Published in final edited form as:

Curr Protoc Neurosci. 2012 January ; CHAPTER: Unit10.5. doi:10.1002/0471142301.ns1005s58.

A Rubric for Extracting Idea Density from Oral Language Samples

Vineeta Chand¹, Kathleen Baynes², Lisa M. Bonnici³, and Sarah Tomaszewski Farias¹

¹University of California at Davis, Department of Neurology, Sacramento

²University of California at Davis, Center for Mind and Brain

³University of Aberdeen, Scotland, School of Language and Literature

Abstract

While past research has demonstrated that low idea density (ID) scores from natural language samples correlate with late life risk for cognitive decline and Alzheimer's disease pathology, there are no published rubrics for collecting and analyzing language samples for idea density to verify or extend these findings into new settings. This paper outlines the history of ID research and findings, discusses issues with past rubrics, and then presents an operationalized method for the systematic measurement of ID in language samples, with an extensive manual available as a supplement to this article (Analysis of Idea Density, AID). Finally, reliability statistics for this rubric in the context of dementia research on aging populations and verification that AID can replicate the significant association between ID and late life cognition are presented.

Keywords

idea density; Alzheimer's Disease; semantics; language; cognitive reserve

Introduction

Idea density (ID) is a conceptually appealing means of measuring the semantic content of language, and has been used in longitudinal and diagnostic cross-sectional studies of cognitive aging and Alzheimer's disease (AD). Expressive language in AD is characterized as progressively more semantically 'empty' and syntactically simplified yet, until late stages, grammatically well formed (Kemper et al., 1993), and past studies suggest that ID provides an early life linguistic predictor of late life cognitive decline and dementia.

Using various methods to measure ID, previous research has shown that low ID measured early in life is associated with a host of negative outcomes as people age, i.e. increased risk of cognitive impairment in late adulthood (Snowdon et al., 1996), decreased brain weight and increased degree of cerebral atrophy (Riley et al., 2005), increased AD neuropathology on postmortem exam (Snowdon et al., 1996), and increased all-cause mortality (Snowdon et al., 1999). While the majority of ID findings are based on the longitudinal Nun Study, the correlation between early life low ID and late life cognitive impairment has recently been replicated in a different population which also includes men (Engelman et al., 2010). Higher ID has also been shown to modify the impact of AD pathology on cognition; that is, in those with higher ID, AD pathology had less impact on cognition (Iacono et al., 2009).

Collectively, this research suggests that early life ID may be a proxy for cognitive reserve (Engelman et al., 2010; Snowden et al., 2000): the cognitive reserve theory posits that subjects with greater reserve evidence a compensatory response to disease-based damage and maintain intact cognition for longer than those with low reserve (Stern, 2002; Stern, 2003). From this perspective, ID may have potential as a measurable variable which mediates dementia-based decline trajectories. Yet, while ID merits further exploration, independent replication and expansion of past ID findings is challenging because formal standardized methods by which to collect and extract ID have not been published.

In this paper we discuss how ID has emerged conceptually, how it has been traditionally defined, how it correlates with clinical and pathology measures of AD, and the issues that motivated the development of this rubric, Analysis of Idea Density (AID). We next introduce AID as a detailed method for collecting language samples and extracting ID. By drawing on linguistic insights and by offering extensive coding guidelines, AID allows for easier and more consistent measurement of ID while maintaining a focus on semantic content over verbosity or grammatical structure. We document how AID operationalizes ID coding, discuss theoretical and practical motivations which underlie the AID coding, and then demonstrate AID coding reliability and the results of AID-based ID in association with a cross-sectional study of clinical measures of dementia. This paper is intended as an introduction and overview of AID goals, methods and reliability, while the complete AID rubric is available as a supplement.

Background

The conceptual underpinnings of ID

At the most basic level, the goal in measuring ID is to quantitatively operationalize how much information is conveyed in natural language (either oral or written) per number of words. It is thus a measure of the efficiency in which one communicates ideas.

Previous methods for measuring ID were derived from theories of semantic representation. Kintsch (1974) who believed that word count and syntactic complexity were not psychologically real measures of memory load, developed a theory relating semantic content to memory load to examine how semantic complexity influences recall accuracy, and hence, measure memory. This approach was developed into an explicit rubric—termed propositional analysis—by Turner and Greene (1977) for deriving the propositional content (which they labeled the propositional text base) for prose language. Propositional analysis, based on semantic case grammar (Fillmore, 1968; Fillmore, 1969), captures conceptual units which consist of a relation and a number of arguments connected through that relation. Operationally, propositional analysis compares the propositional content of participants' story retellings to a baseline propositional count derived from the propositional analysis of the original story presented to participants. While memory research in this vein continues (e.g., Long et al., 2005; Miller, 2003) and measurements of Content Information Units (CIUs) collected in response to traditional psychometric picture-description tasks like the Cookie Theft (e.g., Ehrlich et al., 1997; Kave and Levy, 2003), one byproduct of Turner and Greene's rubric is the subsequent reinterpretation of propositional density as a measure of linguistic ability (Kemper, 1992). Propositional density—also known as P-Density and idea density (ID, used hereafter)—analyzes novel language input (i.e. original language samples, not the retelling of a previously presented story or picture) for semantic content to derive ID scores as a ratio of propositions to total words.

Previous Methods for Measuring ID

ID has been operationalized in two distinct rubrics. Kemper's Language across the Lifespan coding manual (Kemper, 1993, hereafter LAL) presents a working version of Turner and

Greene's manual analysis of semantic relations to then derive ID as the ratio of propositions to words. More recently, CPIDR, an automated software program, has been introduced (Brown et al., 2008). It determines propositional content based on part-of-speech tagging: specifically, it approximates semantic content by counting selective content words (verbs, adjectives, adverbs, prepositions and subordinating conjunctions), and then adjusting this count based on post-analysis rules. CPIDR ID scores thus represent the ratio of valid content words to the total number of words in a language sample. It is significant that while past rubrics take very different approaches to operationalizing ID, and both have demonstrated significant correlations with various clinical outcomes, they are potentially capturing very different aspects of language, and have never been directly compared.

Past ID Findings—While one cannot directly compare longitudinal and cross-sectional results, below is an overview of past ID findings related to AD and aging outcomes from both types of studies. ID scores—both derived from early and late life, and from written and oral language—have been examined in association with several demographic, cognitive, genetic and pathological measures within older adult sample populations (Table 1).

