Decoding speech intent from non-frontal cortical areas

Prashanth R. Prakash[1]*, Jason K. Hsieh[2], Robert D. Flint[1], Joshua Rosenow[1], Matthew C. Tate[1], and Marc W. Slutzky[1]

[1]*Northwestern University, Chicago, Illinois, USA;* [2]*Cleveland Clinic Ohio, USA*

**320 E Superior St Chicago IL 60613, E-mail: prashshant.prakash@northwestern.edu*

Introduction: BCIs aiming to decode speech production to restore communication have largely recorded signals from the speech sensorimotor cortices, including ventral pre-central and postcentral gyri and inferior frontal gyrus [1-3]. The temporal and the parietal lobes are important areas of interest for speech and language perception, but thus far there is a lack of evidence of a speech production signal in these areas. If we could decode speech production from these areas, it could potentially be used to restore communication to people with communication disorders, including expressive aphasia, in which the frontal lobe is damaged. Here, we sought evidence for a speech production signal using electrocorticographic (ECoG) signals recorded from the temporal and parietal cortices from 4 participants.

Materials, Methods, and Results: ECoG arrays containing 19 to 64 electrodes were placed on the temporal and/or parietal cortices in participants undergoing resection of epileptic foci or brain tumors. In participants with epilepsy, standard arrays (10-mm interelectrode spacing) were placed according to clinical necessity. In participants with tumors, mini-ECoG arrays (8x8, 4-mm interelectrode spacing) were placed on the temporal and/or parietal lobes intraoperatively. Participants were presented with single words on a screen in random order. They were instructed to read each word silently, hold it in memory while viewing a blank screen, and then cued visually to say it out loud. This enabled us to disentangle the ECoG signatures of speech production from those of reading or comprehension. ECoG high-gamma (HG) band [70-300 Hz] power in 100-ms non-overlapping windows was used as features for decoding speech intent. Each window was labeled according to the respective behavioral state: speech or silence. To avoid bias due to imbalanced classes and only include causal information, for every spoken word we included either a speech window at the voice onset of that word or a silence window from 1.5-1.6 s after the voice offset of the word. We trained a linear support vector machine to classify the behavioral state of each window using a history of 4 windows. We varied the offset between HG power and speech/silence window from -1.5 s to 0.7 s, where negative values indicate HG leading behavior.

Using this technique, we decoded speech vs. silence, using only causal information (last HG bin before speech/silence window onset) with accuracies ranging from 67.2%-80% over participants ($p<0.03$ in all participants, t-test, compared to shuffled labels). To further investigate evidence for a speech intent signal, we used demixed principal components analysis on the HG power from these cortices. We computed the principal components using HG data from -0.5 s to the onset of speech. We observed a separation between the speech and silence behavioral states in a lower dimensional space using the first 2 most significant demixed principal components.

Discussion and Significance: These results suggest that there is a speech production signal encoded within the temporal and parietal lobes. This signal appears approximately 500 ms prior to the onset of intended speech and can be seen in a low-dimensional manifold of HG power as suggested by demixed PCA. The existence of such activity in the posterior superior temporal lobe has been suggested in some linguistic speech production models, but there has been limited evidence for this activity to date [4]. These results inform us about previously understudied cortical areas for spoken language production. This may advance the development of speech BCIs for people with communication disorders including language disorders (aphasia) as well as motor speech disorders (locked-in syndrome).

References:

1. Ramsey, N.F et al Decoding spoken phonemes from sensorimotor cortex with high-density ECoG grids. NeuroImage, 2018. 180: p. 301-311
2. Herff C, et al. Generating Natural, Intelligible Speech From Brain Activity in Motor, Premotor, and Inferior Frontal Cortices. Front Neurosci. 2019
3. Moses, D.A., et al. Neuroprosthesis for Decoding Speech in a Paralyzed Person with Anarthria. N Engl J Med, 2021.
4. Hickok, G. Computational neuroanatomy of speech production. Nat Rev Neurosci 2012