Contrasting static and contextualised embeddings in the use of semantic feature vectors in neurophysiological prediction

Heather Dial[a*], Lokesh Pugalenthi[b*], Nike Gnanateja[c], Rachel Tessmer[d], Maya Henry[d**], Jessy Li[b**]

*:co-first author,**:co-senior author. a: Department of Communication Sciences and Disorders, College of Liberal Arts and Social Sciences, University of Houston; b: Department of Linguistics, College of Liberal Arts and Social Sciences, University of Texas at Austin; c: Department of Communication Sciences and Disorders, School of Health and Rehabilitation Services, University of Pittsburgh; d: Speech Language and Hearing Sciences, Moody College of Communication, University of Texas at Austin

Introduction

- Primary progressive aphasia (PPA)
  - progressive loss of speech and language
- Three subtypes
  - Semantic
  - Logopenic
  - Nonfluent

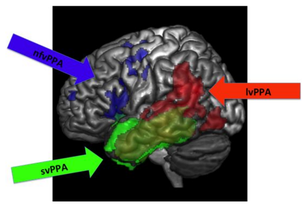


Figure 1: The impact and location of three PPA subtypes

- Temporal Response Function (TRF) Modeling[1]
  - Maps neurophysiological data to stimulus' acoustic/linguistic feature(s)
- Natural Language Processing (NLP) models can be used to derive feature vectors that approximate a word's semantic features

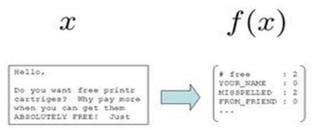


Figure 2: Transforming text into a feature vector

- NLP model's choice of word embeddings impacts feature vectors
- Aim: To compare how two different embedding types affect TRF-based neurophysiological prediction

Methods

- EEG responses obtained while participants listen to audiobook
- *n* = 10 for each PPA subtype + age-matched controls

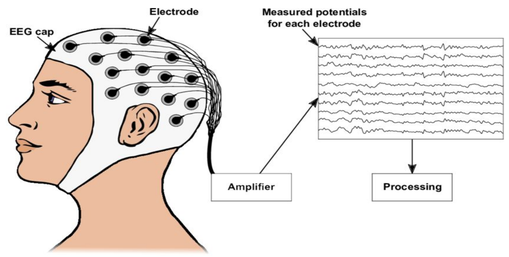
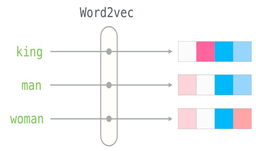


Figure 3: An illustration of the EEG process

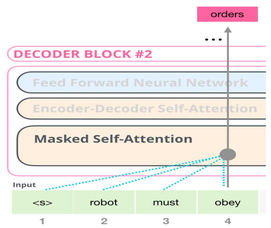
Methods

- Obtained feature vectors for each word in audiobook text using word2vec & GPT2



Word2Vec uses static embeddings[2]

Figure 4: In word2vec, a word's feature vector does not take into account its surrounding words.



GPT2 uses contextualised embeddings

Figure 5: In GPT2, its masked self-attention layer ensures a word's feature vectors takes previous words into account

- Computed each word's semantic dissimilarity value for use in TRF[3]

$$\text{Dissimilarity}(w_i) = 1 - \rho[f(w_i), \text{mean}[f(w_{i-1}), f(w_{i-2}), \ldots, f(w_2), f(w_1)]]$$

Figure 6: Dissimilarity formula for a text's *i*th word. p represents pearson correlation, f represents feature vector

- TRF modeling used to predict EEG responses as a function of semantic dissimilarity. mTRF toolbox used[4]
- TRF trained with LOOCV

