Neuron 51, 845–859, September 21, 2006 ©2006 Elsevier Inc. DOI 10.1016/j.neuron.2006.08.030

Temporal Processing and Adaptation in the Songbird Auditory Forebrain

Katherine I. Nagel[1,*] and Allison J. Doupe[1,*]
[1] Keck Center for Integrative Neuroscience
Department of Physiology
University of California, San Francisco
San Francisco, California 94143

Summary

Songbird auditory neurons must encode the dynamics of natural sounds at many volumes. We investigated how neural coding depends on the distribution of stimulus intensities. Using reverse-correlation, we modeled responses to amplitude-modulated sounds as the output of a linear filter and a nonlinear gain function, then asked how filters and nonlinearities depend on the stimulus mean and variance. Filter shape depended strongly on mean amplitude (volume): at low mean, most neurons integrated sound over many milliseconds, while at high mean, neurons responded more to local changes in amplitude. Increasing the variance (contrast) of amplitude modulations had less effect on filter shape but decreased the gain of firing in most cells. Both filter and gain changes occurred rapidly after a change in statistics, suggesting that they represent nonlinearities in processing. These changes may permit neurons to signal effectively over a wider dynamic range and are reminiscent of findings in other sensory systems.

Introduction

A central problem for all sensory systems is how to represent complex dynamic stimuli over a wide range of intensities.

Songbirds, like humans, recognize their vocalizations whether they are soft or loud (Lohr et al., 2003). How the brain achieves this level-invariant recognition is unclear. At its highest levels, the songbird auditory system contains "feature detectors" that respond selectively to individual learned songs (Margoliash, 1983; Margoliash and Fortune, 1992; Gentner and Margoliash, 2003). These areas receive input from an area called field L, the avian analog of primary auditory cortex (Fortune and Margoliash, 1992, 1995; Wild et al., 1993). Field L responds broadly to many classes of auditory stimuli and shows organized tuning for basic auditory properties such as spectral frequency and temporal modulations (Scheich et al., 1979; Muller and Leppelsack, 1985; Heil and Scheich, 1985; Hose et al., 1987; Lewicki and Arthur, 1996; Gehr et al., 1999; Hausberger et al., 2000; Theunissen et al., 2000; Sen et al., 2001; Grace et al., 2003). Several groups have tried to model how song-selective responses might arise from combinations of field L outputs (Lewicki and Konishi, 1995; Drew and Abbott, 2003). However, few studies have looked at how neural responses in these areas depend on song or stimulus intensity.

In the mammalian auditory brainstem and midbrain, responses to simple tones and noises can depend on intensity in highly nonlinear ways (Young and Voigt, 1982; Sachs and Young, 1980; Nelken and Young, 1994; Nelken et al., 1997). In the cortex, the picture is less clear. Many studies have shown nonmonotonic and nonlinear responses to tones, intensity modulations, and ripple stimuli (Phillips and Hall, 1987; Phillips et al., 1994; Calhoun and Schreiner, 1998), while other studies have stressed the linearity of cortical responses (Kowalski et al., 1996a, 1996b; Escabi et al., 2003; Barbour and Wang, 2003). How postthalamic auditory neurons encode stimulus features across intensities remains controversial.

The goal of the present study was to understand how the responses of field L neurons to complex dynamic stimuli depend on the intensity of the stimulus. We focused on coding of amplitude modulations in time, because these are a prominent feature of both song and speech and can carry a great deal of the information present in these signals (Shannon et al., 1995; Theunissen and Doupe, 1998).

To describe the processing of naturalistic amplitude modulations over different intensity ranges, we developed a set of synthetic stimuli that capture many aspects of song's amplitude modulations but sample the space of possible modulations more thoroughly. We then used reverse-correlation techniques (Eggermont et al., 1983; Eggermont, 1993; Epping and Eggermont, 1986; Kim and Young, 1994; Brenner et al., 2000; Chander and Chichilnisky, 2001; Fairhall et al., 2001; Depireux et al., 2001; Kim and Rieke, 2001; Miller et al., 2002; Escabi and Schreiner, 2002; Baccus and Meister, 2002) to extract a linear receptive field (filter) and a nonlinear gain function from the responses to these stimuli. Together, the filter and nonlinearity allowed us to characterize the feature selectivity, threshold, and gain of each cell and to predict responses to novel sounds.

We found that our stimuli permitted robust estimates of filters and nonlinear gain functions for field L neurons. These models made good predictions of responses to novel synthetic stimuli and revealed neural sensitivity to a broad range of stimulus timescales and features. We then demonstrated that filters and gain functions depend in systematic and specific ways on the mean and variance of the stimulus amplitude. These changes are very reminiscent of those observed in the early visual system (Enroth-Cugell and Lennie, 1975; Chander and Chichilnisky, 2001; Baccus and Meister, 2002) and have been shown mathematically to improve coding efficiency (Atick, 1992). Finally, we examined the time course of these changes and found that they occurred on two timescales: a fast change in filter shape and gain, followed by a slower change in threshold.

Our findings suggest that common computational strategies underlie sensory processing in multiple domains, but challenge current models of song recognition.

*Correspondence: knagel@phy.ucsf.edu (K.I.N.), ajd@phy.ucsf.edu (A.J.D.)

