Temporal Harmonization: Improved Detection of Mild Cognitive Impairment from Temporal Language Markers using Subject-invariant Learning

Bao Hoang, Siqi Liang, Yijiang Pang, Hiroko Dodge, Jiayu Zhou

Illidan Lab @ University of Michigan

Harvard Medical School

DISCLOSURE OF RELEVANT FINANCIAL RELATIONSHIPS

The authors **have no** relevant financial relationships with ineligible companies to disclose.

Applications	Biomarkers	Disease Progression	Drug Discovery
Methodology	Transfer Learning Multi-task, Few-Shot,Adaptation	Multi-Modality Fusion	Robust Learning against Missing and Noisy Data
Infrastructure	Privacy Protection, Federated Learning (Data Heterogeneity, Resource Heterogeneity, Availability), Fairness, Distributed Optimization		

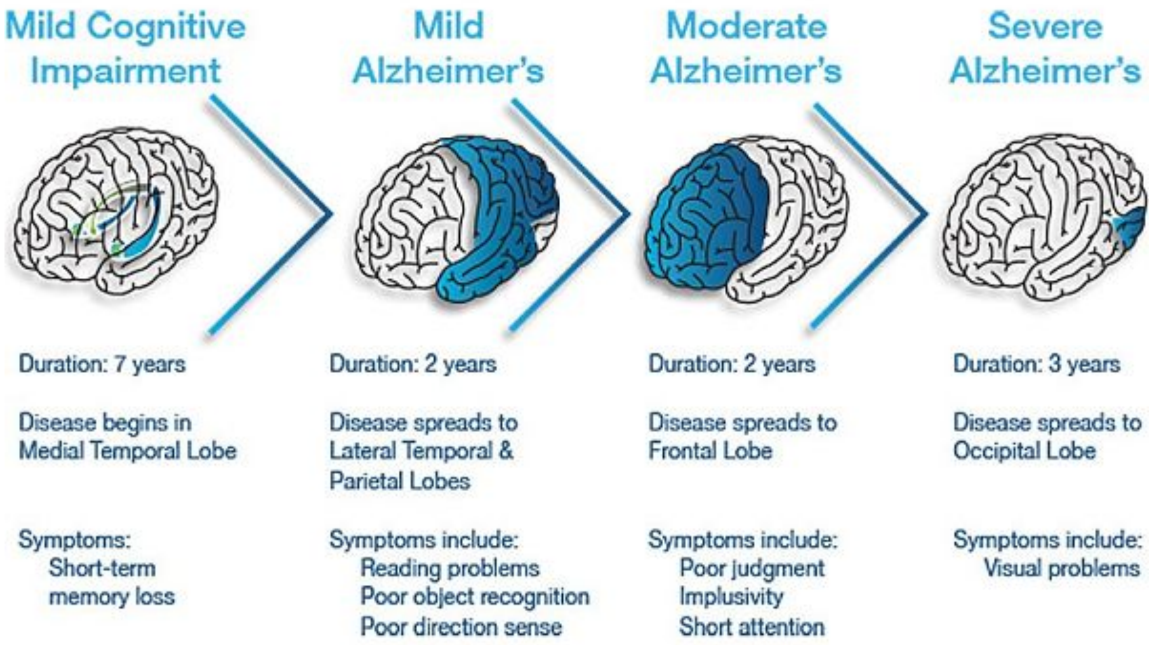
Alzheimer’s disease is officially listed as the seventh leading cause of death in the United States in 2021 and 2022 with 31.0% and 28.9% death rates, respectively.

