LETTERS

Ultra-fine frequency tuning revealed in single neurons of human auditory cortex

Y. Bitterman[1,2], R. Mukamel[3,4], R. Malach[5], I. Fried[4,6] & I. Nelken[1,2]

Just-noticeable differences of physical parameters are often limited by the resolution of the peripheral sensory apparatus. Thus, two-point discrimination in vision is limited by the size of individual photoreceptors. Frequency selectivity is a basic property of neurons in the mammalian auditory pathway[1,2]. However, just-noticeable differences of frequency are substantially smaller than the bandwidth of the peripheral sensors[3]. Here we report that frequency tuning in single neurons recorded from human auditory cortex in response to random-chord stimuli is far narrower than that typically described in any other mammalian species (besides bats), and substantially exceeds that attributed to the human auditory periphery. Interestingly, simple spectral filter models failed to predict the neuronal responses to natural stimuli, including speech and music. Thus, natural sounds engage additional processing mechanisms beyond the exquisite frequency tuning probed by the random-chord stimuli.

Sounds are decomposed to different frequency bands by the auditory periphery. Tonotopic ('by frequency') organization is kept throughout the auditory pathway, at least up to and including primary auditory cortex. In vision and somatosensation, the resolution of the peripheral sensors to a large degree determines overall behavioural discrimination capabilities. However, in the auditory system, frequency just-noticeable differences in well-trained subjects may be 30 times smaller than the presumed bandwidth of the peripheral filters ('critical bands', typically about a sixth of an octave in humans, as measured in psychoacoustical tests). Electrophysiological correlates of critical bands have been suggested[4–6], and frequency just-noticeable differences can be derived by integrating information over a large population of neurons[7], but there are currently no reports of a significant population of single neurons the bandwidth of which corresponds to the behavioural just-noticeable differences. Does the high-frequency resolution expressed behaviourally have explicit neural representation? If so, can high-frequency resolution explain the response patterns to complex sounds?

Responses of neurons in human auditory cortex were recorded from four patients with intractable epilepsy monitored with intracranial depth electrodes to identify seizure foci for potential surgical treatment[8]. Using clinical criteria, electrodes were implanted bilaterally in the transverse gyri of Heschl, loci of the auditory cortex (see Methods). Patients were presented with artificial random-chord stimuli at a resolution of six tones per octave (two patients) or 18 tones per octave (one patient), and with segments from the popular English-speaking western film ''The Good, the Bad and the Ugly'' (three patients, see Methods). Thus, for many neurons, the stimulus ensemble included both artificial stimuli and more structured stimuli. The artificial stimuli were designed to sample evenly the spectral range of the movie soundtrack. Results are based on 95 units recorded in four patients.

Figure 1 displays raster responses of one unit to the different frequencies in the six-tones-per-octave random-chord stimulus. Each frequency appeared simultaneously with two other frequencies selected essentially randomly. Only one of the 41 possible frequencies elicited excitatory responses in this unit. Furthermore, when a tone burst of that frequency appeared in the stimulus, a sustained response outlasting tone duration was elicited with high reliability. The lack of excitatory response to the two adjacent frequencies implies that this unit was more selective than the frequency resolution of the stimulus (six tones per octave).

Of 31 units from the two patients presented with the six-tones-per-octave random-chord stimulus, 27 had a narrow, well-circumscribed frequency response area. About half (14/31) showed reliable responses to tone bursts at a single frequency, with no consistent excitatory response to any other frequency. Thirteen units responded to two to three adjacent frequencies. The rest (4/31) exhibited more complex responses. The resolution of six tones per octave was thus too coarse directly to measure the spectral bandwidth of most units.