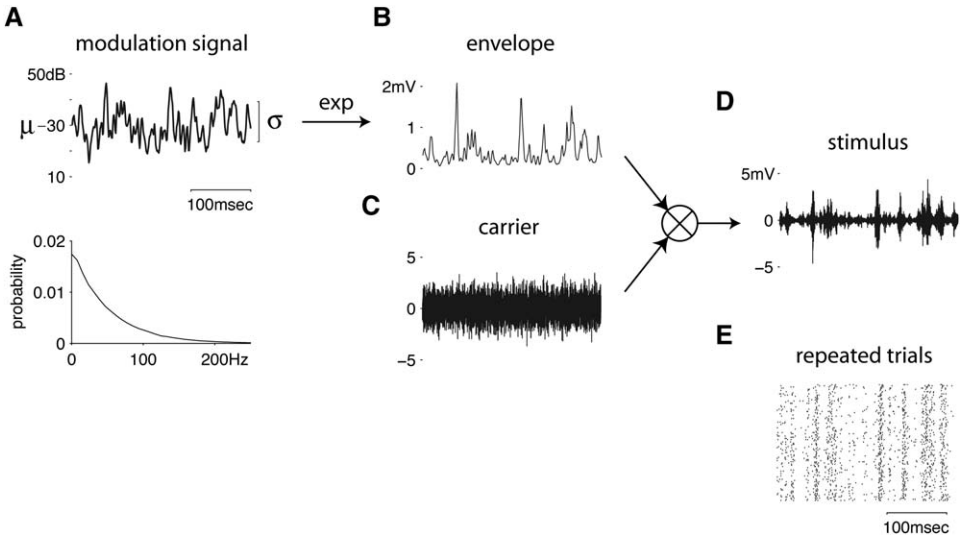

Figure 1. Construction of a Stimulus with a Power Spectrum and Amplitude Distribution Similar to Natural Sounds

The modulation signal ([A], top panel) specifies the local amplitude of the stimulus in decibels and consists of filtered Gaussian noise with a power spectrum $P(f) = \exp(-f/50\ \text{Hz})$ ([A], bottom panel). Exponentiating the modulation signal creates a linear voltage envelope (B) with an exponential distribution of amplitudes. The envelope modulates a carrier (C) consisting of narrow- or broadband noise. The stimulus (D) is generated by multiplying together the envelope (B) and the carrier (C). (E) shows responses of a single unit to 100 repeats of the modulation signal shown in (A), each paired with a different carrier. Each row represents the spike times during one segment repeat. Because the carrier varied across trials, columns of spikes represent responses to the modulation signal alone. Response peaks follow peaks in the aligned stimulus example above.

Results

Using Amplitude-Modulated Noise Stimuli to Drive Field L Neurons

To study processing of amplitude modulations in field L, we developed a stimulus composed of two parts: a slowly varying modulation envelope and a rapidly varying noise carrier.

The envelope was designed to capture the temporal frequency and amplitude distributions of natural sounds (Attias and Schreiner, 1998; Singh and Theunissen, 2003). Because they are dominated by slow changes in amplitude, natural sounds have power spectra that decrease as a function of temporal frequency (Singh and Theunissen, 2003). Natural communication sounds also contain many silent periods, giving rise to an exponential distribution of amplitudes but a more Gaussian distribution of log (dB) amplitudes (Singh and Theunissen, 2003).

To create a stimulus with these temporal frequency and amplitude properties, we first created a Gaussian noise signal with a decreasing exponential distribution of temporal frequencies (Figure 1A, lower panel). This modulation signal specifies the loudness of the stimulus in decibels at each point in time (Figure 1A, upper panel). We then exponentiated the modulation signal to produce a pressure envelope with an exponential distribution of linear amplitudes (Figure 1B). Finally, we multiplied this envelope with a narrow or broadband noise carrier (Figure 1C) to generate the full stimulus (Figure 1D, see Experimental Procedures for full details).

Although this stimulus consisted of both fast (carrier) and slow (envelope) fluctuations, we analyzed responses only with respect to the slowly varying modulation signal (Figure 1A). Recent studies have suggested that this log envelope is the property most linearly encoded by neurons in field L (Gill et al., 2006). To isolate responses to the slowly varying signal, we repeated the same modulation signal but used different noise segments to form the carrier each time. Figure 1E shows the response raster of a cell to 100 such repeats: the columns of spikes represent reliable spike patterns driven by the slow modulation signal.

Half of the 5 s long segments in our stimulus were randomly assigned to be repeats of the same modulation signal; the remaining signals were unique. Unique signals were used to broadly sample the space of possible amplitude modulations and to estimate model parameters. Repeated trials were used to test the model. Because we used different data to fit and test the model, we ensured that the quality of our predictions was not due to over-fitting.

Responses of Field L Neurons Are Well Modeled Using a Linear-Nonlinear Model

To characterize the temporal response properties of each cell, we modeled its responses using a linear-nonlinear model (Brenner et al., 2000; Chander and Chichilnisky, 2001; Fairhall et al., 2001; Baccus and Meister, 2002).

