

A Commentary on Linguistic Representation of Vowels in Speech Imagery EEG

San Chu*

Department of Neuroscience, Doshisha University, Kyoto, Japan

Commentary

Received: 28-Nov-2023, Manuscript No. Neuroscience-24-121499;
Editor assigned: 30-Nov-2023, Pre QC No. Neuroscience-24-121499 (PQ); **Reviewed:** 14-Dec-2023, QC No. Neuroscience-24-121499;
Revised: 20-Aug-2024, Manuscript No. Neuroscience-24-121499 (R);
Published: 27-Aug-2024, DOI: 10.4172/neuroscience.8.4.001

***For Correspondence:** San Chu, Department of Neuroscience, Doshisha University, Kyoto, Japan; **E-mail:** chu802@edu.cn

Citation: Chu S. A Commentary on Linguistic Representation of Vowels in Speech Imagery EEG. RRJ Neuroscience. 2024;8:001. 2024;10:001.

Copyright: © 2024 Chu S. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

DESCRIPTION

Progress is being made in the field of neural decoding for direct communication in Brain-Computer Interfaces (BCIs). Current research is dedicated to identifying spoken signals through the analysis of multi-channel Electrocorticograms (ECoGs) in the brain cortex. If we could instead detect linguistic information from Electroencephalograms (EEGs), then BCIs could enjoy much wider practical applications, for instance improving the Quality of Life (QoL). Efforts to aid Amyotrophic Lateral Sclerosis (ALS) patients face numerous unresolved challenges. Although research on spoken Electroencephalograms (EEGs) can utilize motor command data for identifying speech-related signals, the absence of this advantage in imagined speech EEGs (i.e., EEGs recorded during silent, unspoken speech) necessitates the exclusive identification of linguistic representations from within the EEG.

In this paper, we propose a model of encoding and decoding for linguistic information $L(k)$, k ; frequency. The encoding process convolves an input spectrum of random signal $W(k)$ and $L(k)$ and outputs an EEG spectrum $X(k)$. In the decoding procedure, the EEG spectrum $X(k)$ undergoes analysis through an inverse filter $H(k)$ within a feedback loop, incorporating $L(k)$. Linear Predictive Analysis (LPA) is applied to analyze imagined speech EEGs around the Broca area. The LPA spectrum patterns are converted to line-spectra that become closer to symbolic forms.

Because there is no reference EEG signal corresponding to the exact moment the speech was imagined (that is, spoken silently in the human subject's mind), our goal is to identify the locations and patterns within the multi-channel EEG signal that represent phones or phrases. Through an examination of numerous EEG line-spectra related to phones, words, and sentences, we have discovered that integrating (or pooling) multi-channel data reveals distinct chunks of open syllables. Specifically, consonant-vowel combinations, or CVs, with durations ranging from 7 to 9 frames (56 to 72 (ms)), become evident. Additionally, these CVs can be categorized as monosyllables.

Principal Component Analysis (PCA) is employed to search and reconstruct a collection of vowel spectra $\{X(k)\}$, providing a visualization of linguistic information through eigen-vectors $\phi(m)$. Additionally, the Subspace Method (SM) is utilized to explore the suitable spectra of vowels, facilitating the recombination of $\{X(k)\}$ and the redesign of the eigen-vector set. This iterative search is repeated 4 times, then the last eigen-vector set of each vowel $\Psi(v, m)$; $v=a, e, i, o, u$ is fixed. The eigen space $\Psi(v, m)$ contains the linguistic representation of vowels, and we can observe it in a visible form by calculating the referencing vector

Research & Reviews: Neuroscience

$G(v)$ that is the accumulated spectrum with the weight $\lambda(m)/\lambda(1)$.

The magnitude of the eigen-value $\lambda(m)$ signifies its contribution to $G(v)$. Notably, the presence of two spectral peaks (P1, P2) in the higher frequency range of $G(v)$ brings to mind the two formant frequencies (F1, F2) observed in audio spectra of spoken vowels. It is worth mentioning that the five vowels in the P1-P2 scatter plot approximately align along a line, whereas cardinal vowels in an F1-F2 plot for spoken speech typically form a quadrilateral.

We employed a jack-knife technique, utilizing four out of the five human subjects as training data, while the remaining one human subject served as the test data. This process of training and testing was iteratively repeated by alternating training and test data for all human subjects, leading to cross-validation. Across 5 human subjects (that is, $1425 \times 4=5700$ samples for training, and $1425 \times 1=1425$ samples for testing). Each human subject imagined speech data set of 57 Japanese short syllables. The classifier, based on Convolutional Neural Networks (CNNs), achieved an average recognition accuracy of 72.6% for vowels in imagined speech.

To enhance the interpretation of linguistic patterns in EEG signals, our objectives include (a) extracting consonantal data, (b) enhancing the accuracy of vowel and consonant recognition, partially achieved through the expansion of imagined speech datasets, and (c) developing decoding modules for isolated words and/or connected phrases, catering specifically to applications in Brain-Computer Interface (BCI) technology.