



Published in final edited form as:

*Aphasiology*. 2013 July 1; 27(7): . doi:10.1080/02687038.2012.757760.

## Formulaic Language in Alzheimer's Disease

**Ms. Kelly Ann Bridges, M.A. [Ph.D. Student]** and

New York University, Communicative Sciences and Disorders, 665 Broadway, New York, 10012 United States & The Nathan S. Kline Institute for Psychiatric Research, Geriatrics Department, 140 Old Orangeburg Rd., Orangeburg, 10962 United States

**Dr. Diana Van Lancker Sidtis, PhD., CCC/SLP**

New York University Steinhardt, Department of Communicative Sciences, 665 Broadway, New York, 10012 United States & Nathan Kline Institute, Geriatric Division, 140 Orangeburg Road, Orangeburg, 10962 United States

### Abstract

**Background**—Studies of productive language in Alzheimer's disease (AD) have focused on formal testing of syntax and semantics but have directed less attention to naturalistic discourse and formulaic language. Clinical observations suggest that individuals with AD retain the ability to produce formulaic language long after other cognitive abilities have deteriorated.

**Aims**—This study quantifies production of formulaic expressions in the spontaneous speech of individuals with AD. Persons with early- and late-onset forms of the disease were compared.

**Methods & Procedures**—Conversational language samples of individuals with early- (n = 5) and late-onset (n = 6) AD and healthy controls (n = 5) were analyzed to determine whether formulaic language, as measured by the number of words in formulaic expressions, differs between groups.

**Outcomes & Results**—Results indicate that individuals with AD, regardless of age of onset, used significantly more formulaic expressions than healthy controls. The early- and late-onset AD groups did not differ on formulaic language measures.

**Conclusions**—These findings contribute to a dual process model of cerebral function, which proposes differing processing principles for formulaic and novel expressions. In this model, subcortical areas, which remain intact into late in the progression of Alzheimer's disease, play an important role in the production of formulaic language. Applications to clinical practice include identifying preserved formulaic language and providing informed counseling to patient and family.

### Keywords

Formulaic language; Alzheimer's disease; Dual-process model

---

Alzheimer's disease (AD) is a neurological degenerative disease that leaves those affected confused and often unable to communicate appropriately. In addition to loss of memory, AD is known for a decline in productive language (Blair, Marczinski, Davis-Faroque & Kertesz, 2007; Kemper, Greiner, Marquis, Prenovost & Mitzner, 2001; Kempler, 1995). Various components of language have been examined in order to specify the nature of the expressive language decline, with the literature predominantly showing a decline in semantically intact

communication reflected in reduced productive vocabulary (Altmann & McClung, 2008). In contrast, syntax has been identified as an aspect of language that is relatively preserved in AD (Kempler & Zelinski, 1994), although some controversy in this area remains (Bates, Harris, Marchman & Wulfeck, 1995).

Less well researched are anecdotal clinical reports of preserved formulaic language production in the naturally occurring conversation of people with Alzheimer's disease. Formulaic language, comprised of conversational speech formulas, idioms, pause-fillers, and other fixed expressions known to the native speaker, plays a very important role in everyday language use (Van Lancker, 1973, 1988; Wray, 2002; Van Lancker Sidtis, 2004, Fillmore, 1979; Wray & Perkins, 2000; Pawley & Syder, 1983). Persons in the late stages of AD, who are no longer able to communicate verbally about themselves or their lives, are commonly observed to utilize formulaic expressions such as "It's nice to see you again" and "Excuse me" with normal fluency, leading to an impression of conversational competence. Better understanding of a selective preservation in AD of this important component of communicative competence will contribute meaningfully to the clinical picture of AD and will shed light on brain-behavior models of language.

## Formulaic Language in Alzheimer's disease

Formulaic language constitutes a significant proportion of expressive language, with estimates for conversational speech at 24% (Van Lancker Sidtis & Rallon, 2004; Kuiper, 2009). Formulaic language is comprised of fixed expressions that are known to the native speaker, and includes idioms, proverbs, speech formulas/conventional expressions, expletives and pause-fillers. The use of formulaic expressions in neurologically disordered populations is of increasing interest in recent years with a focus on aphasia secondary to stroke (Code, 1989; Van Lancker Sidtis & Postman, 2006; Sidtis, et al., 2009). Studies of formulaic language in left- and right-hemisphere damage indicate that formulaic language production is abundant in those persons with left-hemisphere damage, but reduced with right-hemisphere lesions (Van Lancker Sidtis & Postman, 2006; Sidtis, et al., 2009). Empirical studies have not yet been extended to the Alzheimer's population, but in persons with AD, descriptive study and clinically observations suggest that production of formulaic expressions is relatively preserved. In one study, formulaic language was implicated as a contributing factor to deficient semantic content in AD language (see Nicholas, Obler, Albert & Helm-Estabrooks, 1985). Davis and colleagues (Davis, 2005; Maclagan, Davis, & Lunsford, 2008; Davis & Maclagan, 2009) provide descriptive accounts of the preserved ability to produce formulaic expressions including pause fillers, fixed expressions, discourse markers, and other formulaic sequences in Alzheimer's disease for a variety of functions. Documenting the status of formulaic language in AD speech may contribute to an accurate linguistic profile for this population.