Figure 2 illustrates this process for two cells. The linear filter (Figures 2A and 2D) tells us what feature of the stimulus best drives the neuron. It is extracted from the data by computing the average modulation signal surrounding a spike, then removing the influence of stimulus correlations from the resulting waveform (Experimental Procedures). The spike occurs at zero on the x axis; the filter has structure only to the left of (before) the spike, as the cell responds causally to features of the stimulus. Small error bars (dotted lines,

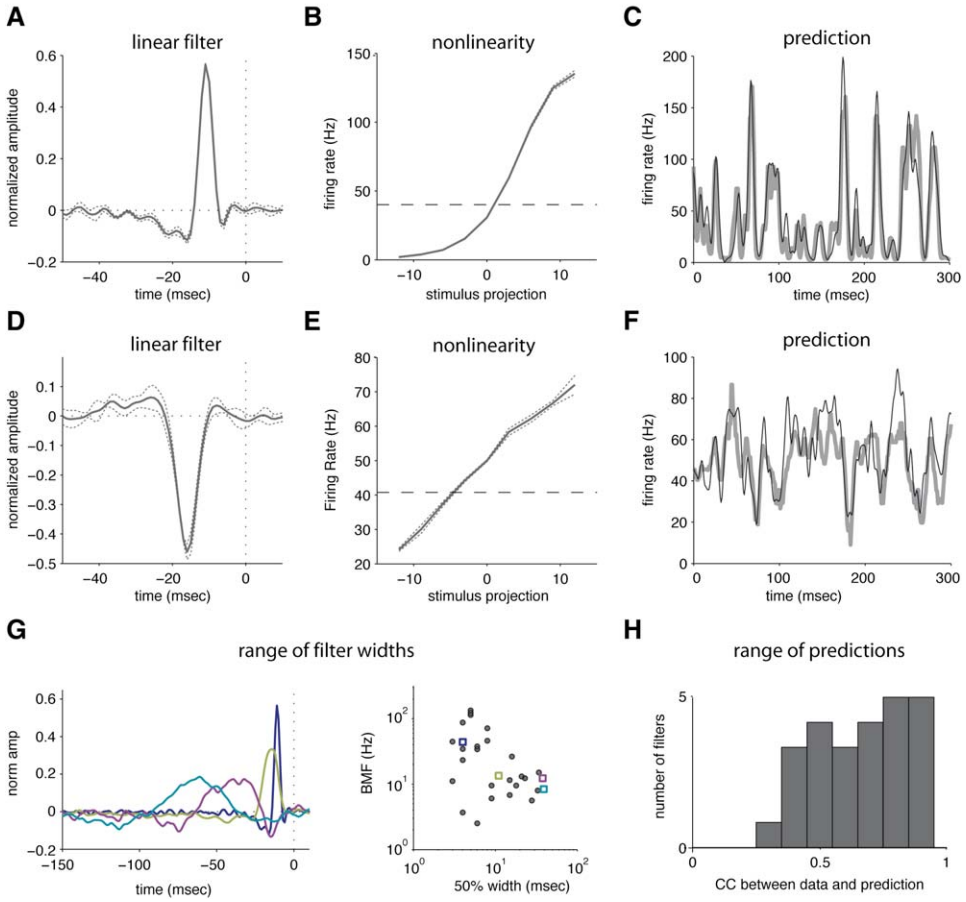

Figure 2. A Linear-Nonlinear Model Successfully Predicts a Large Fraction of the Response to Our Stimulus

(A and D) Two examples of linear filters derived from the responses of single units to the nonrepeated modulation signals (30 dB mean, 6 dB variance). The amplitude of each filter is normalized so that the dot product of the filter with itself is 1. Filters thus describe the features that drive each cell, but not the magnitude of the cell's response. Dashed lines represent the standard deviation of five independent estimates of each filter's shape.

(B and E) Nonlinear gain functions show the relationship between the firing rate of the neuron and the output of the filters in (A) and (D). The x axes show the projection of the mean-subtracted modulation signal onto each filter and indicate the similarity of the stimulus to the filter. Dotted lines indicate the standard deviation of five jackknife estimates of the gain function. The flat dashed line in each nonlinearity plot represents the average spontaneous firing rate of each neuron recorded prior to stimulus playback.

(C and F) Predicted (wide gray) and actual (narrow black) PSTHs of the response to repeated modulation signal segments. The correlation coefficients between data and prediction were 0.89 for the cell in (A)–(C) and 0.67 for the cell in (D)–(F).

(G) (Left panel) Filters from four different units show a broad range of temporal frequency preferences. (Right panel) Distribution of the best modulation frequency (BMF) versus the 50% width of the absolute value of the filter (50% width) for all significant filters (33 from 36 cells). The four cells shown in the left-hand plot are indicated by colored boxes.

(H) A histogram of correlation coefficients between predicted and actual PSTHs for all the significant filters (33 out of 36 cells recorded with a 30 dB mean and 6 dB standard deviation stimulus) indicates that the linear-nonlinear model performed well. The mean correlation coefficient was 0.64 ± 0.18 (standard deviation).

Experimental Procedures) indicate that our estimates of filter shape were robust.

The shape of the filter indicates the cell's preferred temporal feature. Filters with a single dominant peak behave as low-pass "integrators" and produce responses that look like a smoothed, delayed version of the log stimulus envelope. A response of this type (from the cell in Figure 2A) is shown in Figure 1E, where a column of spikes follows each peak in the stimulus above (Figure 1D). The sign of the filter peak determines whether the cell's firing rate is elevated or depressed by peaks in the stimulus. Filters with a biphasic shape behave as band-pass "differentiators" and respond to either onsets or offsets in the stimulus, depending on the order of the positive and negative components in the filter. Most of our cells behave as combinations of integrators and differentiators, making a strict classification of cell types difficult. The filter shown in Figure 2A has a small negative peak followed by a large positive peak; it behaves mostly as a positive integrator, but is sensitive to onsets as well. The filter shown in Figure 2B behaves mostly as a negative integrator, but also responds weakly to offsets.

