Optimizing Sound Features for Cortical Neurons

R. Christopher deCharms, *et al.*
Science **280**, 1439 (1998);
DOI: 10.1126/science.280.5368.1439

Optimizing Sound Features for Cortical Neurons

R. Christopher deCharms, David T. Blake,
Michael M. Merzenich*

The brain's cerebral cortex decomposes visual images into information about oriented edges, direction and velocity information, and color. How does the cortex decompose perceived sounds? A reverse correlation technique demonstrates that neurons in the primary auditory cortex of the awake primate have complex patterns of sound-feature selectivity that indicate sensitivity to stimulus edges in frequency or in time, stimulus transitions in frequency or intensity, and feature conjunctions. This allows the creation of classes of stimuli matched to the processing characteristics of auditory cortical neurons. Stimuli designed for a particular neuron's preferred feature pattern can drive that neuron with higher sustained firing rates than have typically been recorded with simple stimuli. These data suggest that the cortex decomposes an auditory scene into component parts using a feature-processing system reminiscent of that used for the cortical decomposition of visual images.

Feature processing by neurons in the primary visual cortex has been studied extensively since the discovery by Hubel and Wiesel that the "right" kinds of stimuli for visual cortical neurons are moving, oriented bars within a spatial receptive field (1)—a finding later confirmed and extended in detail by reverse-correlation methods (2). The fundamental feature-processing characteristics within the spectral receptive fields of neurons in the awake primary auditory cortex are less completely resolved, except in species with particularly well understood auditory behavior, such as bats (3), owls (4), and song-birds (5), where stimuli have been se-

Keck Center for Integrative Neuroscience, University of California, San Francisco, San Francisco, CA 94143–0732, USA.

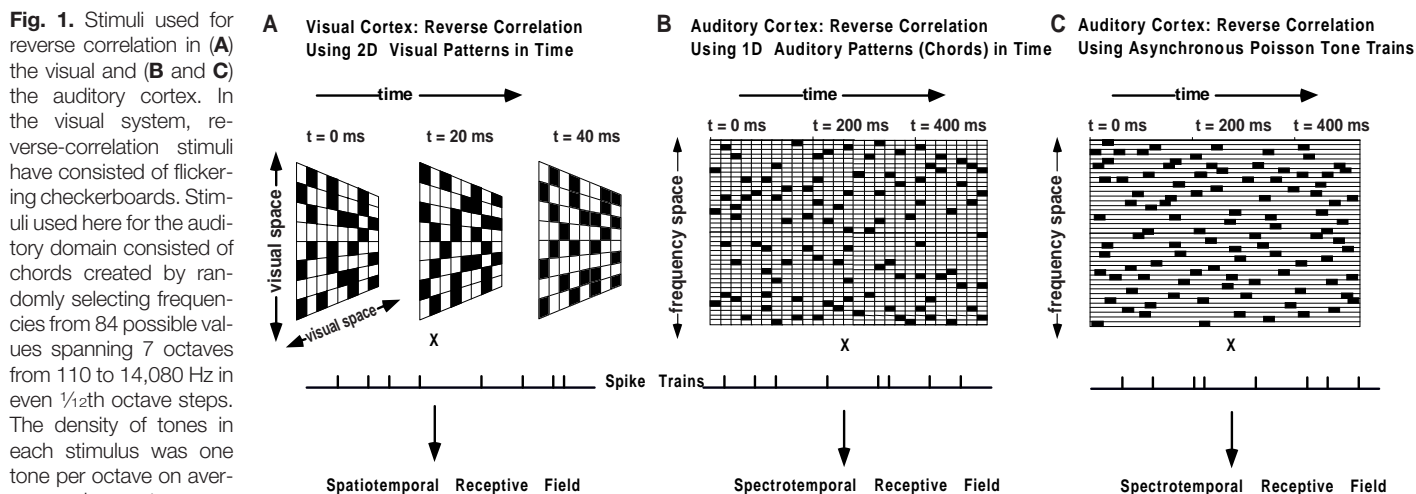*To whom correspondence should be addressed. E-mail: merz@keck.ucsf.edu

lected on the basis of ethological principles (6). Our understanding of auditory cortex physiology in other mammalian species results largely from studies of anesthetized animals (7), which has demonstrated that auditory neurons are tuned for a number of independent feature parameters of simple stimuli including frequency (8), intensity (9), amplitude modulation (10), frequency modulation (11), and binaural structure (12). However, auditory responses to multiple stimuli can also enhance or suppress one another in a time-dependent manner (13), and auditory cortical neurons can be highly selective for species-specific vocalizations (14), indicating complex acoustic processing by these cells. It is not yet known if or how these many independent selectivities of auditory cortical neurons reflect an underlying pattern of feature decomposition, as has been

suggested (15). Further, because sustained firing-rate responses in the auditory cortex to tonal stimuli are typically much lower than visual responses to drifting bars (16), it has been suggested that the preferred type of auditory stimulus may still not be known (17). For these reasons, we investigated whether a reverse-correlation method similar to that used in the visual cortex would discern the full spectral and temporal feature-response profile of auditory cortical neurons in the awake animal.

