Supplementary Methods

For both experiments, all participants gave informed consent according to procedures approved by the Ethical Committee of the Geneva University Hospital.

Functional images were acquired on a 1.5T Philips system with a GRE EPI sequence using BOLD (Blood Oxygenation Level Dependency) contrast. Image volumes consisted of 32 contiguous 3.4 mm axial slices (TE/Flip = 40ms/90°, FOV = 220mm, matrix = 128x128) parallel to the inferior occipito–temporal plane, and were acquired using a sparse imaging procedure with a repetition time (TR) of 5 sec including an effective acquisition time (TA) of 2.5 sec interleaved with a silent gap of 2.5 sec. Auditory stimuli (750ms duration) were delivered binaurally (different on right and left side in Experiment 1, identical on both sides in Experiment 2) via MRI–compatible headphones, and were always presented with a varying jitter during the silent gap between each EPI volume (24 different events for each experimental conditions, plus 48 null events without auditory stimulation as a baseline condition, for each experiment).

Functional scans were realigned, corrected for slice timing, normalized to the MNI template (resampled voxel size: 3x3x3 mm³), and spatially smoothed (8mm isotropic FWHM Gaussian kernel) using standard procedures in SPM2 (www.fil.ion.ucl.ac.uk). Each event type was modelled as a separate regressor convolved with a canonical hemodynamic response function, and movement parameters from realignment corrections were entered as additional covariates of no interest. Statistical parametric maps were generated from linear contrasts between conditions in each participant. A second–stage random–effect analysis (RFX) was then performed using one–sample t–tests on contrast images obtained in each subject for each comparison of interest (df = 13). We report activations of at least 10 voxels that survived a threshold of $P < 0.001$ at the voxel level, corrected at $P < 0.05$ for the brain volume showing a significant main effect of neutral speech sounds vs silent baseline (at $P < 0.001$ uncorrected), unless mentioned otherwise for specific regions of interest. Additional standard ANOVAs were performed on parameter estimates of activity extracted from SPM peaks in both hemispheres using SPSS.