Kochanek KD, Murphy SL, Xu JQ, Arias E. Mortality in the United States, 2022. NCHS Data Brief, no 492. Hyattsville, MD: National Center for Health Statistics. 2024. DOI: https://dx.doi.org/10.15620/cdc:135850

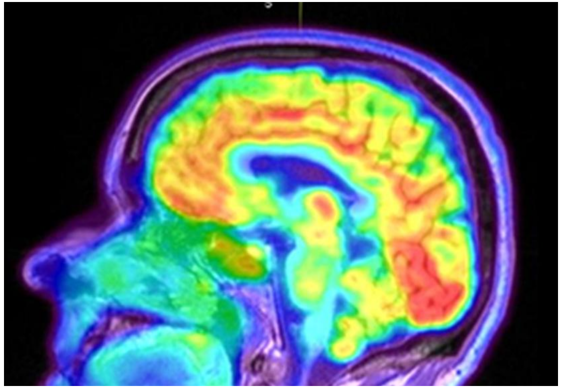
Neurodegenerative Disease - Alzheimer’s

Early Stage

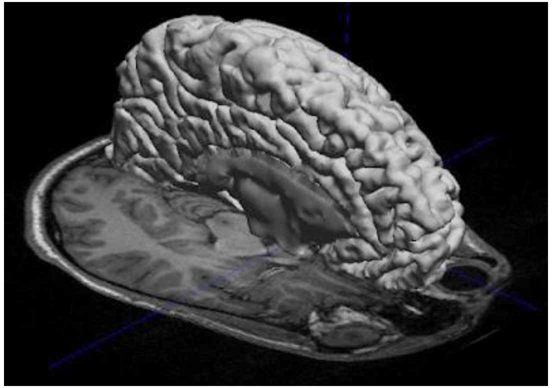
MCI



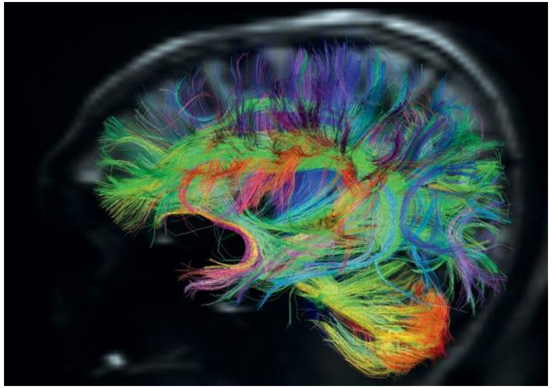
Early Diagnosis and Biomarkers



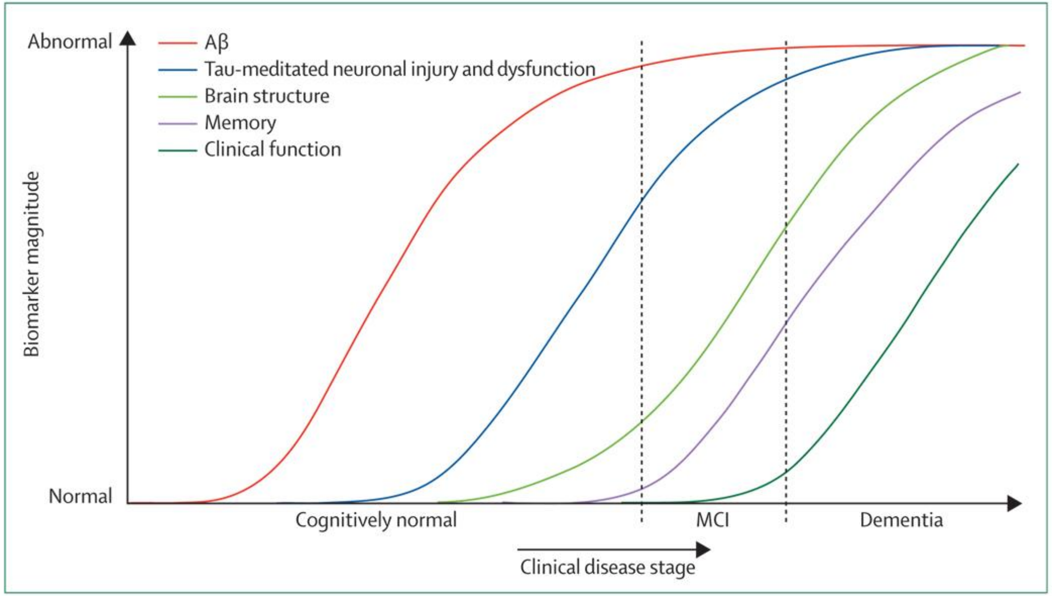
Huang, Shuai, Jing Li, Liang Sun, Jieping Ye, Adam Fleisher, Teresa Wu, Kewei Chen, Eric Reiman, and Alzheimer's Disease NeuroImaging Initiative. "*Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation*." NeuroImage 50, no. 3 (2010): 935-949.



