

A Morphological Approach to the Longitudinal Detection of Dementia

Najoung Kim

School of Computing, Korea Advanced Institute
of Science and Technology
291 Daehak-ro, Daejeon, Republic of Korea
njkim@nlp.kaist.ac.kr

Jong C. Park*

School of Computing, Korea Advanced Institute
of Science and Technology
291 Daehak-ro, Daejeon, Republic of Korea
park@nlp.kaist.ac.kr

ABSTRACT

The impact of cognitive impairment on linguistic abilities has been a topic of continuous interest in dementia studies. However, there is a lack of systematic agreement on the longitudinal association between dementia progression and the patients' morphological capacity, and the role of morphological phenomena other than inflection has been relatively underreported. We present a longitudinal study of writings by Iris Murdoch (diagnosed of Alzheimer's Disease after her death) and Arthur Conan Doyle (no known record of dementia diagnosis), using two novel measures to account for the usage of complex morphology and lexical innovation. The results imply an association between lexical innovation and cognitive decline caused by dementia, as observed in Murdoch's works beginning from her mid-fifties, in contrast to a milder tendency in Doyle's works. Our findings contribute to a potential for facilitating early diagnosis of dementia through automated language processing approaches.

Author Keywords

Lexical Innovation; Morphology; Dementia; Alzheimer's Disease; Mild Cognitive Impairment, Natural Language Processing; Automatic Text Analysis.

ACM Classification Keywords

I.2.7 Natural Language Processing: Text analysis.

INTRODUCTION

The goals of this study are twofold. First, we investigate whether we could establish a statistically significant connection between the progression of dementia and the subjects' command of morphology. Whether the trends found in previous dialog-based experiments are similarly observed in text analysis or not, and whether including non-inflectional morphology yields a different result would be the focal points of this research. Second, if it is indeed the case that a significant correlation can be established between the subjects' command of morphology and dementia progression, we would like to propose an automation guideline in order to facilitate future language processing applications for the detection of early symptoms of dementia or signs of a mild cognitive impairment (MCI) phase.

Motivation

Dementia, such as Alzheimer's Disease (AD), is known to have a long preclinical phase; for example, a decade or longer for AD [22]. Prior research suggests that the patients go through a transitional stage of mild cognitive impairment (MCI) from healthy aging to dementia, in which the symptoms are often non-obvious or mild [16, 17, 20]. For this reason, identifying early signs of cognitive decline in an individual has been an important issue for clinical dementia research [23].

Studies have been conducted to apply a text processing approach to this issue, examining dementia patients' linguistic characteristics manifested in writings in a longitudinal manner [5, 6, 9, 12, 15, 18]. However, the role of morphology has not been investigated in such studies. The effect of dementia on morphology has been explored mostly in oral, task-based experiments involving a real-time speech production of the subjects. As noted in [1], there is a lack of research on the association between morphological competence¹ and diachronic progression of the disease. Moreover, most existing studies have exclusively considered inflectional morphology [1].

Building on from these previous studies, we attempt a longitudinal examination of potential declines in morphological capacity reflected in the subjects' writings. We furthermore expand the types of morphological phenomena accounted for by capturing non-inflectional paradigms through novel quantification methods. We also investigate whether these methods can be developed or integrated into an automated language processing system that detects preclinical stages of dementia.

Related Work

A similar longitudinal study has been conducted in [9] and [15] using the same dementia-diagnosed subject – Iris Murdoch. [9] examined novels by Murdoch in terms of diachronic lexical and syntactic changes. However, as pointed out by [15], this analysis was limited to 3 of her novels and showed problems in methodology. [15] presented a reinforced, larger-scale study of Murdoch's novels and additionally analyzed the works of P.D. James

¹ The term “competence” in this article is not used in the context of any particular linguistic theory where the term bears a contrastive meaning to “performance.”

* Corresponding Author

(healthy control) and Agatha Christie (potential undiagnosed dementia patient) for comparison. Their results support the argument that certain linguistic changes observable in dementia patients' writings over time are correlated with diachronic progression of the disease. This tendency was clearer in lexical decline than in syntactic decline.