This question is of interest also because previous studies have shown that basal ganglia strokes lead to a significant diminishment in formulaic expressions (Speedie, Wertman, Tair & Heilman, 1993; Sidtis, Canterucci & Katsnelson, 2009), implicating a role of subcortical nuclei in normal function. Supportive data come from preliminary studies of persons with Parkinson's disease, in whom spontaneous speech samples reveal a significantly lower proportion of formulaic expressions (Rogers, Sidtis, & Sidtis, 2009). In contrast to these cases of subcortical stroke and basal ganglia dysfunction due to Parkinson's disease, Alzheimer's disease affects the cortical expanse sparing basal ganglia until late in the disease progression. Confirmation of preserved formulaic language in AD could lend support to a dual process model of language, whereby novel, newly generated language is modulated by the left hemisphere and formulaic expressions rely in large part on right hemisphere-subcortical circuitry (see Van Lancker Sidtis 2004, 2012).

## Early-onset and late-onset AD

Our speech samples obtained from previous research studies included the possibility of comparing early with late onset AD. Comparisons of older and younger individuals with AD are only sparsely represented in scientific studies. There is some evidence that people who acquire AD at an early age may exhibit different language (Filley, Kelly, & Heaton, 1986) and psychological (Toyota, et al., 2007) symptoms when compared to people who acquire AD at a later age. Arising primarily from evidence of biological and genetic differences between individuals with early- and late-onset AD, there is debate about the true nature of these two groups (see Cummings, Vinters, Cole & Khachaturian, 1998, for a review).

## Summary

The use of formulaic expressions has gained scientific interest in the past decade, and most effort has been directed toward healthy speakers. Less is known about disordered language. Recent studies of persons with neurological impairment implicate a right hemisphere-subcortical circuit in the production of formulaic expressions (Van Lancker Sidtis, 2012). In AD, clinical observations point strongly to preserved production of formulaic expressions, which, when unrecognized, may mask veridical cognitive and language deficiencies. Outside of anecdotal reports, little is known about the production of formulas in people with AD. Still less about how production may differ between early- and late-onset AD. Examination of the spontaneous speech of individuals with AD is also of interest to models of language processing in the brain, because a major portion of basal ganglia and frontal lobes remains relatively intact as the cortical disease progresses. Empirical comparisons of language competence between the early- and late-onsets groups are few, and some studies indicate group differences, but how much and in what ways they differ linguistically has not yet been fully examined.

## Aims of this study

The primary aim of this retrospective study was to empirically assess how people with Alzheimer's disease compare with healthy control participants in their production of formulaic expressions. A secondary aspect of this study was to compare people with early- and late-onset AD on formulaic language to explore the proposal that early-onset AD entails greater verbal deficits than late-onset AD. Based on clinical reports, it is hypothesized that individuals with AD will use more formulaic expressions than healthy adults. For early- and late-onset predictions, it is expected that greater verbal deficits will emerge for participants with early-onset AD with proportionally more formulaic language than those with late-onset AD.

## Method

### Participants

This study evaluated the content of the speech produced by a group of participants with probable Alzheimer's disease (AD) who were the subject of previous reports by Glosser and her colleagues (Glosser & Kaplan, 1989; Glosser & Deser, 1990; Glosser & Friedman, 1991). Eleven participants diagnosed as having probable AD and no other known medical or mental health problems (3 males, 8 females), and five healthy control participants (3 males, 2 females;  $M$  age = 63.60 years,  $SD$  = 13.79) with no medical or mental health problems were studied. The diagnosis of probable AD (McKhann, et al., 1984) was supported for all participants by thorough medical, neurological, neuropsychological, and neurodiagnostic evaluations that assured dementia symptoms could not be attributed to any other neurological, psychiatric or medical problem. All participants were followed by physicians

for one year following diagnosis to confirm progressive gradual decline in cognition and memory characteristic of AD, and no participants had a history of drug or alcohol abuse, or other psychiatric, neurologic or cerebrovascular disease (Glosser & Kaplan, 1989; Glosser & Deser, 1990; Glosser & Friedman, 1991). Speech and language function was evaluated by certified speech-language pathologists. Testing was consistently accomplished by members of the research team of clinicians conducting this project. Institutional review board approval was obtained prior to collecting speech samples. Further information about participant language abilities was not provided with the speech samples; however, thorough linguistic analysis of syntax and word frequency was conducted during another phase of this study. Using text-frequency analysis and a clausal complexity-based analysis, lexical frequency and syntactic complexity were examined. AD subjects ( $n = 11$ ) were found to use significantly fewer low-frequency (highly specific) words than HC,  $t(14) = 2.28, p = .039$ , but the groups did not differ on the production of syntactically complex structures,  $t(14) = -0.77, ns$ . Both of these findings are in concordance with the majority of the literature on language in AD (Garrard, Maloney, Hodges & Patterson, 2005; Murdoch & Chenery, 1987; Kempler & Zelinski, 1994; Cummings & Benson, 1992), supporting, for our subject selection criteria, the claim that this is a representative sample of the AD population.

