# Connected Language in Late Middle-Aged Adults at Risk for Alzheimer's Disease

**Kimberly Diggle Mueller**[1,5], **Rebecca L. Koscik**[1], **Lyn S. Turkstra**[5], **Sarah K. Riedeman**[5], **Asenath LaRue**[1], **Lindsay R. Clark**[1,3], **Bruce Hermann**[1,4], **Mark A. Sager**[1], and **Sterling C. Johnson**[2,1,3]

[1]Wisconsin Alzheimer's Institute, University of Wisconsin School of Medicine and Public Health, Madison, WI

[2]Geriatric Research Education and Clinical Center, William S. Middleton Memorial VA Hospital, Madison WI 53705

[3]Wisconsin Alzheimer's Disease Research Center, University of Wisconsin School of Medicine and Public Health, Madison, WI

[4]Department of Neurology, University of Wisconsin School of Medicine and Public Health, Madison, WI

[5]Department of Communication Sciences & Disorders, University of Wisconsin, Madison, WI

## Abstract

Connected language is often impaired among people with Alzheimer's Disease (AD), yet little is known about when language difficulties first emerge on the path to a clinical diagnosis. The objective of this study was to determine whether individuals with psychometric (preclinical) evidence of amnestic Mild Cognitive Impairment (MCI) (pMCI) showed deficits in connected language measures. Participants were 39 pMCI and 39 cognitively healthy (CH) adults drawn from the Wisconsin Registry for Alzheimer's Prevention, who were matched for age, literacy, and sex. Participants completed a connected language task in which they described the Cookie Theft picture from the Boston Diagnostic Aphasia Examination. Language samples were analyzed across 3 language domains: content, syntactic complexity, and speech fluency. Paired t-tests were used to compare CH and pMCI groups on all variables, and Cohen's *d* effect sizes were calculated for each comparison. The CH and pMCI groups differed significantly on measures of content (e.g., CH group produced more semantic units, more unique words and had larger idea density, on average, than the pMCI group). The picture description findings are consistent with previous retrospective studies showing semantic language differences in adults with autopsy-confirmed AD. Given that these comparisons are between cognitively healthy and pMCI individuals (before a clinical MCI diagnosis), these findings may represent subtle language difficulty in spontaneous speech, and may be predictive of larger language changes over time.

Corresponding author: Kimberly D Mueller, MS, Wisconsin Alzheimer's Institute, Wisconsin Alumni Research Foundation – 9th Floor, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin 53792, USA. Phone: 608-829-3300; fax: 608-829-3315, kdmueller@wisc.edu.

## INTRODUCTION

A major goal of neuropsychological research in Alzheimer's Disease (AD) is early detection of cognitive decline. To date much of this research has focused on identifying individuals who meet criteria for mild cognitive impairment (MCI), because of the high risk of converting to dementia in this group [1, 2]. MCI diagnostic criteria established by a joint National Institute on Aging and Alzheimer's Association consensus panel include the following: concern about cognitive decline by the person or significant other, preservation of independence in functional abilities, and impairment in one or more cognitive domains [2]. Although episodic memory impairment is often the main cognitive symptom of Alzheimer's Disease, considerable research evidence suggests that the neuropathological changes of Alzheimer's Disease develop years or decades before the onset of such cognitive symptoms [3-5]. Therefore the quest to find sensitive neuropsychological measures that are effective in detecting very early decline, that can be differentiated from typical changes associated with aging, has been ongoing.

In addition to episodic memory decline, declines in semantic memory have also been described early on the continuum of Alzheimer's Disease. Semantic memory refers to knowledge of concepts, objects, people and words. Barbeau et al., 2012 [6] found that a group with MCI performed significantly worse than controls on several semantic memory tasks, including naming famous faces, cultural knowledge, and famous public events. The authors argue that the semantic deficits exist without temporal gradient (i.e., deficits exist no matter how old or recent the learning occurred), and thus may not be caused by episodic memory deficits. The authors also found that the semantic impairments were correlated with anterior, extra-hippocampal, temporal lobe brain structures as well as to the hippocampus. Didic et al., 2011 [7] describe semantic memory as being "context-free"; that is, the recall of previously learned facts, words or names is dissociable from the original context in which it was learned, and is less dependent upon the hippocampus and more dependent upon anterior medial temporal structures which appear to be affected earlier in Alzheimer's disease. Didic et al., 2011 argue that even the more sensitive tests designed to assess episodic memory function, such as word list learning and recall, are "context-free" by nature and difficulties may be reflective of semantic, as well as episodic memory problems. Venneri et al., 2016 theorize that assessments targeted toward semantic memory may be more sensitive in distinguishing normal aging processes from preclinical AD; that is, normal aging typically results in a gradual decline in episodic memory, while semantic memory is usually spared, if not improved, in normal aging [8].

Semantic memory difficulties, within the context of everyday functioning, may be manifest by word retrieval difficulties and naming impairments, which are common complaints in both MCI and AD. Faber-Langendoen et al. [9] found that 36% of mild AD patients and

100% of severe AD patients had aphasia (i.e., impairment in some aspect of language function) concurrent with memory impairments.

In order to characterize word finding problems, clinicians and researchers often use confrontational naming tasks as a means of quantifying this difficulty. While declining performance on typical confrontational naming tasks is evident in Alzheimer's Disease [10], and confrontational naming tasks have been useful in staging more moderate to severe AD [11], such tests have not been useful in reliably detecting differences in preclinical AD or MCI[12, 13] .

Word finding may also be evaluated by using verbal fluency tasks, in which patients are asked to name words of a category or words beginning with a specified letter, and are often used to assess problems with both executive function and language in Alzheimer's Disease [10]. Specifically, studies of verbal fluency in MCI have shown lower scores in those with MCI and AD compared to age-matched controls [14-16]. Recently, our group found differences in verbal fluency abilities between persons identified as "cognitively healthy" and those identified as 'psychometric MCI', a construct using robust norms across three longitudinal visits and analogous to early or 'preclinical' MCI, which may possibly identify very subtle cognitive dysfunction [17].

