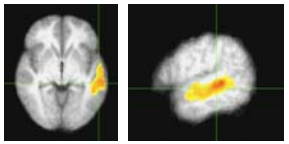
Changes in functional organization following naturalistic training on complex auditory categories

R. Leech[1], L. Holt[2], J. Devlin[1] & F. Dick[1,3]

1) Birkbeck/UCL Centre for NeuroImaging, London, UK; 2) Carnegie Mellon University, USA; 4) CRL, UC-San Diego, USA.

Introduction

- How do we acquire complex auditory categories such as phonetic categories?
- What accounts for 'speech-selective' brain activation?



Speech > environmental sounds (From speech localiser - see Methods).

For example, brain shows greater activation for spoken words over and above short, spectrotemporally complex, meaningful non-linguistic environmental sounds (e.g., a dog barking), in the left STS (Price et al, 2005; Dick et al., 2007).

Is this region particularly adapted to processing the acoustical or informational characteristics of speech sounds per se?

Or alternatively, might at least some of the speech-sensitivity of the left-STS be a consequence of higher-order auditory categorization processes?

The current study

We asked whether expertise with artificial complex auditory stimuli (via extensive training) can elicit a shift to more speech-like patterns of neural activation.

We also investigated if there are other training-related shifts in brain activation that resemble those reported for training studies of novel speech exemplars (e.g., Golestani and Zatorre, 2004).

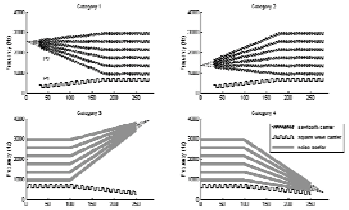
Thirdly we investigated whether relative activation in speech-selective regions distinguishes between good and bad learners of auditory categories.

Here, subjects were scanned before and after incidental training on novel, artificial spectrotemporally complex sounds using a computer game (Wade & Holt, 2005).

Methods

Stimuli

Four categories of complex sounds with spectral peaks at two locations - P1 and P2 - corresponding to formant resonances of the human vocal tract. P1 and P2 had either onset or offset gradients. To discriminate between categories subjects needed to simultaneously pay attention to multiple cues (i.e., solve a non-linear mapping).



Experimental Design

Participants

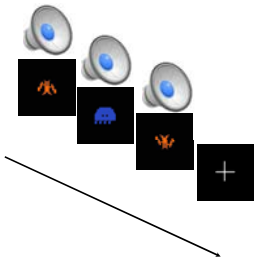
17 young healthy English speakers with normal hearing

Pre-training scan ==>> computer game (5 hrs) ==>> Post-training scan

- Visual oddball detection task in the scanner (oddball = upside-down alien).
- In the training computer game subjects killed or captured aliens; at harder levels of the game subjects needed to use the auditory categories to progress to the next level.
- In the post-scan we also ran a speech localiser (speech > environmental sounds) using the same functional scanning protocol as the artificial sounds.
- Following post-scan, outside the scanner each subject performed a short behavioural test to match each sound with each category.

fMRI scanning:

- *1.5 T Siemens Avanto MRI scanner w/12-element head coil*
- *27 Oblique axial slices, 3.5mm^3*
- *EPI image sequence: TR=3400 ms; TE=50ms, FOV=224 mm; Matrix 64*64 Two blocks of 180 TRs.*
- *T1-weighted MPRAGE (1mm^3) in first session*
- *Rapid sparse-sampling design.*



fMRI analysis:

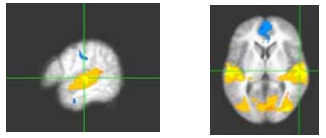
All analyses were performed using FSL version 4.0. *1st level analysis:* indivudal EPIs motion-corrected, high-pass filtered, pre-whitened and smoothed (6mm) w/ motion parameters added in as regressors. The design matrix was convolved with a double gamma function with temporal derivatives and was temporally filtered. The first-level contrast was **alien sounds** vs **rest**.

Higher-level analyses:

- Whole group pre- vs post-training paired t-test
- Post- minus pre-training activation regressed against individual subjects' perfrormance on behavioural post-test (i.e,. ability to match aliens to sounds)
- Post- minus pre-training activation masked by an ROI defined by the speech localiser (speech > environmental sounds)

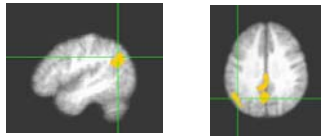
Results

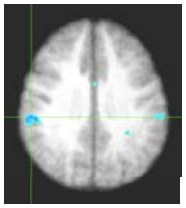
Alien sounds vs rest
Whole group, corrected for multiple comparisons at clusterwise p <0.05

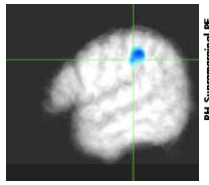


Whole group, post- > pre-training
(p <0.05 clusterwise)

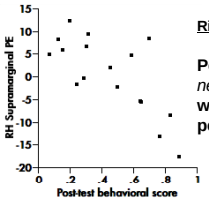
Increases in activation in bilateral precuneus and right supramarginal gyrus.





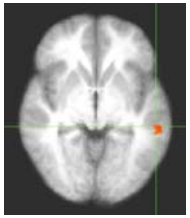


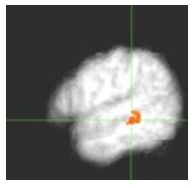
MNI, 59 -32 34



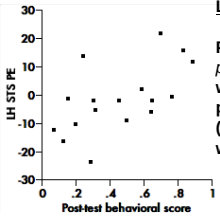
Right Supramarginal Gyrus

Post- > pre-training *negatively* correlated with sound-category performance score





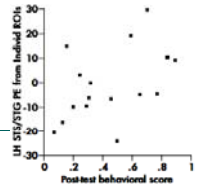
MNI, -58 -34 -7



Left mid STS

Post- > pre-training *positively* correlated with sound-category performance score (p<0.05 clusterwise within speech mask

The performance-related increase in left 'speech-selective' regions was also confirmed by ROI analyses in native space, using each individual subject's Speech-Env functional localizer (p <0.05)



References

Golestani, N. & Zatorre, R. J. 2004. Learning new sounds of speech: reallocation of neural substrates. Neuroimage, 21, 494–506.
Gauthier, I. et al. (1999) Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects *Nat. Neurosci.* 2, 568–573
Narain C, et al, 2003 Defining a left-lateralised response specific to intelligible speech using fMRI. *Cereb Cortex.*13:1362-1368.
Price, C., et al, 2005. Speech-specific auditory processing: where is it? *Trends Cogn. Sci.* 9, 271–276.
Wade, T., and Holt, L. L. 2005. Incidental categorization of spectrally complex non-invariant auditory stimuli in a computer game task, *J. Acoust. Soc. Am.* 118, 2618–2633.

Contact: r.leech@bbk.ac.uk

Supported by: MRC NIA G0400341