Age: While every longitudinal study which has compared early and late life ID scores—regardless of language medium (i.e. oral vs. written)—confirms a significant correlation between ID scores across the lifespan (e.g., Snowdon et al., 1996), the specific relationship between aging and change or stability in ID is disputed. Oral samples collected from a cross-sectional age-stratified cohort demonstrate higher ID scores in early versus late life (Kemper and Summer, 2001). Complementing this, a longitudinal analysis of written language samples across a 60 year time span demonstrates that 1) ID scores gradually decline in a uniform linear fashion over the lifespan, 2) participants with a late life clinical dementia diagnosis have significantly lower early life ID scores than those who do not go on to have a clinical dementia diagnosis and 3) the former demonstrate a shallower slope of decline in ID than the latter (Kemper et al., 2001). However, findings comparing early and late life written ID may be limited by attrition: some late life dementia participants were unable to complete the writing task (Kemper et al., 2001).

Education: Snowdon *et al.* (1996) find a significant positive correlation between early life written ID and years of education, but qualify this finding by demonstrating that educational attainment accounts for less variation in ID than late life MMSE scores. Early life written ID does not, however correlate with educational achievement (based on high school English and Math grades) (Kemper et al., 2001). Late life oral ID does not correlate significantly with years of education (Mitzner and Kemper, 2003; Reed et al., 2010) and late life written ID demonstrates conflicting results: Mitzner & Kemper (2003) do not find an association with education, while Kemper *et al.* (1993) find a significant positive correlation between ID and years of education: collectively, these suggest that education and ID may provide unique information.

Late life Cognition and Physical Condition: Higher early and/or late life written ID (defined as the upper two thirds of scores) demonstrate a significant positive correlation with late life cognition (Iacono et al., 2009; Kemper et al., 2001; Kemper et al., 1993; Mitzner and Kemper, 2003; Riley et al., 2005; Snowdon et al., 1996), while late life oral ID shows a non-significant trend towards a positive correlation with cognition (Mitzner and Kemper, 2003). No relationship has been found between late life oral ID and late life physical condition (Mitzner and Kemper, 2003), however late life written ID demonstrates a significant positive correlation with late life physical condition (Mitzner and Kemper, 2003)—here, a floor effect, where it is impossible to collect written samples beyond a certain level of physical disability, is also documented.

Clinical Diagnosis: Comparing early life written ID to clinical diagnoses, there is a significant correlation linking low ID scores to increased risk for MCI diagnosis (Iacono et al., 2009; Riley et al., 2005) and AD diagnoses (Engelman et al., 2010; Iacono et al., 2009; Kemper et al., 2001; Reed et al., 2010), while Kemper *et al.* (1993) found that low late life written ID also demonstrates a significant increased risk for AD diagnosis and a negative linear relationship with AD severity measured through the Clinical Dementia Rating scale (CDR, Morris, 1993).

Genetics and Post-Mortem Pathology: Comparing ID with increased genetic risk of AD, no correlation has been found between early life written ID and the presence of an APOE ϵ^4 allele (Riley et al., 2005). Low early life written ID also correlates significantly with increased all-cause mortality (Snowdon et al., 1996), while low early life ID significantly increases risk for lower post-mortem brain weight (Riley et al., 2005) and autopsy-confirmed AD pathology (Riley et al., 2005; Snowdon et al., 2000; Snowdon et al., 1996). Further, high early life written ID demonstrates a significant positive correlation with asymptomatic AD confirmed upon autopsy (Iacono et al., 2009).

In sum, while the majority of research on ID has used the Nun Study cohort to examine early life written language, more scarce late life oral and written language samples also indicate that lower ID scores increase the risk of late-life cognitive decline, clinical AD diagnosis and AD pathology post-mortem, and the correlation between early life written ID and clinical AD diagnosis has recently been replicated in an entirely different sample (Engelman et al., 2010). Collectively these studies of cognitive aging and dementia suggest that ID is a measure of early neurocognitive development and/or cognitive reserve. ID is also conceptually appealing for clinical investigation because it is less subject to floor effects than, e.g. measures of grammatical complexity (Kemper et al., 2001), and because it has good ecological validity, as a means of investigating natural language data. While the homogeneity and narrow range of populations studied thus far require replication in different samples, ID merits further exploration across the lifespan.

Problematizing Existing ID Rubrics—While LAL-based measures of ID have shown promising correlations with clinical and pathological indicators and do not correlate strongly with educational level or age, the LAL rubric has not been taken up by other researchers for several possible reasons. First, no published explicit and standardized rules on how to extract ideas from written or oral language samples exist. Second, the LAL rubric, as primarily based on Turner & Greene's propositional analysis (1977), requires a more sophisticated knowledge of linguistic structure and formal analytic linguistic frameworks than is likely to be available in clinical settings. Third, there are no formal guidelines on how to address oral language in narrative and conversational genres—that is, how one can analyze natural oral language. Oral language is structurally distinct from written language (Chafe, 1985; Chafe and Tannen, 1987), while informal oral narratives also tend to be much 'messier' than written language or even more formal oral genres. For example, informal oral language is more likely to include more revisions, pauses, repetitions and other dysfluencies. Informal oral samples thus require more explicit guidelines to achieve sufficient coding consistency and to accurately capture semantic content.

Recognizing the pitfalls associated with qualitative human coders and seeking to make the coding of ID less cumbersome, CPIDR (Brown et al., 2008; Covington, 2007) takes a different approach to propositional analysis, relying on a part-of-speech tagger and subsequent recoding rules to ascertain ID scores. However, by relying on an automated part-of-speech tagger (which assigns each word to a syntactic category, e.g. noun, verb) and refinement rules (which broadly group complex phrases as single propositions), several issues arise.

First, this type of analysis can fail to successfully disambiguate between words which can function in more than one syntactic category. As Covington *et al.* (2007) admit, CPIDR fails to code 'orange' as a noun within the context 'an orange and a banana'. Without accuracy scores for CPIDR's part-of-speech assignment, it is unclear whether CPIDR coding decisions and subsequent ID scores reflect a meaningful parse of the narrative.

Second, by focusing on part-of-speech assignment at the word level and then only selectively reconsidering propositional content at the phrasal level within the readjustment rules, CPIDR analytically reframes the concept of ID without reference to the larger semantic context of the narrative—or even utterance. This choice may obscure differences in conciseness and semantic novelty, both of which increase progressively with AD progression.