A high-resolution random-chord stimulus with 18 tones per octave was presented to a third patient. Of 16 units recorded in this patient, 14 exhibited a highly elevated firing rate in response to a single frequency, with additional weaker, although significant, responses to only one or two adjacent frequencies. The average bandwidth of these units can be conservatively estimated at about a twelfth of an octave, in agreement with the results presented above (Fig. 2a). Figure 2b displays typical spectro-temporal receptive fields (called 'artificial STRFs' below) derived from responses to the random-chord stimuli by spike-triggered averaging. The best frequencies, defined as the frequency that elicited maximal response, ranged from 250 to 2 kHz in this population (Fig. 2c). It is generally accepted that the frequency tuning curve of the auditory periphery in humans has a width of about a sixth of an octave[3]. Therefore, when presented with random chords, the great majority of auditory cortical neurons showed substantially better frequency selectivity than the auditory nerve.

The frequency discrimination performance based on responses in single trials was estimated using receiver operating characteristic (ROC) analysis. We compared the empirical spike count distributions elicited by the different frequencies and determined the lowest discrimination threshold for each of the 47 units tested with the random-chord stimuli. Performance was quantified by the probability of correct decision in a two-interval, two-alternative forced choice test. Discrimination threshold was set at 70.7%, as typically done in auditory psychophysics. In more than 60% of the excitatory cells (25/42) discrimination was above threshold for the smallest possible frequency difference tested, the spectral resolution of the stimulus (20/27 units tested with six tones per octave and 5/15 units tested with 18 tones per octave; see for example Fig. 3).

[1]Department of Neurobiology, Life Science Institute, [2]Interdisciplinary Center for Neural Computation, Hebrew University, Jerusalem 91904, Israel. [3]Ahmanson-Lovelace Brain Mapping Center, David Geffen School of Medicine, [4]Division of Neurosurgery, David Geffen School of Medicine and Semel Institute for Neuroscience and Human Behaviour, University of California Los Angeles (UCLA), Los Angeles, California 90095, USA. [5]Department of Neurobiology, Weizmann Institute of Science, Rehovot 76100, Israel. [6]Functional Neurosurgery Unit, Tel Aviv Sourasky Medical Center and Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel.