The Alzheimer's (AD) group was separated into 5 early-onset Alzheimer's (EO) participants ( $M$  age at testing = 53.80 years,  $SD = 5.54$ ) and 6 late-onset Alzheimer's (LO) participants ( $M$  age at testing = 85.50 years,  $SD = 3.94$ ) groups. In order to characterize language comparing early- and late-onset AD participants, additional analyses of lexical frequency and syntactic complexity were performed using mixed ANOVAs. Results of an ANOVA of lexical frequency revealed a significant interaction between subject group (EO vs. LO vs. HC) and lexical frequency (high vs. low),  $F(2,13) = 4.33, p = .036$ , with post-hoc analysis indicating that the LO used significantly fewer low-frequency words than the healthy control group ( $p = .046$ ), with the EO group falling between LO and HC. An additional ANOVA exploring syntactic complexity (simple vs. complex) and subject group (EO vs. LO vs. HC) did not reveal significant group differences.

The Alzheimer's group as a whole ranged in cognitive severity from mild- to moderately-impaired based on their *Mini Mental State Examination* (MMSE) scores, as participants ranged from 13 to 20 out of a total possible 30 (see Folstein, Folstein & McHugh, 1975), with the exception of one participant with an MMSE of 7, indicating severe impairment. Years post-diagnosis ranged from 1.5 to 4 and education ranged from 8 to 16 years. Healthy control participants were recruited as volunteers for other psychological experiments or were relatives of the individuals with AD. Healthy participants were interviewed to assure no history of psychiatric illness, alcohol or drug abuse, neurological or medical conditions, and to rule out that they were not taking psychoactive medications (Glosser & Kaplan, 1989; Glosser & Deser, 1990; Glosser & Friedman, 1991). For demographic information providing subject information and group mean values for gender, years of education, years post-diagnosis, and MMSE scores see Table 1.

## Materials and Procedures

The speech samples were obtained retrospectively through an archived study conducted in the 1980s in the Northeastern USA by the team of psychologists, speech pathologists, and neurologists referred to above. Participants were audiorecorded during an interview (following the guidelines in the Boston Diagnostic Aphasia Examination, Goodglass & Kaplan, 1983) that typically lasted between 5 and 20 minutes, in which participants talked about their family and/or careers without time restriction (Glosser & Kaplan, 1989; Glosser & Deser, 1990; Glosser & Friedman, 1991). Total word counts, varying between 235 and 1090 words, are presented in Table 1. Participation by the interviewer was minimal,

containing occasional utterances to encourage further conversation (for example, “uh huh” and “what else did you do”).

### Formulaic Language Analysis

Formulaic language is comprised of fixed expressions that are known to the native speaker, often serve a social function in naturalistic conversation, and include idioms, proverbs, speech formulas/conventional expressions, expletives and pause-fillers. They are characterized by stereotyped form (words, word order, prosody), conventional meanings (usually nonliteral and with nuances and connotations), and high relevance to social context. A method for identifying and quantifying formulaic expressions using formal and functional criteria has been established in previous published studies (see Van Lancker Sidtis & Rallon, 2004; Sidtis et al., 2009). For this study, utterances were categorized as conversational speech formulas (*let me see, you know, first of all*), idioms (*he's at the end of his rope*), proverbs (*a stitch in time saves nine*), discourse elements (*so, and, oh*), pause-fillers (*uh, um*), expletives (*shoot, darn*), sentence stems (fixed expressions that start a sentence: *I guess, I think*), or formulaic distortion errors (*If I can't say anything pleasant, just keep quiet*, an erroneous production of, *If you can't say anything nice don't say anything at all*) by two trained, native English-speaking researchers blind to the demographic data of the participant groups and the purpose of the study. When categorizations differed between the two raters, a third rater mediated the decision.

The total number of words in formulas in proportion to the total number of words in the sample was calculated for each participant and used for between-group comparisons. In order to accommodate the varying sizes of the speech samples, the measure for this phase of the study was the proportion of words in formulaic expressions in each speech sample (the number of words in formulas divided by the total number of words in the sample). Further subcategories were analyzed (conversational speech formulas, idioms/proverbs, discourse elements, pause-fillers, sentence stems, and formulaic distortion errors) and compared between groups using a metric of the percent of words in each formula type out of the total number of words in formulaic expressions.

## Results

### Demographic Comparisons

Data were collected from 11 individuals with AD [5 with early-onset (EO) and 6 with late-onset (LO)] and 5 healthy adults (HC). For all demographic data, nonparametric comparisons were used to determine between-group differences, as the normality and homogeneity of variance assumptions were violated for some variables and subject groups.