Third, CPIDR is not able to distinguish on-topic concise language from oral dysfluencies and compensation strategies common to AD progression such as increased lexical and non-lexical fillers, off-topic verbosity, metadiscourse, revisions and distant repetitions (Bayles et al., 1992; Causino Lamar et al., 1994; Croisile et al., 1996; Davis and MacLagan, 2009; Ehrlich, 1994; Garcia and Joannette, 1994; Garrard et al., 2005; Guendouzi and Müller, 2006; Kave and Levy, 2003; Mortensen, 1992; Müller and Guendouzi, 2005; Ramanathan, 1997).

Fourth, given the lack of any strong correlation uncovered in past research linking clinical or pathological criteria for AD with a decline in grammatical complexity (Snowdon et al., 1996), and instead, the relatively late preservation of grammatical structure and complexity in severe AD, CPIDR's consistent scoring of subordinating conjunctions as propositions may capture grammatical complexity when no relational propositional content is present (e.g. *They ran that I went to school*), and/or may unequally impact ID scores across age groups given that semantic competence degrades at a different rate than grammatical competence. These may limit the validity of CPIDR as a measure of novel semantic content and may also affect the degree to which CPIDR-based ID scores can successfully capture decline in ID that is associated with global cognitive decline.

While the relative effectiveness of different ID scoring methods remains an empirical question, it is troubling that propositional content as based on part-of-speech tags may inadvertently ignore linguistic patterns which are hallmarks of AD-based decline and which are also not semantically meaningful and novel. For these reasons, our rubric builds upon LAL to offer more explicit scoring guidelines which can be applied to both oral and written language.

AID

The AID manual was developed to set forth a series of principles and examples that can be used to guide the analysis of late life oral narratives, and is available as an online supplement to this article. Our four guiding principles were: 1) to develop methods that would be clear and easily applied by both linguists and non-linguists, 2) to capture novel semantic content, and not grammatical structure, 3) to provide systematic rules and guidelines in addition to realistic examples to promote consistent and transparent scoring decisions and 4) to base analysis on the textual narrative, and not on a speaker's possible and/or hypothetical intentions. This work is built upon the manuals that preceded it and to the extent that is successful, we are highly indebted to the authors of those manuals.

The AID analytic framework is less dependent upon any formal training in logic or linguistics, and was developed out of a collaborative effort among researchers across several disciplines including neuropsychology and linguistics. This was done to help ensure that the decisions we made were correct relative to current conventions of linguistic structure and

discourse analysis while maintaining a dependence on intuitions that were clear to a native or fluent English speaker. In areas where consistent consensus has been difficult, we have occasionally made arbitrary rules that allow us to facilitate good inter-rater agreement—these are highlighted as such within the manual. While recognizing that some of these decisions are based on syntactic principles and relations, we have tried to maintain a focus on semantic complexity over grammatical relations.

AID thus attempts to remain true to the concept of ID and builds on the LAL guidelines to provide more extensive coding, transcription and word count guidelines, which can facilitate reliable and systematic scoring. However, this attention to semantic content over syntactic structure has forced us to occasionally diverge from LAL and return to Turner and Green's conception of a semantic proposition. At issue here is not only the notion that grammatical complexity is relatively independent of ID and is maintained much further along AD progression (Snowdon et al., 1996), but also that lexical choices can motivate—or require—different structures (Biber et al., 1999).

In order to approach semantic content—and hence ID—without favoring required accompanying grammatical structure, we have drawn on descriptive and corpus linguistic findings of structural requirements across lexical sets in order to develop coding guidelines which are internally consistent, analytically transparent and focused on semantic content over syntactic structure. Further, given the lack of an association between late life oral ID and education (Mitzner and Kemper, 2003; Reed et al., 2010), the positive correlation between late life oral ID and cognition in an ethnically and educationally heterogeneous sample (Reed et al., 2010), the trend towards a positive correlation between high late life ID and late life cognition in a different sample (Mitzner and Kemper, 2003), and the unavailability of similar early life language samples for most populations, AID's focus on the analysis of late life oral data allows further exploration of ID through contemporary narrative data. This focus on late life data collection, oral samples and the non-significant relationship between ID and education may prove especially useful in examinations of less educated and/or more heterogeneous populations for whom standardized cognitive measures are less successful (Baird et al., 2007; Boone et al., 2007; Byrd et al., 2006; Cosentino et al., 2007; Gasquoin, 1999; Manly and Echemendia, 2007). Below, we sketch out our process and principles for oral language data collection, transcription and ID scoring for use with older adults at risk of or already developing dementia and illustrate their application in a sample narrative (Example 1). Appendix A provides a full ID analysis of Example 1.

Example 1. Sample Late life oral narrative based on the prompt “Tell me about your hometown”

I I spent most of my uh growing up years in uh, in growing up in, in the Bronx which now has a somewhat ah se-seedy reputation. But uh, this wasn't. Well this was really middle class, um, people, school teachers, firemen, policemen, that range of uh income. Uh no one acted up there. And it was it was great. They have the. This was I think the largest housing project in the United States 2500 families. Uh and it's pretty good stores uh you know a commercial area and stuff like that. It was like a little city. Very very [second very is stressed] well maintained. We'd play in the playgrounds. And then they they sponsored um um, baseball teams and basketball teams. So I played on teams.

Data Elicitation

While linguistic analyses of syntactic category, word frequency and phrasal level syntactic structure ratios vary based on medium (oral vs. written language), formality, and topic, it is unknown whether or to what degree ID scores may also vary across data types. The progressive decline in memory common to AD, where distant memories are better preserved

than more recent memories, suggests that ID scores may vary based on what type of memories are elicited. For this reason, to elicit speech samples we have chosen to use prompts which 1) tap remote and well-rehearsed memories, 2) are topically uniform and 3) are presented to participants in a uniform fashion—these are important components of best practices approach to collecting and analyzing oral language ID in the a standardized fashion within the context of dementia progression and/or risk.

Transcription

There are multiple ways to represent oral language through text (Edwards, 1993), and transcription is thus the first phase of analysis (Pavlenko, 2008). The degree of detail encoded within a transcript must be negotiated with analytic goals and how analytic choices may transform language (Fox, 2004; Liu, 1999). Detailed transcripts can encode, e.g. phonetic pronunciation, pause length and intonation, but are time-consuming to create. In turn, this level of detail, within the context of measuring semantic content, can affect the global readability of a transcript, can over-emphasize dysfluencies, and may even encourage bias against marked and non-standard language practices. Under-detailed transcripts which fail to encode, e.g. lexical and non-lexical fillers and intonation, are much faster to create, but can mask dysfluencies, and may misrepresent the speaker's original output in various ways. For example, failing to include salient intonation may mask elements which are intended as ironic and may occlude differences between words or phrases which are repeated and stressed for emphasis versus elements which are repeated, but do not convey emphasis. Heeding these potential issues, the AID transcription guidelines have two specific goals: transcripts should maintain a reasonable degree of faithfulness with their original audio, and they should present all information pertinent to narrative content for ID analysis.

