Commentary
Towards early prediction of Alzheimer's disease through language samples
Jed A. Meltzer\!
\textsuperscript{a,b,c}

\textsuperscript{a} Rotman Research Institute, Baycrest Health Sciences, 3560 Bathurst St, Toronto, ON, Canada
\textsuperscript{b} University of Toronto, Dept. Psychology, Toronto, ON, Canada
\textsuperscript{c} University of Toronto, Dept. Speech-language Pathology, Toronto, ON, Canada

Although accurate diagnosis of Alzheimer's Disease (AD) remains a priority for research, even more research interest currently focuses on the prediction of the disease years or decades before its onset. Because the neurodegeneration caused by the disease is likely irreversible, a better treatment strategy would be to identify those undergoing the early changes linked to eventual disease onset and to administer a mitigating treatment (yet to be developed) at that time. One biomarker of intense interest is naturalistic language samples, as they are easy to acquire, completely noninvasive, and, compared to most neuropsychological assessments, easily repeated on a regular basis without practice effects. However, the analysis is complicated, laborious, and potentially subjective. In recent years, advances in machine learning and natural language processing have been applied to language samples for the detection of dementia, and researchers have achieved considerable success in distinguishing the speech of individuals with and without dementia \cite{1–3}.

Despite these advances, the predictive power of language samples is largely unproven, given that very few studies have been able to examine participants years before an eventual diagnosis of AD, to compare the language output of those who do and do not go on to develop the disease \cite{4}. A prospective study of this topic would require a very large sample, take many years to complete, and would have a relatively low yield of positive findings for the effort. Fortunately, as the potential value of speech and other cognitive measures as a biomarker has come to increased attention, large-scale prospective studies of health in general have begun to include them in their assessment batteries. As published in \textit{EClinicalMedicine}, Elif Eyigoz and colleagues present an analysis \cite{5} of written language samples collected in the Framingham Heart Study (FHS), one of the world's largest and best-known prospective health studies. Founded in 1948, the FHS began to incorporate a neuropsychological test battery in 1999, including a brief written picture description. Critically, in the years since this was introduced, some participants went on to develop dementia while many more did not, allowing a retrospective comparison based on data fortuitously acquired years earlier. Elif Eyigoz et al. applied state of the art analysis procedures to samples from 270 participants finding that dementia onset before age 85 could be identified with 70–75% accuracy. This is well above chance performance, and slightly below the best performance seen in studies comparing current dementia patients with controls \cite{6}. The authors then demonstrate that their model picks up on many of the same linguistic trends seen in previous comparisons of AD vs. controls.

This finding is exciting, being among the first to show predictive value of language samples well before the onset of dementia, while its limitations point the way for future work. The accuracy rates of 70–75% are encouraging but not yet satisfactory for a realistic clinical tool, but this is likely to be an inevitable consequence of the limited language samples available. The samples are descriptions of a single picture (“Cookie Theft”), and limited both in length to a few dozen words, and in content to what is shown in the picture. Longer samples covering more extensive topics would surely provide a more sensitive view of linguistic changes. More significantly, the samples are only written. Although written samples have a history of predictive value, speech is a more natural form of communication giving access to several important quantitative variables, especially those related to the ease of word finding, including overall speech rate and pausing \cite{7}.

It remains to be seen how high the predictive accuracy of language samples can be pushed given more extensive data sources, and it is important that such data be collected. The most advanced machine learning algorithms cannot overcome the limitations of sparse input. Fortunately, collection of speech data is simple to implement and simple to include within a larger natural history study like the FHS. A number of longitudinal health studies now include detailed speech measures and can be expected to yield new insights into the earliest stages of the evolution of dementia \cite{8,9}.

Above all, this study illustrates the value of open science — a simple measure that was not the original focus of the FHS provided valuable new knowledge when shared with the larger scientific community. Given the complexity and expense of longitudinal
studies of dementia, this level of data sharing needs to become the norm. Researchers should take pains to harmonize data collection and archiving procedures, while addressing practical concerns such including consent and privacy (especially with speech, as the human voice is inherently identifiable). Additionally, attention should be paid as to which kinds of speech elicitation tasks provide the most informative samples, as there are many different options, including picture description, story retell (e.g. Cinderella), autobiographical interviews, and dyadic conversation [10].

Declaration of Competing Interest

Dr. Meltzer is a shareholder and advisor of Winterlight Labs.

References

[1] Orimaye SO, Wong JS, Golden KJ, Wong CP, Soyiri IN. Predicting probable Alzheimer’s disease using linguistic deficits and biomarkers. BMC Bioinformatics 2017;18(1):34.

[2] Fraser KC, Meltzer JA, Rudzicz F. Linguistic features identify Alzheimer's disease in narrative speech. J Alzheimers Dis 2015;49(2):407–22.

[3] Fraser KC, Meltzer JA, Graham NL, Leonard C, Hirst G, Black SE, et al. Automated classification of primary progressive aphasia subtypes from narrative speech transcripts. Cortex 2014;53:43–60.

[4] Snowdon DA, Kemper SJ, Mortimer JA, Greiner LH, Weikstein DR, Markesbery WR. Linguistic ability in early life and cognitive function and Alzheimer’s disease in late life. Findings from the Nun Study. JAMA. 1996;275(7):528–32.

[5] Eygzo E, Mathur S, Santamaria M, Cecchi G, Naylor M. Linguistic markers predict onset of Alzheimer’s disease. EClinical Medicine 2020. doi: 10.1016/j.eclinm.2020.100583.

[6] Petti U, Baker S, Korhonen A. A systematic literature review of automatic Alzheimer's disease detection from speech and language. J Am Med Inform Assoc 2020.

[7] Guo Z, Ling Z, Li Y. Detecting Alzheimer’s disease from continuous speech using language models. J Alzheimers Dis 2019;70(4):1163–74.

[8] Mueller KD, Kosick RL, Hermann BP, Johnson SC, Turkstra LS. Declines in connected language are associated with very early mild cognitive impairment: results from the wisconsin registry for alzheimer’s prevention. Front Aging Neurosci 2018;9:437.

[9] Farhan SM, Bartha R, Black SE, Corbett D, Finger E, Freedman M, et al. The ontario neurodegenerative disease research initiative (ONDRI). Can J Neurol Sci 2017;44 (2):196–202.

[10] Boschi V, Catricala E, Consonni M, Chesi C, Moro A, Cappa SF. Connected speech in neurodegenerative language disorders: a review. Front Psychol 2017;8:269.