Assessment of isolated language functions, however, might not capture the magnitude of problems experienced in everyday communication contexts. For evaluation that more closely approximates language demands of everyday life, researchers have turned to assessment of spontaneous, connected speech, both as a diagnostic resource as well as a research tool toward understanding the subtle deficits reported by patients in their natural settings. Connected language is elicited by having an individual produce spoken or written language in response to a particular question or stimulus, including picture description tasks[18-21], familiar story re-telling [22] (e.g., "Cinderella" story) , wordless picture book descriptions [23], and general autobiographical interviews [24]. Language is then analyzed using measures such as efficiency (e.g., idea density), conciseness (e.g., how many words required to express an idea), complexity (e.g. syntax complexity), and speech fluency (e.g., the number of repetitions, circumlocutions, filled pauses). Analysis of connected language has revealed subtle differences years or decades before diagnosis of AD [20, 21, 25, 26] . These language differences were a salient finding in the Nun Study[25], in which idea density in early life writing samples was an accurate predictor of later-life progression to AD. Links between content of language samples and preclinical AD also were reported by Cuetos [20], who examined spontaneous language in a group of 19 carriers of the E280A mutation in the Presenilin 1 gene at the preclinical stage of AD (mean age = 43.2). Carriers showed deficits in the production of semantic categories in a picture description task during the pre-clinical phase of disease. There were no differences between carriers and non-carriers in average sentence length or total number of sentences, which implies that clinical examination of language productivity would not necessarily detect group differences. More recently, Ahmed [27] examined longitudinal profiles of 15 adults (mean age 71.2) with autopsy-confirmed AD and found decreasing linear trends in syntactic complexity and semantic and lexical content across stages of disease, including the MCI stage of disease. Given evidence of subtle language changes in MCI, it was of interest to see if connected language differences

could be a cognitive marker in very early declines that are prior to a diagnosis of MCI. Specifically, we hypothesized that participants who were cognitively healthy would differ on connected language outcomes from a subset with psychometric evidence of aMCI (i.e., the subset did not have a clinical diagnosis of MCI, but did meet statistical criteria for early decline as described in the methods). While evidence for early language change exists based on retrospective analysis [26, 28], to our knowledge no studies have examined connected speech and language prospectively in an at-risk cohort.

In the current study, we examined connected language samples in a group of adults (ages 40-65 years at baseline) who were at risk for AD, defined as "psychometric MCI (pMCI)." The psychometric MCI construct refers to the presence of memory declines across at least two of three longitudinal study visits based on internal norms (described further in the methods section). We hypothesized that adults who were clinically cognitively intact but met psychometric criteria of MCI (see Koscik[29]) would have lower scores on measures of content, syntactic complexity, and speech fluency than age- and education-matched adults defined as cognitively healthy.

## MATERIALS AND METHODS

### Participants

Participants were recruited from the Wisconsin Registry for Alzheimer's Prevention (WRAP) study (N>1500), an ongoing longitudinal cohort study enriched for positive parental history of AD. WRAP study design and assessment protocols are described in detail elsewhere [30, 31]. In brief, participants in the positive parental history group have at least one biological parent with either autopsy-confirmed or probable AD as defined by NINCDS-ADRDA (National Institute of Neurological and Communicative Disorders and Stroke - Alzheimer's Disease and Related Disorders Association) research criteria [32]. At baseline, WRAP participants are cognitively intact and in early to late middle-age (M=54 years), and return for follow-up assessments 4 years after baseline and every 2 years thereafter. Participants complete extensive neuropsychological testing at each study visit. Collection of speech samples was added to the third visit protocol and collection has been ongoing across all visits since that time. All study procedures were approved by the University of Wisconsin-Madison's Institutional Review Board, and procedures were done in accord with the ethical standards of the Human Research Protection Program of the UW-Madison and the 1975 Declaration of Helsinki.

### Selection of Cases and Controls for the Present Analyses

**Psychometric MCI (pMCI)—**The primary predictor in our analysis was the psychometric MCI (pMCI) status group (i.e., cognitively healthy (CH) versus pMCI). The construct of pMCI, described in detail elsewhere [29], requires a participant to have episodic memory scores more than 1.5 SD below robust norms on at least 2 of 3 study visits; robust norms were developed by using a reference group's baseline scores to develop more sensitive prediction limits based on age, gender, and WRAT-3 reading scores[29].

**Inclusion criteria**—For this study, individuals with neurological diagnoses including meningitis, stroke, epilepsy, multiple sclerosis and Parkinson's disease, as well as individuals who were non-native English speakers, were not eligible. After those exclusions were applied, we identified 40 participants who had speech samples and who also met psychometric MCI (pMCI) criteria. Once eligible cases were identified, one cognitively healthy participant ("CH"; no episodic memory or executive function scores more than 1.5 SD below robust norms across 3 visits) was randomly selected for each pMCI case from the subset of CH in the same age range (no difference greater than 3 years), gender, and literacy level (CH were from the same decile, using age- and sex-adjusted Wide Range Achievement Test -3 (WRAT-3) scores [33]) as the pMCI case.

The sample size of 40 for each of two study groups was targeted initially based on feasibility of analyzing speech samples and sensitivity analyses that suggested power of .8 to detect effect sizes of ~.40 to .50 (assuming a correlation between the pMCI and CH groups of .10-.15). One speech sample was unusable, so the final sample size was 39 per group.

**Additional predictors and covariates**—Other variables examined included depressive symptoms as measured by the Centers for Epidemiologic Studies Depression scale (CES-D, [34, 35]), parental family history of AD (FH+), the presence of one or more *APOE* ε4 allele, obtained as described elsewhere [35], standardized language measures (Boston Naming Test[36]; Verbal Fluency tasks), and measures of episodic memory and executive function (Table 1).

## Language Sample Procedure

Participants provided informed consent to have speech samples recorded. Participants were provided with the "Cookie Theft" picture from the Boston Diagnostic Aphasia Examination [36], and were instructed to "Tell me everything you see going on in this picture." The test administrators were instructed to provide minimal feedback, but were instructed to give the scripted prompt, "do you see anything else going on?" if responses were unusually brief. Duration of the picture description speech sample ranged from approximately one to three minutes in total length (including examiner prompts). All responses were recorded using an Olympus VN-6200PC digital audio recorder.

