Vowels in the brain

Keith Johnson

University of California, Berkeley

Working with a team at Univ. of California, San Francisco we have been uncovering some patterns in neural response during speech perception and production [1-7]. This talk will be about the latest installment in this research program. In work done with Yulia Oganian, Ilina Bhaya-Grossman, and Edward Chang [8] we derived two-dimensional speech receptive fields in electrocorticography recordings from the human auditory cortex. During neurosurgery for intractable epilepsy, fifteen native Spanish-speaking patients listened to natural continuous speech from the DIMex corpus of read sentences, and to synthetic vowel tokens spanning the acoustic vowel space. A high-density grid of electrodes (256 electrodes, 4mm spacing) on the surface of the cortex recorded the high-gamma (70-150 Hz) neural activity at each electrode. Electrodes over the superior temporal gyrus were active during listening, and contained enough information to be able to discriminate between the vowel categories. The response patterns indicate some key features of neural representation of vowels. We found that neural activity at single sites on STG was highly selective to particular zones in formant space, and that this tuning shifted dynamically to adjust for speaker-specific context. When using the population of formant-encoding electrodes, vowel categories could be accurately decoded.

Bibliography

[1] Chang, Edward; Rieger, Jochem; Johnson, Keith; et al. (2010) Categorical speech representation in human superior temporal gyrus. Nature Neuroscience 13, 1428-1432. https://doi.org/10.1038/nn.2641

[2] Bouchard KE, Mesgarani N, Johnson K, Chang EF. (2013) Functional organization of human sensorimotor cortex for speech articulation. Nature 495, 327-32. https://doi.org/10.1038/nature11911

[3] Mesgarani Nima; Cheung, Connie; Johnson, Keith; Chang Edward F. (2014) Phonetic feature encoding in human superior temporal gyrus. Science 343, 1006-10. https://doi.org/10.1126/science.1245994

[4] Cibelli, Emily S.; Leonard, Matthew K.; Johnson, Keith; Chang, Edward F. (2015) The influence of lexical statistics on temporal lobe cortical dynamics during spoken word listening. Brain and Language 147, 66-75. https://doi.org/10.1016/j.bandl.2015.05.005

[5] Cheung, Connie; Hamilton, Liberty; Johnson, Keith; Chang, Edward F. (2015) The auditory representation of speech sounds in human motor cortex. eLife 2016;5, e12577. https://doi.org/10.7554/eLife.12577

[6] Chartier, Josh; Anumanchipalli, Gopala K.; Johnson, Keith; Chang, Edward F. (2018) Encoding of articulatory kinematic trajectories in human speech sensorimotor cortex, Neuron 98,1042-1054. https://doi.org/10.1016/j.neuron.2018.04.031

[7] Sjerps, Matthias J.; Fox, Neal P.; Johnson, Keith; Chang, Edward F. (2019) Speaker-normalized sound representations in the human auditory cortex. Nature Communications 10, 2465. https://doi.org/10.1038/s41467-019-10365-z

[8] Oganian, Y. Bhaya-Grossman, I., Johnson, K., Chang, E.F. (2022) Vowel and formant representation in human auditory speech cortex. bioRxiv. https://doi.org/10.1101/2022.09.13.507547