

Speech Synthesis from neural decoding of spoken sentences

Gopala K. Anumanchipalli, Josh Chartier, Edward Chang

Technology that translates cortical activity into speech would be transformative for people unable to communicate as a result of neurological impairment. Decoding speech from neural activity is challenging because speaking requires extremely precise and dynamic control of multiple vocal tract articulators on the order of milliseconds. Here, we designed a neural decoder that explicitly leverages the continuous kinematic and sound representations encoded in cortical activity to generate fluent and intelligible speech. A recurrent neural network first decoded direct cortical recordings into vocal tract movement representations, and then transformed those representations to acoustic speech output. Modeling the articulatory dynamics of speech significantly enhanced performance with limited data. Naïve listeners were able to accurately identify and transcribe decoded sentences. Additionally, speech decoding was not only effective for audibly produced speech, but also when participants silently mimed speech. These results advance the development of speech neuroprosthetic technology to restore spoken communication in patients with disabling neurological disorders.

Contact Us
UCSF Main Site

© 2013 The Regents of the University of California

Source URL: <https://speechneuro.ucsf.edu/speech-synthesis-neural-decoding-spoken-sentences>