Reverse correlation stimuli used in visual cortex experiments have consisted of rapidly presented two-dimensional checkerboard spatial patterns (Fig. 1A). The auditory stimuli used here consisted of rapidly presented random chords or asynchronous random tone progressions that evenly spanned a portion of the one-dimensional receptor surface of the cochlea (Fig. 1, B and C). Rather than measuring tuning for a preselected parameter, this unbiased method constructs the average full auditory stimulus preceding spikes from a neuron, whatever form it may take (18). Alternatively, this can be viewed as the average response of the neuron driven by each separate stimulus component, which is numerically identical (but is reversed in time, and is expressed in units of mean density of stimulus components instead of the mean density of spikes). Neuronal reverse-correlation techniques were originally developed for characterizing auditory neurons in the periphery (19), but while these methods have been applied to more peripheral structures and visual cortical cells, researchers have only recently begun to succeed in applying this method to auditory cortical neurons (20). Data presented here are taken from extensive characterizations of 206 iso-

Fig. 1. Stimuli used for reverse correlation in (**A**) the visual and (**B** and **C**) the auditory cortex. In the visual system, reverse-correlation stimuli have consisted of flickering checkerboards. Stimuli used here for the auditory domain consisted of chords created by randomly selecting frequencies from 84 possible values spanning 7 octaves from 110 to 14,080 Hz in even $\frac{1}{12}$th octave steps. The density of tones in each stimulus was one tone per octave on average, and seven tones per

A Visual Cortex: Reverse Correlation Using 2D Visual Patterns in Time

B Auditory Cortex: Reverse Correlation Using 1D Auditory Patterns (Chords) in Time

C Auditory Cortex: Reverse Correlation Using Asynchronous Poisson Tone Trains

chord, but the stimuli were stochastic so that a given chord could be composed of a variable number of tones of randomly selected frequencies. Rates of 10 to 100 chords/s have been used, as well as stimuli with random, nonaligned onset times in each frequency band. Data presented here used 50 chords/s or the equivalent number of nonaligned tones.

The complete stimulus set lasted for 10 min, thereby including 30,000 individual chords. The stimuli were 20-ms long pure tones with 5-ms cosine onset and offset ramps, and were presented at either 70 or 50 dB, or both. Fewer frequencies are shown in the diagram than were actually used.

lated single neurons in the primary auditory cortex (AI) of two awake owl monkeys (*21*).

Neurons in the auditory cortex have traditionally been analyzed by counting spikes evoked by individual pure-tone stimuli to compute frequency tuning at differing intensity levels (*22*). Figure 2A shows the standard frequency response area of a single AI neuron at a range of intensities for comparison. The remainder of Fig. 2 shows plots of the full linear average pattern of feature selectivity in spectral content and in time for seven representative AI neurons in the form of spectrotemporal receptive fields, measured at a 70-dB sound pressure level (SPL), including this same neuron from (A) presented in (B) and (C). The spectrotemporal receptive field (STRF) is effectively a sonogram of the linear estimate of the optimal stimulus and is computed during the same amount of recording time as a traditional tuning curve, 10 min. Figure 2, B to I, shows the effect of average firing rate on each neuron driven by each individual component of the complex stimulus as the pixel's color (spikes per second). The reverse-correlogram is thus presented as a stimulus-triggered spike rate average, analogous to a standard peristimulus time histogram but reversed in time, and is identical to the estimated optimal stimulus for the cell (a spike-triggered stimulus average in units of mean stimulus density). To demonstrate the typical consistency of the

method to reproduce the complex details of receptive fields, Fig. 2C presents a second estimate of the optimal stimulus for the neuron in (B) using an entirely different stimulus set and using asynchronous rather than synchronous tone onsets. Considerable structure in the regions of both increased and decreased firing rate is apparent in the STRF that is not evident in the traditional tuning curve.
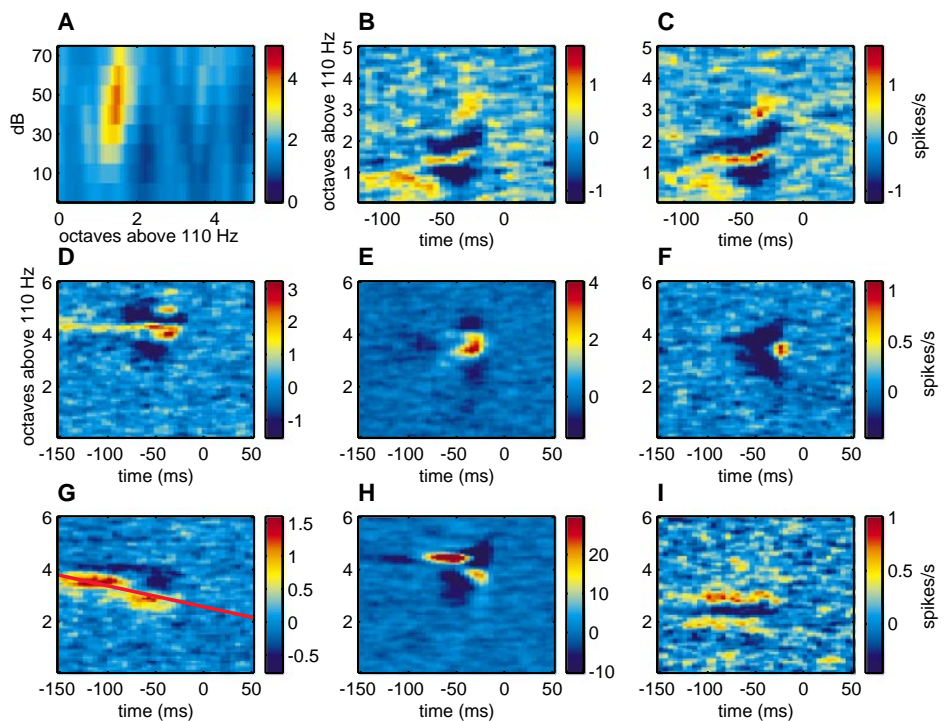
Only a minority of neurons in AI (5% of total, n = 206 neurons) had spectrotemporal receptive fields with just a single region of increased rate and no inhibition, which corresponds to simple selectivity for a particular frequency region, with no additional feature-processing structure. We have found that cells of this type (*23*) are less common than cells with complex multipartite receptive fields that include regions of both increased and decreased firing rate as well as temporal structure. We will refer to these regions as excitatory and inhibitory, although these are measures of deviations from mean ongoing rate during characterization stimuli and are not necessarily diagnostic of an underlying synaptic mechanism.

