Is cognitive impairment associated with reduced syntactic complexity in writing? Evidence from automated text analysis

Fredrik Sand Aronsson, Marco Kuhlmann, Vesna Jelic & Per Östberg

To cite this article: Fredrik Sand Aronsson, Marco Kuhlmann, Vesna Jelic & Per Östberg (2020): Is cognitive impairment associated with reduced syntactic complexity in writing? Evidence from automated text analysis, Aphasiology, DOI: 10.1080/02687038.2020.1742282

To link to this article: https://doi.org/10.1080/02687038.2020.1742282

© 2020 The Author(s). published by Informa UK Limited, trading as Taylor & Francis Group

View supplementary material

Published online: 09 Apr 2020.

Submit your article to this journal

Article views: 220

View related articles

View Crossmark data
Is cognitive impairment associated with reduced syntactic complexity in writing? Evidence from automated text analysis

Fredrik Sand Aronsson\textsuperscript{a,b}, Marco Kuhlmann\textsuperscript{c}, Vesna Jelic\textsuperscript{d,e} and Per Östberg\textsuperscript{a,b}

\textsuperscript{a}Department of Clinical Science, Intervention and Technology, Division of Speech and Language Pathology, Karolinska Institutet, Stockholm, Sweden; \textsuperscript{b}Speech and Language Pathology, Medical Unit, Karolinska University Hospital, Stockholm, Sweden; \textsuperscript{c}Department of Computer and Information Science, Linköping University, Linköping, Sweden; \textsuperscript{d}Department of Neurobiology, Care Sciences and Society, Division of Clinical Geriatrics, Karolinska Institutet, Stockholm, Sweden; \textsuperscript{e}Theme Aging, Karolinska University Hospital, Stockholm, Sweden

ABSTRACT

Background: Written language impairments are common in Alzheimer's disease and reduced syntactic complexity in written discourse has been observed decades before the onset of dementia. The validity of average dependency distance (ADD), a measure of syntactic complexity, in cognitive decline needs to be studied further to evaluate its clinical relevance.

Aims: The aim of the study was to determine whether ADD is associated with levels of cognitive impairment in memory clinic patients.

Methods & procedures: We analyzed written texts collected in clinical practice from 114 participants with subjective cognitive impairment, mild cognitive impairment, and Alzheimer’s disease during routine assessment at a memory clinic. ADD was measured using automated analysis methods consisting of a syntactic parser and a part-of-speech tagger.

Outcomes & results: Our results show a significant association between ADD and levels of cognitive impairment, using ordinal logistic regression models.

Conclusion: These results suggest that ADD is clinically relevant with regard to levels of cognitive impairment and indicate a diagnostic potential for ADD in cognitive assessment.

ARTICLE HISTORY

Received 27 September 2019
Accepted 10 March 2020

KEYWORDS

Syntactic complexity; average dependency distance; Alzheimer’s disease; mci; sci

Introduction

Alzheimer’s disease (AD) is a progressive neurodegenerative disease and the most common cause of dementia in the population over 65 years of age. Central nervous system damage in the form of neurofibrillary tangles and senile plaques containing amyloid beta accumulates in a fixed pattern during several decades. Neurofibrillary changes appear first in the locus coeruleus, a noradrenergic nucleus in the brainstem that supports attention, memory, and arousal, before progressing to limbic structures in the temporal lobes and on to neocortical association areas (Braak & Del Tredici, 2015), causing impairments in...
several cognitive domains. Impairments in episodic memory, the ability to recall earlier experiences, are a hallmark of the disease as well as impairments in working memory. Language impairments are also common, even in early stages of the disease, and the decline can be described as hierarchical: more complex structures are often the first to deteriorate (Emery, 2000). The language impairment primarily affects semantic abilities, including word-finding difficulties and semantic paraphasias, whereas syntactic and phonological abilities are considered to be relatively preserved (Ferris & Farlow, 2013; Taler & Phillips, 2008; Verma & Howard, 2012; Vuorinen et al., 2000). The semantic deficit might be due to pathological changes in the hippocampal region and the posterior association cortex (Eichenbaum, 2000; Manns et al., 2003). The notion that syntactic abilities are somewhat buffered against the effects of the disease in the early stages are in line with studies showing a sparser syntax with fewer subordinate clauses but with few grammatical errors; however, comprehension of complex syntax is clearly impaired in AD (Croisile, 1999; Emery, 2000; Kemper et al., 1993; Lyons et al., 1994). Possibly, the reduction in syntactic complexity in the language production of AD patients has been underestimated, especially in written language. Thus, writing skills become markedly impaired during the course of the disease (Croisile, 1999; Harnish & Neils-Strunjas, 2008; Lambert et al., 2007).