To determine group differences for age and education, a series of Kruskal-Wallis tests were conducted. Significant group differences were found for age,  $H(2) = 11.47, p = .003$ , and education,  $H(2) = 9.61, p = .008$ . Post-hoc tests for age using Mann-Whitney's  $U$  with alpha adjusted to .0167 for multiple comparisons indicate that the LO group was significantly older than the EO group,  $U = 0.00, z = -2.74, p = .004$ , and the HC group,  $U = 0.00, z = -2.74, p = .004$ , but the EO and HC groups did not differ,  $U = 5.50, z = -1.47, ns$ . Additional post-hoc tests with Mann-Whitney's  $U$  with alpha adjusted to .0167 for multiple comparisons revealed that for years of education, EO and LO did not differ significantly,  $U = 4.00, z = -2.11, p = .052$ , nor did EO and HC,  $U = 4.00, z = 0.183, ns$ , but LO had significantly fewer years of education than HC,  $U = 1.00, z = -2.65, p = .009$ . Pearson's chi square revealed that gender did not differ significantly between the groups,  $\chi^2(2) = 1.78, ns$ . Within individuals with AD, Mann-Whitney's  $U$  was used to determine whether the EO and LO groups differed on *MMSE* scores and time duration since AD diagnosis. EO and LO did

not differ significantly for mean *MMSE* scores,  $U = 6.00$ ,  $z = -1.69$ , *ns*, or for duration since AD diagnosis,  $U = 12.5$ ,  $z = -0.50$ , *ns*. For means, standard deviations and other subject group data, see Table 1.

### Quantitative results for formulaic language

Results of the Shapiro-Wilk test of normality and Levene's test of homogeneity of variance indicate that the EO, LO, and HC groups' data on percent of words in formulaic expressions are normally distributed,  $W(5) = 0.88$ , *ns*,  $W(6) = 0.98$ , *ns*,  $W(5) = 0.81$ , *ns*, respectively, and variances are homogeneous, Levene's  $F(2,13) = 0.34$ , *ns*, supporting parametric statistics for between-group comparisons. To determine if individuals with AD produce more formulaic expressions than healthy adults, a one-way *ANOVA* was performed. Results revealed that there was a significant and strong effect of subject group on the percent of words in formulaic expressions,  $F(2,13) = 12.10$ ,  $p = .001$ ,  $\eta^2 = .651$ . Post-hoc comparisons using Tukey's HSD indicate that both EO participants ( $M = 37.43\%$ ,  $SD = 6.42$ ),  $p = .001$ , and LO participants ( $M = 33.97\%$ ,  $SD = 4.68$ ),  $p = .006$ , used significantly more words in formulaic expressions than healthy controls ( $M = 21.79\%$ ,  $SD = 4.81$ ), but the EO and LO groups did not differ significantly from each other. (see Figure 1 and Table 2).

A series of one-way *ANOVAs* revealed no group differences for the production of any of the specific formula types (speech formulas, proverbs/idioms, discourse units, pause fillers, expletives and sentence stems). See Table 2 for means and standard deviations.

### Qualitative results for formulaic errors

As a qualitative observation, several distinctive formulaic errors appeared in the AD speech samples. These errors were distortions of formulaic expressions (*If I can't say anything pleasant, just keep quiet*; likely morphed from, "If you can't say anything nice, don't say anything at all"; and *Put down my mind to it*, from, "Put my mind to it"). These errors from the speakers with AD differ from speech error blends on formulas reported for healthy adults (e.g., "*Talks to my heart strings*," Kuiper, Van Egmond, Kempen & Sprenger, 2007, p. 337) and have not been reported before. Healthy control participants in this study did not have any instances of formulaic expression error blends or formula distortion errors. It can be speculated that while participants with AD have intact basal ganglia, allowing retention of formulaic expressions long into the progression of their disease, cortical dysfunction affects the shape of the formulaic expressions in production. This perspective conforms to a process model of language function, whereby formulaic expressions are modulated by a basal ganglia-right hemisphere-cortical circuit. It is a reasonable speculation that right hemisphere cortical dysfunction contributes to the unusual formulaic distortions.

The following is an excerpt of speech that is selected to be representative of the current study's sample of speakers with AD.

(Discussing a son's job) *But uh...but he does work in the business. Well he...he's kind of uh...not kind of but he...he is a manager. And uh, oh, he...he does...uh. He does a lot of going around and see that the stores are stocked with...have what they have to. And you know. And that, that sort of thing..*

The speech sample above is typical of the speech samples of individuals with AD in this study, as it shows relatively preserved syntax, high frequency vocabulary, and a large proportion of formulaic expressions, here in the form of sentence stems (*but, and*), pause-fillers (*uh, oh, you know*), conventional expressions (*that sort of thing, what they have to*) and discourse elements (*well*). Of the 60 words in this sample, 34 (57%) are words in formulaic expressions.

## Discussion

This study sought to clarify how formulaic language is affected by AD by specifically assessing the proportion of formulaic language produced. Results support clinical observations and the hypothesis under study, indicating that people with AD use more formulaic language than healthy adults. Finally, based on the results of this study, individuals with early- and late-onset AD do not differ linguistically in their use of formulaic language. Results from this study of formulaic expression production do not lend support to a previous claim that early-onset AD is characterized by greater linguistic impairment (Filley, et al., 1986).

