

# Neural responses in human fusiform gyrus support a model of heteromodal representation of familiar speakers

*Jasmine L. Hect<sup>1</sup>, Kyle Rupp<sup>1</sup>, Ariane Rhone<sup>2</sup>, Matthew Howard, III<sup>rd</sup>, Taylor J. Abel<sup>1,3</sup>*

<sup>1</sup>Department of Neurological Surgery, University of Pittsburgh, Pittsburgh PA

<sup>2</sup>Department of Neurosurgery, University of Iowa, Iowa City, IA

<sup>3</sup>Department of Bioengineering, University of Pittsburgh, Pittsburgh PA

[jasmine.hect@pitt.edu](mailto:jasmine.hect@pitt.edu)

## Abstract (max 300 words)

Voice and face processing are theorized to occur through convergent neural systems to facilitate speaker recognition. Prior neuroimaging studies suggest processing of a familiar voice engages the bilateral fusiform gyri (FG). However, it remains unknown what role the FG may play in voice processing, and at which point it becomes engaged during voice perception. The purpose of this study was to investigate neural responses to familiar voices and faces in human FG, using direct electrophysiologic recordings. In this exploratory analysis, we tested the hypothesis that neural populations in extrastriate visual cortex respond to familiar voice. Recordings were acquired from n=5 epilepsy surgery patients during a speaker identification task using visual and auditory stimuli of familiar speakers (U.S. presidents Barack Obama, George W. Bush, and Bill Clinton). Patients were presented with pictures of presidents or clips of their voices and asked to identify the portrait/speaker. Non-familiar faces were presented in a separate task as a post-hoc control, and a subset of patients (n=3) also completed a passive listening task with isolated words spoken by an unfamiliar speaker. We found that extrastriate visual cortex, including FG, exhibits responses to voice and establish the temporal dynamics of this response. These responses are relatively delayed (300-500ms) and predominately occur at sites responsive to visual stimuli of familiar faces. We found that responses to familiar voice were relatively lower in magnitude compared to familiar faces. In a subset of patients, we demonstrate these sites show responses to familiar voice, but not to words spoken by an unfamiliar talker. These findings suggest a model of heteromodal representation in extrastriate cortex, a region traditionally considered to be a part of the unimodal visual processing hierarchy. Furthermore, the task/stimulus dependence (naming of familiar voices) and latencies of voice responses in FG suggest a role in higher-order identity processing.

[Word Count: 300]