Transcript integrity is addressed by including several elements from the oral presentation, including lexical (e.g. *well this was really* and *uh you know a commercial area*) and non-lexical fillers (e.g. *most of my uh growing up years in uh*), metadiscourse, incomplete words (e.g. *somewhat ah se- seedy reputation*), and intonation in specific contexts (e.g. [second *very* is stressed]), which can each contribute to assessing and coding whether language is semantically on-topic and novel (that is, not simply a repetition). Notes by the transcriber, in addition to metadiscourse, are included in the transcript, but are partitioned and not analyzed for ideas or words. Narratives are transcribed following standard English lexical orthography, and additional transcription guidelines to handle variation have been developed specifically to promote analytic consistency. For example, numbers can be verbalized in several ways (e.g. *twenty five hundred* vs. *two thousand five hundred*), but are consistently transcribed numerically (i.e. *2500*) in AID.

Autobiographical narratives are by nature comprised of several interconnected ideas in which a range of linguistic and paralinguistic cues can be used both to convey content and to maintain the speaking position. The goal of AID is to document propositional content and quantify ID, not to evaluate adherence to written standards or value specific discourse strategies over others. Thus, when separating a narrative into smaller units, we employ the term *utterance*, not *sentence*. The latter is a set construct for chunking written text, while *utterance* refers to a unit of oral speech, which may or may not correspond to a written sentence (e.g. the utterance *Very very well maintained* does not correspond to a written sentence because it lacks a subject and verb, while it is considered a separate utterance in AID). Utterances are operationally defined through narrative segmentation, which relies on several factors, including topic, pauses, intonation, content and the larger context to determine utterance boundaries.

Transcripts of grossly unequal length may themselves introduce analytic bias. Patients further along the AD progression are more likely to produce shorter narratives, and it is

unclear how narrative length itself may affect ID scores—raw word count has not been explored or even documented in previous research on ID. Thus, in order to avoid introducing analytic bias, the transcription of longer samples is halted at the end of the utterance containing the 400th word. Based on two rater's independent transcription of six narratives following the AID transcription guidelines, inter-coder utterance segmentation consistency is very high (Interclass correlation coefficient (ICC) = .9745).

Word Count

While word count criteria are taken seriously in descriptive, psycholinguistic and corpus linguistic research (Biber et al., 1999; Brysbaert and New, 2009; Gales and Chand, 2010), there is less awareness of word criteria as variable and hence worthy of definition in clinical linguistic research. Further, total word count is an essential component to deriving ID scores as the proportion of ideas to words. Word count issues include decisions on how and whether to count contractions and possessives, multi-word numbers, metadiscourse, acronyms, lexical fillers, proper nouns, multi-word idiomatic phrases and compound words (some which have multiple acceptable orthographic forms). Different criteria have a clear effect on total word count: for the narrative in Example 1, AID-based word count tallied 118 words (and 11 non lexical fillers), CPIDR Regular Mode (version 3.2.2785.24603) tallied 130 words (which matches a word count from Microsoft Word 2002) and CPIDR Spoken mode from the same release tallied 121 words (see annotation at the end of this paragraph)

Meanwhile, Textpad, a freeware text editor, calculates 133 words for the same text. Given that ID scores are derived as the ratio of ideas to words, and that some of the elements counted by CPIDR and other document editing software are not considered words, and that nonlexical fillers are not transcribed in all ID rubrics, having systematic word count criteria are important. AID specifies how each word count is handled, and based on two rater's independent tally of the total number of words across six narratives following the AID word count guidelines, word count consistency is both very high (ICC = .9810) and is in keeping with corpus linguistics conventions.

Examining the CPIDR Spoken version word count more closely, some elements which are clearly 'words' are not counted while non lexical fillers are inconsistently interpreted as words. For example, in the utterance **And then they they sponsored um um, baseball teams and basketball teams**, CPIDR does not mark the second repeated **they** as a word, while it does mark the second repeated **um** as a word.

Idea Coding

Broadly, AID separates and systematically counts meaningful phrasal and utterance level syntactic groupings which *functionally* convey new on-topic content to the listener (Armstrong, 2005; Martin and Rose, 2003). By focusing on the phrasal level, we are able to consistently handle phrases with different internal syntactic structures and word counts, and to more directly tap semantically meaningful units rather than syntactic constructions. For example, prepositions serve a variety of syntactic and semantic functions: they can introduce a location or direction (e.g. *in the Bronx*) and are then counted as a separate idea, while they can also, in conjunction with a verb, functionally instantiate a single predicative meaning (e.g. prepositional particles are not counted as a separate idea when they are functionally part of a predicate, *pass out*, i.e. to faint, and *no one acted up there*, i.e. no one misbehaved). Our analysis provides a rubric for determining whether a word or set of words contributes distinct semantic content, or is a syntactically required component of a larger structure, and thus should not be counted as a separate idea. The AID manual also presents straightforward

rules to handle items in which complex syntax may obscure a simple meaning and vice versa.

We follow Turner and Greene (1977) in organizing our manual around three semantic subtypes: Predication, Modification, and Connectives. However, these terms are used loosely: their value is as a first step to uncovering and quantifying novel semantic units in natural discourse. As these three semantic subtypes are discussed, examples of lexical or phrasal forms which semantically fit each functional role are drawn from the narrative in Example 1, while the full AID analysis of Example 1 is presented in Appendix A.

Predication—Predication is the expression of the basic argument of a statement. It expresses an action or a state of being (e.g. *played*, *were*), and can include the related actor and/or object. We group verbal clauses as consisting of the subject (regardless of whether the subject is overtly or covertly expressed, given that subject ellipsis is not rare in conversational genres, e.g. Biber et al., 1999: 1048), the verb, any tense, modal or aspectual markers and a single predicative object. This permits coding consistency across variously preferred syntactic structures and uniformly handles predicates which can manifest in multiple syntactic structures, e.g. *gave the letter to her*, with a prepositional phrase and a noun phrase following the verb versus *gave her the letter* with two noun phrases following the verb.

