Title	音声知覚における脳内の音声エンコーディング･デコーディングのプロセスの研究
Author(s)	周, 迪
Citation	
Issue Date	2023-09
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/18781
Rights	
Description	Supervisor: 鵜木祐史, 先端科学技術研究科, 博士

Abstract

Research Background

In recent years, there has been a growing trend towards using naturalistic setups in speech research. These setups involve presenting participants with continuous speech, such as stories or conversations, instead of isolated words or sentences. This provides a more realistic understanding of speech processing and comprehension in real-world conditions. However, fMRI studies have found that naturalistic speech processing elicits more widespread brain activation compared to traditional setups. This poses a challenge for traditional neurocognitive models of language, which are based on isolated words or simple sentences and have a strong left hemisphere bias. These models fail to explain how naturalistic speech is processed in the brain.

Research Methodology and Purpose

This study is guided by the theory of systems, which allows us to explore and assess the functioning of the brain system by estimating its encoding function. This study focuses on the process of semantic processing during speech perception and investigates the extraction of the temporal amplitude envelope (TAE) from the speech signal. This TAE carries crucial semantic information and is encoded by specific brain regions involved in semantic processing. This study aims to describe the encoding of TAE in the brain using a linear time-invariant (LTI) model. Given the linearity of the system, it can reverse this transformation and recover the original semantic information of TAE. The research philosophy is based on the belief that by accurately estimating the encoding function and capturing the information at the semantic processing stage, it may be possible to infer what is being heard solely from the decoding process, without directly hearing the sound.

To validate this philosophy, this study aims to <u>(1) identify the brain regions responsible for semantic processing</u>, <u>(2) determine the semantic representations in the brain regions</u>, and <u>(3) explore the possibility of recovering the original semantic information from these representations</u>. In the experiment, this study exposed participants to both normal speech, which contained semantic information, and time-reversed speech, which disrupted the semantic content. The aim was to compare the brain activity patterns between the two conditions and identify the specific brain regions responsible for processing semantic information. Then, it could proceed to reconstruct the original TAE information from these identified brain regions. And this study employed a hyper-alignment method to map EEG signals from the scalp to the cortex level to overcome the spatial limitation of EEG.

Results

(1) Semantic processing engages frontal, temporal, and cingulate brain areas, it involves a widespread distribution of brain regions beyond the traditional temporal and frontal areas.

(2) Semantic representations can be found in these brain areas. Moreover, this study has observed a stronger correlation between neural oscillations in the delta and gamma frequency bands and semantic processing.

(3) Through the reverse decoding process, this study successfully reconstructed the TAEs of speech from brain activity and recovered the semantic information using the NVS technique.

Research significance

Our study is unique in that it combines both temporal and spatial dimensions to offer new insights into the natural language processing that occurs in the brain. Using an innovative methodology, we have been able to estimate encoding functions across various brain regions - a feat that has previously been challenging in both fMRI and EEG research. By challenging traditional models, we've emphasized the extensive involvement of multiple brain regions and the dynamic nature of their encoding capabilities. Our findings suggest that the desynchronization between different subnetworks, especially within the frontal and temporal areas, plays a crucial role in the brain's semantic information processing mechanism. In our study, we were able to successfully reconstruct the TAEs of speech from brain activity and recovered semantic information using noise-vocoded speech (NVS). While restoring speech intelligibility on unknown data remains a challenge due to noise and inter-subject variability, we were able to achieve a perfect fit of TAEs and recovered semantic information to a certain level of speech intelligibility in the known training dataset. This contributes to a deeper understanding of language processing and has potential implications for technologies like speech-brain interfaces in the future.

Keywords: electroencephalography (EEG), speech encoding/decoding, temporal response function, source localization, neural entrainment