Writing is a complex task, not only requiring linguistic abilities but also several cognitive capabilities such as working memory and executive functions. Flower and Hayes (1981) suggested a model of writing consisting of three separate processes: planning which includes goal setting and organization of ideas, translation of these plans into written text, and finally reviewing. These are not separate phases but rather processes between which the writer constantly shifts attention. This looping or recursive writing process differs from a view of writing as being linear with separate, temporally ordered, phases. To control and coordinate the three processes, the monitor plays a vital role in the initial model. It also requires the incorporation of information from long-term memory, including topical knowledge, and information from what Flower and Hayes (1981) called the task environment – everything external to the writer that may contribute to the writing process such as the specific assignment. Several revisions have been made to the original model (see Hayes, 2012, for an overview). Transcription of the ideas into written text was not originally represented in the model as it was thought to be highly automatized and thus did not compete for the limited cognitive resources during writing. Several studies in the 1990 s (for example, Berninger et al., 1994) did, however, show that transcription does indeed tax working memory and reduce the resources available for other aspects of writing. Motivation is also included in the revised model as it affects not only the quality of texts – partly owing to the fact that highly motivated writers seem to spend more time editing the output – but also because it determines how likely a person is to write in the first place (Hayes et al., 1990). Of special interest for our study is the inclusion of working memory in the revised model of Hayes. As stated above, writing is a complex activity with several separate processes competing for cognitive resources (Kellogg, 2001). Kellogg has further demonstrated that the use of an outlining strategy where writers organize and plan their ideas and disposition of the text beforehand allowed them to devote more resources to the translation of these ideas (Kellogg, 1990). As working memory capacity is a limited resource, these models suggest that effective strategies and automatization might help reduce the cognitive overload.
The writing impairments in AD are multifactorial; misspelling, particularly for phonetically ambiguous words, and an increase in irrelevant information in written discourse, as well as impaired motor aspects and spatial organization are seen during the progression of the disease (Croisile, 1999; Henderson et al., 1992). Written language impairment in AD can also be observed on the discourse level, taking the structural information of the text into consideration including measures of syntactic complexity (see, e.g., Kemper et al., 1993; Pakhomov et al., 2011; Snowdon et al., 1996). A significant decrease in syntactic complexity has thus been observed in patients with AD compared to healthy controls both in spoken language (Ahmed et al., 2013; Forbes-McKay et al., 2014; Lundholm et al., 2018; Roark et al., 2011) and in written texts (Kemper et al., 1993; Tsantali et al., 2013). Snowdon et al. (1996) measured the idea density in letters from applicants to an order of nuns. They showed a correlation between idea density in early adulthood, measured in the application letters, and the presence of AD in late adulthood. In addition to the cognitive impairments, Riley et al. (2005) found a significant association between low idea density and neuropathology, including more severe neurofibrillary pathology and cerebral atrophy. Kemper et al. (2001) examined changes in grammatical complexity and idea density in a group of healthy elderly as well as in a group of elderly with AD. There was an age-related decline in grammatical complexity and idea density, especially from the mid-70s. Participants with a higher initial level of complexity showed a slightly faster rate of decline. In the dementia group, the decline was accelerated by the disease regardless of age. These results are important as they show that the decline in complexity with increasing age first reported by Snowdon et al. (1996) continues in late adulthood, although some studies have not observed this decline in complexity with increasing age, possibly due to the different measures used (Nippold et al., 2014).