We follow previous models in not counting tense (e.g. *acted up*) or aspect (e.g. *We'd play*) separate from the verb, given that English verbs are syntactically required to each have a tense or aspect which can be conveyed overtly or covertly (e.g., *people say* covertly conveys either a perfective or habitual aspect, depending on the context, and is overtly marked as present tense). This coding decision allows for analytic consistency, given that it is sometimes very problematic to determine whether conditionality versus tense is being expressed through a modal. For example, *people will say* can be interpreted as conveying future tense or the speaker's degree of certainty that the event will occur.

Modification—Ideas of modification provide further detail regarding the specification of an item, process or state. They can restrict the meaning of an idea by specifying a particular quality of it (*good stores* rather than *stores*, *played on teams* rather than *played*), specifying a quantity or modifying the quality of an element (*most of my uh growing up years*, *pretty good stores*) or modifying another idea or relation through negation (*this wasn't*). Prepositional and adjectival words or clauses are counted separately when they do not function as the predicative object. Adverbial words or clauses, meanwhile, never fill this functional role, and are counted separately. Negation is interpreted as a means of modifying a predicate or another modifier, and coded as an idea. Determiners (also known as articles, e.g. *a*, *the*) are often functionally ambiguous—they can be interpreted as syntactically required and semantically meaningless (or at least, conveying more syntactic than semantic information) or alternately interpreted as providing additional information on whether the referent is explicit versus vague (*a project* versus *the projects*) and the quantity of the referent (e.g. *a commercial area* versus *some commercial areas*). Meanwhile, they are also optional for some lexical items and in some syntactic structures (e.g. singular count nouns tend to require an article, while plural forms often can appear bare without an article, e.g. in utterances like *We'd play in (the) playgrounds*, *the* is variably present). Given that AID is fundamentally focused on capturing how content-full a narrative is, not on capturing required syntactic structures, explicit rules for handling determiners and other functionally ambiguous elements are presented. If anything, these rules may err on the side of under attributing ideas to potentially information-bearing units, but they do so in a consistent fashion, which allows for transparency across coders and language samples.

Connectives—Connective ideas are words or phrases that show a relationship between multiple ideas expressing conjunction, causality or contrast. Connectives can semantically express temporal or conjunctive relationships between other clausal units (e.g. *they sponsored um um, baseball teams and basketball teams. So I played on teams*), while they also can serve an interactional function, e.g., to maintain the speaking position, in the case of habitual utterance initial conjunctions (e.g. *And it was it was great*). Functional role and frequency across the narrative are used to separate semantically meaningful connectives which convey temporal sequencing or dependency from syntactically or metadiscursively functioning connectives: the latter are not counted. For example,

Additional Linguistic Measures in AID

As narratives are coded for ideas, several other measures are tallied for each utterance. Each of the following have been hypothesized as contributing to global semantic dysfluencies, but their respective frequencies and possible inter-relationships within oral narratives and their relationship with cognition and decline are unclear, given past emphases towards analyzing written language, and the lack of quantitative research on their patterns (though qualitative examinations of AD discourse have cited these as evident in and salient to distinguishing AD discourse, c.f. Guendouzi and Müller, 2006; Müller and Guendouzi, 2005; Ramanathan, 1997). In order to more narrowly define what elements of narratives contribute to ID and what elements may more drastically change through AD progression, we define and tally following categories through the lens of oral language structure: utterance completeness, lexical filler count and non-lexical filler count per utterance, the number of repeated ideas per utterance, the presence of utterance initial conjunctions per utterance and a word count for semantically empty extender clauses (e.g. *and stuff like that*) per utterance. Criteria for each of these is offered in the manual: explicit discussion of and/or guidelines for handling most of these features are not present in LAL, while they are foreshadowed as elements which may affect ID scores in past ID research (e.g., Brown et al., 2008 exclude lexical fillers and inconsistently exclude repeated words, with an unknown effect on propositional count).

AID Findings

Given that we have departed from other methods of measuring ID in some ways, and the lack of substantial research on late life oral samples and cognition, we next present the results of the AID rubric with late life oral narratives collected from an aging population of healthy and dementing individuals. Inter-rater reliability results from this population on transcription and word count were presented above, while inter-rater reliability for idea count are next presented. We also tested the rubric in two additional ways. Two oral samples were collected from a subset of older adults across the diagnostic spectrum within a short time span to examine whether AID-based late life oral scores are stable across a short time span. Further, a pilot study was conducted to examine whether AID-based ID scores can distinguish healthy and demented older adults. Significantly, no previous research on ID has examined short term ID stability, while it follows from the highly significant correlations between ID and clinical diagnostic outcomes.

Test Data—All participants were evaluated at the University of California at Davis Alzheimer's Disease Center (UCD ADC) as part of an ongoing longitudinal study of cognitive impairment in an educationally and ethnically diverse sample of older adults. The design of this longitudinal study has been described in detail (Mungas et al., 2010; Mungas et al., 2005).

AID Inter-rater Reliability

We follow LAL in defining ID scores as the number of ideas expressed per 10 words: $[(\text{ideas/words}) \times 10]$. Using the AID rubric, we have demonstrated excellent inter-rater reliability ($ICC = .9662$) for idea count based on two coders' independent analysis of six narratives. In conjunction with the high word count inter-rater reliability, AID-based ID scores are thus highly consistent between raters.

AID Within Subject Test Retest Reliability—Two oral language samples were collected from 30 elderly patients from the UCD ADC cohort across the AD diagnostic spectrum (normal through probable AD, cognition rated according to standardized CRD guidelines) within a short time span (Mean = 67.4 days; Range = 2:134 days; SD = 43.64 days) to investigate the short term stability of ID scores. Descriptive statistics for word count are Mean = 253.32; Range = 39: 471; SD = 119.35 and for ID scores are Mean = 3.46; Range = 1.07:5.16; SD = 0.65.

In order to investigate within subject difference, not simply group averages, and given that 1) a normal distribution cannot be assumed based on either past literature or on this data set 2) that there was no hypothesis regarding whether ID scores were more likely to increase versus decrease within this short time span, planned comparisons using the non-parametric matched pairs test, a (two-tailed) Wilcoxon Signed Rank test, revealed no significant difference ($p = 0.1274$) between test and retest ID scores by subject. It is salient to note that subjects with a probable AD diagnosis are as stable for ID across this short time span as cognitively normal subjects.