An effect of age on the results cannot be fully ruled out, as the healthy control group was younger than the older onset AD group. We know of no research comparing older and younger subjects on formulaic language proportions in spontaneous speech, as designed in this study. However, a related, recent study concerning age differences in healthy individuals suggested a mild reduction in idiom production tasks in older persons (Conner et al., 2011). In our study, however, no significant differences in proportion of formulaic expressions were seen between the early- and late-onset groups, who also differed significantly in age. Further, the early-onset (younger) AD participants yielded a (not significantly) higher proportion of formulaic expressions (which is the metric of deficit in this study). Therefore, a major effect of age is not apparent. This is an area of interest for future research.

Given that formulaic expressions number in the hundreds of thousands (Kuiper, 2009; Jackendoff, 1995), there is likely great variability in how and to what extent this ability is manifest in any individual person. As has been observed in persons with aphasia, we might expect individuals with AD to have radically varied repertoires of routinized expressions. The varied ranges of expressions available for use may obscure the extent to which persons with language disturbances are utilizing formulaic language in the clinical setting. From his original observations in persons with severe aphasia, Jackson (1874/1958) records such brief yet routinized expressions as, *Oh dear, Bless my life, Take care, and Goodbye*. Riese (1949) (cited in Critchley, 1970, p. 269) described an aphasic patient who regained speech ability, but spoke mainly using long Shakespearean expressions: *Me thinks the lady doth protest too much*. In a detailed chronicle of her husband's decline with AD, Hoblitzelle (2008, p. 208) described numerous instances of the production of lengthy poetic passages or literary expressions, which, as an English professor, he had previously memorized and often used. These lengthy, stored utterances continued to be produced even when verbal ability for propositional language was gravely diminished in the end stage of the disease:

*Be near me when my light is low, When the blood creeps, and the nerves prick And tingle, and the heart is sick, And all the wheels of Being slow.* (Alfred Lord Tennyson)

The preservation of formulaic language in AD may be accounted for by considering the role of the basal ganglia in action forms and sequencing based on procedural memory, as formulaic language can be considered routinized semi-automatic verbal motor gestures. Further research is necessary to discover other procedural and holistic memory functions that may remain unscathed in AD, while found to be impaired in cases of subcortical damage. Comparable procedural functions are piano playing, games, and sports. Beatty et al. (1994) found that severely impaired persons with AD not only retained the ability to play musical instruments, but also engaged in contract bridge, dominoes and canasta.

Finally, results on preserved production of formulaic language in Alzheimer's speech support hypotheses regarding the neurological substrates of formulaic language.

Neuropsychology has long associated the left hemisphere with the production of novel speech, while newer studies show that formulaic language (see Van Lancker Sidtis, 2004; 2012) is modulated in the right hemisphere and in subcortical nuclei, regions responsible for modulation of movement, emotion, and the establishment of procedural memory. As noted in the introduction, individuals who have subcortical damage due to stroke or Parkinson's disease show reduced incidence of formulaic language in their speech compared to matched healthy controls (Speedie et al., 1993; Van Lancker Sidtis, 2004; Sidtis et al., 2009, 2012; Rogers et al., 2009). Alzheimer's disease is a cortical dementia primarily affecting the temporal and parietal lobes and then progressing to the frontal lobes (Cummings & Benson, 1992), sparing subcortical regions until the late stages. The participants with AD in the present study had a significantly higher proportion of formulaic language than healthy controls, which provides further evidence for a subcortical involvement in the production of non-novel, formulaic language. Alzheimer cortical disease likely contributes to the unusual distortions of formulaic expressions described above. The functions of procedural memory and routinized gestures, known to reside in the basal ganglia (Graybiel, 1998), are compatible with a significant preservation of formulaic expressions in contrast to impairment of informational, propositional language.

The overuse of formulaic language by individuals with AD may account in part for the other signature linguistic characteristics associated with this group. One is semantic deficiency or the empty speech described for AD. Formulaic expressions are commonly used by speakers for social purposes and do not necessarily add propositional meaning. Further, the prevalence of formulaic expressions may contribute to conclusions about syntactic preservation in these persons, which are prevalent in the literature (Murdoch & Chenery, 1987; Kempler & Zelinski, 1994; Garrard, Maloney, Hodges & Patterson, 2005). Many (not all) conventional expressions, speech formulas, idioms and proverbs are produced with apparent grammatical structure and therefore may yield an appearance of syntactic well-formedness. The presence of formulaic utterances is characteristic of healthy speech, but at approximately 25% of total talk, compared to approximately 35% in the AD groups. Having formulaic utterances preserved in increased proportion may give an impression of intact communication, while substantial novel information is reduced.