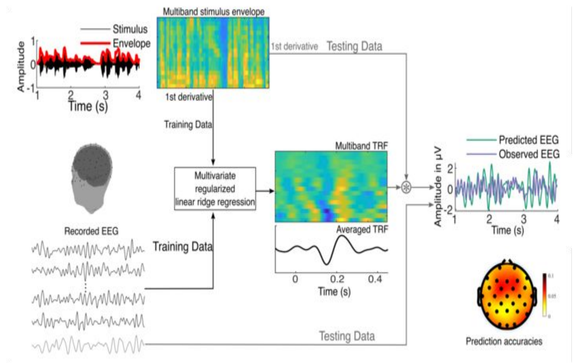
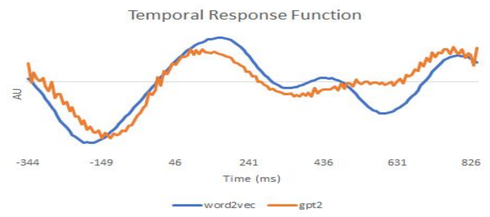


Figure 7: Diagram depicting sample temporal response function modeling for spectrotemporal features.

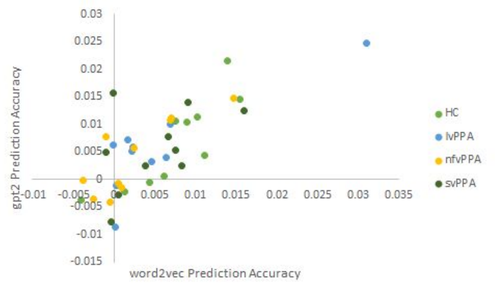
Results

Figure 8: Comparing TRF produced w/ word2Vec and GPT2



Similar TRF model from GPT2 and word2Vec

Figure 9: Correlating GPT2 and word2Vec's predictive accuracy. HC refers to healthy control.



- Similar range on both axes indicates similar predictive accuracy
- Highly correlated across PPA subtypes + controls
- Little difference in predictive accuracy across groups

Discussion + Future steps

- Unexpectedly, contextualised embeddings did not provide a superior TRF wrt. neurophysiological prediction
- However, this can only be said for the TRF model's **current** parameters

Immediate future:

- Need to determine existence of parameters where contextualised embeddings outperform static embeddings **as expected**

Future steps

There are two TRF parameters to initially experiment with:

- Change time lags
  - Currently -500 ms−1000ms
  - TRF noisy at the edges, try trimming time lags to -100 ms−700ms
- Change EEG frequency range from 1−4 Hz to 1−15Hz

Long-term future:
Evaluate whether TRF modelling can differentially diagnose PPA subtypes

Acknowledgements

- Texas Advanced Computing Center (TACC) for large computing and storage resources
- NRSA Postdoctoral Fellowship F32DC016812 (Heather Dial)
- NIH/NIDCD R03DC013403 and R01DC016291 (Maya Henry)

References

- [1]Dial HR, Gnanateja GN, Tessmer RS,Gorno-Tempini ML, Chandrasekaran B and Henry ML (2021) Cortical Trackingof the Speech Envelope in LogopenicVariant Primary Progressive Aphasia.Front. Hum. Neurosci. 14:597694
- [2]The Stanford AI Lab Blog. 2021. BERT, ELMo, & GPT-2: How Contextual are Contextualized Word Representations?. Available at: https://ai.stanford.edu/blog/contextual/.
- [3]Broderick MP, Anderson AJ, Di Liberto GM, Crosse MJ, Lalor EC. Electrophysiological Correlates of Semantic Dissimilarity Reflect the Comprehension of Natural, Narrative Speech. Curr Biol. 2018 Mar 5;28(5):803-809.e3. doi: 10.1016/j.cub.2018.01.080. Epub 2018 Feb 22. PMID: 29478856.
- [4]Crosse MJ, Di Liberto GM, Bednar A and Lalor EC (2016) The Multivariate Temporal Response Function (mTRF) Toolbox: A MATLAB Toolbox for Relating Neural Signals to Continuous Stimuli. Front. Hum. Neurosci. 10:604.