Zhou, Jiayu, Jun Liu, Vaibhav A. Narayan, Jieping Ye, and Alzheimer's Disease Neuroimaging Initiative. "*Modeling disease progression via multi-task learning*." NeuroImage 78 (2013): 233-248.



Wang, Qi, Liang Zhan, Paul M. Thompson, Hiroko H. Dodge, and Jiayu Zhou. "*Discriminative fusion of multiple brain networks for early mild cognitive impairment detection*." In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pp. 568-572. IEEE, 2016.

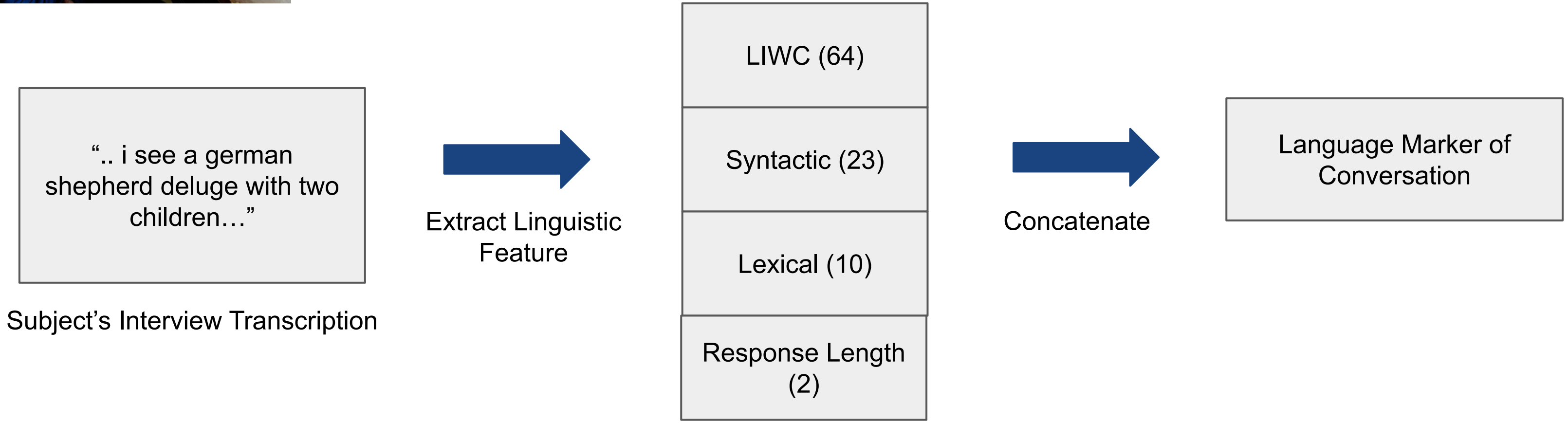


Biomarkers are **effective**, but it is already **too late** when brain markers are obtained from a patient.

Jack Jr, Clifford R., David S. Knopman, William J. Jagust, Leslie M. Shaw, Paul S. Aisen, Michael W. Weiner, Ronald C. Petersen, and John Q. Trojanowski. "Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade." *The Lancet Neurology* 9, no. 1 (2010): 119-128.

Language Markers

- An early detection approach of MCI that is **affordable and accessible.**
- Has been shown diagnostic efficacy in detecting early MCI [1].
- Extract language markers from conversations to build predictive models.
- Semantic, Syntactic, and Lexical features are used for language markers.



[1] Hoang B, Pang Y, Dodge HH, Zhou J. Subject Harmonization of Digital Biomarkers: Improved Detection of Mild Cognitive Impairment from Language Markers. Pac Symp Biocomput. 2024;29:187-200. PMID: 38160279; PMCID: PMC11017207.

> Interspeech. 2021 Aug-Sep:2021:3830-3834. doi: 10.21437/interspeech.2021-2002.

