# A closed-loop brain-computer interface that synthesizes intelligible words for a clinical trial participant with amyotrophic lateral sclerosis

Miguel Angrick[1], Shiyu Luo[1], Qinwan Rabbani[1], Daniel N. Candrea[1], Samyak Shah[1], Griffin W. Milsap[2], William S. Anderson[1], Chad R. Gordon[1], Kathryn R. Rosenblatt[1], Lora Clawson[1], Nicolas Maragakis[1], Francesco V. Tenore[2], Matthew S. Fifer[2], Hynek Hermansky[1], Nick. F. Ramsey[3], Nathan. E. Crone[1]

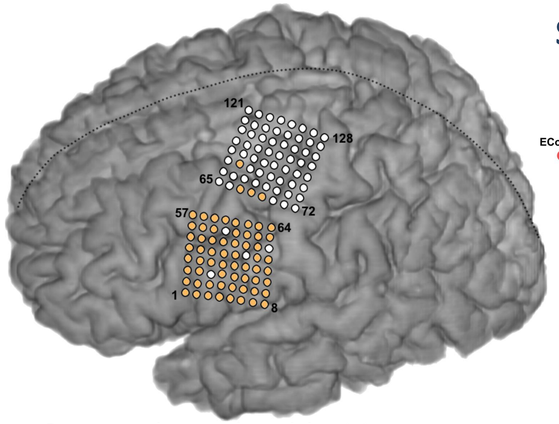
[1]The Johns Hopkins University, Baltimore, MD, USA; [2]Johns Hopkins Applied Physics Laboratory, Laurel, MD, USA; [3]UMC Utrecht Brain Center, Utrecht, The Netherlands

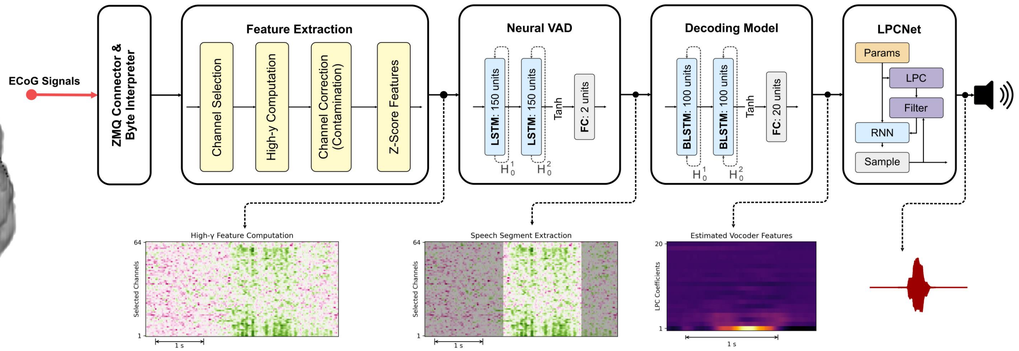
Preprint

## Motivation

- Neurological disorders such as ALS can weaken muscles critical for speech
  - Reduced intelligibility or loss of speech altogether
- Implantable Brain-Computer Interfaces (BCIs) can potentially restore acoustic speech
  - By translating neural signals from speech production directly to acoustic waveforms
  - If speech is still possible, can use it for acoustic target in machine translation
- Many studies and models focus on epilepsy patients being monitored in EMU
  - Recover acoustic speech signals retrospectively
  - Not exposed to same challenges when it comes to real-time usage
  - Different neurological disorder
- Online speech synthesis from participant with amyotrophic lateral sclerosis (ALS)
  - Study participant enrolled in ongoing clinical trial (ClinicalTrials.gov, NCT03567213)
  - Approved by the Johns Hopkins Institutional Review Board (IRB) and the FDA under an investigational device exemption (IDE)
  - Speaking capabilities gradually deteriorating, individual words often intelligible

## Speech synthesis approach



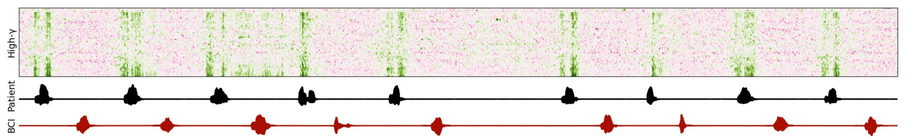

**Deep neural network model training:**

- Extraction of time-aligned neural and acoustic frames across training days
- High-gamma (High-Y) activity features (70-170Hz) from ECoG signals
- Voice activity detection (VAD) information for each frame for neural VAD
- Linear Predictive Coding (LPC) coefficients for LPCNet vocoder

**Inference:**

- Real-time streaming system on top of node-based computational graph
- ECoG signals are streamed via BCI2000 and converted to High-Y activity
- Neural VAD identifies segments of spoken speech in High-Y activity
- Decoding model estimates vocoder features used for waveform generation

## Experimental Design

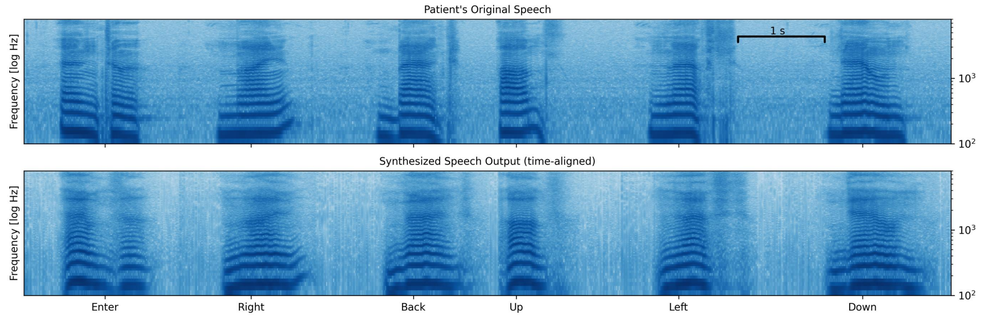

**Data Collection:**

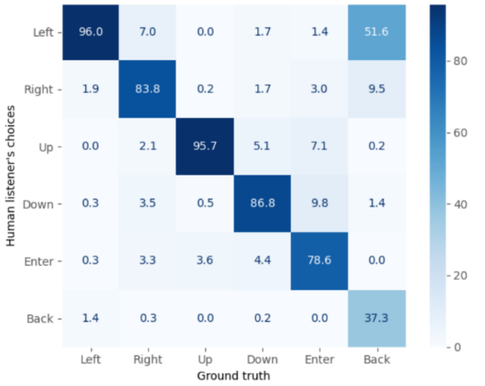
- Single words from a 6-word vocabulary
- Presented on screen in front of participant
- Participant instructed to read them out aloud
- Data normalized on day-specific calibration data

**Online speech synthesis**

- Participant repeats experiment
- Synthesized audio is played back
- Simultaneously, stored on disc for offline evaluation

## Results




- Evaluation of the intelligibility of the synthesized speech indicates that many words can be correctly recognized by human listeners (80%), share similar characteristics to original speech
- Our results provide further evidence that previous speech BCI demonstrations with epilepsy patients may be generalizable to speech-impaired individuals, particularly those with ALS.

Research reported in this publication was supported by the National Institute Of Neurological Disorders And Stroke (NINDS) of the National Institutes of Health (NIH) under Award Number UH3NS114439. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.