Intracerebral investigation of the neural representation of voice in human auditory cortex using voice-like acoustic stimuli

Jasmine Hect[1], Kyle Rupp[1], Emily Harford[1], Kanupriya Gupta[1], Hope Reecher[1], Frederic Dick[2], Lori Holt[4], Taylor Abel[1,5]

[1]Department of Neurological Surgery, University of Pittsburgh
[2]Department of Psychological Sciences, Birkbeck College
[3]Department of Experimental Psychology, UCL
[4]Department of Psychology, Carnegie Mellon University
[5]Department of Bioengineering, University of Pittsburgh

Jasmine.hect@pitt.edu

Abstract (max 300 words)

Voice perception engages characteristic regions of auditory cortex. However, it remains unclear to what extent these regions rely on shared or unique mechanisms for processing voice and non-voice sounds (1–5) and for assessing for sound patterns specific to voice (6–9). We used direct intracerebral recordings (from auditory cortex of four patient-participants with epilepsy undergoing chronic monitoring) to test the hypothesis that temporal voice areas (TVAs) rely on shared representations of voice and other natural sounds across superior temporal gyrus (STG) and superior temporal sulcus (STS). Data were recorded while participants listened to 1) voice and non-voice stimuli (Voice Localizer (10)) and 2) synthetic sounds generated from modulated noise (11), called Gaussian Sound Patterns (GSPs). The GSPs stimuli mirror spectrotemporal features of natural sounds, while remaining perceptually distinct, in line with prior fMRI work (1). We used a CNN to classify GSPs into sound categories (12) and selected stimuli most and least likely to be classified as voice (250 total). We extracted broadband high-gamma activity (HGA; 70-150 Hz) and identified sound-responsive channels (two-sample t-test, FDR-corrected, $q < 0.01$). We tested decoding (80% train, 20% test) of stimulus category from HGA prior to cross-task decoding, used to examine similarities in the neuronal representation between voice and GSPs within sound-responsive channels. Decoding of voice from non-voice was significant for all patients (61-76%, $p < 0.001$). Decoding accuracy of GSPs was significantly above chance for two patients (63% and 68% accuracy, $p < 0.001$), and nonsignificant for two (46% and 55% accuracy). We found cross-task decoding did not perform above chance when GSPs (41-48%) or Voice Localizer (41-52%) were used as the training set. These preliminary data suggest TVAs may employ unique representations of voice, even when spectrotemporal properties are controlled between artificial and natural sound stimuli.

[Word count: 294]

References

1. Staib M, Frühholz S. Cortical voice processing is grounded in elementary sound analyses for vocalization relevant sound patterns. Prog Neurobiol. 2021 May 1;200:101982.

2. Norman-Haignere S, Kanwisher NG, McDermott JH. Distinct Cortical Pathways for Music and Speech Revealed by Hypothesis-Free Voxel Decomposition. Neuron. 2015 Dec 16;88(6):1281–96.

3. Peretz I, Vuvan D, Lagrois MÉ, Armony JL. Neural overlap in processing music and speech. Philos Trans R Soc B Biol Sci. 2015 Mar 19;370(1664):20140090.

4. Zatorre RJ, Belin P, Penhune VB. Structure and function of auditory cortex: music and speech. Trends Cogn Sci. 2002 Jan 1;6(1):37–46.

5. Staib M, Frühholz S. Distinct functional levels of human voice processing in the auditory cortex. Cereb Cortex. 2022 Mar 26;bhac128.

6. Chevillet M, Riesenhuber M, Rauschecker JP. Functional Correlates of the Anterolateral Processing Hierarchy in Human Auditory Cortex. J Neurosci. 2011 Jun 22;31(25):9345–52.

7. Hickok G, Poeppel D. The cortical organization of speech processing. Nat Rev Neurosci. 2007 May;8(5):393–402.

8. Staeren N, Renvall H, De Martino F, Goebel R, Formisano E. Sound Categories Are Represented as Distributed Patterns in the Human Auditory Cortex. Curr Biol. 2009 Mar 24;19(6):498–502.

9. Fontolan L, Morillon B, Liegeois-Chauvel C, Giraud AL. The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex. Nat Commun. 2014 Sep 2;5(1):4694.

10. Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. Voice-selective areas in human auditory cortex. Nature. 2000 Jan;403(6767):309–12.

11. McDermott JH, Wrobleski D, Oxenham AJ. Recovering sound sources from embedded repetition. Proc Natl Acad Sci. 2011 Jan 18;108(3):1188–93.

12. Gemmeke JF, Ellis DPW, Freedman D, Jansen A, Lawrence W, Moore RC, et al. Audio Set: An ontology and human-labeled dataset for audio events. In: 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2017. p. 776–80.