AID-based Late Life Oral ID and Clinical Diagnosis—ID was measured based on the AID rubric from oral narratives collected from 87 older adults at the UCD ADC. ID scores were examined in association with global cognitive function, measured through the CDR rating scale. Separate random effects models tested the effect of ID on demographic and diagnostic ratings, with the latter adjusted for age and years of education. While the study design and CDR scoring criteria have been described (Reed et al., 2010), demographically, the average age of the sample was 74.2 (SD = 6.76), and the average years of education was 14.5 (SD = 3.56). Fifty-two percent of the sample was female. The majority of the sample (75%) was Caucasian, 13.6% were African American, 9% were Hispanic and 2% were of another ethnic background. 38 participants were cognitively normal, 35 had MCI and 14 had a dementia syndrome.

A simple Pearson's correlation coefficient was used to estimate the relationship between ID and demographic variables. Within such, neither age nor gender was significantly related to ID (respectively, $ps = .96$ and $.38$). There was a very modest (also not statistically significant) trend towards a positive relationship between ID and education ($r = .14$, $p = .14$). We next examined whether ID varied as a function of diagnostic group (cognitively normal, MCI or dementia). Planned comparisons revealed that the dementia group showed statistically significantly lower ID than either the normal group ($P < .0001$) or the MCI group ($p < .0001$). However, the MCI group and the cognitively normal group, which together constituted 84% of the sample, did not differ from one another in terms of ID ($p = .211$). The mean ID score (of ideas per 10 words) for the dementia, MCI, and normal groups were 3.77, 4.04 and 4.03, respectively.

These results provide evidence of external validity for the AID rubric (i.e. the separation of groups), replicate some existing ID findings, and also expand on past ID findings. In this population, late life oral ID distinguishes healthy elderly adults from elderly clinically-diagnosed AD patients: healthy adults have significantly higher ID. Confirming the results in Mitzner & Kemper (2003), late life oral ID is not associated with age. Expanding on past

ID results, this study is the first gender-balanced (52% female) study and first exploration within an ethnically and educationally heterogeneous population of the association between ID and clinical AD diagnosis. It demonstrates no relationship between ID and gender and also suggests that ID may serve as an index of reserve even in the presence of early disease.

Discussion

The AID inter-rater reliability results for utterance segmentation, word count and idea count, in conjunction with the short term AID-ID test/retest reliability, the lack of an association between AID-ID and age or gender, the non-significant relationship between ID and education, and the strong correlation between AID-ID and clinical diagnosis suggest that AID is a replicable, reliable rubric for continued investigation of ID in dementia research on aging populations. In addition, the new results presented here demonstrating no significant association between gender and ID and the significant correlation between late life oral ID and clinical diagnosis suggest that late life oral samples, which 1) can be collected later through the diagnostic spectrum, 2) have strong ecological validity as a measure of natural language, and 3) are less subject to attrition than written samples, can profitably contribute to investigating AD decline trajectories, the effects of and thresholds for cognitive reserve, and the relationship between semantic content, language, and aging more broadly.

Conclusion

This paper provides an overview of AID, a systematic rubric for collecting, transcribing and scoring ID in late life oral language samples. ID shows promise as an ecologically valid measure of changes in language practices over the lifespan, and of AD risk, cognitive reserve and/or AD-based cognitive decline, yet it is not widely used. While the Nun Study findings have not been replicated across a range of demographics and in large-scale samples, the availability of the AID rubric, which has strong inter-rater reliability and extensive coding guidelines, will allow researchers to continue exploration of how ID relates to cognitive reserve, AD risk, and heterogeneous AD decline trajectories.

Until standardized ID data collection methods and scoring criteria are used across several populations and investigative groups, the probability of past ID findings being a false positive are unclear (Ioannidis, 2005). By making available the AID rubric, as a transparent and detailed scoring guideline which has demonstrated reliability and success in a sample cohort, our goal is to encourage additional research on ID which can confirm or clarify past findings. Future examinations of ID may contribute to understanding cognition across the lifespan, and may alternatively or in collusion also help understand and predict dementia based decline trajectories and risk factors for late-life dementia.

Acknowledgments

Funding Sources: This work was supported in part by the National Institute on Aging [AG021511 to S.T.F, AG031563 to B. Reed and AG10129 to C. DeCarli].

Literature Cited

- Armstrong E. Language disorder: A functional linguistic perspective. *Clinical linguistics & phonetics*. 2005; 19:137–153. [PubMed: 15823952]
- Baird AD, Ford M, Podell K. Ethnic differences in functional and neuropsychological test performance in older adults. *Archives of Clinical Neuropsychology*. 2007; 22:309–318. [PubMed: 17317093]
- Bayles KA, Tomoeda CK, Trösset MW. Relation of linguistic communication abilities of Alzheimer's patients to stage of disease. *Brain and language*. 1992; 42:454–472. [PubMed: 1377076]