Neurons with multipartite excitatory and inhibitory receptive fields can serve as detectors of stimulus edges in both sound frequency and in time. The neuron shown in Fig. 2D had a receptive field structure within its frequency-responsive area showing a long,

narrow region of excitation flanked by inhibition, suggesting that this cell would extract information about sound components containing a continuous-frequency edge at a precise tonal location, or very spectrally narrow sounds at this frequency (observed selectivity is shown in Fig. 3). The neuron shown in Fig. 2E showed a brief region of excitation, with its dominant feature being symmetrical lateral inhibitory components above and below, also suggestive of detecting stimulus edges and strong tuning for sound-component width (*24*). The neuron shown in Fig. 2F had an STRF that predicted selectivity for stimuli with little sound energy in the receptive field, followed by strong energy—the edge of a stimulus in time rather than in frequency. This type of neuron would be predicted to respond to successive stimulus transients separated by times greater than a characteristic minimum time (*25*).

Neurons were also observed that responded with increased rates to one frequency range at one time, and to a shifted frequency range at a later time, with regions of inhibition showing a similar shift. The neuron shown in Fig. 2G is an example of this type. This pattern of response within the receptive field is strongly analogous to motion energy detectors (*25*), which detect a moving stimulus edge, and this cell was indeed selective for the direction and rate of tones sweeping in frequency. The red line on the panel

Fig. 2. The spectrotemporal receptive fields of neurons in the primary auditory cortex of the awake primate show the pattern of sound features selected for by particular neurons. The traditional tuning curve (**A**) was computed by counting the number of spikes elicited from a single auditory cortical neuron by 100-ms pure-tone pips presented at 84 frequencies and eight intensities (scale bar in spikes per stimulus, intensities in decibel SPL, data convolved with 2-pixel-wide Gaussian). The spectrotemporal receptive field (**B**) shows the full time-frequency structure of the same neuron's sound-feature selectivity (note the different axes), which is proportional to its estimated optimal stimulus including multiple excitatory and inhibitory regions, in units of mean spikes per second. Spectrotemporal receptive fields (**B** to **I**) were computed as described in the text. The spectrotemporal receptive field in (C) was recomputed for the same neuron with a completely different stimulus set. In (B) the stimuli were made up of random chords with synchronous onset times for each component in the chord, and in (C) all individual components were presented at completely random and asynchronous times. Receptive fields are all from well-isolated single-neuron recordings. Receptive-field structures correspond to the average rate of spikes from the neuron at time zero driven by each stimulus component frequency at the lag time shown; this is the standard peristimulus spike rate value triggered on each stimulus component but reversed in time. Proportionally equivalent, the spectrotemporal receptive field is also the average stimulus preceding each neuronal spike at time zero, an average of the stimulus components triggered on the spike occurrence. Therefore, scale bars shown (in spikes per stimulus) are directly proportional to the probability of occurrence of each stimulus component (in stimuli per spike). Spectrotemporal receptive fields were convolved with a 1-pixel-wide Gaussian to reduce noise.

illustrates the slope of the sweeping tonal stimulus that drove the strongest response from this cell (observed selectivity is shown in Fig. 3). The neurons shown in Fig. 2, B, H, and I, had complex multimodal receptive-field structures, with multiple excitatory or inhibitory subregions. Response areas of this type are indicative of selectivity for stimulus-feature conjunctions, sometimes quite complex, and perhaps related to sounds that the animal must process within its learned environment. In addition, the neuron in (I) was chosen to illustrate a neuron with an STRF that was weaker and more diffuse, a feature often found in spatially nonlinear V1 neurons characterized as complex cells (*26*).

To conservatively estimate the statistical significance of each response component, we divided spectrotemporal receptive field data for each neuron into 10 segments of 1 min each and smoothed the data with a 1-pixel-wide Gaussian. Response subregions were deemed significant only if they were at least three standard errors above or below the overall mean value and were greater than one pixel in size. Of 206 single units recorded, 13.1% did not yield a statistically significant STRF in 10 min of recording time, suggesting that these cells were not responsive to the noise stimuli or were strongly nonlinear. Of cells with a significant STRF, 46.4% had multiple excitatory subregions, and 48.6% had multiple inhibitory subregions. The percentages of cells with different numbers of response subregions are presented in Table 1. Response areas were highly localized in temporal and spectral extent, suggesting decomposition of sounds into small-component features in the primary sensory cortex. The median bandwidth of total response area was 1.8 octaves (*27*), and the median temporal duration of response was 64 ms (*28*).