Numerous non-longitudinal linguistic studies have also been conducted to discover the effects of dementia and also healthy aging on morphology. [25] suggested that the ability to produce regular past-tense verbs was retained compared to irregular verb production in patients with AD and posterior aphasia, whereas grammatical rules were impaired in Parkinson's Disease and anterior aphasia patients. [4] examined past-tense verb productions in AD and semantic dementia patients, as well as in healthy older adults, finding that consistency and regularity of the past-tense forms affected performance. [24] reported an impaired production of verbal inflection and argument structures in the agrammatic variants of dementia, but not in logopenic and semantic variants. However, [8] did not find significant differences in use of inflected verbs between subject groups with progressive nonfluent aphasia, groups with semantic dementia and the control group. Although there is no systematic agreement on the effect of dementia (and their subtypes) on morphology, the results from these previous studies suggest that consistently predictable, generalizable morphological paradigms may be retained in certain types of dementia, such as AD.

Moreover, previous studies have reported impoverished lexical diversity in dementia patients [10, 13]. We hypothesize that a decrease in lexical innovation could be partially contributive to this phenomenon, alongside with the deficit in lexical access, which is generally agreed to be a cognitive effect of dementia.

Hypothesis

We present two novel measures to account for two linguistic properties in this study: the usage of complex morphology and lexical innovation. Drawing from the currently available experimental works on the effects of dementia on morphology, we predict that complex morphology – mostly regular and concatenative in English – would remain intact in Murdoch's data (taking into account that the subject had dementia of the AD type), whereas lexical innovation would show a declining trend. As it is possible that healthy aging might also be a factor in linguistic decline, we will additionally examine the diachronic patterns observed in Doyle's data and provide comparative analysis.

METHODOLOGY

We processed and analyzed novels written by British authors Iris Murdoch and Arthur Conan Doyle spanning their literary careers. The analysis was conducted in terms of two measures: complex morphology and lexical innovation.

Data

12 novels by Murdoch and 13 novels by Doyle were chosen as experiment data, with regards to their genre and the author's age when the novels were published. Murdoch's novels analyzed in this study were published between ages 35 and 76 (died 80), and Doyle's between 27 and 66 (died 70). As Doyle's novels consist of two distinct genres, historical and detective/adventure fictions, we focused on the 8 novels from the latter genre only in the lexical innovation measure to reduce potential lexical biases [14]. Doyle's novels that are collections of short stories rather than a full-novel narrative were also excluded. The full list of novels used in this experiment, with the respective titles and the authors' age, is given in the appendix.

Linguistic Measures

In order to observe the possible effect of dementia on the patients' command of morphology and lexical productivity, we devised two linguistic measures that have not been applied to similar longitudinal studies to the extent of our knowledge.

The first measure quantifies the usage of complex morphology. Any token that consists of more than a single morpheme is considered complex. Thus, this measure captures both inflectional and derivational morphology – phenomena most frequently represented by regular concatenative paradigms. The proposed measure is to calculate the number of complex word tokens over the total number of word tokens. We also designed a multimorpheme-sensitive version of this measure in order to observe the effects of word tokens that contain more than two morphemes. In this version, we assign weights to each complex token count instead of treating every token count equally. The weight is proportional to the total length of the concatenated morphemes. For non-concatenative morphology such as common irregular past tense verbs (e.g., 'meet'-'met', 'sit'-'sat'), absolute values of length differences are used as weights for complex words shorter than their simple forms, and a default weight of 2 is used for complex words of unchanged length. To improve the accuracy of morphological analysis, we first tagged the data with a part-of-speech tagger based on a maximum entropy model [21], and used this information to retrieve the simple lemmatized form of a given word token using the Morphy module of WordNet [3] with Porter Stemmer [19] as backoff.

We predict that the subject datasets would not exhibit a significant correlation between the ratio of complex morphology and the authors' ages, based on the previous experimental results that support regular inflectional morphology retention in dementia patients [4, 8, 11, 25]. The measure's ability to capture applications of regular derivational morphological rules such as adjectival and adverbial suffixes (e.g., '-ly', '-ic') might affect the results in an unexpected manner, but considering the fact that these are also regular rules, it is most likely that no particular pattern would be observed over time or between individuals.