The width of the filter sets an upper limit on the temporal frequencies to which the cell can respond. Because the exact frequency response depends both on the width of the filter and its shape, preferred temporal frequency is best characterized in the frequency domain

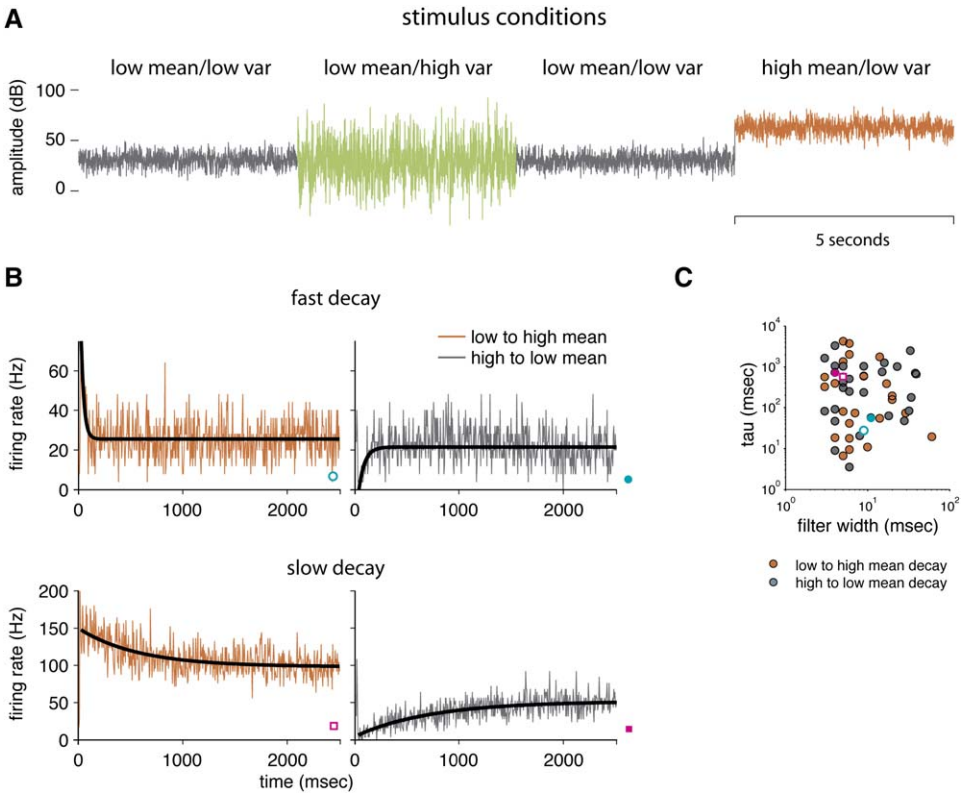

Figure 3. Changes in Stimulus Statistics Lead to Slow Changes in Firing Rate

(A) To explore how coding depends on the statistics of the stimulus, we altered the mean and standard deviation of the modulation signal every 5 s. Statistical conditions were presented in a fixed order: low mean/low variance (gray: 30 dB ± 6 dB), low mean/high variance (green: 30 dB ± 18 dB), low mean/low variance, and high mean/low variance (red: 63 dB ± 6 dB). These colors will be used throughout the paper to indicate responses to each condition.

(B) Cells adapt at different rates to a change in mean stimulus amplitude. (Top) PSTH of a cell showing rapid decays after the switch from low to high mean (red: tau = 28 ms) and after the switch from high to low mean (gray: 56 ms). (Bottom) PSTH of a cell showing slow decays. Tau = 574 ms for the switch from low to high mean, 730 ms for the switch from high to low mean. PSTHs are binned in 5 ms windows. Only the first 2.5 s of the response after each switch are shown.

(C) Filter width under high mean/low variance versus decay time for both low to high (red) and high to low (gray) transitions. Filter timescales and decay timescales are poorly correlated (correlation coefficients = −0.2 for low to high mean and 0.04 for high to low mean), suggesting that filter width and decay time represent independent forms of temporal sensitivity. The two example cells are shown by a blue circle and red square; open symbols represent low to high mean transitions, and filled symbols represent high to low mean transitions.

(Figures 4A–4C, second column). Figure 2G shows the 50% width of the absolute value of each filter (50% width) versus the peak of each filter's power spectrum (its "best modulation frequency," or BMF). Each parameter has an approximately uniform distribution along its log axis, suggesting a greater than 10-fold spread in temporal frequency sensitivities among cells.

The nonlinearity (Figures 2B and 2E) describes the relationship between filter output and firing rate. It is calculated by passing the mean-subtracted modulation signal through the filter—measuring the similarity of the stimulus to the filter at each point in time—then comparing the filter output to the neuron's actual firing rate (Experimental Procedures). Most of the cells we recorded had nonlinearities with a flat, subthreshold region at negative values and a linear coding region at positive values, when the stimulus most closely resembled the filter. Some, like that in Figure 2B, showed saturation at high positive values.

To assess the quality of our models, we used the filter and nonlinearity pair to generate a prediction of each neuron's response to the repeated modulation signal (Figures 2C and 2F). We then compared this prediction (thick gray lines) to the actual PSTH (thin black line). The correlation coefficient between prediction and data for the neuron in Figures 2A–2C was 0.89, making this one of our best models. The neuron in Figures 2D–2F had a correlation coefficient of 0.67, the median for our population. The distribution of correlation coefficients is shown in Figure 2H (n = 33 significant filters from 36 recorded cells). The strength of these correlations suggests that our model captures a significant portion of the behavior of neurons in our population.