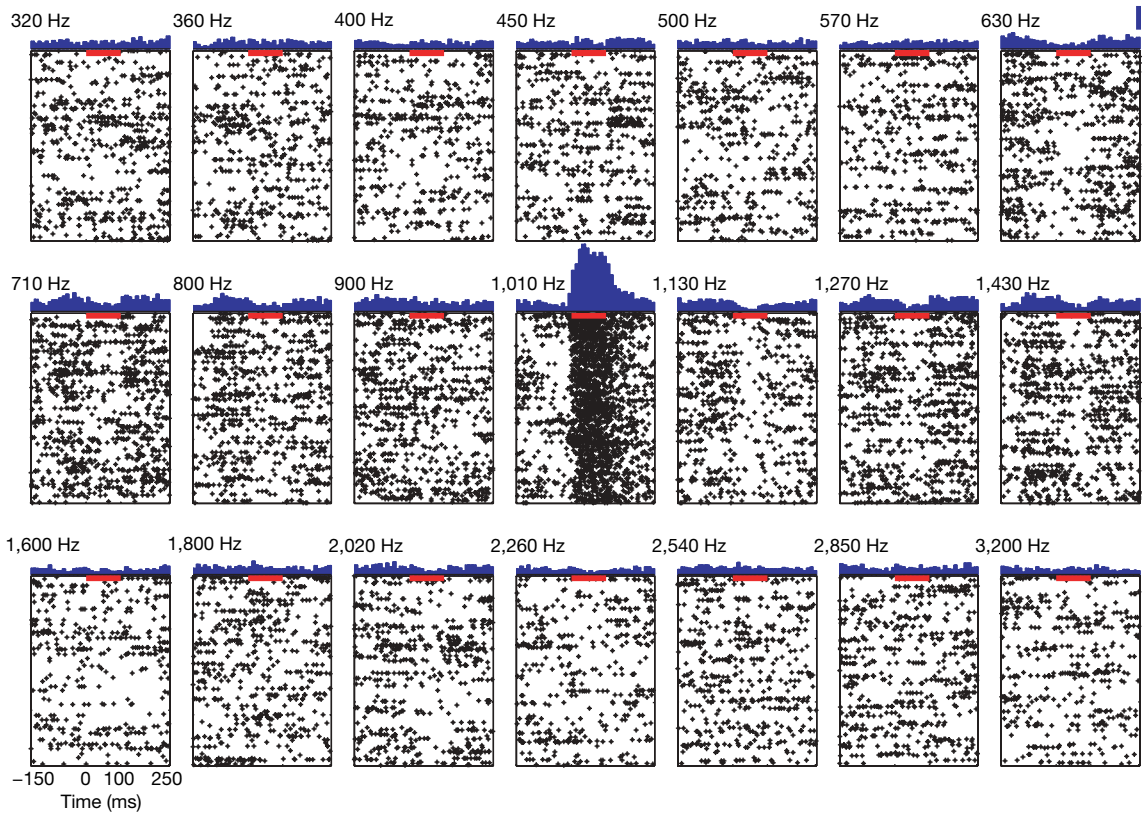

Figure 1 | Response selectivity. Raster plots of responses of one unit to chords containing the frequency specified above each panel (270 repetitions in each panel) and peristimulus time histograms (PSTH, blue; bin width 10 ms) based on these raster plots (the scale line at the top right PSTH corresponds to a firing rate of 16 spikes per second; maximum firing rate at preferred frequency: 47 spikes per second). Red bars mark 100 ms (duration of one chord) from the beginning of the response to the preferred frequency. The frequency table contained 20 additional frequencies (below 320 Hz and above 3,200 Hz); no other frequency elicited significant responses.

For these units, we linearly interpolated spike count distributions to simulate possible distributions at intermediate frequencies that were not actually tested (see Methods). Thresholds were again estimated by the smallest frequency interval that could be discriminated using these intermediate distributions. These thresholds are underestimates because maximum slopes of frequency response curves are bounded from below by linear interpolation. Even so, this procedure revealed units that had discrimination thresholds that matched and even exceeded the behavioural performance of naive human subjects[9] (Fig. 3e).

Do units also respond as narrow spectral filters when presented with natural sounds? We analysed responses elicited by nine-minute clips from the soundtrack of the feature film "The Good, the Bad and the Ugly", shown twice in each recording session. The soundtrack contained approximately equal-duration segments of dialogue, music and background noise. The average firing rate was not significantly different between responses to the random-chord stimuli and responses to the soundtrack (paired t-test, $t = 1.04$, degrees of freedom d.f. = 13, not significant), suggesting the soundtrack was, on average, as successful as random chords in driving neuronal responses, with comparable reproducibility (see Supplementary Information).

We estimated STRFs from responses to the soundtrack (called 'natural STRFs' below) using generalized reverse correlation techniques following ref. 10. The exquisite spectral filtering clearly apparent in the artificial STRFs was partially lost—natural STRFs were noisier and appeared to have richer structure (Fig. 4a). Nevertheless, there were similarities between natural and artificial STRFs estimated for the same unit. For the units recorded with both stimuli, the best frequency of the artificial STRF and the best frequency of the natural STRFs were highly correlated ($r = 0.7$,