## Transcriptions

Transcriptions of the speech samples for this study were performed by a trained speech-language pathology graduate student (SKR) using Codes for Human Analysis of Transcripts (CHAT;[37]). Utterance boundaries were determined by the "T-unit" classification system, defined as a main clause and all of its modifiers [38]. We coded the samples for automatic analyses by the Computer Language Analysis (CLAN) program [37]. Common codes included those for fillers, revisions, repetitions, pauses, semantic units (for description of these measures, see below in "language measures"), as well as for various non-verbal behaviors (e.g., laughing). Semantic units, parts of speech, total words, total utterances, and other quantifiers were automatically extracted by the CLAN program.

Ten percent of samples were re-transcribed and coded by a second rater, a trained speech-language pathologist (KDM), and inter-rater agreement was 98.9% for transcriptions and 82.5% for semantic unit coding. Disagreements for semantic unit coding were discussed, a manual was developed based on consensus, and semantic units were re-coded. Both the examiners and coders were blinded to the group status (CH or pMCI) of participants at the time of data collection and transcription.

## Language Measures

The measures used for analysis and illustrative examples are summarized in Tables 1a and 1b. Language measures were chosen based on previous literature that focused on discourse analysis in MCI or preclinical AD [18-20, 27, 28].

## Semantic Content and Lexical Richness

**Total semantic units—**The total number of semantic units for the "Cookie Theft" picture description task was obtained using the semantic unit classification described by Croisile [39] and Ahmed [19], that included 23 semantic units, consisting of people, objects, actions and attitudes.

**Semantic unit idea density—**Semantic Unit Idea Density was calculated by dividing the total number of semantic units by the number of words.

**Propositional idea density—**An additional form of propositional idea density was automatically extracted from the CLAN program, adapted from the Computerized Propositional Idea Density Rater (CPIDER3) [37, 40, 41]. In brief, propositions correspond to verbs, adjectives, adverbs, prepositions and conjunctions divided by the total number of words (excluding repetitions and fillers).

**Unique words and total words—**The number of unique words and total words were calculated.

**Moving Average Type-Token Ratio (MATTR)—**For a measure of lexical diversity we used the Moving-average type-token ratio (MATTR), which is an index based on a moving window that computes an average of type-token ratios for each successive window of a fixed length [40] (we used fixed length of 10).

**Proportion of pronouns—**This measure was calculated by dividing the number of pronouns by the total number of nouns plus pronouns.

**Proportion of verbs—**This ratio was defined as number of verbs divided by total verbs plus nouns.

## Syntactic Complexity

**Mean length of utterance—**As a measure of syntactic complexity for both tasks, we calculated Mean Length of Utterance (MLU), the number of morphemes (units of linguistic meaning) divided by number of utterances.

**Verb index**—Verb index was defined as the ratio of number of verbs to total number of utterances. We calculated this measure as a way to capture the number of clauses per utterance, and because verb deficits have been shown to coincide with disruptions in discourse [42].

## Speech Fluency

**Maze Index**—In order to measure speech fluency, we defined mazes as repetitions, revisions, false starts and fillers (e.g., "uh, um"). Maze index was calculated by dividing the total number of mazes by the number of utterances [43-45].

**Filled pauses**—Filled pauses was the proportion of fillers (e.g., "um," "uh") to number of utterances, and was examined based on previous findings showing an increase in filled pauses [45].

## Analyses

Prior to testing our hypotheses, all language variables were transformed to normally-distributed z-scores (i.e., $\sim N(0,1)$), except for Boston Naming Test. Due to high ceiling effects of BNT in our sample, no transformations yielded a normal distribution. Paired sample t-tests or Chi-Square Tests of Independence were used to compare CH and PMCI groups on all variables. For the Semantic Units measure, two outliers were identified so we applied Winsor adjustments to these values. For the Boston Naming Test (BNT) [36], we used a Mann-Whitney Wilcoxon test due to the high ceiling effects of the BNT in the WRAP sample. Effect sizes were calculated for each comparison using Cohen's *d*: (mean group 1 – mean group 2)/pooled sd. All tests were two-tailed using an alpha of .05. For the discourse measures that differed significantly by pMCI status, we also examined the Pearson correlation coefficients among discourse, verbal fluency, and verbal memory (using the combined sample of n=78). Analyses were performed using SAS v9.3, R version 3.2.2, and SPSS version 22.

## RESULTS

The average age of the participants included in these analyses was 63.1 years (SD=6.3); 22% were female; and the average WRAT reading standard score was 110.6(7.4). The cognitively healthy (CH) and pMCI (psychometric MCI) groups differed significantly in terms of *APOE-4*, in that the CH group had a higher percentage of participants with at least one *E4* allele; family history status was not significantly different. Table 2 shows that the demographic characteristics of the pMCI groups were very similar, as would be expected based on the matched sampling strategy that was used. Table 2 also describes the two samples regarding neuropsychological performance, and is reflective of the defining nature of the psychometric aMCI criteria [10, 29]. Specifically, the verbal and learning scores were significantly different between the two groups, but the executive function scores were not. Both letter fluency and category fluency, which were not part of the psychometric aMCI diagnostic criteria, yielded worse scores for the PMCI group than the CH group (p = .0005 and p<.0001, respectively); the Boston Naming Test was not significantly different between groups.

### Connected language results are presented in Table 3

To facilitate review of all language outcomes in a single graph, the mean(se) of the corresponding z-scores across several language outcome measures are shown in Figure 1 by cognitive status groups. The CH and PMCI groups differed significantly on two content measures: pMCI subjects produced fewer unique words (p = .028) and fewer semantic units (p = .018) than the CH group. Marginal differences were found in Propositional Idea Density (p = .09) and in the Moving Average Type-Token Ratio (MATTR) (p = .06). The CH and PMCI groups did not differ significantly on Maze Index or filled pauses/utterance, nor on any of the syntactic complexity measures.

