Early Auditory Cortical Regions Discriminate Intelligible from Unintelligible Speech

Kayoko Okada, Jonathan Venezia, William Matchin, Kourosh Saberi, John T. Serences & Gregory Hickok

Department of Cognitive Sciences, University of California, Irvine

Department of Psychology, University of California, San Diego

Introduction

Previous functional neuroimaging studies have used conventional subtraction methods to identify cortical regions that are important for speech comprehension. For example, Scott, et al. (2000, 2006) contrasted clear (normal) speech with spectrally rotated speech (which destroys intelligibility while putatively preserving acoustic complexity) and found that regions along the STS were selective for intelligible speech. Early auditory areas, such as A1 and surrounding tissue, did not respond differently to clear v rotated speech indicating that these regions were coding low level acoustic features that were effectively identical across speech and complex acoustic control stimuli.

However, others have questioned the validity of standard signal amplitude-based speech v non-speech contrasts as a means to identify cortical regions critical to speech perception (Hickok & Poeppel, 2007; Okada & Hickok, 2006). Here we use multivariate pattern classification methods to re-assess the sensitivity of early auditory areas to speech v non-speech "control" stimuli.

Methods

Subjects

Twenty healthy volunteers (6 females) were scanned at UC Irvine. All volunteers were right-handed and were native English speakers.

Imaging Parameters

Philips Achieva 3T, fiitted with an 8 channel RF receiver head coil.
Functional: Fast Echo EPI, sense reduction factor=2.4, TR=2.7s, TE=25ms, size=1.95x1.95x2mm

Stimuli & Design

In a block design experiment, subjects were presented with 4 types of stimuli previously used to identify speech selective regions (Scott et. al 2000):

Intelligible Speech
- Clear speech (sentences) [Sp]
- Noise-vocoded speech [Voc]

Unintelligible Speech
- Spectrally rotated speech [Rot]
- Spectrally rotated noise-vocoded speech [Rot_Voc]

There were 4 trials of each type per session, 10 sessions total.
1 trial = 13.5 seconds.

Task: Indicate with a button press: Intelligible? Unintelligible?
Before scanning, subjects were trained to understand the noise-vocoded speech.

A1 localizer
- 1 session, amplitude modulated (8 Hz) broadband noise.
- Ten on-off cycles (13.5 s)

fMRI Results

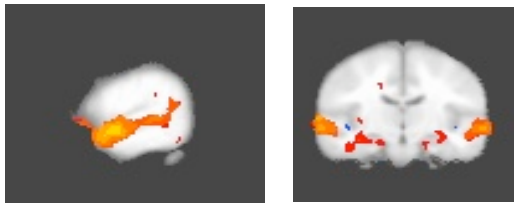
Data analysis

A). Regression analysis in AFNI.
1. Group analysis performed (n=19).
2. ROIs identified in each subject for pattern classification: A1 and STS
   - A1: identified by localizer scan.
   - STS: Anterior and posterior sites were identified using the contrast: intelligible > unintelligible speech (Sp+Voc > Rot+Rot_Voc), p<.0001.

B). Pattern classification: A1 and STS

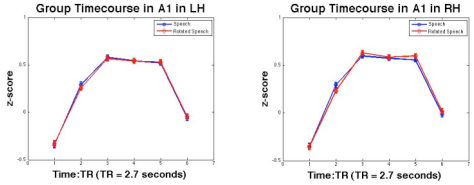
Group Analysis Results

Standard subtraction analysis replicates previous results.
Contrast intelligible v unintelligible speech (Sp+Voc > Rot+Rot_Voc), p<.001



Bilateral STS activation. No activation in Heschl's gyrus/A1

Group Timecourse graphs (n=19) in A1 demonstrating no amplitude differences for intelligibility. Here we focus on Speech v Rotated Speech conditions because rotated speech is often used as the control condition to isolate speech selective regions. (Spectrally rotated speech is unintelligible to the perceiver but is acoustically complex and preserves the overall pattern of spectrotemporal variation of speech).



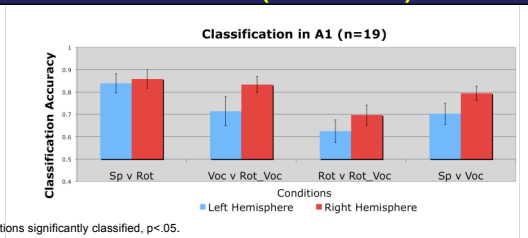
Pattern Classification Results

Classifications:

Sp v. Rot [voxels sensitive to intelligibility, acoustically similar]
Voc v Rot_Voc [voxels sensitive to intelligibility, acoustically similar]
Rot v Rot_Voc [both are unintelligible, acoustically different]
Sp v Voc [both are intelligible, acoustically different]

fMRI Results (continued)

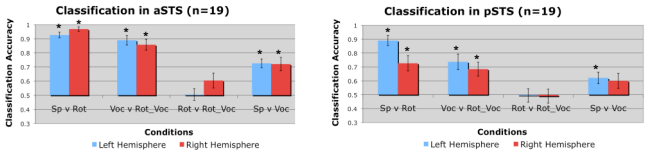
ROI: A1



* All conditions significantly classified, p<.05.

Despite observing no amplitude differences in A1 for intelligible v unintelligible speech, these voxels significantly discriminated intelligible from unintelligible speech.

Exploratory Analysis: anterior STS and posterior STS



* Indicates conditions significantly classified, p<.05.

Conclusions

Standard subtraction analysis replicated previous findings and identified speech selective sites along STS in both hemispheres. Regions in Heschl's gyrus and surrounding areas in the supratemporal plane did not show significant activation in the same contrast.

Using pattern classification analysis, we found that voxels extracted from A1 classified intelligible from unintelligible speech despite observing no BOLD amplitude differences. Voxels in A1 also significantly classified rotated speech and rotated noise vocoded speech -- stimuli which are both unintelligible and are acoustically different. This suggests that acoustic differences may be driving discriminability in early auditory regions.

In STS, as expected, intelligible speech was discriminated from unintelligible speech. In addition, in preliminary analysis, voxels in STS did not significantly classify rotated speech from rotated noise-vocoded speech. We suggest that sub-regions within STS likely demonstrate sensitivity to different types of speech and non-speech acoustic signals and may be coding more abstract speech representations.

Contact: K. Okada okadak@uci.edu or G. Hickok greg.hickok@uci.edu
This research was supported by NIH DC03681 to Gregory Hickok.