Different measures of syntactic complexity have been used in the literature but two of the more prominent ones are idea density, sometimes referred to as propositional density, and average dependency distance (ADD). Other measures used in previous research include Frazier score (Frazier, 1985) and Yngve depth (Yngve, 1960), as well as the number of words, sentence length, and total number of clauses in the text. Idea density was originally constructed as a measure of readability and is based on the number of propositions in a given text (Kintsch & Keenan, 1973; Turner & Green, 1977), whereas ADD is based on the average distance between words and their syntactic dependents in a phrase or sentence. In contrast to constituency grammar, which describes the syntactic structure of an utterance as a hierarchy of phrases, dependency grammar models this structure in terms of asymmetric relations between individual words. The two words in a dependency relation are called the head and the dependent. These relationships can be described as the unequal relationship between a parent word and the dependent (Hudson, 2010, p. 147). As an example, frozen is dependent on pizza in the phrase frozen pizza, it depends on its parent to be used correctly. In contrast to idea density, ADD is theoretically anchored in models of working memory. The multicomponent model of working memory presented by Baddeley and Hitch (1974), and later refined by A. D. Baddeley (2000), is probably the best-established model of working memory. In the multicomponent model, working memory is defined as “...a brain system that provides temporary storage and manipulation of the information necessary for such complex cognitive tasks as language comprehension, learning, and reasoning” (A. Baddeley, 1992, p. 556). The revised model has four components: The phonological loop and visuospatial sketchpad act as short-term storage for verbal and visual information. The episodic
buffer merges information held in short-term storage with information from long-term memory and acts as a system for temporary storage of integrated information. Lastly, the central executive coordinates the process and diverts information to the subsidiary systems. Several studies have shown reduced working memory capacity in AD, particularly related to the central executive (e.g., Baddeley et al., 1991; Huntley & Howard, 2010).

How is working memory capacity related to ADD? Gibson’s Dependency Locality Theory (DLT; Gibson, 1998, 2000) is based on dependency distances and contains two components: a storage cost, that is the storage of a word in anticipation for one of its dependents, and an integration cost, which can be defined as the capacity required to integrate the stored word into the sentence. This integration cost is strongly influenced by the locality of the dependents. Thus, the greater the distance, the greater the cost, whereas the storage cost is related to the number of syntactic categories needed to complete the sentence. A limitation in working memory capacity can, therefore, according to Gibson’s theory, decrease the dependency distances in order to lower the cost of storage and integration.

Estimating the dependency distance in written texts is a time-consuming process and requires previous knowledge of linguistics, which could reduce the applicability in a clinical setting. Therefore, several automated analysis methods have been developed and evaluated using different types of language samples. For example, Lundholm et al. (2018) used automated methods to analyze group differences in syntactic complexity measures as well as other speech-based measures in healthy controls (HC), subjective cognitive impairment (SCI), and mild cognitive impairment (MCI) using transcripts from spoken narratives based on the Cookie Theft picture description task of the Boston Diagnostic Aphasia Examination (Goodglass & Kaplan, 1983). Several complexity measures were used, including ADD. A random forest classifier was used to distinguish the groups from each other. There was no observed significant variation between the three groups on measures of syntactic complexity after correcting for multiple comparisons although some results, including an increase in the number of false starts, approached significance. Despite the lack of significant group differences on individual measures, the complete model showed a discriminative function and was able to distinguish MCI from SCI and MCI from HC, but not SCI from HC. Other studies have reported similar results (Ahmed et al., 2013; Mueller et al., 2018, 2016; Roark et al., 2011).

These cross-sectional studies measure differences between groups with different levels of cognitive impairment. Confounding factors are a potential problem, for example, educational level. One way of measuring the effect of cognitive impairment on syntactic complexity measures while controlling for confounding factors is case studies. Garrard et al. (2004) and Pakhomov et al. (2011) analyzed written texts from Iris Murdoch, an Irish author diagnosed with AD shortly after publishing her last book – Jackson’s Dilemma. Garrard et al. found significantly reduced lexical diversity but there were no differences between the texts on measures of syntactic complexity. In contrast, Pakhomov et al. observed a clear pattern of decline in syntactic complexity. This difference might be attributable to the sensitivity of the measures used to detect these differences as Pakhomov et al. used several computational approaches, for example, ADD, Yngve depth, and Frazier score, whereas Garrard et al. calculated syntactic complexity based on the number of words and clauses per sentence as well as words per sentence-ending marker.
A potential limitation with many of the above-mentioned measures is the lack of theoretical motivation for their application in cognitive disorders. Because ADD is based on models of working memory, this measure should be valid for AD in which working memory impairments, as previously mentioned, are common. There is however a knowledge gap regarding ADD as a marker for cognitive impairment. Although written texts are already being collected during assessment at the memory clinic at which data for this study was collected, ADD is currently not used in the diagnostic process.

