Conjunto de datos del experimento n^{o}2-2001-111TW

Detection of Sounds in the Auditory Stream: Event-related fMRI Evidence for Differential Activation to Speech and Non-speech

Elsa Fernández

April 16, 2008

NOTAS GENERALES

- Sujetos de estudio: 15 adultos diestros voluntarios (6 mujeres, 9 hombres, media edad: 25.5 aos).
- No están disponibles los datos preprocesados por los scaners 3D SPGR y T1 Localizer.
- Los datos funcionales medios han sido proporcionados para cada ejecucin funcional. Las imágenes medias están disponibles para ambos tipos de datos, raw y preprocesados.

ESTIMULOS

- Sounds were presented through insert earphones embedded within 30dB sound attenuating MR compatible headphones using custom presentation software (http://nilab.psychiatry.ubc.ca/vapp).
- Because of the difficulty in accurately measuring absolute intensity values at the exit point from insert earphones, all sounds, speech, complex non-speech and tones, were equated for intensity relative to each other. Sounds were clearly audible above the noise of the scanner, as evidenced by listeners' near-perfect performance.
- The stimuli were of four different types: speech, complex non-speech, high tones, and low tones.
- Speech stimuli consisted of six tokens of a monosyllabic nonsense word ("lif") spoken by a native female English speaker.

- Tokens varied in intonational contour (average minimum and maximum pitch: 202 Hz and 350 Hz respectively), and in duration (525 - 711 ms).
- Complex non-speech stimuli consisted of time varying sine-wave analogues of the speech tokens in which all regions of significant energy were tracked (namely the fundamental frequency and the first three formants; see Figure 1, original article). Sinusoidal waves tracking these energy peaks were created individually in Mathcad 3.1 (Mathsoft Inc., Cambridge, MA).
- Fundamental frequency (corresponding to pitch) was tracked individually for each of the six speech tokens.
- Because the first three formants were virtually identical across the multiple natural repetitions, one set of formants from a representative word token was tracked. This set was composed of the first formant of the initial consonant segment ("l"), and the first three formants of the vocalic segment ("i").
- The sine analogue to ("f") was created using a white noise generator and filtered. This representative set was then added onto the sine wave analogue of the pitch contour of each segment using Signalyze 3.12 (Agora Language Marketplace, Charlestown, MA) to create six different stimuli. Analogues thus retained the duration, pitch contour, amplitude envelope, relative formant amplitude, and relative intensity of their speech counterparts (see Figure 1).
- Low tones were six pure sinusoidal waves of 500 Hz generated using Sound Edit Pro, version 2 (Macromedia Inc, San Francisco, CA) matched in duration to the speech stimuli.
- High tones were six pure sinusoidal waves of 1500 Hz generated using Sound Edit Pro, version 2 (Macromedia Inc, San Francisco, CA) matched in duration to the speech stimuli.

3T (TONE RUNS)

- Participants heard two stimulus runs, an experimental run and a tone run. The background non-target stimulus in both runs was a 1000 Hz tone, occurring with a probability of 0.8.
- The stimuli of interest were presented in a pseudorandom oddball design (separated by 3-5 non-target stimuli).
- In the experimental run, the infrequent sounds consisted of speech (0.1) and complex non-speech (0.1).
- In the tone run, 1500 Hz high tones (0.1) and 500 Hz low tones (0.1) were the infrequent sounds. Each run was 12.5 min, with a 2 s stimulus onset asynchrony (SOA) for a total of 380 total stimuli per run.
- Order of presentation of the stimulus runs was counterbalanced across participants. Participants made a motor response on an MRI-compatible fiber-optic response device (Lightwave Medical, Vancouver, B.C.) using

their left index finger for every infrequent sound they heard. Reaction times were monitored on-line.

WS (Word/Nonword Sine Runs)

- Participants heard two stimulus runs, an experimental run and a tone run. The background non-target stimulus in both runs was a 1000 Hz tone, occurring with a probability of 0.8.
- The stimuli of interest were presented in a pseudorandom oddball design (separated by 3-5 non-target stimuli).
- In the experimental run, the infrequent sounds consisted of speech (0.1) and complex non-speech (0.1).
- In the tone run, 1500 Hz high tones (0.1) and 500 Hz low tones (0.1) were the infrequent sounds. Each run was 12.5 min, with a 2 s stimulus onset asynchrony (SOA) for a total of 380 total stimuli per run.
- Order of presentation of the stimulus runs was counterbalanced across participants. Participants made a motor response on an MRI-compatible fiber-optic response device (Lightwave Medical, Vancouver, B.C.) using their left index finger for every infrequent sound they heard. Reaction times were monitored on-line.

IMAGE PROCESSING

Functional images were reconstructed off-line. Statistical parametric mapping software (SPM99, Wellcome Department of Cognitive Neurology, London, UK) was used for image realignment and normalization into modified Talairach stereotaxic anatomical space (using affine and nonlinear components, as implemented in SPM99). Images were smoothed using a Gaussian kernel (8 mm FWHM) to compensate for intersubject anatomical differences, and to optimize the signal-to-noise ratio. Event-related responses to the stimuli of interest were modeled using a synthetic hemodynamic response composed of two gamma functions and their temporal derivatives (for a discussion of the relative advantages and disadvantages of this modeling method, see Kiehl, Laurens, Duty, Forster, & Liddle, 2001). The peak of the response was modeled at 6 sec poststimulus time, consistent with the results of other event-related fMRI studies (Hickok et al., 1997 ; but see Belin et al., 1999 , for a shorter peaking time in a different acoustic setting). A high-pass filter (cutoff period 89 sec) was incorporated into the model to remove noise associated with low frequency confounds. A low-pass filter (at the Nyquist frequency, with a period of 6 sec) was also applied to remove noise associated with alternations of the applied radio frequency field. Three contrasts were used to create SPMt maps, later transformed into SPMZ maps, for three comparisons of interest:

(a) activation for speech sounds relative to the complex nonspeech stimuli,

(b) activation for the speech sounds relative to simple tones, and

(c) activation for the complex nonspeech sounds relative to the simple tones.

STATISTICAL ANALYSES

Statistical analyses were performed in SPM99 using a fixed-effects model. Because multiple voxels were examined, a correction for multiple comparisons based on the theory of Gaussian fields was employed. The areas of activation reported are significant at the voxel level, with z scores greater than 4.63 corresponding to a corrected significance level of p .05. We further explored hemispheric differences in activation by comparing suprathreshold voxels in each hemisphere for individual listeners. Within the SPM program, we imposed a mask of the MTG on each listener's SPMt map for the main comparison of interest (see (a) above). A custom script was used to extract suprathreshold voxels ($z = 2.63$, p .05 uncorrected) in the left and right hemispheres of every listener. A paired t test was conducted on these hemispheric voxel counts to obtain an index of hemispheric asymmetry.