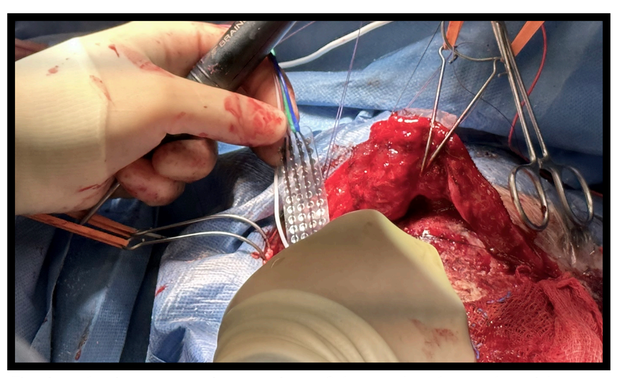
Toward an implanted language neuroprosthesis for severe aphasia

Gross, WL[1,2], Fernandino, L[2], Krucoff, MO[3], Binder, JR[2]

Medical College of Wisconsin, Milwaukee, WI, USA

1. Dept. of Anesthesiology, 2. Dept. of Neurology, 3. Dept. of Neurosurgery

Individuals with persistent severe aphasia after stroke typically have limited rehabilitation options. Because further conventional speech therapy is of limited value in this setting, novel therapeutic approaches are needed. This study explored the potential of an implanted neuroprosthesis for decoding lexical semantic targets from neural activity. Prior brain-computer interfaces (BCIs) proposed as communication assistance devices have largely focused on decoding motor articulation of speech, which is only useful in conditions of paralysis where pre-articulatory phonological skills are preserved, such as ALS. Despite impressive advances in this technology, this approach is not applicable to individuals with severe phonological deficits, such as those with stroke-induced aphasia. Here we propose an alternative approach, decoding neural activity associated with lexical-semantic retrieval rather than motor planning.

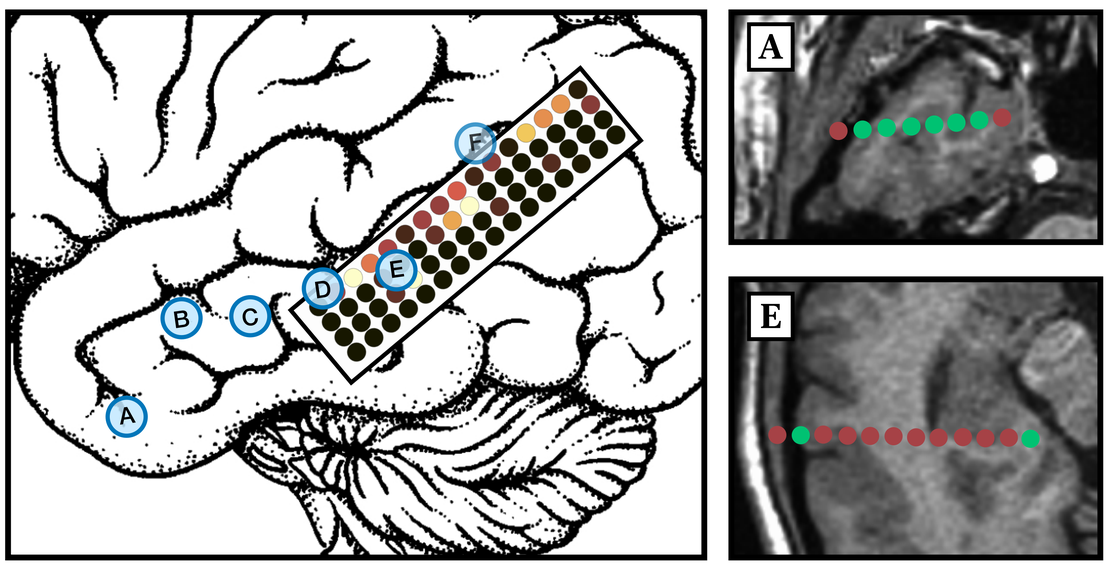


During the intraoperative portion of the study, electrode grids were placed on the brain surface during awake craniotomies.