This method produces an estimate of the feature processing by cortical neurons in frequency and time, but this estimate does not take into account the full range of properties of cortical neurons, particularly their nonlinearities and their processing of stimulus-com-

ponent amplitude. Consequently, as with other characterization methods, the STRF reflects properties of these cells, but should not be taken to quantitatively predict neuronal responses to arbitrary stimuli. However, the feature-processing pattern observed for an auditory neuron's STRF can guide the generation of stimuli used to characterize the cell. Based on the finding of selectivity for stationary or moving spectral stimulus edges, stimuli were designed containing these features. These auditory stimuli are direct analogs of spatial stationary edges and moving gratings.

Individual stimuli are shown in Fig. 3 as color spectrograms where time is plotted along the horizontal axis, frequency content on an octave scale is along the vertical axis, and color corresponds to the intensity of the signal (from blue to red). The static-edge stimuli shown in the bottom of Fig. 3A are 250-ms continuous-noise bursts that contained all frequencies up to a discrete cutoff, with the position of the frequency edge increasing along the series. The neuron presented was highly selective for stationary frequency edges that fell exactly upon its long excitatory region (see the STRF in Fig.
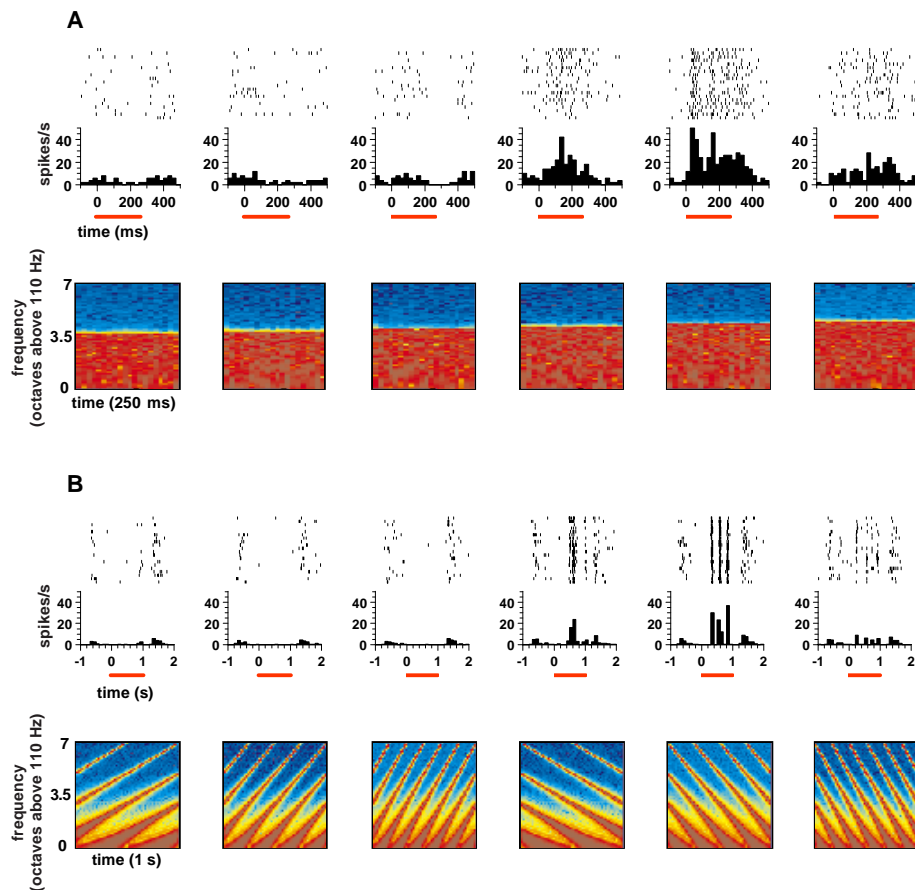

Fig. 3. Detectors of stationary and moving spectral stimulus edges. The six color panels in (**A**) are sonograms of stationary-edge low-pass noise stimuli with increasing frequencies of the upper cutoff. The sonograms shown span 250 ms and 7 octaves, with increasing sound intensity progressing from blue to red. The first stimulus has a cutoff frequency of 1.369 kHz, and each successive stimulus has a cutoff frequency ⅛th of an octave higher. (Top) The average responses to 20 presentations of the stimulus is presented as a peristimulus time histogram, in spikes per second, with 25-ms bins on a time scale in milliseconds, with the duration of the stimulus indicated by the red bar and with records of individual trials presented above. Of 270 total high-pass, low-pass, and band-pass noise stimuli presented, the neuron in (A) was selective for stimuli containing a spectral edge within the excitatory region of its STRF, preferring low-pass edges as shown (best stimulus cutoff, 4.3 octaves above 110 Hz), and including narrow-band stimuli with both upper and lower edges within this region. The sonograms in (**B**) show moving auditory grating stimuli consisting of pure tones spaced at two-octave intervals that sweep upward or downward in frequency at differing rates and are replaced when reaching the frequency range boundary (±4, 8, or 12 octaves/s shown, taken from a set of 24 total variants). These sonograms are 1 s long and span 7 octaves; pure tones appear broader at lower frequencies due to the resolution of the sonograms at low frequency. Peristimulus time histograms are shown in spikes per second with 100-ms bins on a time scale in seconds, with the duration of the stimulus indicated by the red bar and with records of individual trials presented above. The neuron in (B) was selective for descending FM stimuli with a rate of −8 octaves/s. All stimuli were generated digitally and presented at 44.1 kHz from audio compact disc.