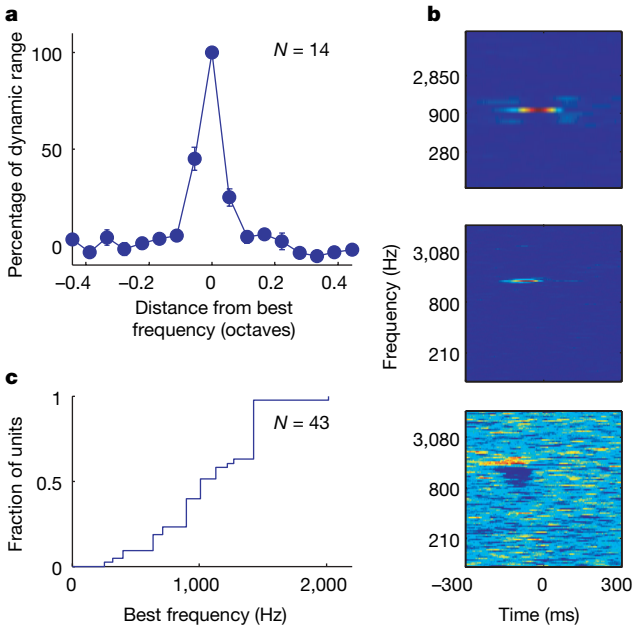

Figure 2 | Frequency tuning in the responses to the random-chord stimulus. **a**, Mean tuning curve (see Methods). Error bars indicate s.e.m. **b**, STRFs of three units estimated from the responses to the random-chord stimulus. The top panel shows a unit tested with six-tones-per-octave resolution that responded to a single frequency (colour scale saturation: 2.5–39 spikes per second). The middle panel shows a unit tested with 18-tones-per-octave resolution that responded predominantly to a single frequency (colour scale saturation: 1–32 spikes per second). The bottom panel shows a unit with complex tuning (colour scale saturation: 0–3.4 spikes per second). **c**, Cumulative distribution of the best frequencies of 43 units with a clear excitatory peak.

d.f. = 16, $P \ll 0.01$; Fig. 4b). This agrees with the general finding that the best frequency is largely independent of auditory context[11].

The first- and second-order statistics characterizing the soundtrack were fully sampled by the random-chord stimuli (verified by comparing the joint distribution of spectral and temporal modulations in the two stimulus ensembles) and the calculation of the natural STRFs corrected for second-order correlations in the stimulus[10]. Thus, if neurons linearly integrate their spectro-temporal input, natural and artificial STRFs should be essentially equivalent. However, the soundtrack also contained higher-order spectral correlations the effects of which on the STRFs could become apparent if the neurons had significant nonlinearities. These effects could be the reason for the additional structure in the natural STRFs.