The second measure quantifies lexical innovation, which is a factor reflective of morphological productivity. This loosely follows the views presented in [2]², in which some complex words are considered to be products of morphological processes rather than retrieved from the lexicon. There is no clear-cut boundary that distinguishes complex words that are produced by rules from ones that are stored, or lexicalized. For the purpose of this study, we define exact word token forms that are unobserved in a reference corpus as lexically innovative. The Brown Corpus [7], a large collection of linguistic data published in 1961, was used as the reference in the experiments described here. Further comments will be made with regard to corpus choice in the discussion section. This primary filtering was reinforced with an online dictionary search, in order to correctly remove existing words that were unobserved in the reference corpus due to their inherently low frequency. British spelling variants of existing words were also removed at this stage (e.g., *colour*, *leveller*, *patronise*). The results after this stage were examined to determine categories of word tokens that may be irrelevant in measuring morphological productivity. The seemingly irrelevant categories were as follows:

- Proper names: Names of existing locations (e.g., *Bayswater*, *Kempsford*), famous people or fictional characters (e.g., *Dmowski*, *Pushkin*, *Prospero*), common names of people (e.g., *Ned*, *Angela*), names of species (e.g., *Atypus*, *atrica*), etc.
- Foreign language: When the author uses a foreign language in full rather than borrowing a single word or a phrase (e.g., *Penché sur elle l'ardent impérateur*; each word token primarily detected as innovative).
- Errors and typos: Errors from the original text data or errors from preprocessing (e.g., *somewhat*, *obsession*, *think-you*).
- Non-morphological creativity: Word tokens that might be considered lexically innovative but are evidently not created by morphological processes: onomatopoeia (e.g., *sssh*, *aaah*), borrowings from foreign languages (e.g., *soto-gari*, *maussade*), etc.
- Dialectic variances: Deliberate changes in spelling in order to reflect characteristics of certain dialects (e.g., *doin'*, *hain't*).

The remaining unique word tokens discarding the tokens from these categories are considered lexically innovative and also morphologically productive. The lexically

² The term lexical “innovation” was chosen as the name of the new measure as it was proposed as a hypernym that includes both productivity (rule-governed process) and creativity (not rule-governed) in [2]. Although our experiments only included seemingly rule-governed processes, a measure that incorporates creativity can easily be obtained by not discarding certain categories described in the methodology section.

innovative tokens involve various types of morphological processes including inflection, derivation, compounding and blending. The actual measure is calculated by dividing the number of unique, lexically innovative word tokens by the number of all word tokens. The number of tokens belonging to the “Errors and typos” category and the “Dialectic variances” category is deducted from the denominator.

The lexical innovation measure was designed presuming a possible correlation with the progression of dementia, but we cannot commit to a particular prediction based on previous experimental results before conducting an actual test, due to the novelty of the proposed method in dementia research.

ANALYSIS

The correlation between age and the results from the two presented morphological measures is tested by Spearman's rank-order correlation coefficient method, following the study in [15]. Line of best fit given by running a simple linear regression is also shown in each figure.

Author	Type	ρ	p -value
Murdoch	Sensitive	0.3479	0.266
Murdoch	Insensitive	0.2867	0.3664
Doyle	Sensitive	-0.3356	0.2622
Doyle	Insensitive	-0.1816	0.5527

Table 1. Correlation between age and complex morphology

For the complex morphology measure, we found a negligible correlation between age and both the multi-morpheme sensitive and insensitive measures in each author's writings. Diachronic change in Iris Murdoch's complex morphology was insignificant, and the correlation between her age and the sensitive/insensitive measures was negligible (MM-Sensitive: $\rho = 0.3479$, p -value = 0.266 /