- Baynes, K.; Chand, V.; Bonnici, L.; Farias, ST. Idea density as a measure of communicative skill in Alzheimer's Disease. *Advances in Basic and Clinical Neuropsychology*; 2007 Midyear Meeting; Bilbao, Spain. Year[*This is a Conference Proceeding]
- Biber, D.; Johansson, S.; Leech, G.; Conrad, S.; Finegan, E. *Longman grammar of spoken and written English*. Longman Publications Group; Harlow: 1999.
- Boone KB, Victor TL, Wen J, Razani J, Ponton M. The association between neuropsychological scores and ethnicity, language, and acculturation variables in a large patient population. *Archives of Clinical Neuropsychology*. 2007; 22:355–365. [PubMed: 17320344]
- 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]
- Brysbaert M, New B. Moving beyond Ku era and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research methods*. 2009; 41:977–990. [PubMed: 19897807]
- Byrd DA, Miller SW, Reilly J, Weber S, Wall TL, Heaton RK. Early environmental factors, ethnicity, and adult cognitive test performance. *The Clinical Neuropsychologist*. 2006; 20:243–260. [PubMed: 16690545]
- Causino Lamar, MA.; Obler, LK.; Knoefel, JE.; Albert, ML. Communication patterns in end-stage Alzheimer's disease: Pragmatic analyses. Bloom, RL.; Obler, LK.; De Santi, S.; Ehrlich, JS., editors. Lawrence Erlbaum; Hillsdale, NJ: 1994. p. 217-235.13
- Chafe, W. Linguistic differences produced by differences between speaking and writing. Olson, DR.; Hildyard, A.; Torrance, N., editors. Cambridge University Press; Cambridge: 1985. p. 105-123.
- Chafe W, Tannen D. The relation between written and spoken language. *Annual Review of Anthropology*. 1987; 16:383–407.
- Cosentino S, Manly J, Mungas D. Do reading tests measure the same construct in multiethnic and multilingual older persons? *Journal of the International Neuropsychological Society*. 2007; 13:228–236. [PubMed: 17286880]
- Covington, MA. CPIDR 3 User Manual. CASPR Research Repor 2007–3. Artificial Intelligence Center, The University of Georgia; Athens: 2007. [*This is a report]
- 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 and language*. 1996; 53:1–19. [PubMed: 8722896]
- Davis BH, MacLagan M. Examining pauses in Alzheimer's discourse. *American journal of Alzheimer's Disease and other dementias*. 2009; 24:141–154. [PubMed: 19150969]
- Edwards, JA. Principles and contrasting systems of discourse transcription. Edwards, JA.; Lampert, MD., editors. Lawrence Erlbaum Associates; Hillsdale: 1993. p. 3-31.1
- Ehrlich, JS. Studies of discourse production in adults with Alzheimer's disease. Bloom, RL.; Obler, LK.; De Santi, S.; Ehrlich, JS., editors. Lawrence Erlbaum; Hillsdale, New Jersey: 1994. p. 149-160.9
- Ehrlich JS, Obler LK, Clark L. Ideational and semantic contributions to narrative production in adults with dementia of the Alzheimer's type. *Journal of Communication Disorders*. 1997; 30:79–99. [PubMed: 9100125]
- Engelman M, Agree EM, Meoni LA, Klag MJ. Propositional density and cognitive function in later life: Findings from the Precursors Study. *Journal of Gerontology Series B: Psychological Sciences and Social Sciences*. 2010; 65:706–711.
- Fillmore, CJ. The case for case. Bach, E.; Harms, RT., editors. Holt, Rinehart & Winston; New York: 1968. p. 1-88.
- Fillmore, CJ. Toward a modern theory of case. Reibel, DA.; Schane, SA., editors. Prentice-Hall; Englewood Cliffs, N.J: 1969. p. 361-375.
- Fox, AA. *Real country: Music and language in working-class culture*. Duke University Press; Durham: 2004.
- Gales, T.; Chand, V. A Corpus-based Approach to Word Frequency and Syntactic Categories: Untangling their relationship within sound change. *Linguistic Society of America*; Baltimore, MD: 2010. [*This is a Conference Paper]

- Garcia, LJ.; Joannette, Y. Conversational topic-shifting analysis in dementia. Bloom, RL.; Obler, LK.; De Santi, S.; Ehrlich, JS., editors. Lawrence Erlbaum; Hillsdale, NJ: 1994. p. 161-183.10
- 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]
- Gasquoine PG. Variables moderating cultural and ethnic differences in neuropsychological assessment: The case of Hispanic Americans. *The Clinical Neuropsychologist*. 1999; 13:376–383. [PubMed: 10726609]
- Guendouzi, JA.; Müller, N. Approaches to discourse in dementia. Lawrence Erlbaum; Mahwah, NJ: 2006.
- Iacono D, Markesbery WR, Gross M, Pletnikova O, Rudow G, Zandi P, Troncoso JC. The Nun Study. Clinically silent AD, neuronal hypertrophy, and linguistic skills in early life. *Neurology*. 2009; 73:665–673. [PubMed: 19587326]
- Ioannidis JPA. Why most published research findings are false. *PLoS Medicine*. 2005; 2:696–701.
- Kave G, Levy Y. Morphology in picture descriptions provided by persons with Alzheimer's disease. *Journal of speech, language, and hearing research*. 2003; 44:341–352.
- Kemper, SJ. Language and aging. Craik, FIM.; Salthouse, TA., editors. Lawrence Erlbaum; Hillsdale, NJ: 1992. p. 213-270.
- Kemper, SJ. Language across the life span: Coding manual. In: Lawrence, KS., editor. Coding Manual. 1993. [*This may be an unpublished manuscript]
- Kemper SJ, 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]
- Kemper SJ, 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]
- Kemper SJ, Summer A. The structure of verbal abilities in young and older adults. *Psychology and Aging*. 2001; 16:312–322. [PubMed: 11405318]
- Kintsch, W. The representation of meaning in memory. Lawrence Erlbaum; Hillsdale: 1974.
- Liu, LH. The question of meaning-value in the political economy of the sign. Liu, LH., editor. Duke University Press; Durham: 1999. p. 13-41.
- Long DL, Baynes K, Prat CS. The Propositional Structure of Discourse in the Two Cerebral Hemispheres. *Brain and language*. 2005; 95:383–394. [PubMed: 16298668]
- Manly JJ, Echেমendia RJ. Race-specific norms: Using the model of hypertension to understand issues of race, culture, and education in neuropsychology. *Archives of Clinical Neuropsychology*. 2007; 22:319–325. [PubMed: 17350797]
- Martin, JR.; Rose, D. Working with Discourse: Meaning beyond the clause. Continuum; London: 2003.
- Miller LMS. The effects of age and domain knowledge on text processing. *Journal of Gerontology Series B: Psychological Sciences and Social Sciences*. 2003; 58:217–223.
- Mitzner TL, Kemper SJ. Oral and written language in late adulthood: Findings from the Nun Study. *Experimental aging research*. 2003; 29:457–474. [PubMed: 12959878]
- Morris J. The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology*. 1993; 43:2412–2414. [PubMed: 8232972]
- Mortensen L. A transitivity analysis of discourse in dementia of the Alzheimer's type. *Journal of Neurolinguistics*. 1992; 7:309–321.
- Müller N, Guendouzi JA. Order and disorder in conversation: Encounters with dementia of the Alzheimer's type. *Clinical linguistics & phonetics*. 2005; 19:393–404. [PubMed: 16019783]
- Mungas D, Beckett L, Harvey D, Farias ST, Reed BR, Owen C, Olichney J, Miller J, DeCarli C. Heterogeneity of cognitive trajectories in diverse older persons. *Psychology and Aging*. 2010; 25:606–619. [PubMed: 20677882]
- Mungas D, Harvey D, Reed BR, Kagust WJ, DeCarli C, Beckett L, Mack WJ, Kramer JH, Weiner MW, Schuff N, Chui HC. Longitudinal volumetric MRI change and rate of cognitive decline. *Neurology*. 2005; 65:565–571. [PubMed: 16116117]
- Pavlenko, A. Narrative Analysis. Moyer, M.; Wei, L., editors. Blackwell; Oxford: 2008. p. 311-325.

