Phoneme ambiguity is reflected very early in primary auditory cortex



Laura Gwilliams^{1,3}, Tal Linzen¹, Kyriaki Neophytou³, Lena Warnke¹, David Poeppel^{1,4} & Alec Marantz^{1,2,3} 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

- i) more activity: feature detectors simultaneously fire in response to relevant aspects of the signal
- ii) less activity: feature detectors respond preferentially to prototypical (unambiguous) examples of the phoneme category

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



- 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

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



responses were measured with magneto-encephalography (208 channel KIT system)



2) Phoneme 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



3) Syllables in Context Preliminary results (n=15): When presenting the same sounds in the context of a word, we also observe an ambiguity effect 60 ms after



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-surfaces around 60 ms after a word's point of disambiguation

References

Chang, E. F., Rieger, J. W., Johnson, K., Berger, M. S., Barbaro, N. M., & Knight, R. T. (2010). Categorical speech representation in human superio temporal gyrus. Nature neuroscience, 13(11), 1428-1432

Connine, C. M., Blasko, D. G., & Hall, M. (1991). Effects of subsequent sentence context in auditory word recognition: Temporal and linguistic constrainst. Journal of Memory and Language, 30(2), 234-250

Di Liberto, G. M., O'Sullivan, J. A., & Lalor, E. C. (2015). Low-Frequency Cortical Entrainment to Speech Reflects Phoneme-Level Processing, Curren Biology, 25(19), 2457-2465

Mesgarani, N., Cheung, C., Johnson, K., & Chang, E. F. (2014). Phonetic feature encoding in human superior temporal gyrus. Science, 343(6174) 1006-1010

contact: laura.gwilliams@nyu.edu

29th Annual CUNY Conference 2016

Supported by the NYU Abu Dhabi Institute under Grant G1001