Phoneme ambiguity is reflected very early in primary auditory cortex

Laura Gwilliams¹,³, Tal Linzen¹, Kyriaki Neophytou², Lena Warnke¹, David Poeppel¹,⁴ & Alec Marantz¹,²,³
New York University Psychology¹ and Linguistics² Department; NYU Abu Dhabi³; Max-Planck-Institute⁴

Introduction
Spoken communication is a game of disambiguation. Its success relies upon the integration of incoming information to update best estimates of the message being received.

Features of a phoneme are recognised ~100 ms after onset in the posterior superior temporal gyrus, through the activation of phonetic feature detectors (Chang et al., 2010; Mesgarani et al., 2014; DiLiberto et al., 2015). However, perceptual commitment is believed to be delayed in the presence of an ambiguous phoneme, such that the percept remains malleable and influenceable by incoming information (Connine et al., 1991).

Hypotheses: Responses to ambiguous sounds
1) more activity: feature detectors simultaneously fire in response to relevant aspects of the signal
2) less activity: feature detectors respond preferentially to prototypical (unambiguous) examples of the phoneme category

Methods
Recorded natural word tokens of a native English speaker, and extracted the first syllable from the spoken words.

Created eleven-step continua of syllables between two unambiguous end-points. VOT: t-d, p-b, k-g; PoA: t-p, p-k.

1) Ambiguity Responses
- Extracted time course of activation in Heschl’s gyrus and the superior temporal gyrus bilaterally
- Coded ambiguity as the distance from each subject’s 50-50 selection point
- Ran spatio-temporal regression over regions from 0-200 ms
- Ran spatio-temporal cluster tests over the same regions from 0-200 ms
- Coded ambiguity as the distance from each subject’s 50-50 selection point
- Extracted time course of activation in Heschl’s gyrus and the superior temporal gyrus bilaterally
- Increased activity for ambiguous tokens around 50 ms after sound onset (p < .001)
- Located in Heschl’s gyrus, left lateralised
- Sensitivity to phonological ambiguity earlier than previously considered

2) Phonoeme Categorisation
- Ran spatio-temporal cluster tests over the same regions from 0-200 ms after phoneme onset
- Largest response to labial plosives (b, p)
- Earlier peak for voiceless stops
- Peaks around 100 ms after onset
- PoA patterns the same for voiced and unvoiced tokens

Discussion
- Phonetic feature detectors fire in response to matching sounds rather than prototypical phonemes
- Sensitivity to phoneme categories becomes active around 50 ms - much earlier than previously considered
- Ambiguity effect re-surface around 60 ms after a word’s point of disambiguation

References