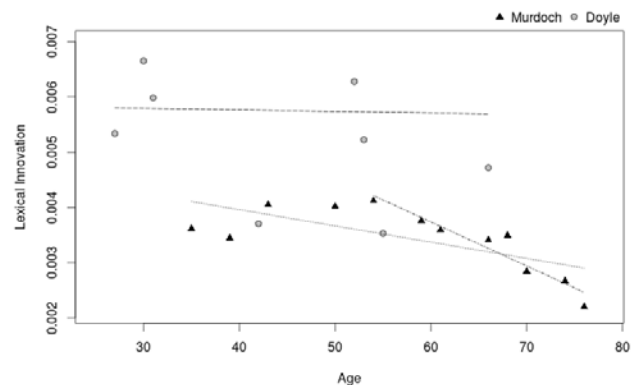


Figure 1. Lexical innovation measure

MM-Insensitive: $\rho = 0.2867$, p -value = 0.3664). Complex morphology measures in Doyle's novels also showed

weak/negligible correlation (MM-Sensitive: $\rho = -0.3356$, p -value = 0.2622 / MM-Insensitive: $\rho = -0.1816$, p -value = 0.5527).

The lexical innovation measure yielded different results for the two authors. In Murdoch's writings, we observed a clear reverse correlation between age and lexical innovation ($\rho = -0.7203$, p -value = 0.011). Additionally, a stronger declining trend could be observed from age 54, when her lexical innovation peaks and then drops continuously to reach an all-time low at age 76. The reverse correlation from age 54 and onwards was particularly strong ($\rho = -0.9762$, p -value = 0.0004). In Doyle's case, there was a reverse correlation of a more moderate kind ($\rho = -0.5952$, p -value = 0.1323), but no clear beginning point of a declining trend as in Murdoch's data could be identified.

Author	Age range	ρ	p -value
Murdoch	35-76	-0.7203*	0.011
Murdoch	54-76	-0.9762*	0.0004
Doyle	27-66	-0.5952	0.1323

Table 2. Correlation between age and lexical innovation
(* $p < 0.05$)

We also calculated cumulative lexical innovation figures for each approximately 10 to 15-year interval. The results showed a much steeper decline in Murdoch's writings over time compared to Doyle's.

DISCUSSION

The analysis results obtained from applying the two measures seem to bear out our original presumptions. The complex morphology measure did not show a significant correlation in either of the authors' dataset, implying that complex morphology is retained over time and also over the progression of dementia. On the other hand, the lexical innovation measure showed a strong reverse correlation with age in Murdoch's works but not in Doyle's,

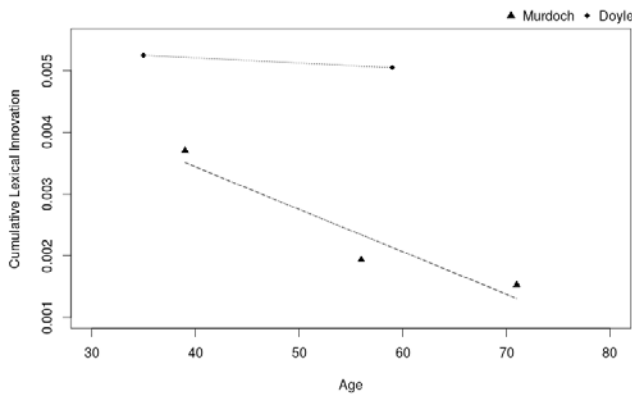


Figure 2. Cumulative lexical innovation

suggesting that dementia progression could be associated with a sharp decline in lexical innovation. Moreover, the

declining trend was clearly observable beginning from her mid-fifties. It is possible that this starting point of the decline is related to the preclinical stage of the disease, and this is more than 15 years before the signs of her linguistic decline were thought to be evident in her last novel *Jackson's Dilemma*, published in 1995. Of course, we cannot state for sure that there is a causal relationship between a declining trend in lexical innovation and the progression of preclinical dementia; we can only suggest a possible association. Further claims must be supported by more experimental evidence. If it is indeed the case, we could develop an automated language analysis system that detects the decline, or integrate this factor into an existing system for better reliability. The following section proposes a brief proposal for such an automation.

Automation

Although the complex morphology measure was calculated in a fully automatic manner, the lexical innovation measure involved a manual filtering effort. We now suggest an automated process that preserves the statistical tendencies manifested in the hand-filtered output.