Table 1. Percentage of single neurons with different numbers of statistically significant excitatory and inhibitory subregions. Abbreviations: Ex, number of excitatory subregions of response area; Inh, number of inhibitory subregions of response area. Of the total, 13.1% did not have a significant STRF by our criteria and are not included in these percentages.

n = 206	Ex = 1 (%)	Ex = 2 (%)	Ex = 3 (%)
Inh = 0	5.0	1.7	—
Inh = 1	23.5	19.0	2.2
Inh = 2	22.3	16.8	2.2
Inh = 3	2.8	3.9	0.6

2D), illustrating the behavior of a static-edge detector. Figure 3B shows drifting grating stimuli that contained pure-tone frequencies separated by two octaves that drifted continuously upward or downward in time, being replaced when the end of the chosen sound-frequency range was reached. The neuron in this example (STRF shown in Fig. 2G; the red line corresponds to −8 octaves/s) illustrates selectivity for the direction and the rate of the repeated moving-frequency edges of an auditory grating stimulus (29).

The STRF method can also be used to create estimated optimized stimuli for neurons with arbitrarily complex patterns of feature selectivity. Complex sound stimuli were designed to match the linear estimate of an optimized stimulus for a neuron and then played back to the same neuron to measure responsiveness. Figure 4A shows the stable, isolated action potential waveforms from an auditory cortical neuron at the time when we collected initial characterization data, and 3 hours later when we presented a set of optimized stimuli. Figure 4D shows the strong response of this neuron to its estimat-

ed optimal stimulus, consisting of 10 repetitions of the linear average stimulus derived from the STRF (lasting a total of 1 s) (30), with individual trial responses shown above. This neuron was driven to a sustained rate of greater than 40 spikes/s by its predicted optimal stimulus. In Fig. 4, B to F, variant stimuli were generated that had the same average spectral content and total power as the predicted optimal stimulus, but were contracted or expanded in time to measure tuning in the local stimulus space, corresponding to a stretching or shrinking of the spectrogram's horizontal axis. The shorter to longer stimulus variants were also repeated 10 times in 1 s. The neuron was strongly selective for the predicted optimal stimulus, compared with stimuli made from components that were stretched by ¼, ½, 2, or 4×. In Fig. 4, H to L, stimuli were generated that had the same total sound energy in each time bin and the same center frequency, but frequency content corresponding to a stretching or shrinking of the spectrogram's vertical frequency axis, making the stimuli broader or narrower in spectral content. The neuron preferred the predicted optimal stimulus over four morphed versions, as well as over a pure tone at the center frequency (shown in Fig. 2G). The STRF for this cell was presented in Fig. 2H. This outstanding example illustrates that auditory cortical neurons can be capable of very high sustained rates when presented with optimized stimuli, even when conventional stimuli are ineffective. When optimized stimuli are used with higher repetition rates than those presented here, the most active primary auditory cortical neurons can achieve sustained rates as high as 60 spikes/s, more comparable to rates found in the primary visual and somatosensory cortices.

From the infinite space of possible sounds, the predicted feature conjunctions for a neuron generated by reverse correlation produced highly effective stimuli for about half of auditory cortical cells. The full distribution of sustained rate responses for all 40 cells that were carried through this long protocol are shown in Fig. 4M. The continuous firing-rate responses to optimized stimuli are plotted along the vertical axis, against the integral of the excitatory area of the neuron's STRF, along the horizontal axis. Of the cells that responded poorly to optimized stimuli, all had STRFs showing only weak excitation. Figure 4 (N and O) shows population averages for the tuning of neurons to their predicted optimal stimuli compared with variants in temporal and spectral structure. The feature conjunctions predicted by reverse correlation on average produced the strongest responses from these cells, with monotonically decreasing responses to perturbations from this stimulus ($P < 0.007$ sign
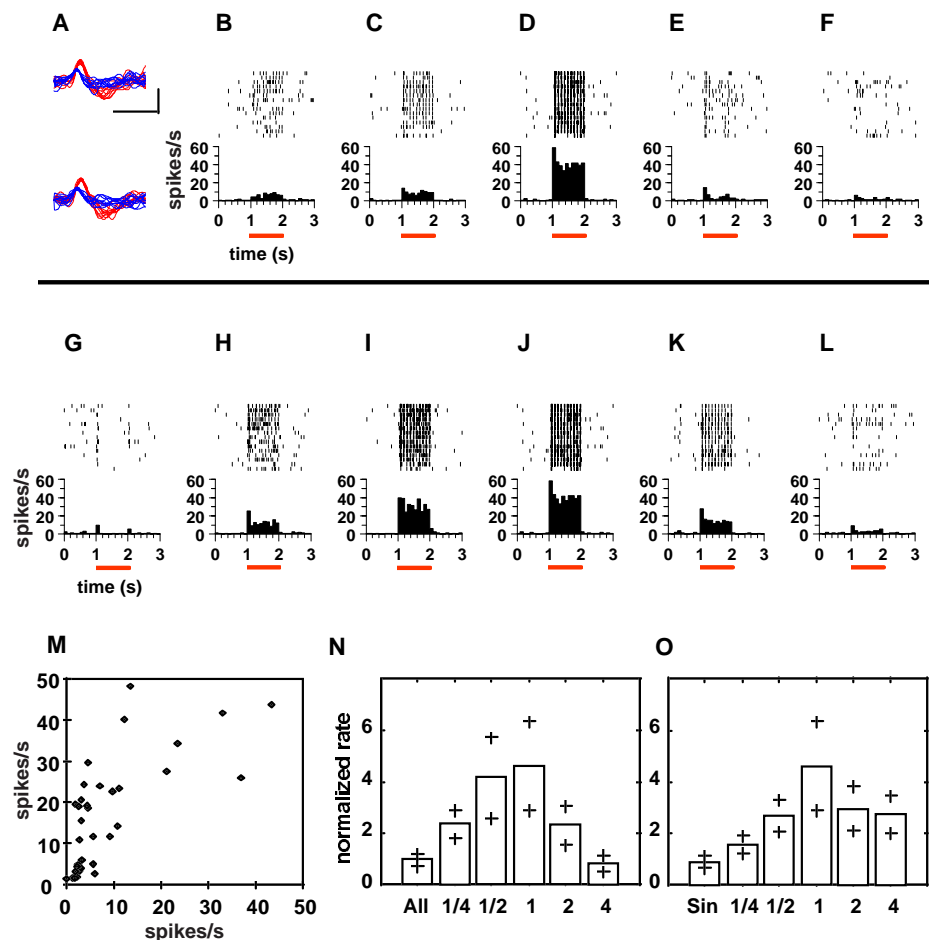