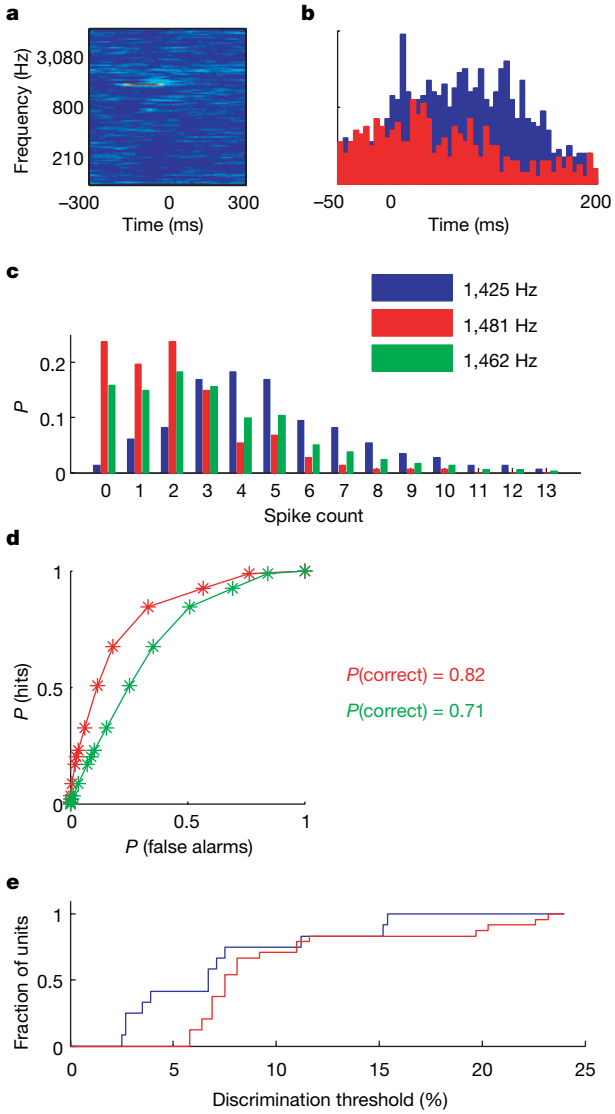

Figure 3 | Frequency discrimination based on single-trial responses. **a**, STRF of an excitatory unit estimated from the responses to the high-resolution random-chord stimulus (colour scale saturation: 6–36 spikes per second). **b**, PSTHs (bin width: 5 ms); blue is the response to the best frequency and red is the response to the adjacent frequency. The ordinate represents firing rate, scale: 0–40 spikes per second. **c**, Empirical spike count distributions of best-frequency responses (blue), of the responses to the adjacent frequency (red) and an estimated distribution of responses to an intermediate frequency (green). The ordinate represents the probability P of observing each spike count. **d**, ROC curves generated from pairs of distributions in **c**. Red: 1,425 and 1,481 Hz (interval: 3.9%). Green: 1,425 and 1,461 Hz (interval: 2.5%). **e**, Cumulative distribution of just-noticeable differences for units tested with random-chord stimuli at six tones per octave (red, $N = 27$) and 18 tones per octave (blue, $N = 15$).

We addressed this by comparing the predictive power of the STRFs within and across context (random-chord stimuli or film soundtrack). If artificial STRFs predict responses to the soundtrack as well as (or better than) natural STRFs, or vice versa, we can conclude that the potential nonlinear mechanisms that are not captured by the STRFs have only a small effect on the neuronal responses. Alternatively, if each STRF predicts the responses to new sounds from the ensemble used to estimate the STRF better than does the STRF derived from the other sound ensemble, then it can be inferred that there are significant nonlinearities in the responses, with the natural sounds possibly engaging processing mechanisms different from those engaged by the artificial sounds.

For units recorded with both stimuli, predicted responses to one-minute segments of the soundtrack were generated with both artificial and natural STRFs (the natural STRF was estimated without using the responses to the segment whose responses were predicted). Predictive power was quantified by the correlation coefficient between the prediction and the actual response of the unit. The expected maximum correlation (estimated as the average correlation between responses to two presentations of the soundtrack) was 0.3 (ref. 12). The predictive power of the artificial STRFs on the soundtrack was notably low: 0.13 ± 0.14 (mean $\pm$ s.d.), about 40% of the expected maximum. More importantly, correlations were significantly higher within context: a natural STRF typically predicted the actual responses to a soundtrack segment better than did an artificial STRF, with an average correlation coefficient of 0.25 ± 0.14 (Fig. 4c), over 80% of the expected maximum correlation. A three-way analysis of variance (ANOVA) on STRF type x predicted segment x neuron showed a highly significant main effect of STRF type, $F_{1,229} = 72$,