The irrelevant tokens that were manually removed can be classified into the five categories as presented in the methodology section. We could employ an automation method that identifies the given word categories by utilizing a Named Entity Recognizer for proper names, a language detector for foreign language tokens, an edit-distance based algorithm for errors/typos and so on. This multi-module approach would yield results closer to the hand-filtered gold standard. However, we can devise a method that is even simpler by focusing on the characteristics of the tokens that are not filtered out. The relevant tokens are products of morphological rules, and would be composed of two or more valid morphemes or words. Thus, we test an automation method that considers lemmatizable word tokens as valid. Of course, this is oversimplification that does not take into account the role of external world knowledge. For example, because the lemmatizer used in the experiments does not know that 'Bayswater' is an existing location, it returns 'Bayswat' as its stem, erroneously judging this word token as lemmatizable. This system cannot account for errors and typos either. Nevertheless, the results show that the statistical tendencies shown in the gold standard are preserved in the automated results, though slightly weakened. Table 3 shows the results for the same dataset used in Table 2, calculated from word tokens obtained using automated filtering. We additionally tested the correlation including both genres of Doyle's novels. The correlation tendencies are preserved in all datasets, though the distinction has become less projecting. This implies that our simple lemmatizability-based automation method is a viable baseline, and the results would be more refined as we add more world knowledge to the system.

Author	Age range	ρ	p -value
Murdoch	35-76	-0.7343*	0.0091
Murdoch	54-76	-0.8571*	0.0107
Doyle	27-66	-0.6905	0.0694
Doyle	27-66 (all)	-0.5790	0.0525

Table 3. Correlation between age and lexical innovation calculated from automated filtering (* $p < 0.05$)

Limitations and future work

In this section we discuss some limitations of the presented study and suggest reinforcements and expansions for future studies.

First, we need more subjects, larger-scale test results and controlled experiments in order to validate the proposed association between dementia progression and our new morphological measures. Examining the works by more dementia-diagnosed authors, such as Terry Pratchett, would be appropriate. Non-author subjects and non-literary writings could be studied as well. More supportive evidence from task-based experiments or biological studies would reinforce our proposal. Moreover, although we have shown the correlation between age (representative of dementia progression, following prior research) and our measures in the demented subject, other factors such as early education may have also influenced the results. Such possibilities must be investigated in future controlled studies. Considering Doyle's results, healthy aging also seems to be mildly responsible for the declining trends. Whether different variants of dementia affect lexical innovation in a distinct manner would also require further discussion.

Second, as the test data are published writings that underwent editing processes, the first measure for complex morphology may not accurately reflect the authors' true morphological capacities, based on the assumption that obvious mistakes in morphology would have been corrected. However, it is still meaningful to observe that the rate of complex morphology is diachronically regular assuming a low rate of morphological errors in the input, as this implies that neither of the subjects avoided using constructions requiring complex morphological forms. Moreover, the measure for complex morphology might be more useful when applied to non-edited inputs such as transcriptions of spontaneous speech or casual writing, where mistakes in morphology would actually surface.

Furthermore, there were some potentially erroneous assumptions we have made in the experimental process. For example, we do not know the exact ages of the authors at the point of composition for each literary work, so we used the ages from the year before the publication date, following [15]. The choice of the reference corpus was another issue; the assumption was that all the words that were present in the corpus were lexicalized and therefore not innovative. This also is oversimplification (and unlikely

to be true), and we wish to investigate the effects of varying the reference corpus and varying the definition of lexicalization. The time difference between the year that the literary data used in our experiments and the reference corpus were published may also affect performance. Potential solutions may be to use a corpus built from sources from a longer timespan, or to use different criteria for judging productive tokens such as token frequency in the given data.

Finally, if we are to expand the application of this detection process to actual patients, we must consider how to account for non-standard styles of writing, spoken language, or spoken-language style writing. In this study we discarded dialectal variations assuming that they do not reflect morphological productivity, but is this a valid assumption? How should we model other types of potential innovations such as zero derivations, spelling alternations, acronyms and more? We hope to explore such issues in future studies, as well as searching for more undiscovered syntactic and semantic factors that are indicative of preclinical stages of dementia.

