 DOI: 10.3217/978-3-99161-050-2-076

Daily independent conversational speech decoding from the intracortical neural activity of a man with ALS

N. S. Card[1*], H. Peracha[1], T. Singer-Clark[1], C. Iacobacci[1], M. Wairagkar[1], X. Hou[1], L. R. Hochberg[2,3,4], D. M. Brandman[1+], S. D. Stavisky[1+]

[+]Co-senior authors

[1]*U.C. Davis Dept. of Neurological Surgery;* [2]*Sch. of Engin. and Carney Inst. for Brain Sci., Brown Univ., Providence, RI;* [3]*VA RR&D Ctr. for Neurorestoration and Neurotechnology, Providence VA Med. Ctr., Providence, RI;* [4]*Dept. of Neurol., Massachusetts Gen. Hosp., Boston, MA*

[*]202 Cousteau Pl Suite 160, Davis, CA, USA. E-mail: nscard@ucdavis.edu

Introduction: Communication is a priority for the millions of people living with dysarthria due to brain injuries or neurological disorders such as stroke and amyotrophic lateral sclerosis (ALS). BCIs can enable rapid, intuitive communication for people with paralysis by transforming the cortical activity associated with attempted speech into text. Recent advancements in speech BCIs have demonstrated that attempted speech can be accurately decoded during prompted speech tasks [1, 2, 3]. However, work remains to achieve accurate speech decoding in a conversational setting that enables the user to communicate their own thoughts and needs with those around them. Here, we report a speech BCI that was accurate enough to enable the user to have extensive conversations with family, friends, and colleagues independently and on a daily basis.

Material, Methods, and Results: A 45-year-old man ('T15') with ALS and severe dysarthria was enrolled into the BrainGate2 clinical trial. Four microelectrode arrays were placed in his left precentral gyrus to record neural activity from 256 intracortical electrodes. We trained a recurrent neural network to decode sequences of phonemes (i.e., the building blocks of words) from T15's neural signals as he attempted to speak. The recurrent neural network was continuously finetuned with new data to enable stable decoding over weeks of use. Predicted phoneme sequences were assembled into the most likely words being spoken by a language model with a potential output vocabulary size of 125,000 words and displayed on a screen in real time. At the end of a sentence, the decoded words could be played aloud using a text-to-speech tool that was programmed to sound like T15's pre-ALS voice, or entered as text on his personal computer. We developed a "conversation mode" user interface that detects when T15 attempts to speak, decodes accordingly, and allows him to make simple corrections to decoded sentences. T15's care partners were trained to connect T15 to the neural recording system whenever he wanted to use it, and the BCI computer system was streamlined so it could be turned on with just a few button presses.

Throughout 550 days since T15's implant surgery, he has independently used the speech BCI as his primary form of communication on more than 300 individual days. He has used the BCI system for >2,500 hours to utter >150,000 sentences (>1,400,000 words) at an average rate of 64.6 words per minute. At the end of each sentence he uses an eye tracker to rate the correctness of the decoded sentence, yielding a distribution of 65.8% of all sentences being "completely correct", 26.1% being "mostly correct", and the remainder being "incorrect". In recent months, sentence-level accuracy has increased to up to 85% of sentences being "completely correct". Furthermore, we have conducted periodic decoding accuracy benchmarks where T15 attempts to say prompted sentences, resulting in an average word error rate of 2.5% maintained throughout 550 days post-implant.

Conclusion: We demonstrate a speech BCI that has enabled a severely dysarthric man with ALS to engage in conversations with his friends, family, and colleagues. The BCI now serves as his preferred method of communication.

Disclosures: M. Wairagkar: Inventor on IP owned by UC Davis; L.R. Hochberg: Consultant to Neuralink, Synchron, Axoft, Precision Neuro, Reach Neuro; D.M. Brandman: Inventor on IP owned by UC Davis; surgical consultant to Paradromics; S.D. Stavisky: Inventor on IP owned by Stanford University (licensed to Blackrock Neurotech and Neuralink) and UC Davis; advisor to Sonera.

References:

[1] Card, NS *et al*. An accurate and rapidly calibrating speech neuroprosthesis. *New Engl. J. Med.* 391, 609–618 (2024).

[2] Willett, FR *et al.* A high-performance speech neuroprosthesis. *Nature* 620, 1031–1036 (2023).

[3] Metzger, SL *et al.* A high-performance neuroprosthesis for speech decoding and avatar control. *Nature* 620, 1037–1046 (2023).

 Published by Verlag der Technischen Universität Graz