DOI: 10.3217/978-3-99161-050-2-099

Advancing speech BCIs towards conversational speeds in people with paralysis

Kamran Hussain[1,2,*], Erin M. Kunz[2], Nick Hahn[2], Akansha Singh[2], Leigh R. Hochberg[3,4,5], Francis. R. Willett[2], Jaimie M. Henderson[2]

[1]*University of California Santa Cruz, Santa Cruz, CA, USA;* [2]*Stanford University, Stanford, CA, USA;* [3]*Brown University, Providence, RI, USA;* [4]*Massachussets General Hospital, Boston, MA, USA;* [5]*Dept. Veteran Affairs Medical Center, Providence, RI, USA;* *E-mail: kamranh@stanford.edu

Introduction: Intracortical speech Brain-Computer Interfaces (BCIs) have shown promise for restoring rapid communication for individuals with paralysis by decoding the neural representation of speech into text [1, 2, 3, 4]. However, current speech BCI systems still operate at speech rates below that of typical speech, possibly because BCI participants naturally set their own pace at slower speeds (20–70 words per minute) that are approximately half that of natural conversational speech (160 words per minute) [1, 2, 3]. We hypothesize that participants are speaking slower as a natural result of their dysarthria during articulation and vocalization. It remains unknown whether people with dysarthria can comfortably increase their speech rate when instructed to do so, and if so, whether decoding accuracy is affected by this change in preferred speech behavior.

Material, Methods and Results: We tested the maximum speech rates achievable by BrainGate2 clinical trial participant T12, who has two 64-channel microelectrode arrays in speech motor cortex, using various verbal behaviors. Imagined speech reached conversational speeds (120-160 words per minute), while even attempted speech at higher instructed rates surpassed T12s typical pace (100 words per minute compared to her usual 60). We then tested how decoding performance varied across four attempted speaking rates using non-vocalized, mimed speech (attempted mouthed behavior): 15, 30, 60, and 120 words per minute, using an open-loop karaoke-style task (Figure 1b) to help T12 pace her speech. In offline evaluations with a limited 50-word vocabulary, we found that decoding accuracy decreased with increasing speech rate, with the optimum decoding accuracy occurring at 30 words per minute. (Figure 1c). Finally, we examined the neural correlates of speech rate and found sentences spoken at different rates were largely stretched or compressed versions of a consistent neural activity template.

Conclusion: Encouragingly, these results suggest that, at least for the participant studied here, speaking rates can be increased substantially when instructed, and the resulting neural activity remains decodable (although accuracy declines). Additionally, the maximum speech rate for each verbal behavior indicates imagined behaviors can achieve average conversational rate while potentially being less tiring and uncomfortable, though trade-offs with decoding accuracy will need to be addressed.

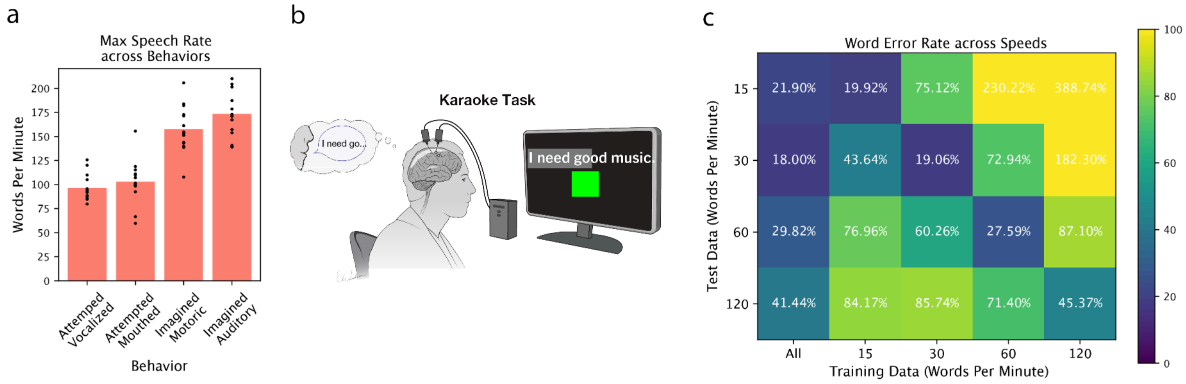

Figure 1: Maximum speech rate across behavior (a), karaoke task design (b), and word error rate for offline decoding when trained and tested at different instructed speeds (c).

Disclosures: The MGH Translational Research Center has a clinical research support agreement (CRSA) with Axoft, Neuralink, Neurobionics, Precision Neuro, Synchron, and Reach Neuro, for which LRH provides consultative input. FRW is an inventor on intellectual property licensed by Stanford University to Blackrock Neurotech and Neuralink. JMH is a consultant for Neuralink and Paradromics, serves on the Medical Advisory Board of Enspire DBS and is a shareholder in Maplight Therapeutics. He is also the co-founder of Re-EmergeDBS, and an inventor on intellectual property licensed by Stanford University to Blackrock Neurotech and Neuralink.

References:

[1] Willett, F.R., Kunz, E.M., Fan, C. et al. A high-performance speech neuroprosthesis. Nature 620, 1031–1036 (2023). https://doi.org/10.1038/s41586-023-06377-x

[2] Metzger, S.L., Littlejohn, K.T., Silva, A.B. et al. A high-performance neuroprosthesis for speech decoding and avatar control. Nature 620, 1037–1046 (2023). https://doi.org/10.1038/s41586-023-06443-4

[3] Card, N.S, et al. An Accurate and Rapidly Calibrating Speech Neuroprosthesis. N Engl J Med 2024;391:609-618, doi: https://doi.org/10.1056/NEJMoa2314132

[4] Kunz, E.M., Meschede-Krasa, B., et al. Representation of verbal thought in motor cortex and implications for speech neuroprostheses. bioRxiv 2024.10.04.616375; doi: https://doi.org/10.1101/2024.10.04.616375

Published by Verlag der Technischen Universität Graz