**Secondary Analyses**—Since the Cognitively Healthy group had significantly more *APOE-4* carriers, we ran secondary analyses using ANCOVAs with *APOE-4* as a covariate (Table 4). The pMCI group continued to show significantly lower Semantic Units (p = .008) than the CH group (adjusted means(se): CH=14.5(.44); pMCI=12.7(.44)). In addition, when adjusting for *APOE-4* status, the pMCI group showed significantly lower propositional idea density than the CH group (p = .028) (adjusted means(se): CH=.46(.01); pMCI=.44(.01)), and fewer unique words (p = .03) (adjusted means(se): CH=71.6(3.3); pMCI=61.1(3.3)). No other differences were statistically significant. To further remove the confounding risk factor of APOE-4, we re-ran the analysis for each APOE subset (i.e., APOE-4 non-carriers and carriers). In the APOE-4 non-carrier subset, the pMCI group (n=31) continued to show reduced semantic units (p=.05), fewer unique words (p=.05), and increased use of filled pauses (p=.028) compared to the CH group (n=20). In the APOE-4 carrier subset, non-significant differences were observed for comparisons of pMCI (n=8 ) and CH (n= 19) in all measures; power to detect significant differences was limited in this subset.

We also examined the relationship between the connected language variables and the more standard verbal fluency measures and the R-AVLT in the neuropsychological test battery (Fig 2). Bivariate Pearson correlations for the set of 78 participants were moderate for unique words and letter fluency (r=.41), unique words and AVLT (r=.39); idea density was moderately correlated with letter fluency (r=.40); Semantic Units was moderately correlated with AVLT (r=.34), but not with the verbal fluency tasks.

## DISCUSSION

In this study, spontaneous language samples were examined from two groups of late-middle-aged participants drawn from a cohort enriched for family history of AD: those who were evidencing psychometric memory decline based on internal robust norms (pMCI) and those who were classified as cognitively healthy. Psychometric memory decline was identified in participants who showed episodic memory scores that were more than 1.5 SD below robust norms on at least 2 of 3 study visits. We hypothesized that the pMCI group would perform worse on language variables characterizing semantic content, syntactic complexity, and speech fluency, than those who were cognitively healthy.

The pMCI group performed worse on both phonemic (letter) and semantic (category) fluency, two measures that were not included in the psychometric criteria for determining

the groups. The confrontational naming task (Boston Naming Test) was not significantly different between groups, consistent with previous MCI literature [10].

Our results suggested that the pMCI group expressed fewer semantic units and fewer unique words than the cognitively healthy group, which is a similar finding to that seen in studies with persons with MCI and AD [20, 21, 27, 28]. Recently, Berisha [28] found that former President Ronald Reagan's use of unique words declined over time, prior to his diagnosis of Alzheimer's Disease. Similarly, while the pMCI group in our sample exhibited verbal learning and memory declines across longitudinal visits that were significant according to robust norms, they were not diagnosed with clinical MCI because they were not sufficiently impaired according to general test norms. Therefore, the pMCI group did not meet formal diagnostic criteria for MCI and are categorized as having a psychometric condition; these findings may suggest and reinforce the idea that language symptoms may be present antecedent to meeting formal diagnostic criteria for the clinical condition (aMCI). The semantic unit findings may represent subtle differences in semantic processing, particularly within the context of the demands of producing spontaneous speech, and may be predictive of larger language changes over time. That the Semantic Units measure was not significantly correlated with either verbal fluency tasks, but was correlated with the Rey Auditory Verbal Learning Test, suggests that it may tap into additional language or other cognitive processes, and may contribute unique knowledge regarding preclinical language changes.

The differences in semantic units between the two groups could support the hypothesis that semantic memory declines may be reflective of the underlying early neuropathology in the transentorhinal, mediotemporal networks, while the hippocampus may still be spared [8]. Further, when the task is embedded within spontaneous speech, a more qualitative analysis affords an informative characterization of problems within a functional context. The task itself is relatively quick and noninvasive for patients to complete, and may be informative not only for very early diagnosis, but also for disease course monitoring in both pharmacological and nonpharmacological clinical trials.

Secondary analyses were conducted to adjust for the higher percentage of participants with *APOE-4* in the cognitively healthy group. Since APOE E-4 is considered to be a risk factor for cognitive decline even in normally aging populations [46], we were interested in adjusting for this possible confound in our control group. This adjustment of controlling for APOE-4 strengthened the results of the Semantic Units and unique words findings, and also revealed that the pMCI group's language samples had lower propositional idea density than the CH group. While several studies have shown that lower propositional density in early life predicted future AD risk [25, 47], and one small study found that young APOE-4 carriers had lower propositional density than noncarriers [48], the effect of APOE4 status on language abilities in aging and in AD risk is not well understood. Whether APOE4 status exerts an effect on future language declines will be addressed in longitudinal follow-up studies on this cohort.

No differences were detected in syntactic complexity; this may be due to the possibility that syntactic complexity remains preserved at preclinical or early phases, as evidenced previously [49]. However, Le et al. [50] analyzed several novels of British writers over time

and found some decreases in the use of passive voice, consistent with other previous findings [51]. Therefore it is possible that our samples of the picture description task are simply too short to detect syntactic differences. The same can be said for the lack of findings with respect to fillers, revisions, and repetitions in this group. It is also possible that syntactic differences will be more detectable upon longitudinal analysis of change within individuals.

The greatest limitations of this study include the small sample size and not adjusting for multiple comparisons. We powered our study to detect plausible effect sizes using an alpha of .05; adjustments for multiple comparisons would have greatly reduced our statistical power, and increasing the sample size was beyond our resources at the time of analysis. An additional limitation includes the fact that several measures that have been shown to identify impairment in Alzheimer's Disease (e.g., error monitoring (the proportion of revised errors to total errors), timing aspects such as performance deviations per minute or response-time to word-finding delays) were not analyzed, as we based our measures on those that have been examined in MCI or preclinical AD. Nonetheless, the addition of such measures would add clarity to the limited research available in this population. Future studies will include factor analyses of the many discourse measures examined in previous AD literature in order to more concisely characterize aspects of language and to reduce the amount of multiple comparisons.

The strengths of this study include utilizing connected language analysis measures in a unique population: a cohort of clinically healthy individuals who are at-risk for Alzheimer's Disease. Thus far, connected language analysis has proven to be useful in detecting deficits at early stages of autopsy-confirmed AD[19] and in preclinical stages of disease in carriers of the E280A mutation [20]. An additional strength of this study is examining groups that are based upon internally-derived norms [29] which allowed us to identify a "preclinical" MCI group. Further examination of connected language using larger sample sizes of the pMCI and CH groups longitudinally may yield important findings regarding improving the ability to detect early language deficits, aiding in differential diagnosis, and providing direction for non-pharmacological communication interventions.

