Response reproducibility

While the soundtrack was presented twice in each recording session, the random chord stimuli were presented only once, and specific chord combinations were typically not repeated. Thus, a straightforward comparison of reproducibility of the responses was not possible. Moreover, during the presentation of random chords, many units had elevated firing rate when their BF was present in the chord, regardless of the other components simultaneously present. Thus, the responses to the random chords effectively consisted of two classes – high firing rate when the BF was present/background firing rate when the BF was not present in the chord. We generated a similar class of preferred events evoking high firing rate in response to the movie by identifying all the 100 ms long sound segments that elicited an elevated response (>92% of maximal response in spike count) in the first run of the movie. We used the coefficient of variation (cv) as a measure of the reproducibility of the neuron's response to its preferred stimuli. We calculated the cv of the responses to chords that included the BF and compared it with the cv of the responses to the preferred events in the second run of the movie. The cv of the two ensembles were similar, indicating that reproducibility of responses was comparable between the two contexts.

Evaluating the predictive power of the STRFs

The predictive power of the STRFs was evaluated when applied to both random chords and the soundtrack. Since for some units (47/63) only responses to the soundtrack were available, we also generated synthetic artificial STRFs for units whose natural STRFs had a clear dominant excitatory region (31/63 units). The synthetic STRF consisted of a 2-D Gaussian filter centered at best frequency (12 ms temporal std and 1/6 octave spectral std, see Fig. S1). Even though these STRFs were tailored to match the responses to the soundtrack, the predictive power of the synthetic STRFs was again significantly lower than that of the natural STRFs (0.21 ± 0.18 and 0.28 ± 0.16 respectively, 3-way ANOVA on STRF type x predicted segment x neuron, main effect of STRF type, $F(1,586)=26$, $p<<0.01$). Using different parameters, including optimizing the spectral and temporal widths for each neuron separately, did not change the conclusions of this analysis.

In a complementary analysis we compared the predictions of the responses to a minute of random-chord stimulus using natural and artificial STRFs. The artificial STRFs were estimated after omitting the responses to the minute being predicted. Predictive power was again better within-context: 0.42±0.18 when using artificial STRFs compared to 0.25±0.15 when using natural STRFs (3-way ANOVA on STRF type x predicted segment x neuron, main effect of STRF type, $F(1,107)=39$ $p<<0.01$). Since the random chord stimulus was presented only once, we didn't have an independent estimate of the maximum achievable correlation. We conclude that the richer structure of the natural STRFs is at least in part a reflection of additional processing mechanisms that are not engaged by artificial stimuli but do shape the responses to the natural stimuli.

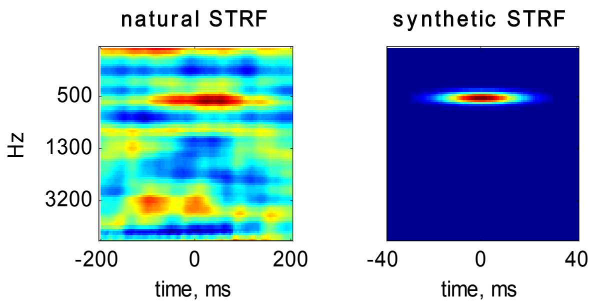

Figure S1 | Natural and synthetic STRFs In cases where units were presented with the soundtrack, but not the random chord stimuli, we constructed synthetic artificial STRFs. The synthetic STRFs consisted of a 2-D Gaussian filter (12 ms temporal std and 1/6 octave spectral std) centered at the excitatory region of the natural STRF.