Fig. 4. Responses of neurons in the primary auditory cortex of the awake primate to predicted optimal stimuli and to variants. (**A**) Voltage waveforms randomly selected throughout recording from one isolated single neuron (red) and other background spikes that crossed a high fixed threshold (blue) are presented at the time of initial characterization and during presentation of our optimized stimulus set 3 hours later. This neuron had a signal-to-noise ratio of >11:1 compared with background noise standard deviation. Horizontal bar, 1 ms; vertical bar, −100 μV. (**B** to **L**) Responses to the predicted optimal stimulus for the neuron in (A) and to variants. In (B) to (F), temporal profiles were expanded or contracted to ¼, ½, 1, 2, 4× the predicted optimal temporal profile. In (G) a pure tone was presented at the predicted best frequency, and in (H) to (L) spectral profiles were expanded or contracted to ¼, ½, 1, 2, 4× the predicted optimal profile. (**M**) The mean firing rate during the predicted optimal stimulus for 40 neurons in the awake primary auditory cortex plotted versus the integral of the positive component of the STRF after thresholding at 2 SDs and dividing by the total time, to yield units of spikes per second. (**N**) The selectivity of this population for the time expansion series and the average response to a set of stimuli optimized for other cells; (**O**) the selectivity of the population for the spectral expansion series and for pure tones. For each response, the increase over background rate was divided by the background rate for that neuron, responses were averaged, and means ± SE are shown. Median responses showed a similar pattern, indicating that the mean values do not merely reflect the contribution of the most pronounced examples.

test, significance of centrally peaked, monotonic distribution), and with much weaker average responses to best-frequency pure tones, or to stimuli designed for other neurons.

These data collectively demonstrate that it is possible to generate linear estimates of the time and frequency feature processing of auditory cortical neurons, although they do not produce a complete quantitative model of selectivity to all stimuli. The method allows the design of stimuli that effectively probe cortical selectivity and feature decomposition. In decomposing visual forms or auditory scenes, the cortex uses detectors with similar characteristics for finding the position of stimulus edges along the sensory receptor surface, finding stimulus edges in time, finding stimulus movements, and finding feature conjunctions. Auditory cortical neurons have responses with similar extents of excitation and inhibition, a response time-course restricted to 20 to 100 ms, and selectivity to direction and velocity of stimulus movement within a temporal range of 2 to 40 cycles/s, which are similar, although not identical, to stimulus decomposition parameters in the visual cortex (31). In addition, these data indicate that by using spectrally and temporally optimized stimuli it is possible to drive many auditory cortical cells to high sustained rates of firing.

REFERENCES AND NOTES

1. D. H. Hubel and T. N. Wiesel, *J. Physiol.* **160**, 106 (1962).
2. J. P. Jones and L. A. Palmer, *J. Neurophysiol.* **58**, 1187 (1987); R. C. Reid, R. E. Soodak, R. M. Shapley, *ibid.* **66**, 505 (1991); G. C. DeAngelis, I. Ohzawa, R. D. Freeman, *ibid.* **69**, 1091 (1993); *ibid.*, p. 1118; *Trends Neurosci.* **18**, 451 (1995); R. C. Reid and J. M. Alonso, *Nature* **378**, 281 (1995); D. L. Ringach, M. J. Hawken, R. Shapley, *ibid.* **387**, 281 (1997).
3. N. Suga, H. Niwa, I. Taniguchi, D. Margoliash, *J. Neurophysiol.* **58**, 643 (1987); M. Kawasaki, D. Margoliash, N. Suga, *ibid.* **59**, 623 (1988); N. Suga, *J. Exp. Biol.* **146**, 277 (1989); H. Riquimaroux, S. J. Gaioni, N. Suga, *Science* **251**, 565 (1991); S. P. Dear, J. Fritz, T. Haresign, M. Ferragamo, J. A. Simmons, *J. Neurophysiol.* **70**, 1988 (1993); D. C. Fitzpatrick, J. S. Kanwal, J. A. Butman, N. Suga, *J. Neurosci.* **13**, 931 (1993); J. A. Simmons, *J. Comp. Physiol. A* **172**, 533 (1993).
4. E. I. Knudsen, S. duLac, S. D. Esterly, *Annu. Rev. Neurosci.* **10**, 41 (1987); E. I. Knudsen and M. S. Brainard, *ibid.* **18**, 19 (1995).
5. D. Margoliash, *J. Neurosci.* **3**, 1039 (1983); M. Konishi, *Brain Behav. Evol.* **44**, 279 (1994); M. S. Lewicki and M. Konishi, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5582 (1995); S. J. Chew, D. S. Vicario, F. Nottebohm, *Science* **274**, 1909 (1996); A. J. Doupe, *J. Neurosci.* **17**, 1147 (1997).
6. N. Suga, *Philos. Trans. R. Soc. London Ser. B* **336**, 423 (1992).
7. J. F. Brugge and M. M. Merzenich, *J. Neurophysiol.* **36**, 1138 (1973); C. E. Schreiner and J. R. Mendelson, *ibid.* **64**, 1442 (1990); J. R. Mendelson, C. E. Schreiner, M. L. Sutter, K. L. Grasse, *Exp. Brain Res.* **94**, 65 (1993); J. J. Eggermont, *Heart Res.* **74**, 51 (1994); J. C. Middlebrooks, A. E. Clock, L. Xu, D. M. Green, *Science* **264**, 842 (1994); X. Wang, M. M. Merzenich, R. Beitel, C. E. Schreiner, *J. Neuro-*