In conclusion, these results suggest that subtle yet detectable changes in connected language may be present during preclinical stages of cognitive decline. For relatively little investment of patient time and effort, language tasks can yield rich information not only about semantic storage and retrieval, but also about the complex cognitive processes involved in producing spontaneous, connected language. Further, these analyses may help to provide tools for differential diagnosis among variants of cognitive decline, such as AD versus Primary Progressive Aphasia [18], or between dementia and other co-morbid conditions such as depression [52]. Future research will examine changes over time, to better understand how such changes correspond to or predict clinical outcomes, document disease progression, and aid in determining strategies and interventions for improved communication between patients and caregivers.

## ACKNOWLEDGMENTS

## REFERENCES

1. Fischer P, Jungwirth S, Zehetmayer S, Weissgram S, Hoenigschnabl S, Gelpi E, Krampla W, Tragl KH. Conversion from subtypes of mild cognitive impairment to Alzheimer dementia. Neurology. 2007; 68:288–291. [PubMed: 17242334]

2. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Gamst A, Holtzman DM, Jagust WJ, Petersen RC, Snyder PJ, Carrillo MC, Thies B, Phelps CH. 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. Alzheimers Dement. 2011; 7:270–279. [PubMed: 21514249]

3. Jack CR Jr. Knopman DS, Jagust WJ, Petersen RC, Weiner MW, Aisen PS, Shaw LM, Vemuri P, Wiste HJ, Weigand SD, Lesnick TG, Pankratz VS, Donohue MC, Trojanowski JQ. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. Lancet Neurol. 2013; 12:207–216. [PubMed: 23332364]

4. Zimmer ER, Leuzy A, Benedet AL, Breitner J, Gauthier S, Rosa-Neto P. Tracking neuroinflammation in Alzheimer's disease: the role of positron emission tomography imaging. J Neuroinflammation. 2014; 11:120. [PubMed: 25005532]

5. Alcolea D, Carmona-Iragui M, Suarez-Calvet M, Sanchez-Saudinos MB, Sala I, Anton-Aguirre S, Blesa R, Clarimon J, Fortea J, Lleo A. Relationship between beta-Secretase, inflammation and core cerebrospinal fluid biomarkers for Alzheimer's disease. J Alzheimers Dis. 2014; 42:157–167. [PubMed: 24820015]

6. Barbeau EJ, Didic M, Joubert S, Guedj E, Koric L, Felician O, Ranjeva JP, Cozzone P, Ceccaldi M. Extent and neural basis of semantic memory impairment in mild cognitive impairment. J Alzheimers Dis. 2012; 28:823–837. [PubMed: 22112550]

7. Didic M, Barbeau EJ, Felician O, Tramoni E, Guedj E, Poncet M, Ceccaldi M. Which memory system is impaired first in Alzheimer's disease? J Alzheimers Dis. 2011; 27:11–22. [PubMed: 21799246]

8. Venneri A, Mitolo M, De Marco M. Paradigm shift: semantic memory decline as a biomarker of preclinical Alzheimer's disease. Biomark Med. 2016; 10:5–8. [PubMed: 26642376]

9. Faber-Langendoen K, Morris JC, Knesevich JW, LaBarge E, Miller JP, Berg L. Aphasia in senile dementia of the Alzheimer type. Ann Neurol. 1988; 23:365–370. [PubMed: 3382172]

10. Taler V, Phillips NA. Language performance in Alzheimer's disease and mild cognitive impairment: a comparative review. J Clin Exp Neuropsychol. 2008; 30:501–556. [PubMed: 18569251]

11. Welsh KA, Butters N, Hughes JP, Mohs RC, Heyman A. Detection and staging of dementia in Alzheimer's disease: Use of the neuropsychological measures developed for the Consortium to Establish a Registry for Alzheimer's Disease. Archives of neurology. 1992; 49:448–452. [PubMed: 1580805]

12. Testa JA, Ivnik RJ, Boeve B, Petersen RC, Pankratz VS, Knopman D, Tangalos E, Smith GE. Confrontation naming does not add incremental diagnostic utility in MCI and Alzheimer's disease. J Int Neuropsychol Soc. 2004; 10:504–512. [PubMed: 15327729]

13. Willers IF, Feldman ML, Allegri RF. Subclinical naming errors in mild cognitive impairment. Demant Neuropsychol. 2008; 2:217–222.

14. Nutter-Upham KE, Saykin AJ, Rabin LA, Roth RM, Wishart HA, Pare N, Flashman LA. Verbal fluency performance in amnestic MCI and older adults with cognitive complaints. Arch Clin Neuropsychol. 2008; 23:229–241. [PubMed: 18339515]

15. Clark LJ, Gatz M, Zheng L, Chen YL, McCleary C, Mack WJ. Longitudinal verbal fluency in normal aging, preclinical, and prevalent Alzheimer's disease. Am J Alzheimers Dis Other Demen. 2009; 24:461–468. [PubMed: 19759254]

16. Price SE, Kinsella GJ, Ong B, Storey E, Mullaly E, Phillips M, Pangnadasa-Fox L, Perre D. Semantic verbal fluency strategies in amnestic mild cognitive impairment. Neuropsychology. 2012; 26:490–497. [PubMed: 22746308]

17. Mueller KD, Koscik RL, LaRue A, Clark LR, Hermann B, Johnson SC, Sager MA. Verbal Fluency and Early Memory Decline: Results from the Wisconsin Registry for Alzheimer's Prevention. Arch Clin Neuropsychol. 2015; 30:448–457. [PubMed: 26025231]

18. Ahmed S, de Jager CA, Haigh AM, Garrard P. Logopenic aphasia in Alzheimer's disease: clinical variant or clinical feature? J Neurol Neurosurg Psychiatry. 2012; 83:1056–1062. [PubMed: 22842206]

19. Ahmed S, Haigh AM, de Jager CA, Garrard P. Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease. Brain. 2013; 136:3727–3737. [PubMed: 24142144]

20. Cuetos F, Arango-Lasprilla JC, Uribe C, Valencia C, Lopera F. Linguistic changes in verbal expression: a preclinical marker of Alzheimer's disease. J Int Neuropsychol Soc. 2007; 13:433–439. [PubMed: 17445292]