The distinction between novel and formulaic language is important in evaluation and treatment of language disorders. This information is valuable for clinicians and caretakers, assisting them in determining optimal interaction, management, evaluation, and treatment for those with AD. Preservation of formulaic language in AD may mask other cognitive and language deficits and contribute to under-recognition of disease at home and in the clinic. Deft and consistent use of formulaic language in social contexts can give an erroneous impression of intact language and cognition. For the clinician, recognizing the overuse or exclusive use of formulaic expressions in the patient with AD is crucial to providing an accurate evaluation of patients' communicative competence and in providing informed counseling to patients and families. Changes in formulaic language use in different neuropathological settings (e.g., subcortical versus cortical) have significant implications for both language theory and therapy and warrant additional study with larger subject groups.

## Acknowledgments

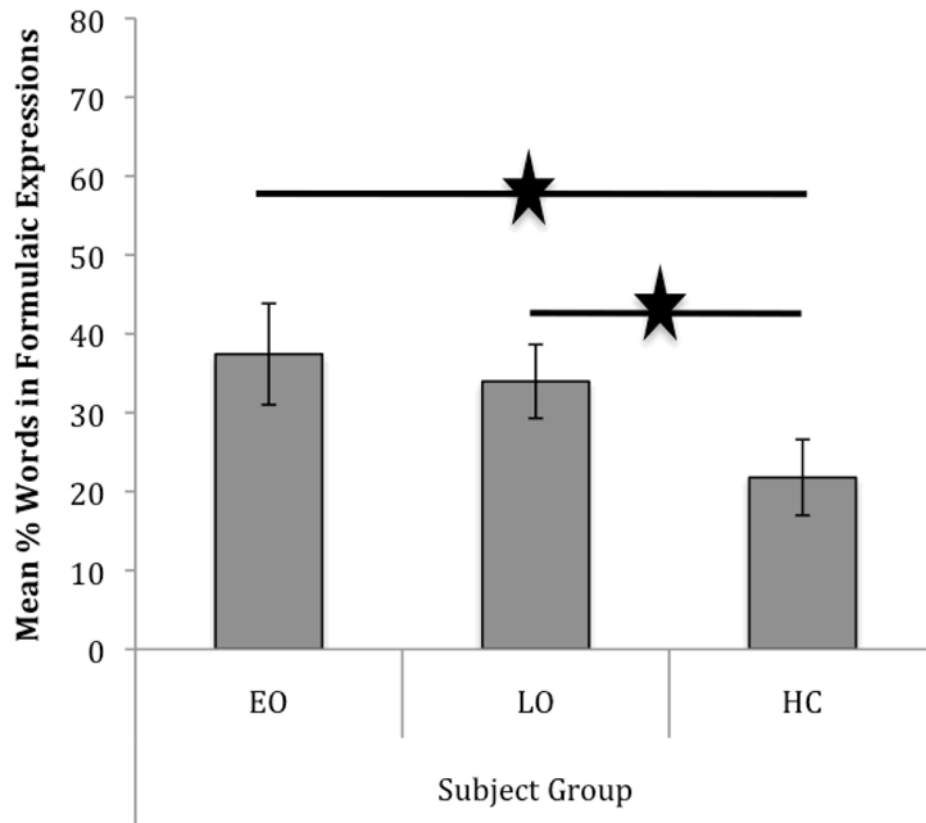
This work was supported by NIDCD grant R01 DC007658. The funding source had no involvement in collection, interpretation, analysis, writing the report, or submitting this report for publication. We appreciate the advice and support of John J. Sidtis throughout this study. Harriet Klein, Sharon Antonucci, Victoria Zeldin, Mila Rotenberg, Zoya Peysakov, and Zepur Dovlatyan contributed generously at various stages of the project.