8. *physiol.* **74**, 2685 (1995); N. Kowalski, D. A. Depireux, S. A. Shamma, *ibid.* **76**, 3503 (1996); *ibid.*, p. 3524.
9. J. F. Brugge and M. M. Merzenich, *J. Neurophysiol.* **36**, 1138 (1973); M. M. Merzenich, P. L. Knight, G. L. Roth, *Brain Res.* **63**, 343 (1973); D. P. Phillips, M. N. Semple, M. B. Calford, L. M. Kitzes, *Exp. Brain Res.* **102**, 210 (1994).
10. J. M. Miller, R. D. Beaton, T. O'Connor, B. E. Pfingst, *Brain Res.* **69**, 101 (1974); M. L Sutter and C. E. Schreiner, *J. Neurophysiol.* **73**, 190 (1995).
11. D. P. Phillips and S. E. Hall, *Exp. Brain Res.* **67**, 479 (1987); C. E. Schreiner and J. V. Urbas, *Hear. Res.* **32**, 49 (1988); J. J. Eggermont, *ibid.* **74**, 51 (1994); D. P. Phillips, M. N. Semple, M. B. Calford, L. M. Kitzes, *Exp. Brain Res.* **102**, 210 (1994); M. L. Sutter and C. E. Schreiner, *J. Neurophysiol.* **73**, 190 (1995); P. Heil, *ibid.* **77**, 2642 (1997); *ibid.*, p. 2616.
12. N. Suga, *J. Physiol. (London)* **179**, 26 (1965); I. C. Whitfield and E. F. Evans, *J. Neurophysiol.* **28**, 655 (1965); J. R. Mendelson and M. S. Cynader, *Brain Res.* **327**, 331 (1985); J. R. Mendelson, C. E. Schreiner, M. L. Sutter, K. L. Grasse, *Exp. Brain Res.* **94**, 65 (1993).
13. J. F. Brugge, N. A. Dubrovsky, L. M. Aitkin, D. J. Anderson, *J. Neurophysiol.* **32**, 1005 (1969); M. N. Semple and L. M. Kitzes, *ibid.* **69**, 462 (1993); *ibid.*, p. 449; J. C. Middlebrooks, A. E. Clock, L. Xu, D. M. Green, *Science* **264**, 842 (1994).
14. D. P. Phillips and M. S. Cynader, *Hear. Res.* **18**, 87 (1985); M. B. Calford and M. N. Semple, *J. Neurophysiol.* **73**, 1876 (1995); M. Brosch and C. E. Schreiner, *ibid.* **77**, 923 (1997); S. A. Shamma and D. Symmes, *Hear Res.* **19**, 1 (1985).
15. Z. Wollberg and J. D. Newman, *Science* **175**, 212 (1972); I. Glass and Z. Wollberg, *Exp. Brain Res.* **34**, 489 (1979); X. Wang, M. M. Merzenich, R. Beitel, C. E. Schreiner, *J. Neurophysiol.* **74**, 2685 (1995).
16. M. M. Merzenich, W. M. Jenkins, J. C. Middlebrooks, in *Dynamic Aspects of Neocortical Function*, G. M. Edelman, Ed. (Wiley, New York, 1985), pp. 397–423; C. E. Schreiner and G. Langner, *Coding of Temporal Patterns in the Central Auditory Nervous System*, G. M. Edelman, Ed. (Wiley, New York, 1988); C. E. Schreiner and J. R. Mendelson, *J. Neurophysiol.* **64**, 1442 (1990); C. E. Schreiner, *Curr. Opin. Neurobiol.* **2**, 516 (1992); X. Wang, M. M. Merzenich, R. Beitel, C. E. Schreiner, *J. Neurophysiol.* **74**, 2685 (1995).
17. R. C. deCharms and M. M. Merzenich, *Nature* **381**, 610 (1996).
18. I. Nelken, Y. Prut, E. Vaadia, M. Abeles, *Hear. Res.* **72**, 237 (1994); *ibid.*, p. 206; *ibid.*, p. 223.
19. This is the linear average response taken from the stimulus space presented. There are a number of alternative forms for estimating an optimized stimulus for a neuron [see F. Rieke, D. Warland, R. R. de Ruyter van Steveninck, W. Bialek, *Spikes: Exploring the Neural Code* (MIT Press, Cambridge, MA, 1997).
20. E. DeBoer and P. Kuyper, *IEEE Trans. Biomed. Eng.* **15**, 169 (1968); E. DeBoer and H. R. deJongh, *J. Acoust. Soc. Am.* **63**, 115 (1978).
21. D. J. Hermes, A. M. Aertsen, P. I. Johannesma, J. J. Eggermont, *Hear. Res.* **5**, 147 (1981); D. J. Hermes, J. J. Eggermont, A. M. Aertsen, P. I. Johannesma, *ibid.* **6**, 103 (1982); J. J. Eggermont, A. M. Aertsen, P. I. Johannesma, *ibid.* **10**, 167 (1983); J. J. Eggermont, P. M. Johannesma, A. M. Aertsen, *Q. Rev. Biophys.* **16**, 341 (1983); M. C. Kao, P. W. Poon, X. Sun, *Biosystems* **40**, 103 (1997); I. Nelken, P. J. Kim, E. D. Young, *J. Neurophysiol.* **78**, 800 (1997); C. E. Schreiner et al., personal communication; S. A. Shamma, D. A. Depireux, J. Z. Simon, *Soc. Neurosci. Abstr.* **23**, 805 (1997).
22. Recordings were made from chronically implanted arrays of up to 45 individually placed ultrafine iridium microelectrodes in the primary auditory cortex. Electrodes had tip lengths of 10 to 25 μm, which yield impedance values of 0.5 to 5 megohm and good isolation of signals from individual neurons. These signals were filtered from 0.3 to 8 kHz, sampled at 20 kHz, digitized, stored to disk, and spike-sorted to yield well-resolved single neurons. Stimuli were presented free field in a double-walled acoustic isolation