21. Forbes-McKay KE, Venneri A. Detecting subtle spontaneous language decline in early Alzheimer's disease with a picture description task. Neurol Sci. 2005; 26:243–254. [PubMed: 16193251]

22. Thompson CK, Mack JE. Grammatical impairments in PPA. Aphasiology. 2014; 28:1018–1037. [PubMed: 25642014]

23. Fergadiotis G, Wright HH, Capilouto GJ. Productive vocabulary across discourse types. Aphasiology. 2011; 25:1261–1278. [PubMed: 22904592]

24. Irish M, Kamminga J, Addis DR, Crain S, Thornton R, Hodges JR, Piguet O. 'Language of the past'–Exploring past tense disruption during autobiographical narration in neurodegenerative disorders. Journal of neuropsychology. 2015

25. Snowdon DA, Kemper SJ, Mortimer JA, Greiner LH, Wekstein DR, Markesbery WR. Linguistic ability in early life and cognitive function and Alzheimer's disease in late life. Findings from the Nun Study. JAMA. 1996; 275:528–532. [PubMed: 8606473]

26. Garrard P, Maloney LM, Hodges JR, Patterson K. The effects of very early Alzheimer's disease on the characteristics of writing by a renowned author. Brain. 2005; 128:250–260. [PubMed: 15574466]

27. Ahmed S, de Jager CA, Haigh AM, Garrard P. Semantic processing in connected speech at a uniformly early stage of autopsy-confirmed Alzheimer's disease. Neuropsychology. 2013; 27:79–85. [PubMed: 23356598]

28. Berisha V, Wang S, LaCross A, Liss J. Tracking discourse complexity preceding Alzheimer's disease diagnosis: a case study comparing the press conferences of Presidents Ronald Reagan and George Herbert Walker Bush. J Alzheimers Dis. 2015; 45:959–963. [PubMed: 25633673]

29. Koscik RL, La Rue A, Jonaitis EM, Okonkwo OC, Johnson SC, Bendlin BB, Hermann BP, Sager MA. Emergence of mild cognitive impairment in late middle-aged adults in the wisconsin registry for Alzheimer's prevention. Dement Geriatr Cogn Disord. 2014; 38:16–30. [PubMed: 24556849]

30. La Rue A, Hermann B, Jones JE, Johnson S, Asthana S, Sager MA. Effect of parental family history of Alzheimer's disease on serial position profiles. Alzheimers Dement. 2008; 4:285–290. [PubMed: 18631980]

31. Sager MA, Hermann B, La Rue A. Middle-aged children of persons with Alzheimer's disease: APOE genotypes and cognitive function in the Wisconsin Registry for Alzheimer's Prevention. J Geriatr Psychiatry Neurol. 2005; 18:245–249. [PubMed: 16306248]

32. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr. Kawas CH, Klunk WE, Koroshetz WJ, Manly JJ, Mayeux R, Mohs RC, Morris JC, Rossor MN, Scheltens P, Carrillo MC, Thies B, Weintraub S, Phelps CH. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement. 2011; 7:263–269. [PubMed: 21514250]

33. Wilkinson, GS. WRAT-3: Wide range achievement test. Wide Range; 1993.

34. Radloff LS. The use of the Center for Epidemiologic Studies Depression Scale in adolescents and young adults. Journal of youth and adolescence. 1991; 20:149–166. [PubMed: 24265004]

35. Johnson SC, La Rue A, Hermann BP, Xu G, Koscik RL, Jonaitis EM, Bendlin BB, Hogan KJ, Roses AD, Saunders AM, Lutz MW, Asthana S, Green RC, Sager MA. The effect of TOMM40 poly-T length on gray matter volume and cognition in middle-aged persons with APOE epsilon3/epsilon3 genotype. Alzheimers Dement. 2011; 7:456–465. [PubMed: 21784354]

36. Goodglass, H.; Kaplan, E. Boston diagnostic aphasia examination booklet. Lea & Febiger; 1983.

37. Macwhinney B, Fromm D, Forbes M, Holland A. AphasiaBank: Methods for Studying Discourse. Aphasiology. 2011; 25:1286–1307. [PubMed: 22923879]

38. Hunt KW. Grammatical Structures Written at Three Grade Levels. NCTE Research Report No. 3. 1965

39. Croisile B, Ska B, Brabant MJ, Duchene A, Lepage Y, Aimard G, Trillet M. Comparative study of oral and written picture description in patients with Alzheimer's disease. Brain Lang. 1996; 53:1–19. [PubMed: 8722896]

40. Covington MA, McFall JD. Cutting the Gordian knot: The moving-average type–token ratio (MATTR). Journal of Quantitative Linguistics. 2010; 17:94–100.

41. Brown C, Snodgrass T, Kemper SJ, Herman R, Covington MA. Automatic Measurement of Propositional Idea Density from Part-of-Speech Tagging. Behavior research methods. 2008; 40:540–545. [PubMed: 18522065]

42. Kim M, Thompson CK. Verb deficits in Alzheimer's disease and agrammatism: Implications for lexical organization. Brain and language. 2004; 88:1–20. [PubMed: 14698726]

43. Nicholas M, Obler LK, Albert ML, Helm-Estabrooks N. Empty speech in Alzheimer's disease and fluent aphasia. J Speech Hear Res. 1985; 28:405–410. [PubMed: 4046581]

44. Bird H, Lambon Ralph MA, Patterson K, Hodges JR. The rise and fall of frequency and imageability: noun and verb production in semantic dementia. Brain Lang. 2000; 73:17–49. [PubMed: 10872636]

45. Wilson SM, Henry ML, Besbris M, Ogar JM, Dronkers NF, Jarrold W, Miller BL, Gorno-Tempini ML. Connected speech production in three variants of primary progressive aphasia. Brain. 2010; 133:2069–2088. [PubMed: 20542982]

46. Bretsky P, Guralnik J, Launer L, Albert M, Seeman TE. The role of APOE-ɛ4 in longitudinal cognitive decline MacArthur Studies of Successful Aging. Neurology. 2003; 60:1077–1081. [PubMed: 12682309]

47. Engelman M, Agree EM, Meoni LA, Klag MJ. Propositional density and cognitive function in later life: Findings from the Precursors Study. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences. 2010; 65:706–711.