CONCLUSION

We have designed two novel measures that account for two types of morphological competence, the use of complex morphology and lexical innovation, and tested their associations with dementia progression, using writings by two British authors Iris Murdoch and Arthur Conan Doyle as experiment data. The results showed that there was a continuous decline in lexical innovation in Murdoch's works, especially noticeable from her mid-fifties. Only mild declining tendency was observed in Doyle's data, without a clear starting point of the trend. This suggests a reverse correlation between lexical innovation and dementia progression, and possibly a weaker reverse correlation between lexical innovation and normal aging. No significant increasing or decreasing tendencies were found in the usage of complex morphology in either of the authors' data; the figures were generally regular over time. Although this retention of complex morphology is in line with some experimental observations [4, 8, 11], we leave open the possibility of a different result when applied to non-edited data such as transcriptions of spontaneous speech or casual writing, or when applied to data produced by patients with different variants of dementia. Furthermore, we have presented a baseline for the automation of the lexical innovation measure (complex morphology measure was already automatic), based on the polymorphemic property of tokens coined by productive morphology. This baseline method, although simple, preserves all tendencies expressed in the manually filtered, gold-standard results. Our findings will potentially contribute to the early diagnosis of preclinical stages of dementia by automatic text processing. As obtaining longitudinal data of individual speakers' linguistic records is becoming easier due to the popularity of social media and digital recording devices, we expect a diagnostic system for dementia based on diachronic linguistic analysis to be increasingly practical in the near future.

ACKNOWLEDGMENTS

We thank the members of the NLP*CL Laboratory at KAIST and the anonymous reviewers for providing helpful comments. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (NSIP) (No. 2010-0028631).

APPENDIX

Title	Year	Age
Under the Net	1954	35
The Bell	1958	39
An Unofficial Rose	1962	43
Bruno's Dream	1969	50
The Black Prince	1973	54
The Sea, the Sea	1978	59
Nuns And Soldiers	1980	61
The Good Apprentice	1985	66
The Book and the Brotherhood	1987	68
The Message to the Planet	1989	70
The Green Knight	1993	74
Jackson's Dilemma	1995	76

Table 4. Novels by Iris Murdoch

Title	Year	Age	Genre
A Study in Scarlet	1887	27	D/A
Micah Clarke	1889	29	H
The Sign of the Four	1890a	30	D/A
The Firm of Girdlestone	1890b	30	H
The Doings of Raffles Haw	1891	31	D/A
Rodney Stone	1896	36	H
The Tragedy of the Korosko	1898	38	H
The Hound of the Baskervilles	1902	42	D/A
Sir Nigel	1906	46	H
The Lost World	1912	52	D/A
The Poison Belt	1913	53	D/A
The Valley of Fear	1915	55	D/A
The Land of Mist	1926	66	D/A

Table 5. Novels by Arthur Conan Doyle

(D/A = Detective/Adventure, H = Historical Fiction)

