DOI: 10.3217/978-3-99161-050-2-183

A pilot study towards synthesizing speech during intraoperative recordings using the Layer 7 cortical interface

Miguel Angrick[1*], Shiyu Luo[1], Daniel N. Candrea[1], Rohit Ganji[1], Katrina Barth[2], Jackie Dister[2], Josh Miller[2], Elton Ho[2], Kurt Lehner[1], Youssef Comair[1], Benjamin Rapoport[2], Nathan E. Crone[1]

[1]The Johns Hopkins University, Baltimore, USA; [2]Precision Neuroscience, New York City, USA

[*]600 N. Wolfe Street, Baltimore, MD, USA. E-Mail: mangric1@jhu.edu

Introduction: Implantable Brain-Computer Interfaces (BCIs) have shown promising results in establishing alternative communication means for people living with a neurodegenerative disease. Previous studies on electrocorticographic (ECoG) signals have shown that increasing the number of contacts and their density can improve speech decoding [1] and synthesis [2] applications. Here, we present a pilot study to test the feasibility of synthesizing speech using the Layer 7 cortical interface (Precision Neuroscience) during an awake craniotomy for glioma resection in the left frontal lobe. Equipped with a total of 1024 contacts placed over ~1.5 cm^2 of face motor cortex, we conducted an experiment in which the patient spoke single words aloud from a list of four after hearing them audibly through a loudspeaker in the operating room. Both neural and acoustic data were simultaneously recorded.

Material and Methods: We developed a BCI system consisting of three components to (1) receive neural and acoustic data in real time and compute time-aligned high-gamma and acoustic features, (2) continuously train a transformer model on those incoming and accumulated features over time, and (3) decode the incoming signals using a continuously updating model. The transformer model maps high-gamma activity into linear predictive coding (LPC) coefficients and utilizes LPCNet to generate the acoustic waveform [3] which is directly written into the soundcard buffer for low-latency playback.

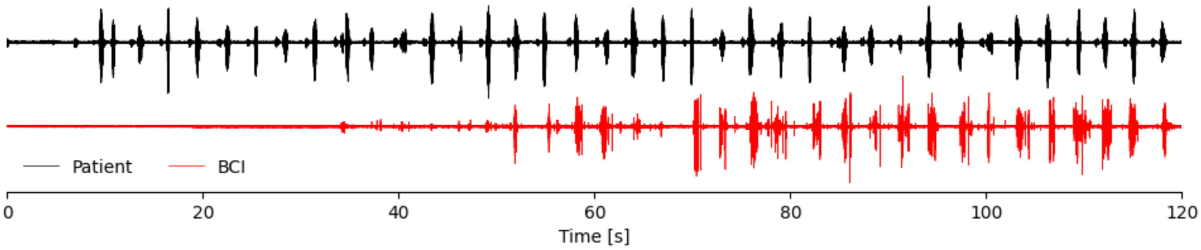

Figure 1. Patient and synthesized speech from the first two minutes in the operating room.

Results: Over time and with accumulating training data, the neural network started to output synthesized speech, at first very quietly, then more audibly. Figure 1 shows a snippet of the actual and reconstructed speech from the first experiment run, indicating that once enough trials have been seen during training, the model is capable of producing speech output at the time when the patient is speaking. Given the short period of time, both for training and data collection, all synthesized samples were not intelligible.

Significance: These preliminary results demonstrate the feasibility of intraoperative speech synthesis, which could be used to verify an effective anatomical placement of the Layer 7 cortical interface in future applications as a fully implantable speech BCI.

Disclosures: KB, JD, JM, EH, and BR have a financial interest in Precision Neuroscience.

References:

[1] Duraivel, S., Rahimpour, S., Chiang, C., Trumpis, M., Wang, C., Barth, K., Harward, S., et al., (2023). High-resolution neural recordings improve the accuracy of speech decoding. *Nature communications*
[2] Metzger, S., Littlejohn, K., Silva, A., Moses, D., Seaton, M., Wang, R., Dougherty, M., Liu, J., Wu, P., et al. (2023). A high-performance neuroprosthesis for speech decoding and avatar control. *Nature*
[3] Wairagkar, M., Card, N., et al., (2024). An instantaneous voice synthesis neuroprosthesis. *bioRxiv*

Published by Verlag der Technischen Universität Graz