48. Medina LD, Rodriguez-Agudelo Y, Geschwind DH, Gilbert PE, Liang L-J, Cummings JL, Ringman JM. Propositional density and apolipoprotein e genotype among persons at risk for familial Alzheimer's disease. Dementia and geriatric cognitive disorders. 2011; 32:188–192. [PubMed: 22134129]

49. Kemper S, LaBarge E, Ferraro FR, Cheung H, Cheung H, Storandt M. On the preservation of syntax in Alzheimer's disease: Evidence from written sentences. Archives of Neurology. 1993; 50:81–86. [PubMed: 8418805]

50. Le X, Lancashire I, Hirst G, Jokel R. Longitudinal detection of dementia through lexical and syntactic changes in writing: a case study of three British novelists. Literary and Linguistic Computing. 2011:fqr013.

51. Bates E, Harris C, Marchman V, Wulfeck B, Kritchevsky M. Production of complex syntax in normal ageing and Alzheimer's disease. Language and Cognitive Processes. 1995; 10:487–539.

52. Murray LL. Distinguishing clinical depression from early Alzheimer's disease in elderly people: Can narrative analysis help? Aphasiology. 2010; 24:928–939.
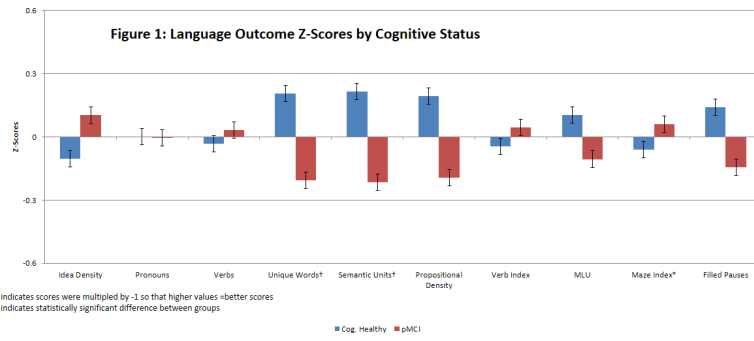
**Figure 1.**
Language Outcome Z-Scores by Cognitive Status

**Table 1a**

Categories of connected language measures.

| Semantic Content and Lexical Richness | Syntactic Complexity | Speech Fluency |
|---|---|---|
| **Semantic Units** | **Mean Length of Utterance (MLU)** | **Maze Index (mazes/utterances)** |
| **Semantic Unit Idea Density** | **Verb Index (verbs/utterances)** | **Filled Pauses/utterances** |
| Unique Words | | |
| Total Words | | |
| Moving Average Type-Token Ratio (MATTR) | | |
| Pronouns (proportion) | | |
| Propositional Idea Density | | |

**Definitions of Measures:** *Semantic Units* refer to a pre-defined set of people, objects, actions and attitudes [19, 39]; *Semantic Unit Idea Density*: total number of semantic units divided by total number of words; *Unique Words:* total number of non-repeated words; *Moving Average Type-Token Ratio*: an average of ratios of unique words to total words for successive windows of a length of 10; *Proportion of Pronouns:* total pronouns divided by nouns plus pronouns; *Proportion of Verbs:* total verbs divided by verbs plus nouns; *Mean length of utterance*: total number of morphemes divided by total utterances; *Verb Index*: ratio of number of verbs to total utterances; *Maze Index:* total number of repetitions, revision, false starts and fillers divided by number of utterances; *Filled pauses:* proportion of fillers to number of utterances.

**Table 1b**

Illustrative examples of aspects of connected language.

|  | **Semantic Units - Content** | **MLU and Verb Index - Syntax** | **Maze Index and Filled Pauses** [*] |
|---|---|---|---|
| **High** | *um well the little boy is standing on a on a stool that's about ready to tip over. he's uh raiding the cookie jar [=! laughs] giving it to his uh looks to be maybe his sister or friend. and uh um [pause] uh looks like she's enjoying receiving a cookie . mom is washing is washing the dishes and the the water is (pause) overflowing onto the onto the floor* | *um and the mother is washing dishes. but she has her back to the children that are trying to get into the cookie jar. and she's not paying attention to the water that's overflowing on the sink and going on the floor.* | *um there's coffee cups on the counter and a plate and um it looks like a very affluent neighborhood [=! laughs] outside the window <and I think she should> and she just looks like a like a Beaver Cleaver June Cleaver [=! laughs] kinda thing. how many woman um work in the kitchen in a in a dress and a shift with an apron ? please I use an apron <but it's not like it> (be)cause I'm too lazy to change my church clothes [=! laughs]. and it's because I'm just make such a disaster in the kitchen.* |
| **Low** | *we have uh a kitchen um with a uh presumably mother and two children. the children as children do are um raiding the cookie jar. and the mother is um washing dishes and has her problem with uh an overflowing sink. and um (pause) that's about it.* | *well the lady's wiping dishes. but the sink is overflowing. and uh kids are robbing a cookie jar .* | *I see a classic nineteen sixties house wife washing dishes wearing an apron. and the sink is overflowing. and water is going onto the floor. I see some children attempting to procure cookies out of the cookie jar with some success. but the young lad is standing on a stool&um and appears to be about to fall over backwards and possibly whack his head on the counter top which mom doesn't see.* |

[*] For Maze Index and filled pauses, higher indicates worse performance.

**Table 2**

Comparison of Demographics and Neuropsychological Data for the Psychometric-aMCI and Cognitively Healthy Groups

