A Comparison of Clinician Assessment of Speech Versus Automated Speech Analysis in Mild Cognitive Impairment and Alzheimer's Dementia

Anthony Yeung, MD[1], Andrea Iaboni, MD DPhil[1,2], Elizabeth Rochon, PhD[1,2], Monica Lavoie, PhD[1], Calvin Santiago, MSc MD[1], Maria Yancheva, MSc[3],Jekaterina Novikova, PhD[3], Mengdan Xu, MSc[3], Jessica Robin, PhD[3], Liam D Kaufman, MSc[3] and Fariya Mostafa, MSc[3]

1. University of Toronto, ON, Canada, 2. Toronto Rehab, University Health Network, ON, Canada, 3. Winterlight Labs, Toronto, ON, Canada

Correspondence: Jessica Robin, PhD <jessica@winterlightlabs.com>

Background

- Language impairment is a core feature of Alzheimer's disease (AD) and other neurodegenerative disorders.[1]
- Prior studies have shown a link between AD symptom severity and declining speech and language capability in picture description tasks.[2]
- Speech and language changes include alterations in speech rate, utterances, frequency of words, word-finding difficulties, and repetitions.[3]
- Despite these pervasive language changes, there is no universally accepted system of terminology used to describe language impairment, and large inter-rater variability can also exist between clinicians.[4]
- In view of current limitations, the role of automated speech analysis is emerging as a novel, and potentially more objective method of assessing language in individuals with neurologic and psychiatric disorders.
- We sought to: (1) define a set of speech and language capability ratings that can be used by clinicians with different areas of specialization, (2) determine if these speech and language ratings are applied consistently in a sample of patients including healthy controls, mild cognitive impairment (MCI), and AD, and (3) use automated speech analysis to identify what acoustic and linguistic variables correlate with clinician ratings of speech and language.

Methods

- Speech samples were obtained via the DementiaBank (DB) dataset through the TalkBank Project, with equal numbers of healthy controls, MCI, and probable AD participants.
- Participants provided a recording of a speech sample which consisted of a verbal description of the Boston Cookie Theft picture.
- The recordings were rated by 5 clinicians (1 geriatric psychiatrist, 1 psychiatry resident, 1 neurology resident, and 2 speech language pathologists) with clinical experience in assessing MCI and AD, according to four characteristics: (1) word-finding difficulty, (2) incoherence, (3) perseveration, and (4) errors in speech; these were rated on a Likert scale (range: 0-3) as being: not present/normal finding, mild, moderate, or severe **(Table 1).**
- Speech recordings were transcribed, and linguistic and acoustic variables were extracted through automated speech analysis using NodeJS and React. Data processing and feature extraction was performed using Python-based standard acoustic and language processing libraries (e.g., spacy) and custom code.
- The correlation between clinician-identified speech characteristics and the acoustic and linguistic variables were then compared using Spearman correlation.
- Exploratory factor analysis (EFA) was then applied to find common factors between variables for each speech characteristic, using R version 3.6.3 and Python version 3.6.

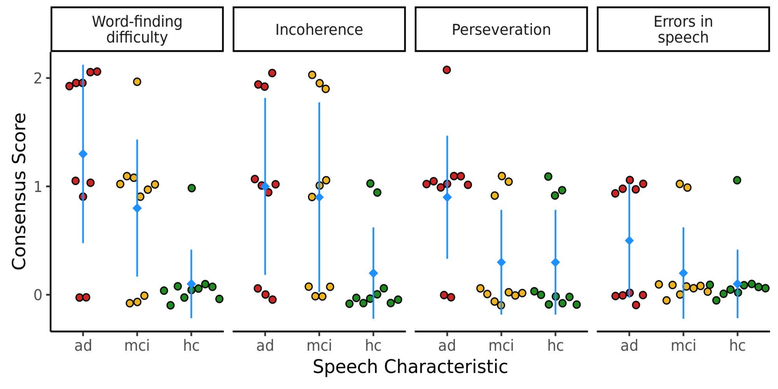
Table 1: Clinician Consensus Table of Speech Characteristics

Speech and Language Characteristic	Clinical Features
Word-finding difficulty	• Reduction in content words, circumlocution, false starts, pauses while searching for words, fluency (rate, phrase length, amount of hesitation), revisions (repetitions of complete words or phrases/elaborations), indefinite terms (fillers).
Incoherence	• Disorganized speech, derailment or sudden topic shifts, tangentiality, flight of ideas, or word salad.
Perseveration	• Repetition of word or phrase even after the stimulus for the behavior (word or phrase) has been taken away; persistence of behavior (word or phrase) despite repeated failure; intrusions (i.e., Inappropriate repetition of prior responses after intervening stimuli).
Errors in Speech	• Phonetic errors (omissions, additions, substitutions, distortions), stuttering, sequences of phonemic approximation.

