 DOI: 10.3217/978-3-99161-050-2-056

Imagined Phoneme Decoding and Protocol Optimization for Non-invasive EEG-based Speech BCIs

Ziyue Zhu[1]*, Rishan Patel[1], Emmanuel Garrison-Hooks[1], Youngjun Cho[1] and Tom Carlson[1]*

[1]*University College London, London, UK*

*Gower Street, London, WC1E 6BT. E-mail: ziyue.zhu.23@ucl.ac.uk, t.carlson@ucl.ac.uk

Introduction: Speech impairments can seriously hinder quality of life. Brain-computer Interfaces (BCIs) provide a potential solution to convert thoughts into text or audio for communication restoration. Non-invasive neuroimaging techniques such as electroencephalography (EEG) offer a safe rehabilitation method [1]. However, abstract neural activities from speech and noise from scalp challenge the imagined speech decoding. Compared with motor imagery, a mature BCI paradigm with standard tasks of 'left/right' hand movement imagery [2], it is also crucial to find a robust paradigm for intuitive speech imagery (SI) BCIs with a distinct word set.

Material, Methods and Results: 10 subjects with English proficiency were asked to imagine silently speaking 6 phonemes (/b/, /p/, /s/, /e/, /i:/, /u:/). 32-channel EEG covering the temporal and primary motor cortex was recorded [3]. Then, a paired two-tailed t-test within the same subjects was done on EEG power and evoked potential to find the difference between imagined phonemes. We used a non-parametric cluster-based permutation test to reduce type I error for multiple comparisons problems [4]. Finally, binary classification tasks were performed for all 15 phoneme pairs. For each subject, we trained a CSP-SVM model and evaluated it with 5-fold cross-validation.

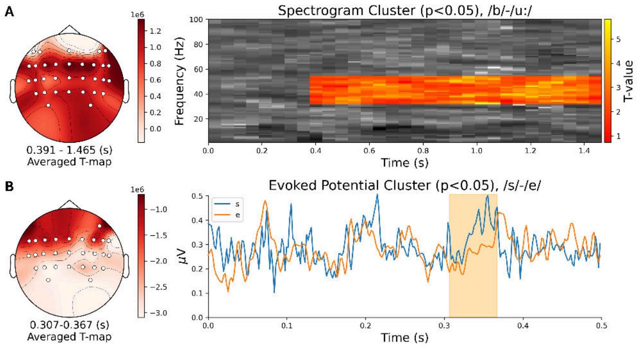

Figure 1: *(A) EEG spectrogram contrast for imagined phoneme /b/ and /u:/. showing significant time-frequency cluster (brightened) with T-values maximized over the significant channels cluster (white circles). (B) EEG evoked potential contrast for imagined phoneme /s/ and /e/ showing significant time window clusters (brightened) with evoked potentials averaged over the significant channel clusters (white circles).*

A significant time-frequency cluster was found for class power contrast between /b/ and /u:/ within the frequency band from 31 to 53 Hz (p=0.026) (Fig. 1(A)), indicating the importance of the gamma band to distinguish between bilabial plosive and rounded vowel phonemes. A significant temporal cluster was also found for evoked potential contrast between /s/ and /e/ around 300ms after the event (p=0.02) (Fig. 1(B)), showing the potential response differing between the fricative consonant and mid-front unrounded vowel. The subject with the best decoding accuracy reached an average accuracy of 67.14% on /b/ and /e/ comparison. The spectral and temporal contrast might indicate that the imagined articulation between consonants and vowels is more distinct and ideal for classification commands [5].

Conclusion: This research studied the EEG pattern of phoneme imagery and provided considerations for command selection for universal speech imagery paradigm optimization. We found significant temporal-spectral differences between some consonant-vowel pairs and the great potential of selecting the distinct phoneme set for reliable SI decoding performance. This ongoing work represents an important step towards real-time decoding with a relatively low-density EEG setup.

Acknowledgments and Disclosures: All authors declare no conflicts of interest.

References:

[1] D. Lopez-Bernal, D. Balderas, P. Ponce, and A. Molina, 'A State-of-the-Art Review of EEG-Based Imagined Speech Decoding', *Front. Hum. Neurosci.*, vol. 16, p. 867281, Apr. 2022.

[2] B. J. Edelman *et al.*, 'Non-invasive Brain-Computer Interfaces: State of the Art and Trends', *IEEE Rev. Biomed. Eng.*, pp. 1–25, 2024.

[3] L. Wang, W. Huang, Z. Yang, X. Hu, and C. Zhang, 'A method from offline analysis to online training for the brain-computer interface based on motor imagery and speech imagery', *Biomed. Signal Process. Control*, vol. 62, p. 102100, Sep. 2020.

[4] M. Meyer, D. Lamers, E. Kayhan, S. Hunnius, and R. Oostenveld, 'Enhancing reproducibility in developmental EEG research: BIDS, cluster-based permutation tests, and effect sizes', *Dev. Cogn. Neurosci.*, vol. 52, p. 101036, Dec. 2021.

[5] N. Mesgarani, C. Cheung, K. Johnson, and E. F. Chang, 'Phonetic Feature Encoding in Human Superior Temporal Gyrus', *Science*, vol. 343, no. 6174, pp. 1006–1010, Feb. 2014.

Published by Verlag der Technischen Universität Graz