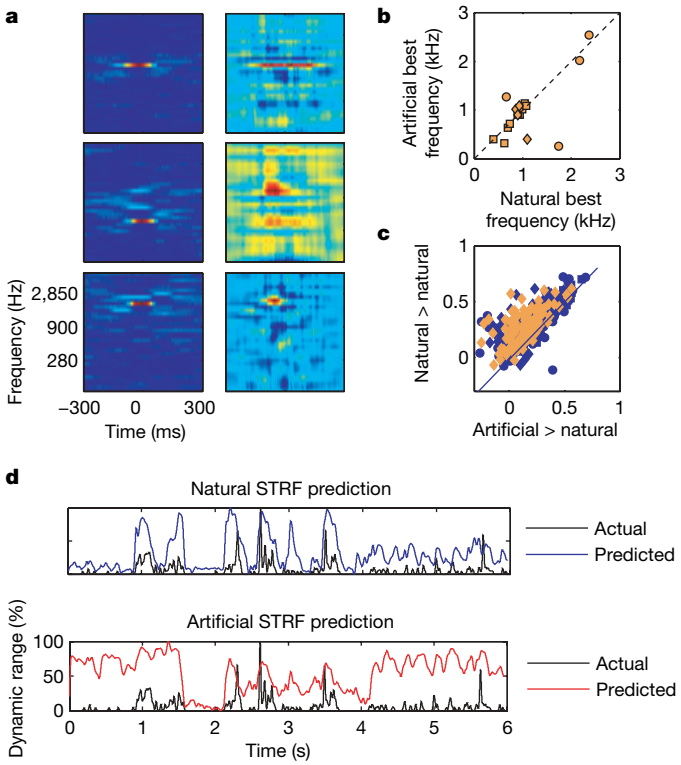

Figure 4 | Natural versus artificial responses. **a**, STRFs of three units based on responses to the random-chord stimulus (left) and to the soundtrack (right). **b**, Best frequency of artificial STRFs versus best frequency of natural STRFs ($N = 16$). **c**, Correlations between predictions and actual responses to one-minute segments from the soundtrack. Abscissa: using artificial STRFs (orange, 14 units) or synthetic STRFs (blue, 31 units). Ordinate: using natural STRFs. **d**, Predictions and response to one minute of the soundtrack by natural (top) and artificial (bottom) STRFs: 121 ms hamming window. Correlation coefficients are 0.46 and 0.18, respectively.

$P \ll 0.01$. The same general result was obtained when we used narrowband filters fitted to each unit instead of the artificial STRFs (see Supplementary Information). Similarly, natural STRFs were substantially less successful in predicting the responses to the random-chord stimuli than artificial STRFs (see Supplementary Information). Thus, stimulus encoding was not entirely determined by frequency selectivity. STRFs exhibited superior predictive power when tested with sounds that belong to the ensemble used to estimate them, suggesting that nonlinear mechanisms participate crucially in shaping the neuronal responses[10].

Our results demonstrate that frequency tuning in the human auditory cortex is substantially narrower than that typically found in the auditory cortex of non-human mammals (except bats). Using pure tones under the commonly used barbiturate anaesthesia, the tuning width at suprathreshold levels was found to be about one octave in cats[13] and about a third of an octave on average in rats[14]. Comparisons of tuning between awake and anesthetized animals within the same species have repeatedly shown that bandwidths are wider in the awake preparation (cats[15], see review[16]; rats[14]). Surveys of tuning in the auditory cortex of the awake macaque reported bandwidths that were typically half to one octave[17], and either very narrowly tuned neurons were rare[17] or bandwidths were wider than a seventh of an octave[18]. In the only other report of a unit in human auditory cortex[1], the width at half-height was at least one octave. The frequency tuning derived from STRFs is typically somewhat narrower than that derived from pure tone responses, but seems to be wider than the data shown here. For example, in deeply anaesthetized cats, the STRF width was about half an octave[19]. Thus, in mammalian responses, the typical selectivity of cortical neurons was worse, not better, than that found on the periphery of the same species. With the caution required by the small sample reported here, we propose that in contrast with animal studies, the spectral selectivity of neurons in human auditory cortex is substantially better than that of the auditory periphery.

These results are relevant to the apparent paradox of frequency hyperacuity demonstrated repeatedly in human psychoacoustics. Subjects with normal hearing, even untrained, successfully detect spectral differences substantially narrower than the presumed bandwidth of single auditory nerve fibres. Our results demonstrate that frequency differences smaller than 3% could be reliably detected from single-trial responses of single units in human auditory cortex. This value is comparable to the minimum detection threshold reported in untrained subjects[9]. Thus, the responses of one of these cortical neurons could, in principle, underlie behavioural performance on a single-trial basis. Tramo *et al.*[20] reported that bilateral lesions of human auditory cortex cause significant elevations in frequency discrimination thresholds, suggesting a functional role for the electrophysiological findings reported here. Remarkably, thresholds (frequency ratios) after the lesions were about 10–20%, matching the peripheral tuning in humans. We therefore suggest that the neural responses we observed in human auditory cortex reflect a readout of information available in the activity of large neuronal ensembles in subcortical stations, and that the auditory cortex is necessary for this readout to be performed, resulting in the behavioural hyperacuity of frequency discrimination in humans.