> **"We demonstrate in this pilot study that machine learning models are able to decode word meaning directly from intracranial electrical potentials. The most critical electrodes were located at known semantic hubs: the anterior temporal pole and the middle temporal gyrus. These findings offer promising evidence for the feasibility of building a language neuroprosthesis based on directly decoding meaning from the semantic system."**

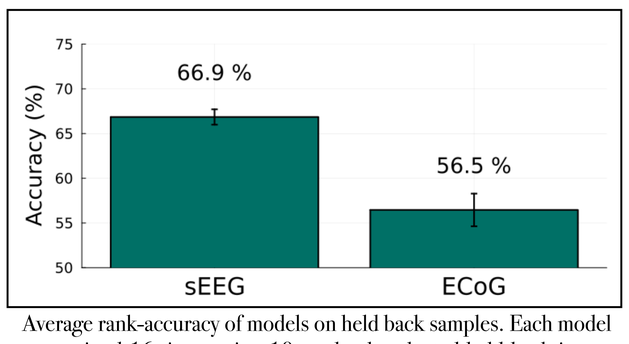
Methods

We recorded signals from two patients: a patient undergoing an awake craniotomy for tumor resection where a high-density ECoG grid was positioned approximately spanning the middle of the MTG to the angular gyrus in the left hemisphere, and a patient undergoing seizure mapping with sEEG electrode placed in the left temporal lobe. Both patients were given picture naming tasks using simple color photos or line drawings of common objects, including animals, tools, vehicles, foods, and musical instruments. The patient undergoing sEEG recording was exposed to 46 stimuli over 181 trials and the patient undergoing intraoperative ECoG to 91 stimuli across 324 trials. Additional training exemplars were created by including copies of trials where the timing was shifted up to 75ms backward and 25ms forward in 0.5ms increments, resulting in 201 versions per stimulus (a total of 36,381 sEEG and 65,124 ECoG trials).

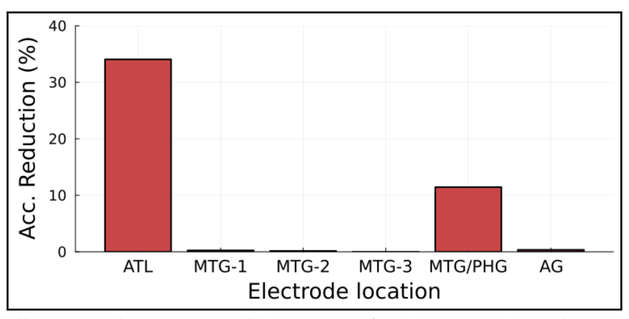


Left panel: location of the electrode grid (large rectangle) and sEEG probes (lettered circles). Grid electrodes that contributed significantly to the prediction model are highlighted in warm colors. Electrodes that contributed significantly in the sEEG patient were only found in probe "A" and "E". Individual electrodes that contributed to model prediction are shown on the axial MRI images in the right panel (green denotes important electrodes).

We extracted high gamma power (HGP) by using the Hilbert transform of the signal filtered with a bandpass of 70Hz to 200Hz. These data were then segmented into 500ms epochs, resampled at 100Hz, and whitened using PCA. We trained the models to predict the vector representation of each word in the experiential model space described by Binder et al. (Cogn Neuropsychol, 2016). The target vectors were also subjected to a dimensionality reduction PCA, from the original 65 features down to 26.



Average rank-accuracy of models on held back samples. Each model was trained 16 times using 10 randomly selected held-back items. Rank accuracy was calculated by making vector predictions and calculating the closeness relative to the entire sample (100% = perfect prediction, 50% = chance).



Percent reduction in model accuracy after lesioning electrodes in either the ATL, MTG, or AG. Model accuracy was most affected by ATL and MTG electrodes, suggesting electrodes in these regions are most critical for concept decoding.

Our machine learning approach used a deep convolutional network comprising six layers of convolutional filters with skip connections, using batchnorm and 50% dropout. The model was trained with ADAM optimization using a learning rate of 0.01 to minimize the MSE of predicted model vectors.

Each model was trained 16 times, up to 50,000 epochs. For model validation, all shifted trials from 10 randomly selected stimuli were held back and rank-accuracy calculated from these stimuli based on the percentage rank of the correct vector. To determine which brain regions contributed most to the model, electrodes were virtually "lesioned" by fixing their value to the mean and measuring the drop in model accuracy.

Results

While both the ECoG ($t(15) = 3.54$, $p=0.003$) and sEEG ($t(15)=19.59$, $p<0.001$) models achieved accuracy significantly greater than chance (50%), the sEEG patient's model was significantly more accurate ($t(30)=5.15$, $p<0.001$). After lesioning both models, electrodes in the MTG, ATL, and parahippocampal gyrus were found to contribute most to model accuracy, with the ATL contributing the most information.

Discussion

In this pilot study, we were able to successfully decode semantic information from intracranial electrodes. While these levels of accuracy are not practical for building a neuroprosthesis per se, they demonstrate that decoding the semantic system directly is possible. Future work with more electrodes and trials may achieve better performance. Consistent with our predictions, the most important electrodes for model performance were regions known to be involved in semantic processing. Interestingly, ATL electrodes contributed the majority of the information for the model.