- Ramanathan, V. Alzheimer discourse: Some sociolinguistic dimensions. Lawrence Erlbaum; Mahwah, NJ: 1997.
- Reed, BR.; Chand, V.; Baynes, K.; Bonnici, L.; Farias, ST. Is idea density in late life a marker of cognitive reserve?. Effects on cognitive decline. 38th Annual Meeting of the International Neuropsychological Society; Acapulco, Mexico. 2010. [*This is a Conference Paper]
- Riley KP, Snowdon DA, Derosiers MF, Markesbery WR. Early life linguistic ability, late life cognitive function, and neuropathology: Findings from the Nun Study. *Neurobiology of aging*. 2005; 26:341–347. [PubMed: 15639312]
- Snowdon, DA.; Greiner, LH.; Kemper, SJ.; Nanayakkara, N.; Mortimer, JA. Linguistic ability in early life and longevity: Findings from the Nun Study. Robine, J-M.; Forette, B.; Francheschi, C.; Allard, M., editors. Springer-Verlag; Berlin, Germany: 1999. p. 103-113.
- Snowdon DA, Greiner LH, Markesbery WR. Linguistic ability in early life and the neuropathology of Alzheimer's disease and cerebrovascular disease: Findings from the Nun Study. *Annals of the New York Academy of Sciences*. 2000; 903:34–38. [PubMed: 10818486]
- 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. *Journal of the American Medical Association*. 1996; 275:528–532. [PubMed: 8606473]
- Stern Y. What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society*. 2002; 8:448–460. [PubMed: 11939702]
- Stern Y. The concept of cognitive reserve: A catalyst for research. *Journal of Clinical and Experimental Neuropsychology*. 2003; 25:589–593. [PubMed: 12815497]
- Turner, A.; Greene, E. Technical Report No. 63. Institute for the Study of Intellectual Behavior, University of Colorado; Boulder: 1977. The construction and use of a propositional text base. [*This is a report]

Appendix A. Sample AID analysis

The below narrative has 50 ideas and 118 words, for an AID-based ID score of 4.24.

I I spent most of my uh growing up years in uh, in growing up in, in the Bronx which now has a somewhat ah se- seedy reputation.

1. spent, I, years
2. years, growing up
3. years, most of
4. in the Bronx
5. 4 which 7
6. now
7. has, the Bronx, a reputation
8. reputation, seedy
9. seedy, somewhat

But uh, this wasn't.

1. was, this, (seedy)
2. NEG n't
3. but

Well this was really middle class, um, people, school teachers, firemen, policemen, that range of uh income.

1. was, this, people
2. people, middle class
3. middle class, really
4. people, teachers
5. teachers, school
6. people, firemen
7. people, policemen
8. (were, people), that range
9. of income

Uh no one acted up there.

1. acted up, no one
2. there

And it was it was great.

1. was, it, great

They have the.

--no ideas--

This was I think the largest housing project in the United States 2500 families.

1. was, this, housing project
2. housing project, the largest
3. in the United States
4. was, it, families
5. families, 2500

Uh and it's pretty good stores uh you know a commercial area and stuff like that.

1. (has), it, stores
2. stores, good
3. good, pretty
4. (has), it, a area
5. area, commercial

It was like a little city.

1. was, it, a city
2. city, like a
3. city, little

Very very [second very is stressed] well maintained.

1. (was, it), maintained
2. maintained, well

3. well, very

4. very, very

We'd play in the playgrounds.

1. would play, we

2. in the playgrounds

And then they they sponsored um um, baseball teams and basketball teams.

1. and then

2. sponsored, they, teams

3. teams, baseball

4. teams, basketball

So I played on teams.

1. so 2

2. played, I

3. on teams

Past Idea Density Findings

The presence of a + demonstrates a significant positive correlation, – demonstrates a significant negative correlation, trends which do not meet significance thresholds are marked in brackets [], and non-correlations with no noticeable trend are marked with 0. Cells which are split into two columns represent conflicting findings for the same measures. Empty cells reflect an absence of published comparisons.

Table 1

	Collected in Early Life		Collected in Late Life			
	Oral Sample	Written Sample	Oral Sample		Written Sample	
Age	(Kemper and Summer, 2001)	0 (Snowdon et al., 1996)	(Kemper and Summer, 2001)	0 (Mitzner and Kemper, 2003)	(Kemper et al., 2001)	0 (Kemper et al., 1993; Mitzner and Kemper, 2003; Snowdon et al., 1996)
Educational Attainment (in years)		+ (Snowdon et al., 1996, however education accounts for less variation than ID in late life MMSE scores)	(Mitzner and Kemper, 2003)	0 (Mitzner and Kemper, 2010)	0 (Mitzner and Kemper, 2003)	+ (Kemper et al., 1993)
Educational Achievement (high school grades)		0 (Kemper et al., 2001)				
Late Life Cognition (various measures, MMSE most typical)		+ (Iacono et al., 2009; Kemper et al., 2001; Riley et al., 2005; Snowdon et al., 1996)	(Mitzner and Kemper, 2003)	[+] (Mitzner and Kemper, 2003)	(Kemper et al., 2001; Kemper et al., 1993; Mitzner and Kemper, 2003)	+
Late Life Physical Condition (ADL)				0 (Mitzner and Kemper, 2003)	+	(Mitzner and Kemper, 2003)
Clinical MCI diagnosis		(Iacono et al., 2009; Riley et al., 2005)			(Kemper et al., 1993)	
Clinical AD diagnosis		(Engelman et al., 2010; Iacono et al., 2009; Kemper et al., 2001)	(Baynes et al., 2007; Reed et al., 2010)		(Kemper et al., 1993)	
AD Pathology		(Riley et al., 2005; Snowdon et al., 2000; Snowdon et al., 1996)				
Asymptomatic AD Pathology (cognitively normal)		+ (Iacono et al., 2009)				

	Collected in Early Life		Collected in Late Life	
	Oral Sample	Written Sample	Oral Sample	Written Sample
Post-Mortem Brain Weight		⁺ (Riley et al., 2005)		
APOE ϵ^{-4} allele presence		⁰ (Riley et al., 2005)		
Late Life Physical Condition (ADL)			⁰ (Mitzner and Kemper, 2003)	⁺ (Mitzner and Kemper, 2003)
Grammatical Complexity of same language sample		⁺ (Snowdon et al., 1996)		⁺ (Snowdon et al., 1996)
Oral vs. Written ID		⁺ (Snowdon et al., 1996)	Written ID scores have greater power to differentiate low and high cognitive ability in late life than oral samples (Mitzner and Kemper, 2003)	
Longitudinal Comparison of ID scores from written samples in early vs. late life		⁺ (Snowdon et al., 1996)		⁺ (Snowdon et al., 1996)