Previous studies in alert human subjects have shown very selective responses in single neurons from other brain areas. Notably, Quiroga *et al.*[21] reported highly specific responses to individual people or landmarks from a subset of medial temporal lobe neurons, suggesting an invariant, sparse code. The high selectivity reported here may be a counterpart of the same phenomenon, resulting in a sparse coding of frequency in auditory cortex. We can only speculate why a low-level cue such as frequency is represented so explicitly and predominantly in single neurons of human auditory cortex but not in the auditory cortex of other terrestrial mammalian species. There is evidence that frequency discrimination in humans is correlated with a number of cognitive skills, including language abilities[22], working memory[23] and learning capabilities[24], but more research is needed to clarify this puzzle.

METHODS SUMMARY

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. In each experimental session, patients 1 to 3 were presented twice in succession with 8:40 min of an audio-visual segment of the film "The Good, the Bad, and the Ugly". Patients 2, 3 and 4 were presented with random-chord stimuli[25,26] accompanied with random visual textures. Each chord had three pure-tone components, selected quasi-randomly out of a frequency table spanning the frequency range of the soundtrack. The tone duration was 100 ms (patients 2 and 3) or 50 ms (patient 4) with 10 ms linear onset and offset ramps. The frequencies were equally spaced along a logarithmic axis from 100 Hz to 10 kHz. The resolution was either a sixth of an octave (41 different frequencies, patients 2 and 3) or 1/18th of an octave (108 frequencies, patient 4). Sequence duration was 3.5 min (patients 2 and 3) or 5 min (patient 4). Data were acquired in ten sessions, all conducted at the patients' quiet bedside using a standard laptop screen and the laptop's built-in speakers (patients 1 and 2) or external speakers (patients 3 and 4). Sound intensity was set to a comfortable hearing level but absolute sound level was not measured. The free-field presentation was most probably accompanied by reverberation. Though unlikely to have influenced the results presented here, these factors represent differences from most studies in anesthetized animals. The data consist of 95 units (20 units from patient 1, 21 from patient 2, 38 from patient 3 and 16 from patient 4). The linear approximation to the response function for each unit in response to the soundtrack was computed using the software package STRFpak[27]. Discrimination thresholds were computed using ROC analysis based on empirical spike count distributions.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 17 June; accepted 14 November 2007.

1. Howard, M. A. III *et al.* A chronic microelectrode investigation of the tonotopic organization of human auditory cortex. *Brain Res.* **724,** 260–264 (1996).
2. Nelken, I. in *Integrative Functions in the Mammalian Auditory Pathway* (eds Oertel, D., Popper, A. N. & Fay, R. R.) 358–416 (Springer, New York, 2002).
3. Moore, B. C. J. *An Introduction to the Psychology of Hearing* Ch. 3 74–114 (Academic Press, London, 1982).
4. Evans, E. F. in *Psychophysics and Physiology of Hearing* (eds Evans, E. F. & Wilson, J. P.) 185–196 (Academic Press, London, 1977).
5. Ehret, G. & Schreiner, C. E. Frequency resolution and spectral integration (critical band analysis) in single units of the cat primary auditory cortex. *J. Comp. Physiol. A* **181,** 635–650 (1997).
6. Ehret, G. & Merzenich, M. M. Complex sound analysis (frequency resolution, filtering and spectral integration) by single units of the inferior colliculus of the cat. *Brain Res.* **472,** 139–163 (1988).
7. Heinz, M. G., Colburn, H. S. & Carney, L. H. Evaluating auditory performance limits: I. One-parameter discrimination using a computational model for the auditory nerve. *Neural Comput.* **13,** 2273–2316 (2001).
8. Fried, I. *et al.* Cerebral microdialysis combined with single-neuron and electroencephalographic recording in neurosurgical patients. *J. Neurosurg.* **91,** 697–705 (1999).
9. Banai, K. & Ahissar, M. Poor frequency discrimination probes dyslexics with particularly impaired working memory. *Audiol. Neurootol.* **9,** 328–340 (2004).
10. Theunissen, F. E., Sen, K. & Doupe, A. J. Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J. Neurosci.* **20,** 2315–2331 (2000).
11. Woolley, S. M., Gill, P. R. & Theunissen, F. E. Stimulus-dependent auditory tuning results in synchronous population coding of vocalizations in the songbird midbrain. *J. Neurosci.* **26,** 2499–2512 (2006).
12. Hsu, A., Borst, A. & Theunissen, F. E. Quantifying variability in neural responses and its application for the validation of model predictions. *Network* **15,** 91–109 (2004).
13. Read, H. L., Winer, J. A. & Schreiner, C. E. Modular organization of intrinsic connections associated with spectral tuning in cat auditory cortex. *Proc. Natl Acad. Sci. USA* **98,** 8042–8047 (2001).
14. Gaese, B. H. & Ostwald, J. Anesthesia changes frequency tuning of neurons in the rat primary auditory cortex. *J. Neurophysiol.* **86,** 1062–1066 (2001).