REFERENCES

1. Auclair-Ouellet, N. "Inflectional morphology in primary progressive aphasia and Alzheimer's disease: A systematic review." *Journal of Neurolinguistics* 34 (2015): 41-64.
2. Bauer, L. *Morphological productivity*. Vol. 95. Cambridge University Press, 2001.
3. Bird, S., Loper, E., and Klein, E. *Natural Language Processing with Python*. O'Reilly Media Inc (2009).
4. Cortese, M. J., Balota, D. A., Sargent-Marshall, S. D., Buckner, R. L., and Gold, B. T. "Consistency and regularity in past-tense verb generation in healthy ageing, Alzheimer's disease, and semantic dementia." *Cognitive Neuropsychology* 23.6 (2006): 856-876.
5. Ferguson, A., Spencer, E., Craig, H., and Colyvas, K. "Propositional Idea Density in women's written language over the lifespan: Computerized analysis." *Cortex* 55 (2014): 107-121.
6. Forbes-McKay, K., Shanks, M., and Venneri, A. "Charting the decline in spontaneous writing in Alzheimer's disease: a longitudinal study." *Acta neuropsychiatrica* 26.04 (2014): 246-252.
7. Francis, W. N., and Kucera H. *Computational Analysis of Present Day American English*. Providence: Brown University Press. (1967).
8. Fraser, K. C., Meltzer, J. A., Graham, N. L., Leonard, C., Hirst, G., Black, S. E., and Rochon, E. "Automated classification of primary progressive aphasia subtypes from narrative speech transcripts." *Cortex* 55 (2014): 43-60.
9. Garrard, P., Maloney, L. M., Hodges, J. R., and Patterson, K. "The effects of very early Alzheimer's disease on the characteristics of writing by a renowned author." *Brain* 128.2 (2005): 250-260.
10. Hier, D. B., Hagenlocker, K., and Shindler, A. G. "Language disintegration in dementia: Effects of etiology and severity." *Brain and language* 25.1 (1985): 117-133.
11. Kavé, G, and Levy, Y. "Morphology in picture descriptions provided by persons with Alzheimer's disease." *Journal of speech, language, and hearing research* 46.2 (2003): 341-352.
12. Kemper, S., Greiner, L. H., Marquis, J. G., Prenovost, K., and Mitzner, T. L. "Language decline across the life span: findings from the Nun Study." *Psychology and aging* 16.2 (2001): 227.
13. Kempler, D. "Language changes in dementia of the Alzheimer type." *Dementia and communication: Research and clinical implications* (1995): 98-114.
14. Lancashire, I., and Hirst, G. "Vocabulary changes in Agatha Christie's mysteries as an indication of dementia: a case study." *19th Annual Rotman Research*

Institute Conference, Cognitive Aging: Research and Practice. 2009.

15. Le, X., Lancashire, I., Hirst, G., & Jokel, R.
"Longitudinal detection of dementia through lexical and syntactic changes in writing: a case study of three British novelists." *Literary and Linguistic Computing* (2011): fqr013.
16. Morris, J. C., and Price, J. L. "Pathologic correlates of nondemented aging, mild cognitive impairment, and early-stage Alzheimer's disease." *Journal of Molecular Neuroscience* 17.2 (2001): 101-118.
17. Morris, J. C., Storandt, M., Miller, J. P., McKeel, D. W., Price, J. L., Rubin, E. H., and Berg, L. "Mild cognitive impairment represents early-stage Alzheimer disease." *Archives of neurology* 58.3 (2001): 397-405.
18. Pakhomov, S., Chacon, D., Wicklund, M., and Gundel, J. "Computerized assessment of syntactic complexity in Alzheimer's disease: a case study of Iris Murdoch's writing." *Behavior research methods* 43.1 (2011): 136-144.
19. Porter, M. F. "An algorithm for suffix stripping." *Program* 14.3 (1980): 130-137.
20. Price, J. L., and Morris, J. C. "Tangles and plaques in nondemented aging and "preclinical" Alzheimer's disease." *Annals of neurology* 45.3 (1999): 358-368.
21. Ratnaparkhi, A. "A maximum entropy model for part-of-speech tagging." *Proc. Conference on empirical methods in natural language processing*. Vol. 1. 1996.
22. Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., ... and Phelps, C. H. "Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease." *Alzheimer's & dementia* 7.3 (2011): 280-292.
23. Tarawneh, R., and Holtzman, D. M. "The clinical problem of symptomatic Alzheimer disease and mild cognitive impairment." *Cold Spring Harbor perspectives in medicine* 2.5 (2012): a006148.
24. Thompson, C. K., Cho, S., Hsu, C. J., Wieneke, C., Rademaker, A., Weitner, B. B., ... and Weintraub, S. "Dissociations between fluency and agrammatism in primary progressive aphasia." *Aphasiology* 26.1 (2012): 20-43.
25. Ullman, M. T., Corkin, S., Coppola, M., Hickok, G., Growdon, J. H., Koroshetz, W. J., and Pinker, S. "A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system." *Journal of cognitive neuroscience* 9.2 (1997): 266-276.