

Summary of experience with and scientific productivity of recordings of cognitive assessments

MGH FTD Unit / ADRC and LEADS

[Brad Dickerson MD](#)

We have been recording digital audio samples of cognitive assessments in cognitively unimpaired younger and older adults and in patients with neurodegenerative disorders in the [MGH FTD Unit](#) since 2008, and in the [LEADS consortium](#) since 2018. The LEADS consortium includes 18 US sites, 14 of which are ADRCs. In LEADS, coordinators at each site are trained on procedures and a digital audio recorder is provided to each site. The following UDS tests are recorded: Craft Story immediate and recall, Category fluency (animals and vegetables), Lexical fluency (Fs and Ls), FTLD battery word reading, sentence repetition, noun and verb naming, and sentence reading; Rey AVLT is also recorded. Essentially no additional time is added to participant visits for these recordings. Our experience in LEADS has been that, once trained, coordinators require a total of approximately 15-20 additional minutes per participant session for file management. At MGH, digital audio recorders were used from 2008 to 2018 and since then secure online platforms have been used to capture audio data. Participants are consented for recording and sharing of digital audio samples (>99% opt in). Digital audio files are stripped of any PHI/PII. At MGH files are stored locally and in secure online storage space. In LEADS, files are stored locally and copies are shared with Indiana University (prime site) for distribution to other investigators.

A brief summary of scientific productivity follows. At MGH, we have reported on findings from patients with Primary Progressive Aphasia, Posterior Cortical Atrophy, and early-stage Alzheimer's disease. Most of these studies employed a picture description task; similar information is present in the retelling of a story (e.g., Craft) and in less structured tasks.

Primary Progressive Aphasia

- In this study, we developed a novel metric of syntactic complexity and applied it to patients with PPA, demonstrating that lexical and syntactic complexity, as measured by their frequencies, are inversely correlated. We built on this finding to show that this syntax-lexicon trade-off is also present in the utterances of healthy speakers, suggesting that it may be a general property of the process by which humans turn thoughts into speech [1].
- In this study, we found that the word types preferred by patients with agrammatic PPA tend to have lower frequencies of occurrence than less preferred words. We then conducted a computational simulation to examine the impact of word frequency on lexical information as measured by entropy, finding that strings of words that exclude highly frequent words have a more uniform word distribution, thereby increasing lexical entropy. To test whether the lexical profile of agrammatism results from their inability to produce long sentences, we asked healthy speakers to produce short sentences during the picture description task. We found that, under this constrained condition, a similar lexical profile of agrammatism emerged in the short sentences of healthy individuals, resulting in lower average word frequency than unconstrained sentences. We extended this finding by showing that, in general, shorter sentences get packaged with lower-frequency words as a basic property of efficient language production, evident in the language of healthy speakers and all primary progressive aphasia variants [2].
- Here we used a computational approach based on language models to measure sentence information through surprisal, a metric calculated by the average probability of occurrence of words in a sentence, given their preceding context. We found that agrammatic PPA patients produced sentences with the same sentence surprisal as healthy controls by using richer words in their structurally impoverished sentences. Furthermore, higher surprisal in agrammatic PPA sentences correlated with the canonical features of agrammatism: a lower function-to-all-word ratio, a lower verb-to-noun ratio, a higher heavy-to-all-verb ratio, and a higher ratio of verbs in -ing forms. We interpret these findings as an alternative account of agrammatic aphasia that regards its word-level features as adaptive, rather than defective, symptoms, a finding that would call for revisions in the

therapeutic approach to agrammatic language production [3].

- Debate continues about whether PPA is best subdivided into three variants and also regarding the most distinctive linguistic features for classifying PPA variants. In this study, we harnessed the capabilities of artificial intelligence (AI) and natural language processing (NLP) to first perform unsupervised classification of concise, connected speech samples from 78 PPA patients. Large Language Models discerned three distinct PPA clusters, with 88.5% agreement with independent clinical diagnoses. Patterns of cortical atrophy of three data-driven clusters corresponded to the localization in the clinical diagnostic criteria. We then used NLP to identify linguistic features that best dissociate the three PPA variants. Seventeen features emerged as most valuable for this purpose. Using these linguistic features derived from the analysis of brief connected speech samples, we developed a classifier that achieved 97.9% accuracy in predicting PPA subtypes and healthy controls. Our findings provide pivotal insights for refining early-stage dementia diagnosis, deepening our understanding of the characteristics of these neurodegenerative phenotypes and the neurobiology of language processing, and enhancing diagnostic evaluation accuracy [4].

Posterior Cortical Atrophy

- We compared the language samples of PCA patients with controls across two distinct tasks: a visually-dependent picture description and a visually-independent job description task. We then conducted an in-depth analysis of the language used in the picture description task to identify specific linguistic indicators that reflect the visuospatial processing deficits of PCA. Patients with PCA showed significant language deficits in the visually-dependent task, characterized by higher word frequency, prolonged utterance latency, and fewer spatial relational words, but not in the visually-independent task. PCA patients struggled to identify certain visual elements as well as the overall theme of the picture. A predictive model based on these language features distinguished PCA patients from controls with high classification accuracy. The findings indicate that language is a sensitive behavioral construct to detect visuospatial processing abnormalities of PCA. These insights offer theoretical and clinical avenues for understanding and managing PCA, underscoring language as a crucial marker for the visuospatial deficits of this atypical variant of Alzheimer's disease [5].

