

# Bridging Blood-Based Biomarkers and Language Markers in Natural Conversations

Yijiang Pang<sup>1</sup>, Liu Chen<sup>2</sup>, Hiroko H. Dodge<sup>2</sup>, Jiayu Zhou<sup>3</sup>

<sup>1</sup> Michigan State University <sup>2</sup> Massachusetts General Hospital, Harvard Medical School <sup>3</sup> University of Michigan

AAIC 2026 Poster #7239 | Biomarkers / Digital health technologies and digital biomarkers



Surrogate CSF biomarker models link biological AD pathology signals to language markers from naturalistic I-CONNECT conversations.

## STUDY GOAL AND COHORT BRIDGE

- Digital speech markers can indicate early cognitive change, but many conversation cohorts lack direct AD biomarker assays.
- The study bridges NACC participants with measured CSF biomarkers and I-CONNECT participants with natural conversation data.
- Shared clinical variables are used to train surrogate biomarker models in NACC and apply them to I-CONNECT.
- I-CONNECT contributes naturalistic video-chat conversations from older adults with normal cognition or MCI.

2,055

Aβ42  
NACC measured

1,875

P-tau181  
NACC measured

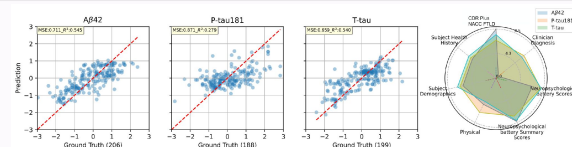
1,986

T-tau  
NACC measured

168

Variables  
shared features

## SURROGATE BIOMARKER MODEL QUALITY

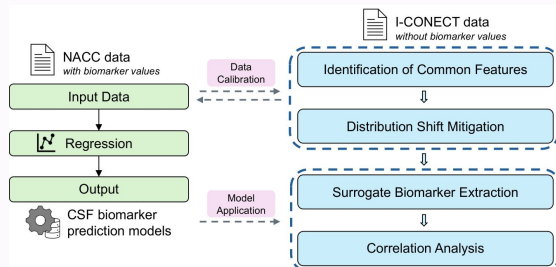


- Aβ42 and T-tau show moderate predictive performance; P-tau181 is weaker and should be interpreted cautiously.
- Top predictors were clinically coherent: orientation, cognitive status, CDR language, delayed recall, and Benson figure delayed drawing.
- The surrogates are driven by cognitive and clinical signals known to relate to amyloid and tau pathology.

## INTERPRETING THE LANGUAGE SIGNAL

- Direction matters: CSF Aβ42 decreases as amyloid burden increases. Positive Aβ42-language correlations mean lower amyloid aligns with richer responses.
- Strongest Aβ42 markers include T-unit count, sentence count, and Root TTR, pointing to reduced syntactic production under burden.
- Tau correlations consistent with subtle language-use changes.
- Correlation magnitudes are modest; natural conversation contains biologically meaningful signal.

## WORKFLOW



## LANGUAGE-BIOMARKER ASSOCIATIONS

| Marker   | Category  | Biomarker | r     | p     |
|----------|-----------|-----------|-------|-------|
| T-unit   | Syntactic | Aβ42      | 0.514 | .0001 |
| Sentence | Syntactic | Aβ42      | 0.502 | .0002 |
| Root TTR | Lexical   | Aβ42      | 0.404 | .0033 |

- Aβ42 associations cluster around syntax, lexical diversity, and response length.

## ACKNOWLEDGEMENTS

This material is based in part upon work supported by the National Science Foundation under Grant IIS-2212174, National Institute on Aging (NIA) R01AG072449 and National Institute of General Medical Sciences (NIGMS) 1R01GM145700.



Intelligent Data Analytics Lab  
UNIVERSITY OF MICHIGAN