METHODS

Patients, electrode locations, and data acquisition. For detailed methods, see ref. 28. In short, extracellular single unit recordings were obtained from four patients with pharmacologically intractable epilepsy, implanted with intracranial electrodes to identify seizure focus for potential surgical treatment. Electrode location was based solely on clinical criteria. All patients had electrodes placed bilaterally in Heschl's gyri. Talairach coordinates for each patient were as follows. Patient 1: [right 45.5, posterior 19.3, superior 10.2]; patient 2: [right 38.8, posterior 12.4, superior 12.3], [left 41.2, posterior 18.2, superior 10.4]; patient 3: channels 1:8 were from right middle superior temporal [42.47 right, 1.01 anterior, 10.54 superior], channels 33:40 were from right posterior superior temporal [39.18 right, 23.14 posterior, 16.68 superior], channels 57:64 were from left superior temporal [48.53 left, 7.03 posterior, 14.01 superior]; patient 4: [right 40.95, posterior 0.83, inferior 0.35]. Each electrode terminated in a set of nine 40 µm platinum-iridium microwires. Signals from these microwires were recorded at 14 kHz for patient 1 and 28 kHz for patients 2, 3 and 4. Raw signal was bandpass filtered between 1 and 9 kHz and recorded using a 64-channel data acquisition system. To verify electrode position computer tomographic scans following electrode implantation were co-registered to the preoperative MRI using Vitrea@ (Vital Images). Patients provided written informed consent to participation in the experiment. The study conformed to the guidelines of the Medical Institutional Review Board at UCLA.

Experimental protocol. In each experimental session, patients 1 to 3 were presented twice with the same audiovisual movie clip, consisting of 8:40 min of an unedited audio-visual segment of "The Good, The Bad, and The Ugly" (starting from minute 38:25 in the original film). Both clip presentations were shown in succession, with a 5–10-minute rest period in between. The patients' task was to follow the plot. Patients 2 and 3 were also presented with a sequence of low-resolution random-chord stimuli lasting 3.5 min accompanied by random visual textures displayed at a rate of 4 Hz. The random-chord stimuli used here were similar to those used in previous studies[25,26] (see Methods Summary). Data was acquired in ten sessions (one session with patient 1, three sessions with patient 2, four sessions with patient 3, and one session with patient 4).

Spike detection and cell selection. For patient 1, the signal from each microwire was high-pass filtered at 300 Hz and at multiples of 60 Hz. Potential spikes were detected by thresholding the filtered signal. Potential spikes were then sorted by template matching the first two principal components. For patients 2, 3 and 4, the raw data was bandpass-filtered between 300 and 3,000 Hz and sorted as in ref. 21.

ROC data analysis. For each unit the empirical spike count distributions elicited by different frequencies were calculated in a relevant response window, starting at chord onset and ending between 50 to 200 ms after chord onset. The temporal window was selected so that it matched the period of elevated firing rate in the PSTH of a unit's response to its best frequency. The ROC curve was computed for the best frequency and each of the other frequencies. The ROC is the plot of the hit rate against false alarm rate for a one-trial, two-alternative classification task, and is controlled by a varying classification threshold. The spike count elicited in a single trial was assigned to the best frequency if the likelihood of that count (the ratio of the probabilities to observe the spike count given the best frequency and the comparison frequency) was above the classification threshold. For each threshold, the hit rate was the fraction of best-frequency trials correctly assigned to the best frequency, and the false alarm rate was the fraction of trials of the comparison frequency incorrectly assigned to the best frequency. The area under the ROC is an estimate of the probability of correct discrimination between the two frequencies in a two-alternative, two-interval forced-choice experiment.

The discrimination threshold was set at 70.7%, corresponding to the threshold tracked by the standard 2-down 1-up adaptive procedure typically used in auditory psychophysics[9]. 'Intermediate' distributions constructed by a weighted sum of two empirical distributions were used as estimates of the spike count distributions at intermediate frequencies that were not sampled in the experiment. Thus, if P1 is the observed distribution of counts in response to frequency f_1 and P2 corresponds to f_2, the interpolated distribution $\lambda P1 + (1-\lambda)P2$ with $0 < \lambda < 1$ was taken to represent the distribution of responses of the unit to a frequency f such that $\log(f) = \lambda \log f_1 + (1-\lambda)\log f_2$. Thresholds were estimated as the smallest frequency interval that could be discriminated using these artificially constructed distributions, and expressed as the interval divided by the geometrical mean of the two frequencies.

Natural STRFs calculation. The linear response function for each unit in response to the soundtrack was computed using the software package STRFpak[27]. The spectro-temporal respresentation of the stimuli that served as input was generated by an auditory nerve simulation implemented by the AIM software package[29].

Predictions. Predictions were generated for stimuli not used in STRF estimation. For each minute of the soundtrack, responses were predicted using both a natural STRF calculated from the responses of the unit to the rest of the soundtrack and an artificial STRF calculated from spike-triggered averaging of the same unit responses to all minutes of random-chord stimuli. For each minute of random-chord stimuli, predictions were calculated using both an artificial STRF calculated from the remaining random-chord responses, and a natural STRF based on the responses of the same unit to all minutes of the soundtrack. The similarity between the predicted response given by the STRF and the actual response, both smoothed with a 121 ms hamming window, was quantified by the correlation coefficient between them, as in refs 10 and 27.

Mean tuning curve. The mean tuning curve was calculated for the excitatory units recorded with the high-resolution random-chord stimulus ($N = 14$). The responses of each unit (mean spike count in a 50 ms bin) to frequencies around the unit best frequency were normalized and aligned on best frequency, and then averaged across all units.