|  | P-aMCI | CH | p-value |
|---|---|---|---|
| *Demographics* |  |  |  |
| Age, mean(sd) years | 63.1(6.4) | 63.1(6.3) | 0.97 |
| Gender, n(%) female | 22(56.4) | 22(56.4) | 1 |
| WRAT reading, mean(sd) | 53.2(3.8) | 53.2(3.6) | 0.95 |
| Education, n(%) >=BA | 29(74.4) | 30(76.9) | 0.79 |
| Family History of AD, n (%) | 29 (74.4) | 31(79.5) | 0.59 |
| APOE-E4, n (%) | 8 (20.5) | 19(48.7) | 0.009** |
| CES-D, mean(sd) [†] | 6.92(7.4) | 6.41(6.1) | 0.74 |
| *Language Production* |  |  |  |
| Boston Naming Test (60) [††] | 57.6(2.7) | 58.5(1.7) | 0.11 |
| Phonemic Fluency (C, F, L) | 42.31(12.1) | 50.18(11.8) | 0.005** |
| Semantic Fluency (animals) | 20.92(4.6) | 25.21(5.7) | < 0.0001*** |
| *Verbal Learning and Memory* |  |  |  |
| AVLT-Total | 39.9(7.3) | 53.13 (6.9) | < 0.0001*** |
| *Executive Function* |  |  |  |
| Trails A | 31.92(12.7) | 28.4(8.4) | 0.16 |
| Trails B | 81.03(45.2) | 69(30.7) | 0.18 |
| Digit Symbol | 51.4(11.7) | 53.7(10.7) | 0.37 |
| Digits Forward | 10.2(2.4) | 11(2.6) | 0.18 |
| Digits Backward | 7.2(2.4) | 7.5(2) | 0.51 |
| Letter-Number Sequencing | 10.3(2.7) | 10.7(2.6) | 0.48 |
| Stroop-Color/Word | 103.4(25) | 103(25) | 0.46 |

[†] Centers for Epidemiologic Studies-Depression scale (Radloff, 1991)

[††] Tested with 2-sample Wilcoxon test due to non-normality in the BNT distribution

**Table 3**

Analysis of Connected Language Variables

| Semantic Content and Lexical Richness | P-aMCI Mean(SD) | CH Mean(SD) | p-value | Cohen's *d* |
|---|---|---|---|---|
| Propositional Idea Density | .44(.05) | .46(.04) | 0.09 | 0.44 |
| Pronouns (proportion) | .35(.1) | .35(.1) | 0.98 | 0.00 |
| Verbs (proportion) | .55(.08) | .54(.08) | 0.78 | 0.13 |
| Semantic Units (possible range: 0-23) | 12.8(3.0) | 14.3(2.9) | **0.02** | **0.46** |
| Semantic Unit Idea Density | .15(.06) | .14(.06) | 0.37 | 0.19 |
| Unique Words (types) | 63.4(23.2) | 73.6(25.8) | **0.03** | **0.42** |
| Total Number of Words | 97.1(46.0) | 115.9(56.4) | 0.1 | 0.37 |
| Moving Average Type-Token Ratio | .94(.03) | .95(.02) | 0.06 | 0.39 |
| *Syntatic Complexity* | | | | |
| Verb Index | 2.4(.6) | 2.4(.6) | 0.71 | 0 |
| Mean Length of Utterance | 11.29(2.8) | 11.81(2.1) | 0.36 | 0.21 |
| *Speech Fluency* | | | | |
| Filled pauses (proportion) | .51(.3) | .59(.3) | 0.21 | 0.27 |
| Total Mazes (proportion) | .8(.6) | .9(1.2) | 0.62 | 0.11 |

Means are raw scores/ratios of all connected language variables.

**Table 4**

Secondary analysis, adjusting for *APOE-4* status.

| | pMCI | CH | p-value | Cohen's *d* |
|---|---|---|---|---|
| *Semantic Content and Lexical Richness* | **Adjusted Mean(SE)** | **Adjusted Mean(SE)** | | |
| Propositional Idea Density | .44(.01) | .46(.01) | **0.028** | **0.58** |
| Pronouns (proportion) | .35(.02) | .35(.02) | 0.96 | 0.06 |
| Verbs (proportion) | .55(.01) | .54(.01) | 0.81 | 0.06 |
| Semantic Units (possible range 0 -23) | 12.7(.44) | 14.5(.44) | **0.008** | **0.65** |
| Semantic Unit Idea Density | .14(.01) | .15(.01) | 0.38 | 0.69 |
| Unique Words ("types") | 61.1(3.3) | 71.6(3.3) | **0.03** | **0.52** |
| Total Number of Words ("tokens") | 97.3(8.4) | 116.8(8.4) | 0.12 | 0.37 |
| Moving Average Type-Token Ratio | .94(.004) | .95(.004) | 0.091 | 0.43 |
| *Syntactic Complexity* | | | | |
| Verb Index (Total verbs/utterances) | 2.4(.1) | 2.4(.1) | 0.99 | 0.3 |
| Mean Length of Utterance | 11.2(.4) | 11.9(.4) | 0.2 | 0.39 |
| *Speech Fluency* | | | | |
| Filled pauses (proportion) | .35(.06) | .43(.06) | 0.33 | 0.22 |
| Total Mazes (proportion) | .70(.15) | .93(.15) | 0.28 | 0.45 |

**Table 5**

Correlations Among Significant Language Variables and AVLT

| | | Semantic Units | AVLT Total Score | Letter Fluency | Category Fluency | Unique Words | Prop.Idea Density |
|---|---|---|---|---|---|---|---|
| Semantic Units | Pearson Correlation | 1 | .331[**] | .257[*] | .177 | .486[**] | −.029 |
| | Sig. (2-tailed) | | .003 | .023 | .122 | .000 | .799 |
| | N | 78 | 78 | 78 | 78 | 78 | 78 |
| AVLT Total Score | Pearson Correlation | | 1 | .472[**] | .491[**] | .393[**] | .124 |
| | Sig. (2-tailed) | | | .000 | .000 | .000 | .278 |
| | N | | 78 | 78 | 78 | 78 | 78 |
| Letter Fluency | Pearson Correlation | | | 1 | .544[**] | .406[**] | .397[**] |
| | Sig. (2-tailed) | | | | .000 | .000 | .000 |
| | N | | | 78 | 78 | 78 | 78 |
| Category Fluency | Pearson Correlation | | | | 1 | .265[*] | .188 |
| | Sig. (2-tailed) | | | | | .019 | .100 |
| | N | | | | 78 | 78 | 78 |
| Unique Words | Pearson Correlation | | | | | 1 | .233[*] |
| | Sig. (2-tailed) | | | | | | .040 |
| | N | | | | | 78 | 78 |
| Prop.Idea Density | Pearson Correlation | | | | | | 1 |
| | Sig. (2-tailed) | | | | | | |
| | N | | | | | | 78 |

Correlation matrix of the discourse variables that were significantly different between groups with category and letter fluency, and Rey Auditory Verbal Learning and Memory Test scores.

[**] Correlation is significant at the 0.01 level (2-tailed).

[*] Correlation is significant at the 0.05 level (2-tailed).