The aim of this study is therefore to evaluate possible associations between the specific measure of ADD in narrative writing and progression of cognitive impairment, from subjective cognitive impairment to AD. For this purpose, the following research question is proposed:

Is there a significant association between ADD in written narratives and the level of cognitive impairment as reflected by diagnostic categories?

Several previous studies investigating syntactic complexity – for example, Kemper et al. (1993) and Snowdon et al. (1996), two seminal studies in the research of language and dementia – have used idea density as the outcome measure. We will therefore assess the association between idea density and diagnosis in a secondary regression model to allow comparison of the two measures.

Methods

Participants

Written text samples were collected in a consecutive series of 138 patients at the Memory Clinic, between 1999 and 2001. All patients underwent a general medical, neurological, and neuropsychological investigation, as well as neuroimaging and standard blood and cerebrospinal fluid diagnostic procedures to exclude other treatable causes of cognitive impairment or dementia apart from a neurodegenerative etiology. Apart from the text samples, background information regarding age, gender, years of formal education, Mini-Mental State Examination score (MMSE, Folstein et al., 1975), and diagnosis was collected from each patient’s medical chart. Language background was not specifically documented, but it can be estimated that a large majority of the participants had Swedish as their first language.

Table 1 summarizes the descriptive statistics for the sample. The patients comprised three diagnostic groups:

(a) Subjective cognitive impairment (SCI). These patients had been referred for subjective cognitive complaints such as forgetfulness but did not fulfill criteria for mild cognitive impairment and thus did not have objective cognitive impairment. SCI is

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>N</th>
<th>Male/Female</th>
<th>Age</th>
<th>Education</th>
<th>MMSE score</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCI</td>
<td>28</td>
<td>9/19</td>
<td>58.7 (9.7)</td>
<td>10.9 (3.1)</td>
<td>28.7 (0.9)</td>
</tr>
<tr>
<td>MCI</td>
<td>41</td>
<td>14/27</td>
<td>62.5 (10.6)</td>
<td>11.4 (4.1)</td>
<td>26.6 (4.2)</td>
</tr>
<tr>
<td>AD</td>
<td>45</td>
<td>18/27</td>
<td>67.4 (8.8)</td>
<td>10.5 (3.9)</td>
<td>22.1 (4.4)</td>
</tr>
<tr>
<td>Total</td>
<td>138</td>
<td>49/89</td>
<td>63.6 (10.3)</td>
<td>10.7 (3.6)</td>
<td>28.7 (0.9)</td>
</tr>
</tbody>
</table>
a predominantly benign condition because a large majority of cases remains cognitively stable or improve on long-term follow-up (Hessen et al., 2017).

(b) Mild cognitive impairment (MCI). Patients with MCI were diagnosed according to modified Petersen criteria (Winblad et al., 2004). They thus had objective impairment (defined as at least 1.5 standard deviation unit below the age-adjusted mean score) in at least one cognitive domain, typically memory, but had essentially preserved activities of daily life and did not fulfill the ICD-10 criteria for dementia.

(c) Dementia in Alzheimer’s disease (AD). Patients with AD were diagnosed according to ICD-10 criteria (World Health Organization, 1992). They had objectively verified memory decline and decline in other cognitive abilities, causing impairment in daily living. The severity of dementia was mild on average.

The original case series also included other cognitive disorders, including variants of frontotemporal dementia, Korsakoff syndrome, Lewy body dementia, mixed dementia, vascular dementia, and Parkinson syndromes; these cases were not included in this study.

Procedure

The texts were written by hand on an A4 piece of unlined paper. The narrative writing task from the Boston Diagnostic Aphasia Examination (Goodglass & Kaplan, 1983) was used and administered by a speech-language pathologist with extensive experience of dementia assessment. The participants were instructed to describe what is happening in the Cookie Theft picture as thoroughly as possible. No strict time limit was set for the task, although most participants finished within the two minutes time frame allotted in the manual.