Responses of Field L Neurons to Changes in Stimulus Statistics

To examine how coding depends on stimulus statistics, we altered the mean and the variance of amplitude modulations in our stimulus (Figure 3A). In our baseline condition (gray), the modulation signal had a mean of 30 dB and a standard deviation of 6 dB. In the high mean condition (red), the mean was raised to 63 dB, while in high

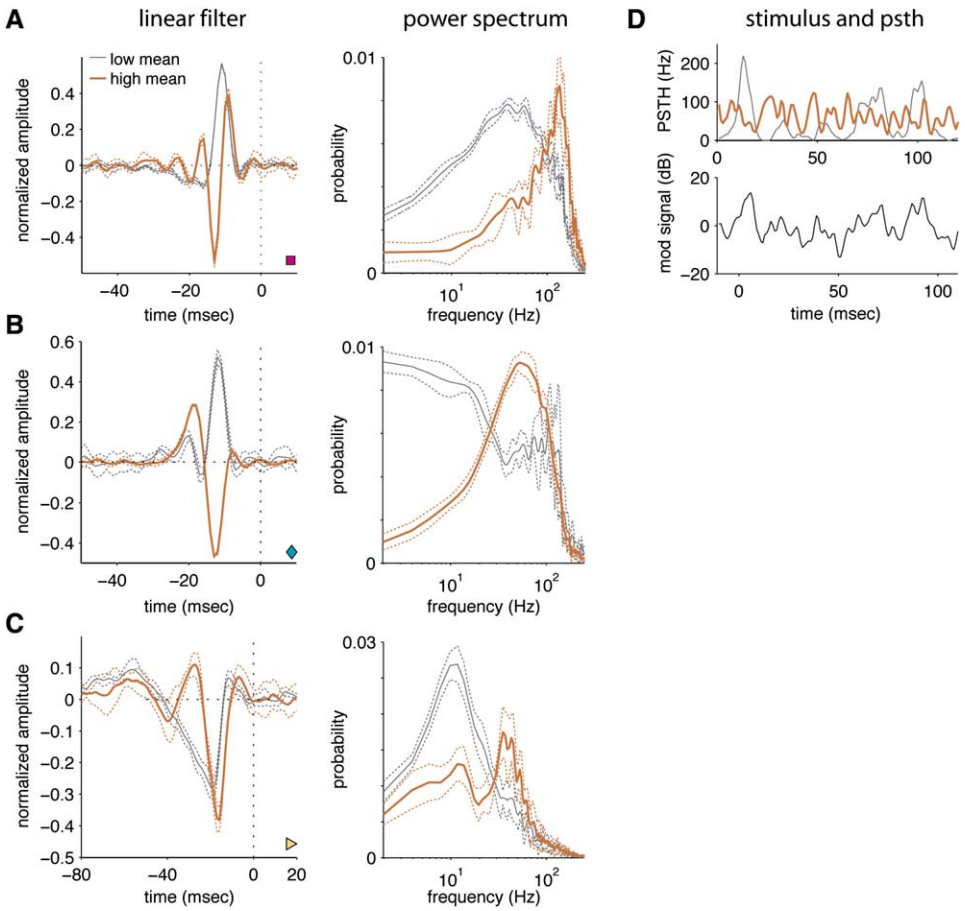

Figure 4. Filter Shape Depends Strongly on Mean Stimulus Amplitude

(A–C) Linear filters (left) and their power spectra (right) for three different cells under the low mean/low variance (gray) and high mean/low variance (red) conditions. For all three cells, the filter obtained under high mean is narrower and has positive and negative components of more similar size. Colored dashed lines surrounding each filter represent the standard deviation of five different estimates of the filter (Experimental Procedures). Black dotted lines at zero represent the time of the spike. Filters are normalized so that the dot product of each with itself is one. (D) Responses of the unit shown in (A) to 120 ms of the repeated modulation signal (black, bottom panel) under low mean (gray) and high mean (red) conditions. The stimulus has been shifted forward by 11 ms, the latency of the peak of the low mean filter, to facilitate comparison.

variance (green) the standard deviation was raised to 18 dB. The high mean and high variance conditions were designed to have approximately the same overall power (67 dB). Stimulus conditions were presented continuously, and the baseline condition was repeated twice in the series, which allowed us to characterize the transitions to and from this condition to the other two.

Changes in Stimulus Mean Produce Systematic Changes in Firing Rate and Filter Shape

Most of the cells we recorded responded strongly to a change in stimulus mean, with a rapid change in firing rate followed by a slow decay (Figure 3B). The rates of decay varied across cells from approximately 4 ms to 4 s and were generally much slower than would be predicted by the width of the filters. Decay times were poorly correlated with filter width (Figure 3D, cc = −0.20 for low to high mean and 0.04 for high to low mean), suggesting that filter width and decay time represent two separate forms of temporal sensitivity—one on the timescale of syllables, and one on the timescale of motifs or bouts. High and low mean decay times were weakly correlated (see Figure S3 in the Supplemental Data available online).

An increase in stimulus mean produced systematic changes in filter shape. Figure 4 shows three examples of filters derived from low mean (gray) and high mean (red) conditions for the same cells. All three show related changes when the mean increases: the filters become narrower, and the sizes of the positive and negative components become more closely matched. These changes in filter shape are also reflected in the power spectrum representation of each filter. The narrowing of the filter in time is reflected in a shift of the power spectrum peak toward higher frequencies, while the increased balance between positive and negative components reduces the response to low frequencies.