## References

- Altmann LJP, McClung JS. Effects of semantic impairment on language use in Alzheimer's disease. *Seminars in Speech and Language*. 2008; 29:18–31. [PubMed: 18348089]
- Bates E, Harris C, Marchman V, Wulfeck B. Production of complex syntax in normal ageing and Alzheimer's disease. *Language and Cognitive Processes*. 1995; 10:487–539.
- Beatty WW, Winn P, Adams RL, Allen EW, Wilson DA, Prince JR, et al. Preserved cognitive skills in dementia of the Alzheimer type. *Archives of Neurology*. 1994; 51:1040–1046. [PubMed: 7945001]
- Blair M, Marcuzinski CA, Davis-Faroque N, Kertesz A. A longitudinal study of language decline in Alzheimer's disease and frontotemporal dementia. *Journal of the International Neuropsychological Society*. 2007; 13:237–245. [PubMed: 17286881]
- Code C. Neurolinguistic analysis of recurrent utterance in aphasia. *Cortex*. 1982; 18:141–152. [PubMed: 6197231]
- Critchley, M. *Aphasiology and other aspects of language*. London: Edward Arnold, Ltd; 1970.
- Connor PS, Hyun J, O'Connor Wells B, Anema I, Goral M, Monéreau-Merry MM, Rubino D, Kuckuk R, Obler LK. Age-related differences in idiom production in adulthood. *Clinical Linguistics & Phonetics*. 2011; 25:899–912. [PubMed: 21728830]
- Cummings, JL.; Benson, DF. *Dementia, a clinical approach*. 2. Woburn, MA: Butterworth-Heinemann; 1992.
- Cummings JL, Vinters HV, Cole GM, Khachaturian ZS. Alzheimer's disease: Etiologies, pathophysiology, cognitive reserve, and treatment opportunities. *Neurology*. 1998; 51:S2–S15. [PubMed: 9674758]
- Davis, BH. So you had two sisters, right? Functions for discourse markers in Alzheimer's talk. In: Davis, BH., editor. *Alzheimer talk, text and context: Enhancing communication*. Hampshire, UK: Palgrave Macmillan; 2005. p. 128-145.
- Davis BH, Maclagan M. Examining pauses in Alzheimer's discourse. *American Journal of Alzheimer's Disease & Other Dementias*. 2009; 24(2):141–154.
- Filley CM, Kelly J, Heaton RK. Neuropsychologic features of early- and late-onset Alzheimer's disease. *Archives of Neurology*. 1986; 43:574–576. [PubMed: 3718284]
- Fillmore, CS. On fluency. In: Fillmore, J.; Kempler, D.; Wang, WS-Y., editors. *Individual differences in language ability and language behavior*. London: Academic Press; 1979. p. 85-101.
- Folstein MF, Folstein SE, McHugh PR. Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*. 1975; 12:189–198. [PubMed: 1202204]
- 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]
- Glosser G, Deser T. Patterns of discourse production among neurological patients with fluent language disorders. *Brain and Language*. 1990; 40:67–88. [PubMed: 2009448]
- Glosser G, Friedman RB. Lexical but not semantic priming in Alzheimer's disease. *Psychology and Aging*. 1991; 6:522–527. [PubMed: 1777139]
- Glosser G, Kaplan E. Linguistic and nonlinguistic impairments in writing: A comparison of patients with focal and multifocal CNS disorders. *Brain and Language*. 1989; 37:357–380. [PubMed: 2804619]
- Goodglass, H.; Kaplan, E. *The assessment of aphasia and related disorders*. Philadelphia: Lea & Febiger; 1983.
- Graybiel AM. The basal ganglia and chunking of action repertoires. *Neurobiology of Learning and Memory*. 1998; 70:119–136. [PubMed: 9753592]
- Hoblitzelle, OA. *Ten thousand joys and ten thousand sorrows*. New York: Penguin Group; 2008.
- Jackendoff, R. The boundaries of the lexicon. In: Everaert, M.; van der Linden, EJ.; Schenk, A.; Schreuder, R., editors. *Idioms, structural and psychological perspectives*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1995. p. 133-169.
- Jackson, JH. On the nature of the duality of the brain. In: Taylor, J., editor. *Selected writings of John Hughlings Jackson*. Vol. 1. London: Hodder and Stoughton; 1874/1958. p. 129-145.