chamber while the animal sat passively in a primate chair and occasionally received juice as a reward, or during performance of a sound frequency discrimination task. Response patterns were largely similar in the two conditions. Spike trains were collected from multiple sites in the cortex simultaneously during the presentation of our characterization stimulus sets, and spike trains were individually sorted and reverse-correlated with the times of occurrence of each of the tonal stimuli. All experiments were performed as approved by the committee on animal research, UCSF.
23. C. E. Schreiner and J. R. Mendelson, *J. Neurophysiol.* **64**, 1442 (1990); D. P. Phillips, M. N. Semple, M. B. Calford, L. M. Kitzes, *Exp. Brain Res.* **102**, 210 (1994).
24. R. C. deCharms, D. T. Blake, M. M. Merzenich, data not shown.
25. When tested with band-pass noise stimuli of different widths at the best center frequency of the cell, the neuron responded most strongly to ½ octave wide noise stimuli matched to the STRF excitatory region, and did not respond well to broad-band stimuli that encroached upon its inhibitory regions (23) [J. P. Rauschecker, B. Tian, M. Hauser, *Science* **268**, 111 (1995)].
26. Neurons in the primary auditory cortex often have low-pass temporal modulation transfer functions, responding only to a limited rate of transient repetition [C. E. Schreiner, J. V. Urbas, *Hear. Res.* **32**, 49 (1988); J. J. Eggermont, *ibid.* **56**, 153 (1991); A. Bieser and P. Muller-Preuss, *Exp. Brain Res.* **108**, 273 (1996); E. H. Adelson and J. R. Bergen, *J. Opt. Soc. Am. A* **2**, 284 (1985)].
27. G. C. DeAngelis, I. Ohzawa, R. D. Freeman, *Trends Neurosci.* **18**, 451 (1995).
28. Total bandwidth range ~0.5 to 3 octaves; median excitatory bandwidth alone 0.7 octaves, range ~0.2 to 1.5 octaves; median inhibitory bandwidth alone 1.1 octaves, range ~0.2 to 3 octaves.
29. Total duration range ~20 to 100 ms; median excitatory duration 37 ms, range ~10 to 70 ms; median inhibitory duration 43 ms, range ~10 to 100 ms.
30. N. Kowalski, D. A. Depireux, S. A. Shamma, *J. Neurophysiol.* **76**, 3524 (1996); *ibid.*, p. 3503.
31. Predicted optimal stimuli were created by taking the average stimulus intensity at each 5-ms time and 1/12th octave frequency bin before an action potential (the STRF in units of mean stimulus intensity), thresholding to capture the positive peaks, convolving this intensity pattern with 5-ms onset and offset ramps, multiplying each component by a random-phase sine wave matched to the appropriate frequency channel, and summing the signals. Perturbations of these stimuli were made by applying the same procedure after stretching the STRF in time or in frequency around the computed centroid frequency of the STRF while maintaining its sum. This produced components that were similar to the predicted optimized stimulus in total frequency spectrum but varied systematically in temporal structure, or were similar in temporal profile of total energy but varied systematically in spectral structure. In each case, after creating the stimulus, multiple copies were added together for 1 s at a repetition rate of 10 per second, and the total energy was adjusted to equal that of a 1-s 50-dB sine wave, which was also presented at the center frequency.
32. J. P. Jones and L. A. Palmer, *J. Neurophysiol.* **58**, 1233 (1987); *ibid.*, p. 1187; C. L. Baker Jr., *Vis. Neurosci.* **4**, 101 (1990); R. C. Reid, R. E. Soodak, R. M. Shapley, *J. Neurophysiol.* **66**, 505 (1991); G. C. DeAngelis, I. Ohzawa, R. D. Freeman, *ibid.* **69**, 1091 (1993); G. C. *ibid.*, p. 1118; *Trends Neurosci.* **18**, 451 (1995); R. C. Reid and J. M. Alonso, *Nature* **378**, 281 (1995).