Discrimination of Overt, Mouthed, and Imagined Speech Activity using Stereotactic EEG

Discrimination of Overt, Mouthed, and Imagined Speech Activity using Stereotactic EEG

P. Z. Soroush[1], S. Y. Cole[1], C. Herff[2], S. K. Ries[3], J. J. Shih[4], T. Schultz[5], and D. J. Krusienski[1]

Abstract Recent studies have demonstrated that it is possible to decode and synthesize acoustic speech directly from intracranial measurements of brain activity. A current major challenge is to extend the ef cacy of this decoding to imagined speech processes toward the development of a practical speech neuroprosthesis for the disabled. The present study used intracranial brain recordings from participants that performed a speaking task consisting of overt, mouthed, and imagined speech trials. In order to better elucidate the unique neural features that contribute to the discrepancies between overt and imagined model performance, rather than directly comparing the performance of speech decoding models trained on respective speaking modes, this study developed and trained models that use neural data to discriminate between pairs of speaking modes. The results further support that, while there exists a common neural substrate across speech modes, there are also unique neural processes that differentiate speech modes.

I. Introduction

Speech is the rst and foremost modality of human interpersonal communication. Brain-Computer Interfaces (BCIs) that decode and synthesize speech could dramatically improve life for individuals unable to speak due to injury or disease. Invasive measurements of brain activity using electrocorticography (ECoG) [1] or stereotactic electroencephalography (sEEG) [2] have recently shown promise for developing such speech BCIs [3], [4], [5], [6], [7].

For those who have lost the ability to speak, the objective is to translate neural processes during imagined speech to acoustic speech. However, the lack of behavioral output during imagined speech makes it extremely challenging to design an effective decoding model [8], [9]. To overcome this challenge, studies often employ neural processes or behavioral output from overt or mouthed (i.e., inaudible articulations without vocalization) speech as a surrogate to study associated neural activity [3], [10] or to train decoding models [5], [7], [11] for imagined speech applications.

While these studies have shown substantial promise, there are clear limitations to using overt speech surrogates for training imagined-speech decoding models. This may be due, in part, to the unique brain regions activated during overt, mouthed, and imagined speech, and the differences between the neural features extracted from these regions [8], [12], [13], [14]. In order better elucidate the unique neural features that contribute to the discrepancies between overt and imagined model performance, rather than directly comparing the performance of speech decoding models trained on respective speaking modes, the present study developed and trained models that used neural data to discriminate between pairs of speaking modes.

*This work was supported by NSF 2011595/1608140 and BMBF 01GQ1602.

[1]P. Z. Soroush, S. Y. Cole, and D. J. Krusienski are with Virginia Commonwealth University, Richmond, VA, USA zanganehp@vcu.edu

[2]C. Herff is with the University of Maastricht, Maastricht, Netherlands

[3]S. Ries is with San Diego State University, San Diego, CA, USA

[4]J. J. Shih is with UCSD Health, San Diego, CA, USA

[5]T. Schultz is with the University of Bremen, Bremen, Germany

Fig. 1: The combined sEEG depth electrode (channel) locations of the 7 participants from different perspectives using an averaged brain model.

II. Methodology

A. Participants and Electrode Locations

sEEG data were collected from 7 native English-speaking participants being monitored as part of treatment for intractable epilepsy at UCSD Health. The demographic information of the participants is provided in Table I. The study design was approved by the Institutional Review Boards of Virginia Commonwealth University and UCSD Health, and informed consent was obtained for experimentation with human subjects. The locations of sEEG electrodes were determined solely based on the participants' clinical needs. A subset of the implanted electrodes for each participant was determined to be in or adjacent to brain regions associated with speech and language processing. Fig. 1 shows the depth electrode locations for the 7 participants, with sEEG electrode (channel) counts provided in Table I.

B. Experimental Design and Data Collection

The experimental setup and trial sequence structure are depicted in Fig. 2. For each trial sequence, a sentence