hermann, B. P., Johnson, S. C., & Turkstra, L. S.** (2018). Declines in connected language are associated with very early mild cognitive impairment: Results from the Wisconsin Registry for Alzheimer's Prevention. *Frontiers in Aging Neuroscience*, 9, 437. <https://doi.org/10.3389/fnagi.2017.00437>
- Mueller, K. D., Koscik, R. L., Turkstra, L. S., Riedeman, S. K., LaRue, A., Clark, L. R., Hermann, B., Sager, M. A., & Johnson, S. C.** (2016). Connected language in late middle-aged adults at risk for Alzheimer's disease. *Journal of Alzheimer's Disease*, 54(4), 1539–1550. <https://doi.org/10.3233/JAD-160252>
- Mueller, K. D., Van Hulle, C. A., Koscik, R. L., Jonaitis, E., Peters, C. C., Betthausen, T. J., Christian, B., Chin, N., Hermann, B. P., & Johnson, S.** (2021). Amyloid beta associations with connected speech in cognitively unimpaired adults. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 13(1), e12203. <https://doi.org/10.1002/dad2.12203>

- Mueller, S. G., Weiner, M. W., Thal, L. J., Petersen, R. C., Jack, C., Jagust, W., Trojanowski, J. Q., Toga, A. W., & Beckett, L. (2005). The Alzheimer's disease neuroimaging initiative. *Neuroimaging Clinics of North America*, 15(4), 869–877. <https://doi.org/10.1016/j.nic.2005.09.008>
- Murdoch, B. E., Chenery, H. J., Wilks, V., & Boyle, R. S. (1987). Language disorders in dementia of the Alzheimer type. *Brain and Language*, 31(1), 122–137. [https://doi.org/10.1016/0093-934X\(87\)90064-2](https://doi.org/10.1016/0093-934X(87)90064-2)
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & 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. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>
- Nicholas, L. E., & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults. *Journal of Speech and Hearing Research*, 36(2), 338. <https://doi.org/10.1044/jshr.3602.338>
- Nicholas, M., Obler, L. K., Albert, M. L., & Helm-Estabrooks, N. (1985). Empty speech in Alzheimer's disease and fluent aphasia. *Journal of Speech and Hearing Research*, 28(3), 405–410. <https://doi.org/10.1044/jshr.2803.405>
- Petersen, R. C. (2016). Mild cognitive impairment. *Continuum: Lifelong Learning in Neurology*, 22(2), 404–418. <https://doi.org/10.1212/CON.0000000000000313>
- Pistono, A., Jucla, M., Barbeau, E. J., Saint-Aubert, L., Lemesle, B., Calvet, B., Köpke, B., Puel, M., & Pariente, J. (2016). Pauses during autobiographical discourse reflect episodic memory processes in early Alzheimer's disease. *Journal of Alzheimer's Disease*, 50(3), 687–698. <https://doi.org/10.3233/JAD-150408>
- Pistono, A., Pariente, J., Bézy, C., Lemesle, B., Le Men, J., & Jucla, M. (2019). What happens when nothing happens? An investigation of pauses as a compensatory mechanism in early Alzheimer's disease. *Neuropsychologia*, 124, 133–143. <https://doi.org/10.1016/j.neuropsychologia.2018.12.018>
- Price, B. H., Gurvit, H., Weintraub, S., Geula, C., Leimkuhler, E., & Mesulam, M. (1993). Neuropsychological patterns and language deficits in 20 consecutive cases of autopsy-confirmed Alzheimer's disease. *Archives of Neurology*, 50(9), 931–937. <https://doi.org/10.1001/archneur.1993.00540090038008>
- Rockwell, N. (1947). *Going and coming* [Oil on canvas]. Norman Rockwell Art Collection Trust, Indianapolis, IN, United States.
- Schneider, J. A., Arvanitakis, Z., Bang, W., & Bennett, D. A. (2007). Mixed brain pathologies account for most dementia cases in community-dwelling older persons. *Neurology*, 69(24), 2197–2204. <https://doi.org/10.1212/01.wnl.0000271090.28148.24>
- Schneider, J. A., Arvanitakis, Z., Leurgans, S. E., & Bennett, D. A. (2009). The neuropathology of probable Alzheimer disease and mild cognitive impairment. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, 66(2), 200–208. <https://doi.org/10.1002/ana.21706>
- Sluis, R. A., Angus, D., Wiles, J., Back, A., Gibson, T., Liddle, J., Worthy, P., Copland, D., & Angwin, A. J. (2020). An automated approach to examining pausing in the speech of people with dementia. *American Journal of Alzheimer's Disease & Other Dementias*, 35, 1533317520939773. <https://doi.org/10.1177/1533317520939773>
- Smith, G. E., & Bondi, M. W. (2013). *Mild cognitive impairment and dementia: Definitions, diagnosis, and treatment*. Oxford University Press.
- Snowdon, D. A., Kemper, S. J., Mortimer, L. H., Greiner, K., Wekstein, D. R., & Markesbery, W. R. (1996). Linguistic ability in early life and cognitive function and Alzheimer's disease in late life: Findings from the Nun Study. *JAMA*, 275(7), 528–532. <https://doi.org/10.1001/jama.1996.03530310034029>
- Tavabi, N., Stück, D., Signorini, A., Karjadi, C., Al Hanai, T., Sandoval, M., Lemke, C., Glass, J., Hardy, S., Lavalée, M., Wasserman, B., Ang, T. F. A., Nowak, C. M., Kainkaryam, R., Foschini, L., & Au, R. (2022). Cognitive digital biomarkers from automated transcription of spoken language. *The Journal of Prevention of Alzheimer's Disease*, 9(4), 791–800. <https://doi.org/10.14283/jpad.2022.66>
- Toga, A. W., Bhatt, P., & Ashish, N. (2016). Global data sharing in Alzheimer disease research. *Alzheimer Disease and Associated Disorders*, 30(2), 160–168. <https://doi.org/10.1097/WAD.0000000000000121>
- Togher, L., Elbourn, E., Kenny, B., McDonald, S., Tate, R., Turkstra, L., Holland, A., Fromm, D., Forbes, M., & MacWhinney, B. (2014). TBI Bank is a feasible assessment protocol to evaluate the cognitive communication skills of people with severe TBI during the subacute stage of recovery. *Brain Injury*, 28(5–6), 723–723.
- Trzepacz, P. T., Hochstetler, H., Wang, S., Walker, B., Saykin, A. J., & Alzheimer's Disease Neuroimaging Initiative. (2015). Relationship between the Montreal Cognitive Assessment and Mini-Mental State Examination for assessment of mild cognitive impairment in older adults. *BMC Geriatrics*, 15(1), 107. <https://doi.org/10.1186/s12877-015-0103-3>
- Vuorinen, E., Laine, M., & Rinne, J. (2000). Common pattern of language impairment in vascular dementia and in Alzheimer disease. *Alzheimer Disease & Associated Disorders*, 14(2), 81–86. <https://doi.org/10.1097/00002093-200004000-00005>
- Wechsler, D. (1987). *Wechsler Memory Scale–Revised*. The Psychological Corporation.
- Yeung, A., Iaboni, A., Rochon, E., Lavoie, M., Santiago, C., Yancheva, M., Novikova, J., Xu, M., Robin, J., Kaufman, L. D., & Mostafa, F. (2021). Correlating natural language processing and automated speech analysis with clinician assessment to quantify speech-language changes in mild cognitive impairment and Alzheimer's dementia. *Alzheimer's Research & Therapy*, 13(1), 109. <https://doi.org/10.1186/s13195-021-00848-x>