DOI: 10.3217/978-3-99161-050-2-182

Phoneme Sequence Encoding in Ventral Motor Cortex

C. Fan[5*], B. Meschede-Krasa[3,6], E.M. Kunz[2,3], F. Kamdar[1], N.S. Card[12], M. Wairagkar[12], C. Iacobacci[12], L. Hochberg[7,8,9,10,11], D.M. Brandman[12], S.D. Stavisky[12], S. Druckmann[6], C. Manning[5], J.M. Henderson[1,3,4], F.R. Willett[1,3]
Stanford Univ., CA: [1]Neurosurg., [2]Elec. Eng., [3]Wu Tsai Neuro Inst., [4]Bio-X, [5]Comp. Sci., [6]Neurobio. | Brown Univ., RI: [7]Eng., [8]Carney Inst. for Brain Sci. | VA Med. Ctr., Providence, RI: [9]VA RR&D Neurorestoration & Neurotech. Ctr. | Mass Gen. Hosp, MA: [10]Neurotech. & Neurorecovery Ctr. | Harvard Med. Sch., MA: [11]Neurol. | UC Davis, CA: [12]Neurol. Surg.
E-mail: stfan@stanford.edu

Introduction: Understanding the neural mechanisms of speech production is fundamental for advancing brain-computer interface (BCI) technologies and neuroscience. Intracortical speech BCIs not only achieve high-accuracy decoding[1,2] but also provide a unique window into how the ventral precentral gyrus (vPCG) encodes continuous speech. Here, we demonstrate that vPCG encodes the entire sequence of phonemes within a word during overt speech production, revealing limitations in traditional models of phoneme sequence representation, such as the competitive queue (CQ)[3] and slot-based models[4].

Material, Methods and Results: Neural spiking activities were recorded from two BrainGate2 clinical trial participants, T12 and T15, while they attempted to speak English sentences. Both participants have amyotrophic lateral sclerosis (ALS), leaving us no reliable phoneme timing information. Additionally, each sentence spoken is unique, making it impossible to eliminate trial-to-trial variability in neural spiking activity by averaging multiple repetitions of the same condition. To address these challenges, we developed a recurrent neural network (RNN) encoding model capable of predicting neural firing rates and phoneme timings for any phoneme sequence. The model leveraged data from multiple sessions (14.1 hours for T12, 12.8 hours for T15) to model how vPCG represents sequences of phonemes, allowing us to investigate the neural mechanisms underlying speech production by interrogating the model.

The RNN model accurately predicted neural activity for novel English sentences (Pearson r: 0.80 for T12, 0.85 for T15). Analyses of model-generated neural activity revealed that vPCG simultaneously encodes up to eight phonemes within a word, with stronger encoding of future phonemes than the past. Using linear regression, we identified that subspaces encoding nearby phonemes were highly correlated, distinct from the orthogonal encoding subspaces observed for encoding actions in a motor sequence in the macaque prefrontal cortex[5]. The encoding magnitude for each phoneme followed a gradient-like pattern similar to the CQ model but showed a gradual decay after articulation. Phoneme representations transitioned dynamically through subspaces, moving from future encoding to output and finally to past subspaces, similar to a slot-based model.

Conclusion: Our findings show that vPCG simultaneously encodes multiple phonemes in correlated subspaces and dynamically transitions these representations through time. The serial order is encoded in both the magnitude and the order-specific subspaces. We believe these findings will advance our understanding of the neural coding of phoneme sequences in the ventral motor cortex.

Disclosures: L.R. Hochberg: Consultant to Neuralink, Synchron, Axoft, Precision Neuro, Reach Neuro. D.M. Brandman: Surgical consultant to Paradromics. S.D. Stavisky: Inventors on IP licensed by Stanford University to Blackrock Neurotech and Neuralink; Advisor to Sonera. J.M. Henderson: Inventors on IP licensed by Stanford University to Blackrock Neurotech and Neuralink; Consultant for Neuralink, Enspire DBS, and Paradromics; equity (stock options) in MapLight Therapeutics; co-founder and shareholder of Re-EmergeDBS. F.R. Willett: Inventors on IP licensed by Stanford University to Blackrock Neurotech and Neuralink.

References:

1. Card Nicholas S. *et al.* An Accurate and Rapidly Calibrating Speech Neuroprosthesis. *N. Engl. J. Med.* **391**, 609–618 (2024).
2. Willett, F. R. *et al.* A high-performance speech neuroprosthesis. *Nature* **620**, 1031–1036 (2023).
3. Bohland, J. W., Bullock, D. & Guenther, F. H. Neural representations and mechanisms for the performance of simple speech sequences. *J. Cogn. Neurosci.* **22**, 1504–1529 (2010).
4. Vousden, J. I., Brown, G. D. & Harley, T. A. Serial control of phonology in speech production: a hierarchical model. *Cogn. Psychol.* **41**, 101–175 (2000).
5. Xie, Y. *et al.* Geometry of sequence working memory in macaque prefrontal cortex. *Science* **375**, 632–639 (2022).

Published by Verlag der Technischen Universität Graz