Abstract

Electrophysiological activity in the brain can be recorded as a subject listens to or imagines speech. Decoding speech information from this brain data means attempting to reconstruct elements of the original speech stimulus. The eventual long-term goal of this type of speech reconstruction would allow for subjects to imagine what they want to say and have a brain-computer interface produce it for them. This would help individuals with diseases such as Amyotrophic lateral sclerosis (ALS) to speak when they no longer have the motor ability to do so. Decoding speech information also helps us to discover more about the brain and how it works in relation to speech. Currently, non-invasive methods of recording brain activity have not achieved as much success in decoding as invasive methods due to the noise and inaccuracy of the data. We set a new baseline for decoding speech information using non-invasive brain recording methods. This involves using Multiway Canonical Correlation Analysis (MCCA) to de-noise the data and a Temporal Response Function (TRF) to decode the brain data into vocoder features. The features were then input into a vocoder to attempt to reconstruct the stimulus speech audio. We also analysed how speech that a subject imagines is decoded in comparison to speech that a subject listens to and proved that our methodologies can work for different types of brain data recording.