Analysis of syntactic complexity

Automated analysis methods were used for the analysis of ADD in the written text samples. The hand-written texts were originally digitalized in a Microsoft Word document by Landfeldt and Söderbäck (2009) using a revision matrix in cases where the texts lacked stylistic elements such as punctuation and capital letters. Each text was converted from Microsoft Word to plain text files and manually tokenized into words and punctuation marks. Some texts had to be revised, mostly because of spelling errors. Instructions for revisions were constructed and used throughout the revision process. In the next step, a pipeline consisting of the part-of-speech tagger Stagger (Östling, 2013) and the syntactic dependency parser MaltParser version 1.9.1 (Nivre et al., 2007) was used to generate dependency trees for the texts. Both the tagger and the parser used the standard Swedish language models available for them; this yields an unlabeled attachment score of 82.37% when evaluated on the gold-standard test section of the Swedish Treebank (Nivre, 2016; https://cl.lingfil.uu.se/~nivre/swedish_treebank). In the last step of the system, a Python script was used to calculate ADD in the output file produced by the dependency parser. See appendix A for a visualization of the analysis process.
Reliability analysis

Six written narratives were selected randomly from each group (SCI, MCI, AD) to evaluate the inter- and intra-rater reliability of the revision template described above. Four independent raters used the revision template to restructure the written narratives to the required format for the syntactic parser. The first author repeated the revision process for the six written narratives and compared them to the original revisions to estimate the intra-rater reliability. Percentage agreement was used to evaluate the two reliability measures. The agreement was calculated as a binary outcome -yes/no- for each separate line of the text files from the raters, with total consensus as the criterion for agreement. To avoid systematic errors due to, for example, the lack of a blank space in one of the texts, this was inserted and coded as an error.

Statistical method

Analyses were performed in Statistica version 13 (TIBCO Software Inc, 2018) and SPSS (IBM Corp, version 25). ANCOVA was used to assess whether group differences in education and years of education were confounders of diagnostic differences in ADD. The alpha level was set at .05 in all analyses. Bonferroni corrections of the alpha level were not used as the overall null hypothesis for all analyses was not of primary interest in the present study; see Armstrong (2014), Perneger (1998), and Nakagawa (2004) for discussion.

One case in the SCI group was excluded from the analysis based on a Cook’s D value >1, classifying it as an outlier. This observation had a very high leverage of almost twice the numerical value on the primary outcome measure as compared to the closest observation. After inspection of the text file, the decision was made to exclude the observation due to a probable measurement error.

Ordinal logistic regression was used to evaluate the association between the dependent variable diagnosis – categorized as SCI, MCI, and AD reflecting an increasing level of cognitive impairment – and the independent variable ADD. As idea density has been used as a measure of syntactic complexity in previous studies, an additional regression model with idea density as the independent variable was used to allow comparison between the two measures. Nagelkerke R² (Nagelkerke, 1991) was used to evaluate the two models. Assumptions for ordinal logistic regression were met, including the assumption of proportional odds as shown by a non-significant test of parallel lines (p = .69).

Ethical considerations

The project was approved by the Regional Ethical Review Board in [location of author’s institution] (registration number 2018/1114-31/2).

Results

A summary of ADD, idea density, and number of words for each diagnostic group is presented in Table 2.
Did age and education influence the comparison between the groups?

To rule out age and education as confounders of diagnostic differences in ADD, a preliminary analysis of covariance was performed with diagnostic group (SCI, MCI, AD) as the independent category variable and age and education in years as covariates. There was a significant effect of diagnostic group on ADD \( [F (2, 105) = 9.09, p < .001] \) but no significant effects of age \( (p = .51) \) or education \( (p = .64) \). Age and education were therefore not included as covariates in the subsequent analyses.

Is ADD associated with diagnosis?

A regression model consisting of ADD as the independent variable and diagnosis as the dependent variable was constructed for the analysis.

The model indicated that ADD is a significant predictor of diagnosis \( [B = -2.05, \text{ Wald} = 22.1, p < .001] \). Goodness-of-fit was assessed using the Deviance test which was non-significant \( (p = .37) \) with approximately 25% of the variation explained by the dependent variable \( (\text{Nagelkerke } R^2 = .25) \). However, the Pearson chi-square, another common test of goodness-of-fit for logistic regression, was significant \( (p < .001) \).

Is idea density associated with diagnosis?

A regression model was constructed using idea density as an independent variable and diagnosis as the dependent variable.