Automatic Detection of Alzheimer's Disease Using Spontaneous Speech Only

Jun Chen [1], Jieping Ye [1], Fengyi Tang [2], Jiayu Zhou [2]

Affiliations + expand

PMID: 35493062 PMCID: PMC9056005 DOI: 10.21437/inte

Article | Open access | Published: 31 March 2020

Scalable diagnostic screening of mild cognitive impairment using AI dialogue agent

Fengyi Tang, Ikechukwu Uchendu, Fei Wang, Hiroko H. Dodge & Jiayu Zhou

Scientific Reports **10**, Article number: 5732 (2020) | Cite this article

Front. Digit. Health, 11 February 2022
Sec. Connected Health
Volume 3 - 2021 |
https://doi.org/10.3389/fdgth.2021.702772

This article is part of the R
Bridging the Gap: Advancing C
Digital World
View all 11 Articles >

The Joint Effects of Acoustic and Linguistic Markers for Early Identification of Mild Cognitive Impairment

Fengyi Tang[1] Jun Chen[2] Hiroko H. Dodge[3] Jiayu Zhou[1]

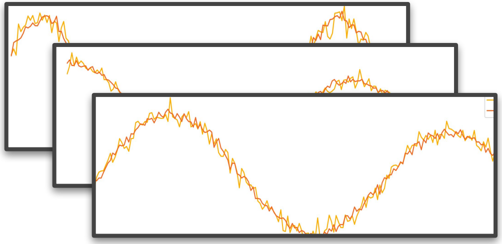
Pac Symp Biocomput. 2023; 28: 7–18.

Detection of Mild Cognitive Impairment from Language Markers with Crossmodal Augmentation

Guangliang Liu,[1] Zhiyu Xue,[1] Liang Zhan,[2] Hiroko H. Dodge,[3] and Jiayu Zhou[1,*]

Temporal Language Markers

- One conversation **rarely** captures enough signal for accurate MCI detection.
- Prior work [1] combines several conversations by majority-voting model outputs, so **loses** detailed changes over time.
- **Our idea:** Model each subject's conversations as a time-series of linguistic markers, thus preserving fine-grained temporal dynamics of cognitive decline.

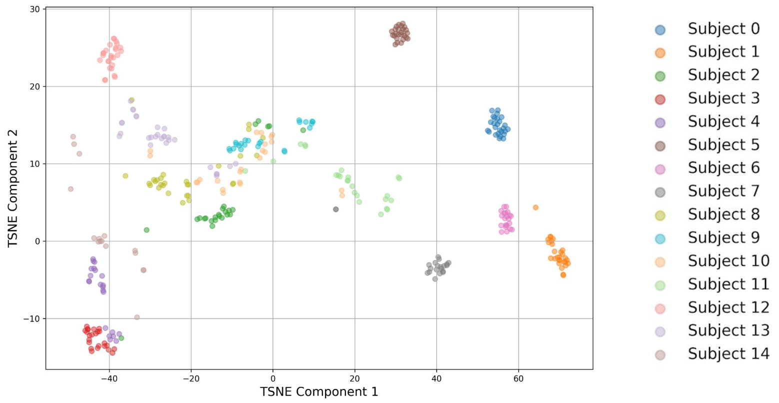


Temporal Linguistic Markers

[1] Hoang B, Pang Y, Dodge HH, Zhou J. Subject Harmonization of Digital Biomarkers: Improved Detection of Mild Cognitive Impairment from Language Markers. Pac Symp Biocomput. 2024;29:187-200. PMID: 38160279; PMCID: PMC11017207.