The cell shown in Figure 4A was typical, with a narrower filter and a larger negative component at high mean, but a similar shape—a negative peak followed by a positive one—in both conditions. Figure 4D illustrates the consequences of this filter change by comparing a segment of the cell's response under low mean (gray) and high mean (red). Under low mean, the

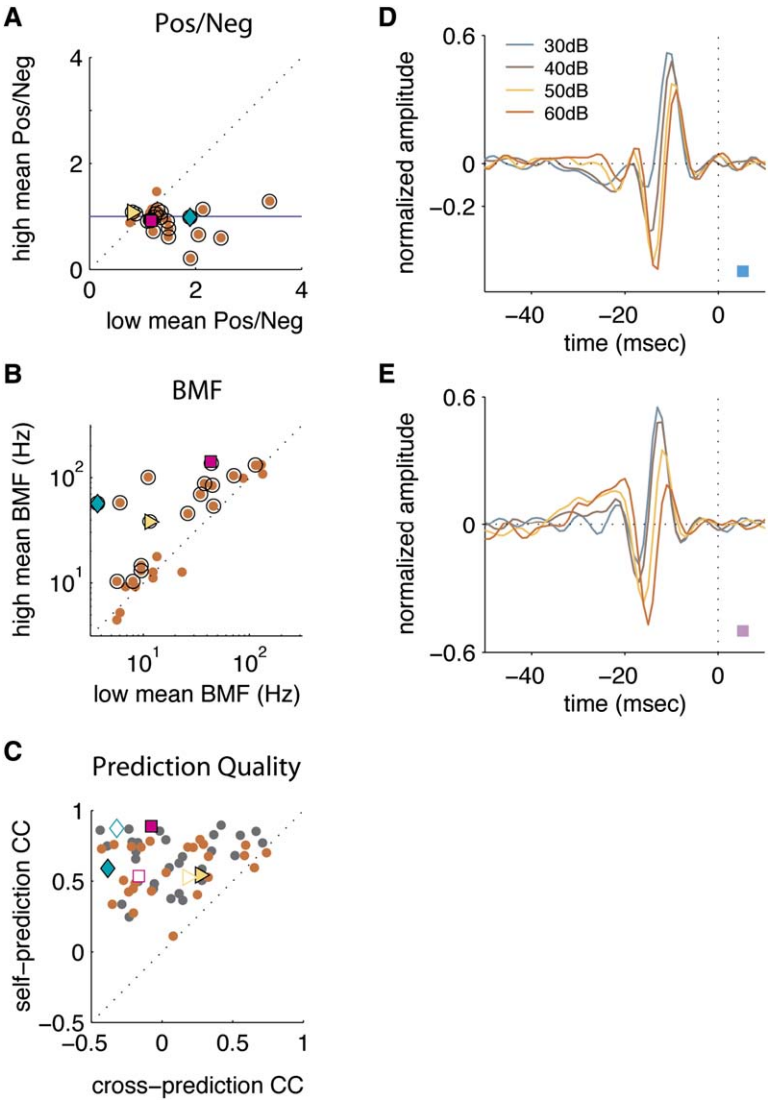

Figure 5. Population Analysis of Mean Effects on Filters

(A) Ratio of positive to negative areas of filters (Pos/Neg) from high mean (y axis) versus low mean (x axis) conditions. Most points lie below the dotted line, indicating that negative components of the filter are larger under the high mean condition. Black circles indicate filters that showed a significant change in Pos/Neg between the two conditions (n = 21/28, $p < 0.05$). The mean Pos/Neg ratio across the population was significantly lower under high mean (0.97 ± 0.04, SE) than under low mean (1.41 ± 0.11, $p = 5.9e^{-4}$) and was not significantly different from one (black line) at high mean, indicating balanced positive and negative components. The three example cells shown in Figure 4 are indicated by the pink square (A), blue diamond (B), and yellow triangle (C).

(B) BMF under high mean/low variance (y axis) versus low mean/low variance (x axis) conditions. Most points lie above the dotted diagonal line, indicating higher BMF under the high mean/low variance condition. Black circles indicate significant changes (n = 16/28, $p < 0.05$). Example cells are indicated by the filled symbols.

(C) Filters derived from data with the same mean make much better predictions (y axis) than filters from data with a different mean (x axis). Red dots indicate correlation coefficients between data and predictions based on high mean filters (mean correlation coefficient = 0.59 ± 0.03 [SE] for same statistics, 0.03 ± 0.07 for different statistics, $p = 4.0e^{-9}$). Gray dots indicate correlation coefficients between data and predictions based on low mean filters (mean correlation coefficient = 0.66 ± 0.03 for same statistics, 0.08 ± 0.07 for different statistics, $p = 2.3e^{-10}$). Data for the example cells are indicated by open (high mean filter predictions) and filled (low mean filter predictions) symbols.

(D and E) Linear filters for two cells at four different mean levels: 30 dB (gray), 40 dB (brown), 50 dB (yellow), and 60 dB (red), with a standard deviation 5 dB. As the stimulus mean increases, the negative component of each filter grows larger and decreases in latency, leading to substantial changes in filter shape.

response resembles a smoothed version of the stimulus: larger peaks occur only when the stimulus is above its mean. Under high mean, the response contains many more rapid peaks, reflecting the faster frequency tuning and narrower shape of the filter under this condition. The peaks in this condition are of similar size throughout the segment, showing that the neuron has largely filtered out the slow modulations and responds more to local changes in stimulus intensity.