Idea density was not significantly associated with diagnosis \( [B = -.25, \text{ Wald} = 1.08, p = .30] \). Again, goodness-of-fit was assessed using the Pearson chi-square \( (p = .31) \) and Deviance test \( (p = .09) \), indicating a good fit of the data to the model with approximately 1% of the variation explained by the dependent variable \( (\text{Nagelkerke } R^2 = .01) \).

Inter- and intra-rater reliability

The inter-rater reliability of the revision template was evaluated using percentage agreement. Over the six included texts, the percentage agreement was 94.7%. Intra-rater reliability, measured before and after the main analysis, was 98.8%.

Discussion

The results of the present study show a significant association between ADD and level of cognitive impairment ordered as SCI, MCI, and AD. Based on Nagelkerke \( R^2 \), approximately 25% of the variation in the independent variable was explained by the dependent

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>ADD</th>
<th>ID</th>
<th>Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCI</td>
<td>2.47 (0.60)</td>
<td>4.12 (0.54)</td>
<td>54.07 (28.05)</td>
</tr>
<tr>
<td>MCI</td>
<td>2.16 (0.40)</td>
<td>4.32 (0.53)</td>
<td>35.44 (15.64)</td>
</tr>
<tr>
<td>AD</td>
<td>1.82 (0.58)</td>
<td>4.01 (0.94)</td>
<td>22.64 (11.08)</td>
</tr>
<tr>
<td>Total</td>
<td>2.10 (0.58)</td>
<td>4.15 (0.73)</td>
<td>34.96 (21.76)</td>
</tr>
</tbody>
</table>
variable. The interpretation of the explained variance is highly dependent on the data and research question. However, we argue that the results are at least strong enough to warrant further studies of the association. In contrast, the explained variance in the model with idea density as the independent variable was markedly lower with only approximately 1% explained variation, and the association non-significant.

In a recent study, Lundholm et al. (2018) used several complexity measures including ADD to study differences in cognitive impairment and a control group. While the model as a whole successfully distinguished MCI from both SCI and healthy controls, there were no significant differences on any specific complexity measure, including ADD, after correcting for multiple comparisons. This discrepancy between our results and those of Lundholm et al. might be partly explained by the task. Although both studies used the Cookie Theft picture to elicit language samples, we used written narratives as opposed to the spoken narratives used by Lundholm et al. As discussed by Croisile (1999), written language might be more sensitive to early changes in linguistic ability than spoken language. Along the same line of reasoning, Vandenborre et al. (2018) showed that written picture descriptions were more sensitive than oral picture descriptions in discriminating individuals with aphasia from healthy controls. Ardila and Surloff (2006) further argue that executive impairment in, for example, dementia might be related to decreased written language complexity because of reduced planning capacity and difficulties in maintaining attention, leading to a less coherent text with lower complexity. Such observations could be related to cognitive writing models such as the one suggested by Hayes (2012). According to the model, writing is recursive so that the focus shifts between the different processes involved. A reduced planning capacity might therefore result in a less coherent text due to a limited ability to structure the text and ideas to be conveyed.

As described above, there was an association between ADD and cognitive diagnosis, but in contrast to previous studies, the relative value of idea density was limited. One possible explanation for this might be the differences in the theoretical framework of the two models. Idea density was constructed as a measure of complexity related to the readability of texts and is primarily used for texts that are longer than those analyzed in the present study. In contrast, ADD is more closely associated with measures of working memory. Gibson’s Dependency Locality Theory (Gibson, 1998, 2000) takes inspiration from Baddeley’s model of working memory and is based on the assumption that longer dependency distances place a greater burden on working memory. As a consequence, longer dependencies increase the storage cost but also the integration cost as more words are stored that need to be connected in written sentences. Therefore, reduced working memory capacity is a possible explanation for the association between ADD and diagnosis as the memory impairments in AD limits the ability to retain the dependencies in the working memory, leading to a lower ADD. According to cognitive writing models, several writing processes compete for the limited working memory capacity. A decrease caused by neurodegeneration might therefore further limit the resources available during writing. This in combination with the previously mentioned planning deficits could further increase the demands on an already reduced working memory capacity, resulting in cognitive overload and lower syntactic complexity. Since working memory impairments are common in AD and the presence of an association between ADD and diagnosis is not that surprising.
Methodological discussion