Problem with Temporal Language Markers

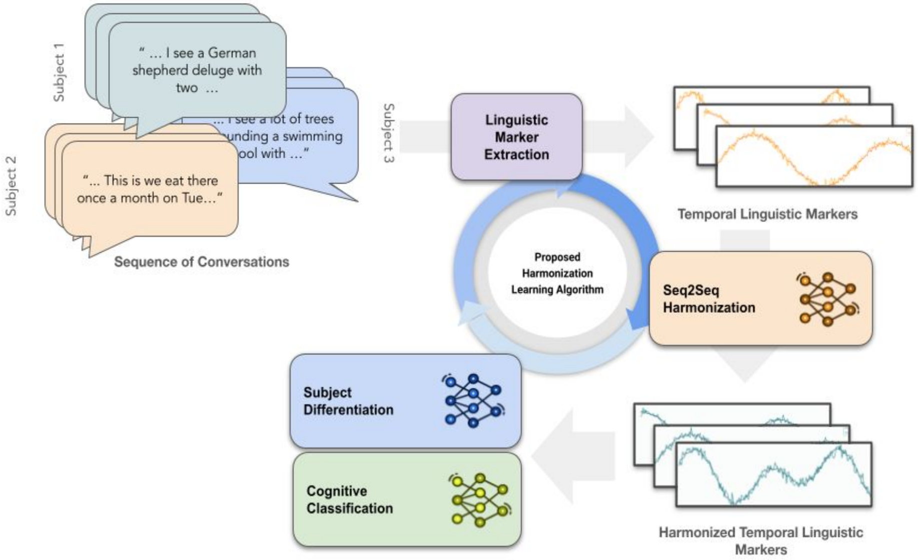
- The way people speak can vary drastically, leading to strong **distributional differences** in temporal language patterns across subjects.
- These differences are much more **outstanding** than the subtle cues that distinguish MCI from normal cognition.
- As a result, models tend to perform poorly on **new, unseen** subjects.



t-SNE plot of RNN's temporal language marker embedding

Proposed Temporal Harmonization Method

- Inspired by domain-invariant representation learning [2], we treat each subject as a domain and remove subject-specific patterns from temporal language markers.



Cognitive Classification

$$\min_{\theta_{TH},\theta_{cog}} \max_{\theta_{subj}} \mathcal{L}_{final} = \mathcal{L}_{cls} - \alpha_1 \mathcal{L}_{adv} + \alpha_2 \mathcal{L}_{ttr} + \alpha_3 \mathcal{L}_{ts}$$

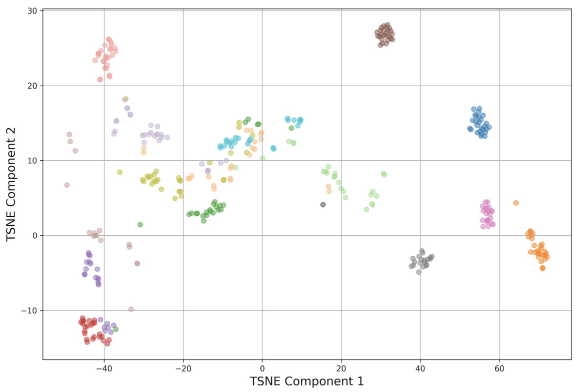
Subject Differentiation

Seq2Seq Harmonization

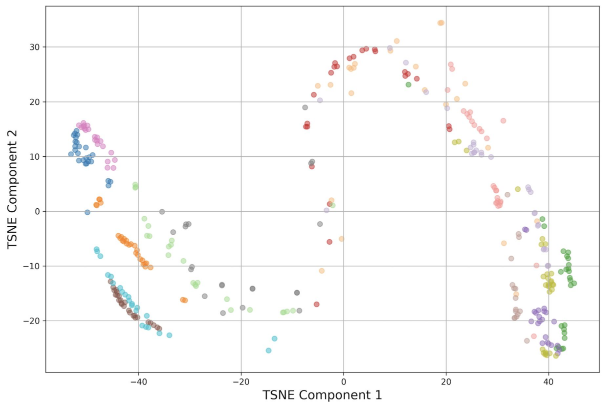
[2] Nguyen, A. T., Tran, T., Gal, Y., & Baydin, A. G. (2021). Domain Invariant Representation Learning with Domain Density Transformations (Version 3). arXiv. https://doi.org/10.48550/ARXIV.2102.05082

Key Result for Temporal Harmonization

Before Harmonization



After Harmonization



t-SNE plot of RNN's temporal language marker embedding

Subject clusters are successfully **destroyed** by harmonization.