15. Qin, L., Kitama, T., Chimoto, S., Sakayori, S. & Sato, Y. Time course of tonal frequency-response-area of primary auditory cortex neurons in alert cats. *Neurosci. Res.* **46,** 145–152 (2003).
16. Moshitch, D., Las, L., Ulanovsky, N., Bar-Yosef, O. & Nelken, I. Responses of neurons in primary auditory cortex (A1) to pure tones in the halothane-anesthetized cat. *J. Neurophysiol.* **95,** 3756–3769 (2006).
17. Recanzone, G. H., Guard, D. C. & Phan, M. L. Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. *J. Neurophysiol.* **83,** 2315–2331 (2000).
18. Schwarz, D. W. & Tomlinson, R. W. Spectral response patterns of auditory cortex neurons to harmonic complex tones in alert monkey (*Macaca mulatta*). *J. Neurophysiol.* **64,** 282–298 (1990).
19. Miller, L. M., Escabi, M. A., Read, H. L. & Schreiner, C. E. Spectrotemporal receptive fields in the lemniscal auditory thalamus and cortex. *J. Neurophysiol.* **87,** 516–527 (2002).
20. Tramo, M. J., Shah, G. D. & Braida, L. D. Functional role of auditory cortex in frequency processing and pitch perception. *J. Neurophysiol.* **87,** 122–139 (2002).
21. Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C. & Fried, I. Invariant visual representation by single neurons in the human brain. *Nature* **435,** 1102–1107 (2005).
22. Benasich, A. A. & Tallal, P. Infant discrimination of rapid auditory cues predicts later language impairment. *Behav. Brain Res.* **136,** 31–49 (2002).
23. Banai, K. & Ahissar, M. Auditory processing deficits in dyslexia: task or stimulus related? *Cereb. Cortex* **16,** 1718–1728 (2006).
24. McArthur, G. M. & Bishop, D. V. Speech and non-speech processing in people with specific language impairment: a behavioural and electrophysiological study. *Brain Lang.* **94,** 260–273 (2005).
25. deCharms, R. C., Blake, D. T. & Merzenich, M. M. Optimizing sound features for cortical neurons. *Science* **280,** 1439–1443 (1998).
26. Schnupp, J. W., Mrsic-Flogel, T. D. & King, A. J. Linear processing of spatial cues in primary auditory cortex. *Nature* **414,** 200–204 (2001).
27. Theunissen, F. E. *et al.* Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. *Network* **12,** 289–316 (2001).
28. Mukamel, R. *et al.* Coupling between neuronal firing, field potentials, and FMRI in human auditory cortex. *Science* **309,** 951–954 (2005).
29. Bleeck, S., Ives, T. & Patterson, R. D. Aim-mat: the auditory image model in MATLAB. *Acta Acustica* **90,** 781–788 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the patients for their cooperation in participating in the experiments. We thank E. Behnke, T. A. Fields, E. Ho and C. Wilson for technical assistance. This work was supported by an ISF grant (to I.N.), a NINDS grant (to I.F.), the US-Israel BSF fund (R.M. and I.F.) and a European Molecular Biology Organization and Human Frontier Science Program fellowship (R.M.).

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to I.N. (israel@cc.huji.ac.il) or to I.F. (ifried@mednet.ucla.edu).

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.