Table 2: Participant demographics by diagnostic group

	Controls (n=10)	MCI (n=10)	AD (n=10)
Age at visit, mean (SD), y	61.2 (9.7)	69.9 (5.9)	64.0 (11.0)
Female (%)	50	50	50
MMSE, mean (SD)	29 (0.9)	24 (2.0)	18 (1.6)
Education, mean (SD), y	14.2 (2.3)	14.0 (1.9)	13.8 (2.2)

Figure 1: Consensus Clinician Ratings for Each Speech Characteristic



Distribution of the consensus clinician ratings for each speech characteristic, by diagnosis group. The mean consensus rating for each group is indicated with a blue diamond and whiskers indicate the standard deviation. For all ratings, a rating of 3 = severe, 2 = moderate, 1 = mild, and 0 = no presence or a normal finding of that speech characteristic.

Presented at Alzheimer's Association International Conference (AAIC) Virtual Conference, July 29, 2020

Results

- The participants demographics/characteristics are described in **Table 2**.
- Clinician rating agreement was high in three of the four speech characteristics (word-finding difficulty: ICC = 0.92, $p < 0.001$; incoherence: ICC = 0.91, $p < 0.001$; perseveration: ICC = 0.88, $p < 0.001$).
- Speech ratings scores were highest (most impairment) in the probable AD group, followed by MCI and controls. Greater impairments in word-finding difficulty and incoherence were more frequent in AD and MCI.
- For **word-finding difficulty**, variables with the highest correlations to clinician ratings were related to the rate of speech, word duration and length and the number of unfilled pauses. Greater severity of word-finding difficulty was associated with slower speech, shorter words and increased pauses.
- For **incoherence**, the variables with the highest correlations were a mix of syntactic, acoustic and lexical variables, reflecting the use of past tense verb phrases, slower speech rate, and words with higher estimated age of acquisition.
- For **perseveration**, variables with the highest correlations were related to the complexity of speech and vocabulary. Greater severity of perseveration was associated with increased repetitiveness of speech, decreased vocabulary richness, and decreased semantic similarity. A large number of acoustic variables also correlated with perseveration.
- For **errors in speech**, the variables with the highest correlations with the consensus clinician ratings included measures relating to the complexity of speech and vocabulary, use of subordinate clauses, and word length.
- EFA showed that between 1 to 4 factors were found to explain each characteristic (data not shown)

Conclusions

- In this exploratory study, variables extracted through automated acoustic and linguistic analysis of MCI and AD speech were strongly correlated to speech and language characteristics rated by clinicians.
- We were able to demonstrate that commonly used clinical terms such as word-finding difficulty, incoherence, perseveration, and errors in speech, can be correlated to features identified through automated speech analysis.
- Strengths of the study include utilizing clinician ratings to provide an objective, understandable, and rational approach to defining speech changes in AD and MCI.
- Limitations include a small sample size and short speech recording sample based on the Cookie Theft task.
- Our work proposes a standardized approach to investigating speech on both a clinical and pathophysiological level. Potential future applications of this method includes the wide scale deployment of speech analysis in resource-limited or remote settings.

References

(1) Forbes-McKay, K., Shanks, M. F., & Venneri, A. (2013). Profiling spontaneous speech decline in Alzheimer's disease: a longitudinal study. Acta Neuropsychiatrica, 25(6), 320-327.
(2) Kavé, G., & Dassa, A. (2018). Severity of Alzheimer's disease and language features in picture descriptions. Aphasiology, 32(1), 27-40.
(3) Slegers, A., Filiou, R. P., Montembeault, M., & Brambati, S. M. (2018). Connected speech features from picture description in Alzheimer's disease: A systematic review. Journal of Alzheimer's Disease, 65(2), 519-542.
(4) Wilson, S. M., Eriksson, D. K., Schneck, S. M., & Lucanie, J. M. (2018). A quick aphasia battery for efficient, reliable, and multidimensional assessment of language function. PloS one, 13(2).