

# Towards Speech reconstruction from MEG signal when audio stimuli

*Tuesday, 25 August 2020 11:30 (30 minutes)*

In recent years, research and development have been conducted on Brain-Machine Interface (BMI), a technology that directly connects the brain and machine using information obtained by measuring brain signals or giving stimulation to the brain.

BMI is expected to contribute not only to the medical field but also to various other fields. BMI technology can be divided into two types (input-type and output-type). Among them, output-type BMI is a technology for sending signals from within the brain to the outside. This technology is being researched and developed to help patients with intractable diseases who are unable to move their bodies as they wish.

Communication with patients with intractable diseases is an essential issue in nursing care. Improving communication can improve a patient's quality of life.

Conventionally, an output-type BMI system that selects words one by one by image stimulation or controlling the cursor from brain activity has been the mainstream. However, they must pay attention to the screen, and input speed is plodding. To reconstruct with simple word information, complex information such as emotions and intonation could not be given. As an innovative method to solve these problems, attention has been focused on a method that converts from brain neural activity to speech.

This method can be divided into two: reconstructing external audio stimuli and reconstructing the subject's own speech. In both methods, other research results show that high-quality speech has been reconstructed. However, in this research, electrodes are surgically implanted in the brain, and the invasive electrocorticography (ECoG) is measured.

In this research, we propose a method to reconstruct external audio stimuli focusing on magnetoencephalography (MEG), which is one of the non-invasive measurement techniques. In this method, the target parameter for reconstruction is a Vocoder parameter (spectral envelope, band aperiodicity, fundamental frequency, voiced-unvoiced (VUV)) used in a speech synthesis system. However, since brain activity data acquired by MEG has limitations such as long-term measurement costs and physical constraints, large-scale data collection for deep learning cannot be performed sufficiently. To learn even small data efficiently, the target parameter is learned with Auto-Encoder.

The target parameters are converted to abstract features by the output of the learned Auto-Encoder middle layer. This research introduces how to perform better learning with limited brain activity data and reconstruct external audio stimuli.

**Primary author:** Mr YAMASHITA, Masato (Kanazawa Institute of Technology)

**Co-author:** Prof. NAKAZAWA, Minoru (Kanazawa Institute of Technology)

**Presenter:** Mr YAMASHITA, Masato (Kanazawa Institute of Technology)

**Session Classification:** Biomedicine & Life Science Applications Session

**Track Classification:** Biomedicine & Life Sciences Applications