- Kemper S, Greiner LH, Marquis JG, Prenovost K, Mitzner TL. Language decline across the life span: Findings from the Nun Study. *Psychology and Aging*. 2001; 16:227–239. [PubMed: 11405311]
- Kempler, D. Language changes in dementia of the Alzheimer type. In: Lubinsky, R., editor. *Dementia and communication: Research and clinical implications*. San Diego: Singular; 1995. p. 98-114.
- Kempler, D.; Zelinski, EM. Language in dementia and normal aging. In: Huppert, FA.; Brayne, C.; O'Connor, DW., editors. *Dementia and Normal Aging*. Cambridge: Cambridge University Press; 1994. p. 331-365.
- Kuiper, K. *Formulaic genres*. UK: Palgrave Macmillan; 2009.
- Kuiper K, Van Egmond M, Kempen G, Sprenger S. Slipping on superlemmas: Multi-word lexical items in speech production. *The Mental Lexicon*. 2007; 2(3):313–357.
- Lieberman P. Human language and our reptilian brain: The subcortical bases of speech, syntax, and thought. *Perspectives in Biology and Medicine*. 2001; 44:32–51. [PubMed: 11253303]
- Maclagan, M.; Davis, B.; Lunsford, R. Fixed expressions, extenders and metonymy in the speech of people with Alzheimer's disease. In: Granger, S.; Meunier, F., editors. *Phraseology: An interdisciplinary perspective*. Amsterdam, The Netherlands/Philadelphia, PA: John Benjamins Publishing; 2008. p. 175-187.
- McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlen EM. Clinical diagnosis of Alzheimer's disease. *Neurology*. 1984; 34:939–944. [PubMed: 6610841]
- Murdoch BE, Chenery HJ. Language disorders in dementia of the Alzheimer type. *Brain and Language*. 1987; 31:122–137. [PubMed: 2437993]
- Pawley, A.; Syder, FH. Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In: Richards, JC.; Schmidt, R., editors. *Language and communication*. London: Longman; 1983.
- Rogers, T.; Sidtis, D.; Sidtis, J. Formulaic language in Alzheimer and Parkinson speech. Poster presented at the 47th Annual Academy of Aphasia Meeting; Boston, MA. 2009.
- Sidtis D, Canterucci G, Katsnelson D. Effects of neurological damage on production of formulaic language. *Clinical Linguistics and Phonetics*. 2009; 23(4):27–284.
- Speedie LJ, Wertman E, Tair J, Heilman KM. Disruption of automatic speech following a right basal ganglia lesion. *Neurology*. 1993; 43:1768–1774. [PubMed: 8414029]
- Toyota Y, Ikeda M, Shinagawa S, Matsumoto T, Matsumoto N, Hokoishi K, et al. Comparison of behavioral and psychological symptoms in early-onset and late-onset Alzheimer's disease. *International Journal of Geriatric Psychiatry*. 2007; 22:896–901. [PubMed: 17343292]
- Van Lancker, D. Language lateralization and grammars. In: Kimball, J., editor. *Studies in syntax and semantics*. Vol. II. New York: Academic Press; 1973. p. 197-204.
- Van Lancker, D. Nonpropositional speech: Neurolinguistic studies. In: Ellis, A., editor. *Progress in the psychology of language*. Hillsdale, NJ: Lawrence Erlbaum; 1988. p. 49-118.
- Van Lancker Sidtis D. When novel sentences spoken or heard for the first time in the history of the universe are not enough: Toward a dual-process model of language. *International Journal of Language and Communication Disorders*. 2004; 39:1–44. [PubMed: 14660185]
- Van Lancker Sidtis, D. Two track mind: Formulaic and novel language support a dual process model. In: Faust, M., editor. *Advances in the neural substrates of language: Toward a synthesis of basic science and clinical research*. London: Blackwell Publishing Ltd; 2012. p. 342-367.
- Van Lancker Sidtis D, Postman WA. Formulaic expressions in spontaneous speech of left- and right-hemisphere damaged subjects. *Aphasiology*. 2006; 20(5):411–426.
- Van Lancker Sidtis D, Rallon G. Tracking the incidence of formulaic expressions in everyday speech: Methods for classification and verification. *Language and Communication*. 2004; 24:207–240.
- Wray, A. *Formulaic language and the lexicon*. Cambridge, UK: Cambridge University Press; 2002.
- Wray A, Perkins MR. The functions of formulaic language: An integrated model. *Language & Communication*. 2000; 20:1–28.



**Figure 1.** Mean percent of words in formulaic expressions for participants with early- (EO) and late-onset (LO) Alzheimer's disease and healthy control (HC) subjects.

Table 1

Demographic data for early- and late-onset Alzheimer's disease groups and healthy control participants.

Participant Group	Age (years)	Gender	Education (years)	Years Post-Diagnosis	MMSE Score	Word Count
Early-onset AD	45	Female	12	3	7	235
	53	Male	11	3	16	314
	55	Female	12	3	13	390
	56	Female	16	3	20	465
	60	Female	13	4	17	305
Early-onset AD Mean (and <i>SD</i> )	53.80 (5.54)	-	12.80 (1.92)	3.20 (0.44)	14.60 (4.92)	
Late-onset AD	81	Female	12	3	16	1090
	82	Male	12	5	20	584
	84	Female	8	1.5	20	652
	86	Female	8	4	18	466
	89	Female	8	3	18	876
91	Male	8	3.5	20	935	
Late-onset AD Mean (and <i>SD</i> )	85.50 (3.94)	-	9.33 (2.07)	3.33 (1.17)	18.67 (1.63)	
Healthy Control	46	Female	18	-	-	362
	56	Male	12	-	-	519
	61	Male	16	-	-	431
	77	Male	16	-	-	242
78	Female	18	-	-	1027	
Healthy Control Mean (and <i>SD</i> )	63.60 (13.79)	-	16.00 (2.45)	-	-	

Note: MMSE scores range from 0 to 30 with "normal" being 24–30, mild cognitive impairment ranging from 20 to 23, moderate cognitive impairment between 10 and 19 and severe cognitive impairment ranging from 0 to 9.

**Table 2**

Mean percent (and standard deviations) of words in formulaic expressions (and formula types taken out of the total number of words in formulas) used by individuals with early- and late-onset Alzheimer's disease (EO, LO) and healthy control (HC) participants

	<b>EO</b>	<b>LO</b>	<b>HC</b>
Total in formulaic expressions	37.43 (6.42)	33.97 (4.68)	21.79 (4.81)
Speech formulas	53.31 (16.03)	56.84 (13.90)	42.84 (14.73)
Proverbs or idioms	0.70 (1.56)	5.30 (7.16)	6.54 (12.17)
Discourse units	10.12 (3.93)	13.13 (5.23)	10.83 (2.28)
Pause fillers	7.24 (6.43)	1.08 (2.41)	4.12 (4.16)
Expletives	0.17 (0.39)	0.92 (2.26)	0.00 (–)
Sentence stems	28.51 (9.51)	21.59 (7.33)	33.27 (13.74)
Formulaic errors	2.64 (1.69)	3.61 (2.67)	0.00 (–)