Table 2. Performance of Subject Differentiation and Confounder Classification Tasks.

Variable	Before Harmonization	After Harmonization (GRU)	After Harmonization (LSTM)
Subject ID (74 classes)	0.990±0.011	0.211±0.041	0.190±0.032
Age (3 classes)	0.991±0.011	0.605±0.057	0.655±0.048
Gender (2 classes)	0.996±0.009	0.630±0.072	0.721±0.053
Years of Education (3 classes)	0.986±0.018	0.629±0.055	0.597±0.040

Classifier cannot predict **subject identity** or **confounder variable** of temporal language markers after harmonization.

Key Quantitative Results

Table 3. Performance of cognitive classification tasks over different harmonization methods.

Method	Model	AUC	F1	Sensitivity	Specificity
No Harmonization	LR	0.591±0.124	0.579±0.126	0.593±0.166	0.568±0.169
	MLP	0.626±0.122	0.593±0.124	0.576±0.153	0.649±0.159
	GRU	0.641±0.121	0.515±0.148	0.517±0.244	0.604±0.272
	LSTM	0.648±0.135	0.486±0.147	0.417±0.174	**0.754±0.184**
Generalized least squares[11]	LR	0.585±0.129	0.529±0.148	0.519±0.187	0.601±0.164
	MLP	0.568±0.122	0.568±0.138	0.565±0.175	0.605±0.175
Subject Harmonization[12]	LR	0.649±0.121	0.592±0.115	0.575±0.157	0.652±0.162
	MLP	0.657±0.113	0.571±0.118	0.546±0.152	0.655±0.152
Temporal Harmonization (Proposed Method)	GRU	0.699±0.124	0.642±0.120	0.671±0.169	0.600±0.174
	LSTM	**0.721±0.111**	**0.658±0.123**	**0.684±0.165**	0.621±0.162

- Demonstrates the **additional benefits** of using temporal sequences of language markers to detect MCI.
- Moreover, applying temporal harmonization **enhances** cognitive detection performance significantly.

Key Quantitative Results

Table 5. Performance of cognitive classification tasks over different sub-populations.

Group	Before Temporal Harmonization			After Temporal Harmonization		
	AUC	Sensitivity	Specificity	AUC	Sensitivity	Specificity
Male	0.600±0.296	0.352±0.314	0.698±0.358	0.641±0.282	0.676±0.313	0.556±0.369
Female	0.674±0.146	0.459±0.228	0.778±0.197	0.754±0.146	0.692±0.207	0.644±0.190
Edu 12-15	0.598±0.255	0.397±0.263	0.690±0.295	0.633±0.241	0.651±0.234	0.513±0.301
Edu 16-18	0.620±0.243	0.365±0.296	0.814±0.243	0.742±0.230	0.654±0.309	0.692±0.255
Edu 19-21	0.917±0.242	0.821±0.359	0.754±0.380	0.907±0.251	0.905±0.273	0.766±0.364
Age 75-80	0.676±0.162	0.395±0.274	0.782±0.197	0.734±0.170	0.685±0.239	0.657±0.180
Age 81-87	0.526±0.300	0.345±0.293	0.852±0.291	0.693±0.296	0.640±0.285	0.695±0.399
Age 88-94	0.672±0.426	0.625±0.451	0.344±0.458	0.844±0.341	0.891±0.299	0.172±0.345

- Temporal harmonization enhances model generalization **across demographic subgroups.**

Discussion and Future Works

- Temporal Harmonization **significantly improves** MCI detection performance across nearly all demographic subgroups.
- Given the limited size and scope of the I-CONECT dataset (74 older adults), future work includes:
  - expanding to **diverse populations**,
  - incorporating **multimodal biomarkers**,
  - and enhancing **explainability** to improve trust and interpretability.

Thanks!

http://illidanlab.github.io

Acknowledgement: This material is based in part upon work supported by the National Science Foundation under Grant IIS-2212174, National Institute of Aging (NIA) 1RF1AG072449, National Institute of General Medical Sciences (NIGMS) 1R01GM145700.