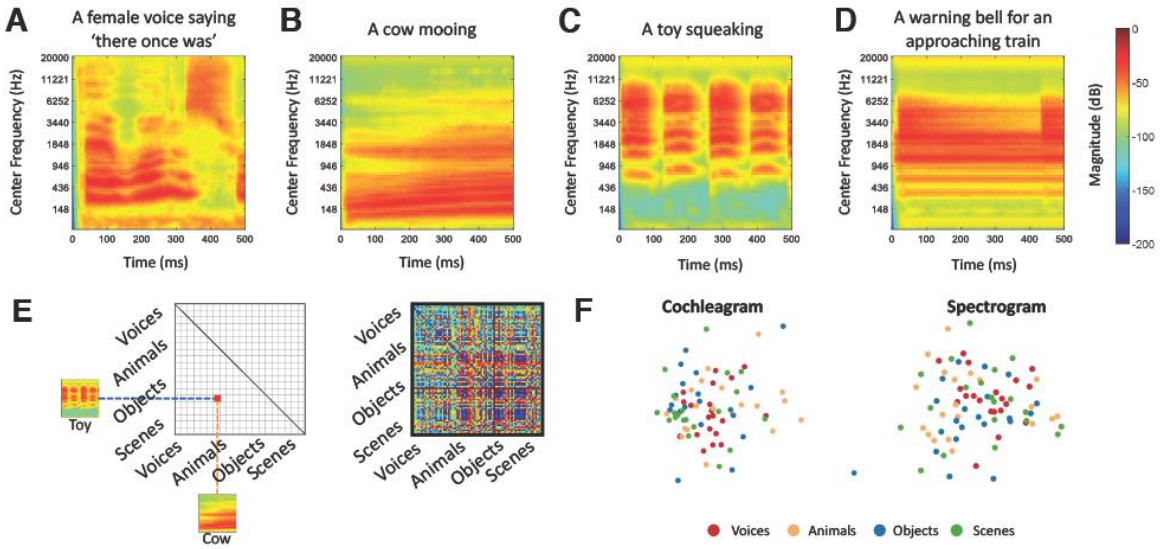
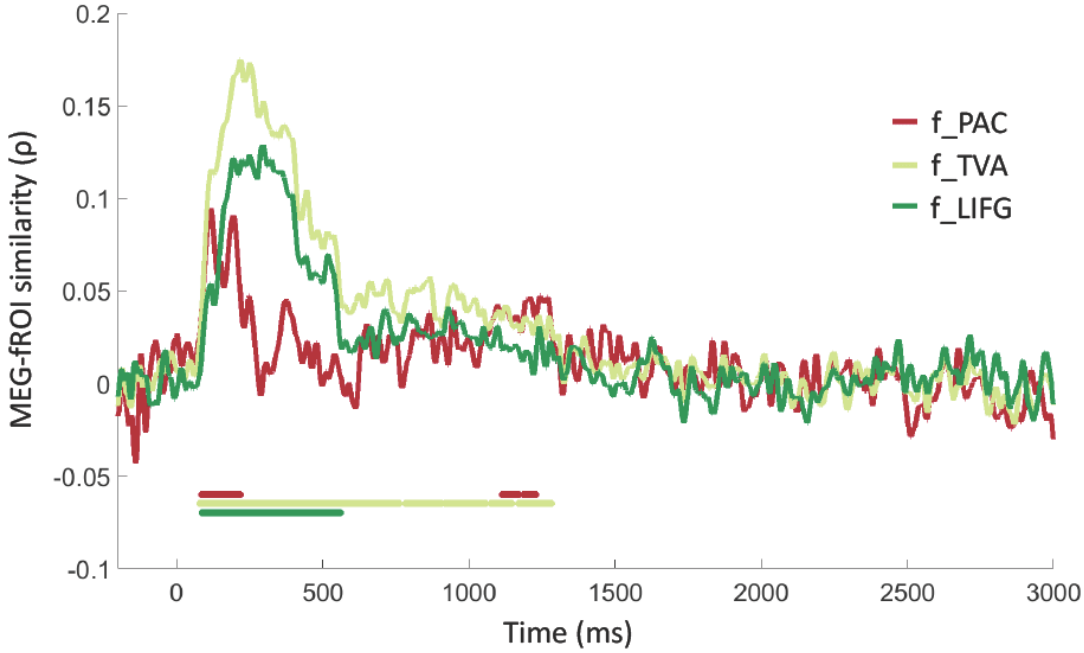
Cochlea to categories: The spatiotemporal dynamics of semantic auditory representations

Supplemental Material



Supplemental Figure S1. Example stimulus cochleagrams from each category: voice (A), animal (B), object (C), scene (D). Spectral similarity (euclidean distance) was used to compare all pairs of stimuli, shown schematically in **E** for a sample pair of stimuli (left), and for the entire 80-item stimulus set (right). The full matrix was used as the Cochleagram model. **F.** Multidimensional Scaling (MDS) visualization of similarity structures for cochleagrams (left) and spectrograms (right), with each dot representing one stimulus exemplar, color-coded by category. To quantify this structure, average between-category distances were compared against between-category distances when category labels were randomly shuffled and permuted 10,000 times. Against a Bonferroni-corrected threshold of p = 0.008, no single category pair differed significantly from randomly permuted category pairs, indicating that categorical structure in brain responses was unlikely to arise from acoustic differences alone.



Supplemental Figure S2. ROI-based fusion of MEG-fMRI in functional auditory ROIs (fROIs). Three regions of interest were functionally defined, primary auditory cortex (f_PAC), temporal voice area (f_TVA), and left inferior frontal gyrus (f_LIFG). Solid horizontal lines represent significant time points observed in MEG-fROI correspondence, indexed by Spearman's ρ and color-coded by region. All statistics, P<0.01, C<0.05, 1000 permutations.