 DOI: 10.3217/978-3-85125-467-9-46

Articulatory gestures are insensitive to within-word context

E. M. Mugler[1*], M. Goldrick[1], M. C. Tate[2], K. Livescu[3], M. W. Slutzky[1, 2, 4]

[1]*Northwestern University, Chicago, IL, USA;* [2]*Northwestern Memorial Hospital, Chicago, IL, USA;* [3]*Toyota Technological Institute at Chicago, Chicago, IL, USA;* [4]*Rehabilitation Institute of Chicago, Chicago, IL, USA*

*303 East Superior, Lurie 8-231, Chicago, IL, USA. E-mail: emily.mugler@northwestern.edu

Introduction: A brain-computer interface for communication that uses speech-related signal as input could drastically improve efficiency and potentially be more intuitive for users [1]. Previously, we have demonstrated patterns of activity within speech motor cortex that represent phoneme production [2]. However, phonology suggests that the basic units of production may be articulatory gestures, such as the tongue closure at the back of the teeth to produce the /t/ phoneme [3]. We therefore investigated speech motor cortex to determine representation of articulatory gestures using electrocorticography (ECoG). Moreover, cortical representation of phonemes varies significantly if the phonemes are at the beginning or the end of a word [4], which poses challenges in speech decoding. If representation of articulatory gestures is independent of context within a word, it would be easier to decode them from speech motor cortex for use in a brain-computer interface. We therefore also analyzed the context-independence of gesture representation, comparing gestures at the beginning of a word to gestures at the end of a word.

Methods: Five subjects undergoing high-density ECoG monitoring during tumor resection participated in our study. Subjects read words aloud at a rate of 1 every 2 seconds. We estimated articulatory gesture onset times by using acoustic-articulatory inversion [5], using microphone-recorded audio synchronized to ECoG recordings (BCI2000). High gamma band activity (70-200 Hz) was z-scored relative to each electrode. We investigated one electrode at a time, decoding the contextual position for any gesture that exhibited some degree of high gamma representation on that electrode. We decoded the position of each gesture in the word (start vs. end of the word) using short time windows (-100 ms to +50 ms) around gesture onset to isolate signal directly related to production. We used linear discriminant analysis to classify the instance of each gesture as either at the beginning or at the end of a word. Cross-validation (90% train, 10% test) was randomized and repeated 100 times. The 95% confidence intervals of resulting decoding accuracy values were calculated. This same process was repeated using randomly shuffled labels to calculate chance performance. Decoding accuracy relative to chance performance revealed whether context could be decoded using gesture information. These results were compared to previous phoneme context decoding during word production [4].

Results: Average high gamma activity for each gesture for each relevant electrode demonstrated a consistent, robust increase in activity irrespective of gesture context in a word. These results differ from phoneme decoding, which reveal substantial differences in high gamma activity related to contextual position in a word. Further, decoding of contextual position of gestures was not significantly different from chance performance using shuffled labels ($p < 0.05$). In contrast, the context of phonemes in a word could be accurately decoded from the high gamma activity ($p < 0.05$).

Discussion and Significance: These results provide evidence that the signals associated with articulatory gestures are more context-independent than those associated with phonemes. Further, these results indicate that articulatory gestures may be the more fundamental building blocks of speech motor cortex activity, largely due to the context-independence across those gestures for which we could determine a relevant electrode with representative information. Accordingly, future speech brain-computer interface devices using ECoG from motor cortex may perform better if they decode articulatory gestures instead of phonemes, as gestures have a more consistent cortical signature. These findings have significant implications outside of brain-computer interface research, as they advance our understanding of how speech is generated from motor cortex activity.

Acknowledgements: We thank our subjects for volunteering for this study and Robert Flint for help with data collection.

References

[1] E. M. Mugler, "Investigation of Speech for Communicative Brain-Computer Interface," University of Illinois at Chicago, 2013.
[2] E. M. Mugler, J. L. Patton, R. D. Flint, Z. a Wright, S. U. Schuele, J. Rosenow, J. J. Shih, D. J. Krusienski, and M. W. Slutzky, "Direct classification of all American English phonemes using signals from functional speech motor cortex," *J. Neural Eng.*, vol. 11, no. 3, p. 035015, Jun. 2014.
[3] C. P. Browman and L. Goldstein, "Articulatory phonology: an overview.," *Phonetica*, vol. 49, no. 3–4, pp. 155–80, Jan. 1992.
[4] E. M. Mugler, M. Goldrick, and M. W. Slutzky, "Cortical encoding of phonemic context during word production," in *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014.
[5] W. Wang, R. Arora, and K. Livescu, "Reconstruction of articulatory measurements with smoothed low-rank matrix completion," *Spok. Lang. Technol.*, 2014.