The cell in Figure 4B shows a more dramatic change in filter shape: under low mean conditions, it has a positive integrator shape, while under high mean conditions, it adopts an offset differentiator shape. The neuron shown in Figure 4C was the least common. Its filter is almost entirely negative at low mean but gains a small positive component at high mean.

Together, these examples suggest that when the mean sound amplitude is low, neurons integrate over a longer time, while when the mean is high, the same cells respond more to local changes. These different response properties can improve the cells' ability to signal effectively in different stimulus regimes (Atick, 1992) and can help prevent saturation when sounds are loud. To determine whether these findings were true across the population, we measured two parameters of the filters: the ratio of positive to negative parts of the filter (Pos/Neg) and the best modulation frequency (BMF).

Figure 5A shows the ratio of positive to negative parts of the filter under high mean (y axis) versus low mean (x axis; n = 28 cells with significant filters under both conditions). Most points lie below the diagonal, indicating that most filters have larger negative components at high mean. The small number of cells that show an increase in Pos/Neg represent filters like the one shown in Figure 4C that are predominantly negative under low mean conditions and gain a positive component under

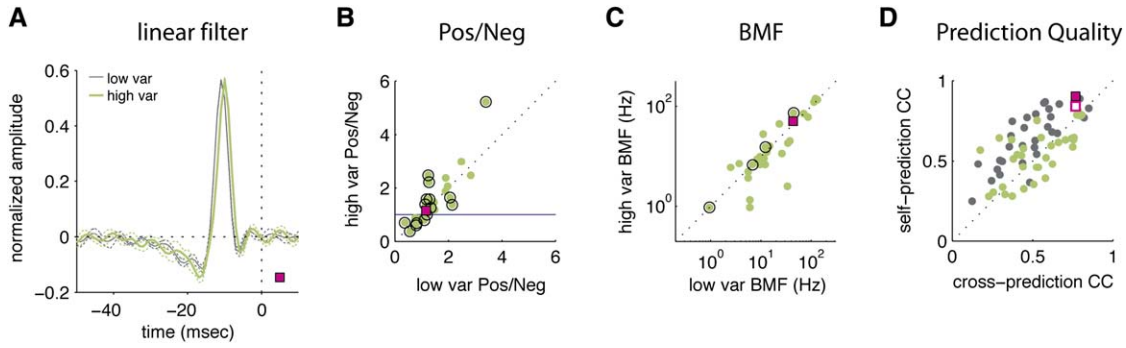

Figure 6. Unlike a Change in Mean, a Change in Variance Does Not Lead to Significant Changes in Filter Shape

(A) Linear filters for the cell shown in Figure 4A under low mean/low variance (gray) and low mean/high variance (green) conditions. Dashed lines surrounding each filter represent the standard deviation of five different estimates of the filter. The increase in variance has no significant effect on filter shape.

(B) The balance of positive and negative components across the population shows no systematic change with variance. Ratio of positive to negative areas in filters derived under low mean/high variance (y axis) versus low mean/low variance (x axis) conditions. Black circles indicate filters that showed a significant change in Pos/Neg between the two conditions (n = 15/32, $p < 0.05$). The mean population ratio of positive to negative areas was not significantly different between high variance (1.53 ± 0.15 SE) and low variance (1.43 ± 0.10) conditions ($p = 0.20$). The example cell shown in (A) is indicated by a red square.

(C) Best modulation frequencies across the population do not change significantly with variance. BMF under low mean/high variance (y axis) versus low mean/low variance (x axis) conditions. Black circles indicate significant changes (n = 4/32). The example cell is indicated by a red square. Black circles here and in Figure 7D indicate nonlinearities that show significant changes, $p < 0.05$.

(D) Filters derived from low variance and high variance conditions make similar predictions. Green dots indicate correlation coefficients between data and predictions based on high variance filters (mean correlation coefficient = 0.54 ± 0.03 [se] for same statistics, 0.56 ± 0.03 for different statistics, $p = 0.41$). Black dots indicate correlation coefficients between data and predictions based on low variance filters (mean correlation coefficient = 0.66 ± 0.03 for same statistics, 0.52 ± 0.03 for different statistics, $p = 6.5e^{-8}$). Data for the example cell are indicated by open (high variance filter prediction) and filled (low variance filter prediction) squares.

high mean. Most points lie close to the line y = 1, which represents balanced positive and negative components under high mean.

To examine whether neurons as a population became sensitive to faster modulations at high mean, we plotted the BMF for all cells under high mean versus low mean (Figure 5B). In this graph, most points lie above the diagonal, indicating that most cells are more sensitive to faster modulations at the high mean. The circled points represent cells that showed a significant increase in best modulation frequency (n = 16/28, one-tailed t test, $p < 0.05$). Together these data suggest that an increase in stimulus mean leads to filters that behave more like differentiators and to a shift in the sensitivity of cells toward high temporal frequencies.

Are the differences we observe in filter shape due to true differences in responses, or do they represent different approximations to the neuron's response? To answer this question, we compared predictions made by filters with the same mean to predictions made by filters with different means. In all cases, data used to test the model were distinct from data used to generate the model.