The data used in the present study was collected consecutively approximately 15 years ago. Even though minor revisions have been made to the diagnostics for neurodegenerative disorders and the specificity might be higher if the data was collected today, this is no major limitation as all participants were tested and diagnosed using the same criteria. The writing task used was not originally designed to measure syntactic complexity specifically and is used in the clinic in conjunction with several different tests. A more challenging writing task might elicit higher levels of complexity, but it might also lead to an inability to complete the task in patients with more severe perceptual or cognitive impairment. Because we still found a substantial association between ADD and level of cognitive impairment in this sample, more difficult writing tasks may prove to have greater clinical relevance but be less suitable for patients with dementia. Another important factor for future studies is differences in motivation. As discussed in the introduction, the motivation can affect how much time and effort the writer devotes to the task. The instructions for the writing task should ideally be formulated to encourage the writer to describe all elements in the text as carefully as possible.

The highest level of cognitive functioning in our sample was represented by the SCI category. Patients with SCI report symptoms such as increased forgetfulness in everyday life but perform within (or above) normal limits on neuropsychological tests. It is a common group in memory clinics (Wahlund et al., 2003). Studies of SCI patients in memory clinics show that it is a primarily benign condition often associated with high stress levels and depressive symptoms (Eckerström et al., 2016), with a majority remaining stable over time or with regression of symptoms over time (Hessen et al., 2017). It would be interesting to explore whether SCI is associated with subjective complaints related to writing skills. This would probably be most prevalent in patients with high-demanding or creative writing occupations.

A specific methodological issue relates to the goodness-of-fit of the ordinal regression model. The Pearson chi-square test for goodness-of-fit was significant for the model with ADD as the independent variable. This might be due to the sample size and magnitude of the independent variable. The numeric values and the spread of values are narrower for ADD as compared to idea density. In such circumstances, very small differences between the assumptions in the given test might have a considerable effect on the p-values. Because of this, and because the Deviance test (another measure of goodness-of-fit commonly used in logistic regression) was non-significant, we chose to report the results from the model.

Manual calculation of syntactic complexity measures such as ADD and idea density are time-consuming to calculate and require a high degree of linguistic knowledge. This reduces the applicability of the measures in a clinical setting. In the present study, we used automated analysis methods of ADD. Using a part-of-speech tagger and a syntactic parser, in addition to a Python script written for the analysis of ADD, we automatized the process. The use of automated analysis methods increases the reliability of the method as all texts are analyzed using the same steps by the computer. On the other hand, the validity of the analysis might be reduced as the part-of-speech tagger and the parser are trained using machine learning and thus are far from perfect. Because we lack gold standard dependency annotations for the patient texts, the quality of the automatic annotation cannot be assessed justly. The written materials on which the parser is trained were collected during the 1990s and consist primarily of government reports and
newspaper articles. Thus, there is a qualitative difference in the text materials and also a potential for words unknown to the parser in the patient texts. However, the picture used for the picture description task collected from the patients are temporally neutral with limited room for new words. The complexity of the patient material is most likely lower than of the material used to train the model, and the applicability needs to be analyzed further in future studies, even though the complexity most likely is not a problem for the model. Despite this fact, the possibility to utilize the measurements of syntactic complexity from written texts already collected routinely in a clinical setting might add to the diagnostic procedure. Future studies need to evaluate the precision of the analysis method compared to gold standard for calculating ADD.

One possible source of variation is the necessary revisions of written narratives to a format accepted by the programs. To aid the revision process, a template was constructed, and an evaluation of the inter- and intra-rater reliability was made. Both the inter-rater and intra-rater reliability were high at 94.7 and 98.8 percent, respectively.

In conclusion, our results showed a significant association between ADD and level of cognitive impairment, whereas idea density showed no significant association. ADD thus seems to account better for the association with diagnosis than idea density, as was also reflected by much higher explained variance. Longitudinal studies are needed to evaluate the predictive potential of ADD for cognitive decline. Further reliability studies are also needed to evaluate the clinical applicability of the automatized analysis methods.

**Disclosure Statement**

No potential conflict of interest was reported by the authors.

**Funding**

This research was supported by Alzheimerfonden (The Swedish Alzheimer's Foundation) and Karolinska Institutet.

**ORCID**

Fredrik Sand Aronsson http://orcid.org/0000-0001-8452-0043

**References**


IBM CorpSPSS statistics for windows, Version 25.0.