Early-stage Alzheimer's disease

- Cross-linguistic analysis was conducted on language samples from 184 English and 52 Persian speakers, comprising both AD patients and healthy controls, to extract various language features. Furthermore, we introduced a machine learning-based metric, Language Informativeness Index (LII), to quantify informativeness. Indicators of AD in English were found to be highly predictive of AD in Persian, with a 92.3% classification accuracy. Additionally, we found robust correlations between the typical linguistic abnormalities of AD and language emptiness (low LII) across both languages. The findings suggest that AD linguistic impairments are attributable to a core universal difficulty in generating informative messages across divergent languages [6].
- We propose that the primary deficit in AD language production is an overreliance on high-frequency words. English has a set of particularly high-frequency verbs that surpass most nouns in usage frequency. Since AD patients tend to use high-frequency words, the byproduct of this word distribution in the English language would be an over-usage of high-frequency verbs. In contrast, Persian features complex verbs with an overall distribution lacking extremely high-frequency verbs like those found in English. As a result, we hypothesize that Persian-speaking AD patients would not have a bias toward the overuse of high-frequency verbs. We analyzed language samples from 95 English-speaking persons with AD (pwAD) and 91 healthy controls, along with 27 Persian-speaking pwAD and 27 healthy controls. Employing uniform automated natural language processing methods, we measured the usage rates of nouns, verbs, and word frequencies across both cohorts. Our findings showed that English-speaking pwAD use higher-frequency verbs than healthy individuals, a pattern not mirrored by Persian-speaking pwAD. Crucially, we found a significant interaction between the frequencies of verbs used by English and Persian speakers with and without AD. Moreover, regression models that treated noun and verb frequencies as separate predictors did not outperform models that considered overall word frequency alone in classifying AD. In conclusion, this study suggests that language abnormalities among English-speaking pwAD reflect the unique distributional properties of words in English rather than a universal noun-verb class distinction. Beyond offering a new understanding of language abnormalities in AD, the study highlights the critical need for further investigation across diverse languages to deepen our insight into the mechanisms of language impairments in neurological disorders [7].

- We employed a pre-trained large language model (Distil-RoBERTa) to identify amyloid- β positivity from a short, connected speech sample. We further use explainable AI (XAI) methods to extract interpretable linguistic features which can be employed in clinical practice. We analyzed data from 74 patients with PPA across its three variants. 51% of the sample was amyloid-positive. Over ten runs of 10-fold cross-validation, the classifier achieved a mean accuracy of 92%. Interpretation models were able to capture the classifier's behavior well, achieving an accuracy of 97% against classifier predictions, and uncovering several novel speech patterns that may characterize amyloid- β positivity [8].

LEADS EOAD

- We transcribed RAVLT recordings from 303 subjects in the Longitudinal Early-Onset Alzheimer's Disease Study. Subjects were grouped by amyloid status and syndrome. Compared with amyloid negative cognitively impaired subjects, EOAD subjects exhibited effects on raw score, primacy, recency, and stopping time. Inter-syndromic differences were noted with raw score, primacy, recency, J-curve, and stopping time. RAVLT measures are sensitive to the effects of amyloid and syndrome in EOAD. Future work is needed to quantify the predictive value of these scores [9].

References

- [1] Rezaii N, Mahowald K, Ryskin R, Dickerson B, Gibson E. A syntax-lexicon trade-off in language production. *Proc Natl Acad Sci U S A*. 2022;119:e2120203119.
- [2] Rezaii N, Ren B, Quimby M, Hochberg D, Dickerson BC. Less is more in language production: an information-theoretic analysis of agrammatism in primary progressive aphasia. *Brain Commun*. 2023;5:fcad136.
- [3] Rezaii N, Michaelov J, Josephy-Hernandez S, Ren B, Hochberg D, Quimby M, Dickerson BC. Measuring Sentence Information via Surprisal: Theoretical and Clinical Implications in Nonfluent Aphasia. *Ann Neurol*. 2023;94:647-57.
- [4] Rezaii N, Quimby M, Wong B, Hochberg D, Brickhouse M, Touroutoglou A, et al. Using Generative Artificial Intelligence to Classify Primary Progressive Aphasia from Connected Speech. *medRxiv*. 2023.
- [5] Rezaii N, Hochberg D, Quimby M, Wong B, McGinnis S, Dickerson BC, Putcha D. Language uncovers visuospatial dysfunction in posterior cortical atrophy: a natural language processing approach. *Front Neurosci*. 2024;18:1342909.
- [6] Bayat S, Santai M, Panahi MM, Khodadadi A, Ghassimi M, Rezaei S, et al. Language Abnormalities in Alzheimer's Disease Arise from Reduced Informativeness: A Cross-Linguistic Study in English and Persian. *medRxiv*. 2024.
- [7] Sanati M, Bayat S, Panahi MM, Khodadadi A, Rezaee S, Ghasimi M, et al. Impaired language in Alzheimer's disease: A comparison between English and Persian implicates content-word frequency rather than the noun-verb distinction. *medRxiv*. 2024.
- [8] Robertson C, Rezaii N, Hochberg D, Quimby M, Wolff P, Dickerson BC. Using explainable artificial intelligence to identify linguistic biomarkers of amyloid pathology in Primary Progressive Aphasia. Submitted.
- [9] Bushnell J, Hammers DB, Aisen P, Dage JL, Eloyan A, Foroud T, et al. Influence of amyloid and diagnostic syndrome on non-traditional memory scores in early-onset Alzheimer's disease. *Alzheimers Dement*. 2023;19 Suppl 9:S29-S41.