Filters generated under both conditions performed similarly when predicting responses to novel stimuli of the same condition: the mean correlation coefficient between data and prediction was 0.59 ± 0.18 (SD) under high mean and 0.66 ± 0.18 under low mean. However, both sets of filters were much worse at predicting responses to the other condition. Figure 5C shows the correlation coefficients between predicted and actual PSTHs using filters with the same mean (self-prediction CC) versus filters with a different mean (cross-prediction CC). Self-predictions are significantly better than cross-predictions for both high and low mean data. This confirms that the changing filter shapes capture a real change in the features to which the neurons respond.

Filters Change Gradually as a Function of Mean

To better understand how changes in filter shape occur—particularly the most dramatic switches from onset to offset selectivity—we varied the stimulus mean randomly among four levels (30, 40, 50, or 60 dB) while holding the variance at 5 dB. Figures 5D and 5E show filters calculated under these four conditions for two cells—one that retains an onset shape across levels and one that switches gradually from onset sensitivity at low mean (gray trace) to offset or acceleration sensitivity (indicated by a triphasic filter) at high mean (red trace). In both cells, a negative peak that is small at 30 dB grows larger and decreases in latency as the mean increases. This suggests that even very striking changes in filter shape might arise through simple mechanisms, such as a change in the relative strength and latency of inhibitory inputs to the cell. Population data from 13 cells (Figure S3) support the finding that filters change gradually and systematically as a function of mean stimulus amplitude.

Changes in Stimulus Variance Do Not Produce Systematic Changes in Filter Shape

In contrast to the effect of mean amplitude, changing the stimulus variance had little effect on mean firing rate or filter shape. Figure 6A shows the effects of an increase in standard deviation from 6 to 18 dB on the cell shown in Figure 4A. Although the change in variance produces

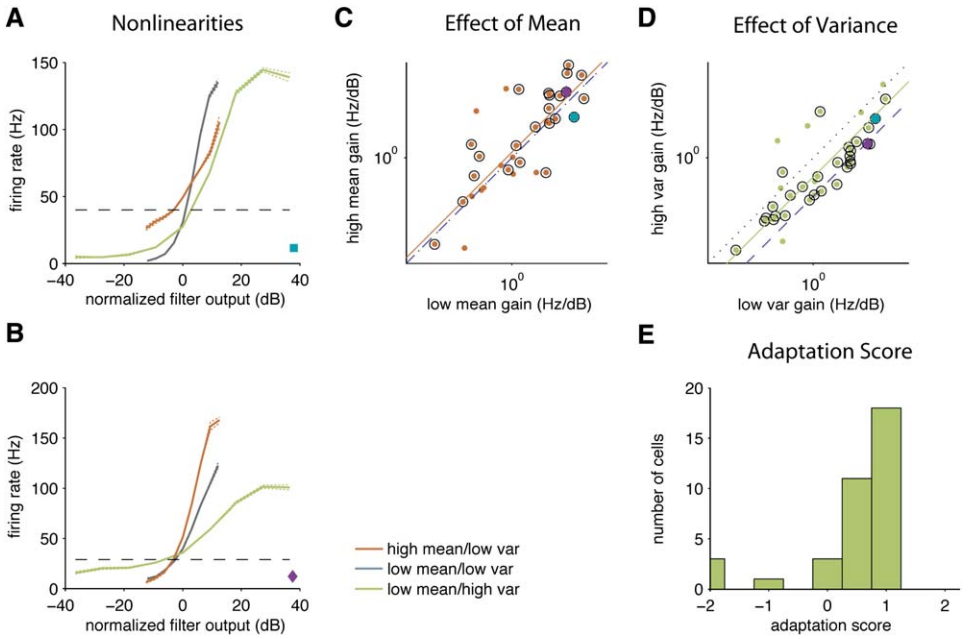

Figure 7. Effects of Mean and Variance on Nonlinearities

(A and B) Nonlinearities for two cells, calculated under three conditions: low mean/low variance (gray), high mean/low variance (red), and low mean/high variance (green). An increase in mean increases gain in one cell (B) and decreases it in the other (A). An increase in variance decreases gain in both cells. Colored dotted lines represent the standard deviation of five jackknife estimates of the nonlinearity (see Experimental Procedures). The black dashed lines represent the spontaneous firing rates of each cell.

(C) An increase in mean produced no net change in gain across the population. Nonlinear gain (Hz/dB) under high mean/low variance (y axis) versus low mean/low variance (x axis) conditions. The dotted diagonal line indicates no change. The red line represents the average population ratio of 1.17. It is not significantly different from one ($p = 0.32$).

(D) An increase in variance produced a systematic decrease in gain across the population. Nonlinear gain under low mean/high variance (y axis) versus low mean/low variance (x axis) conditions. The green line represents the average population ratio of 0.58. It is significantly less than 1 ($p = 3.7e^{-5}$). The dashed blue line represents a decrease in gain proportional to the increase in the standard deviation.

(E) Distribution of adaptation scores (see Experimental Procedures) for all cells comparing low mean/low variance (gray) to low mean/high variance (green, n = 32). An adaptation score of 1 indicates that gain under high variance was exactly 1/3 of gain under low variance. 0 indicates no change in gain between conditions.

the same increase in stimulus power as did the change in mean, the cell shows little significant change in filter shape.

Across our population, correlation coefficients between high variance and low variance filters (mean = 0.84 ± 0.03, SE) were significantly higher than the correlation coefficients between high mean and low mean filters (0.12 ± 0.10, $p = 5.0e^{-8}$), confirming that variance had much less effect on filter shape than did mean amplitude. Changing stimulus variance also had fewer effects on the ratio of positive to negative filter components (Figure 6B) and on best modulation frequency (Figure 6C).

Similar to the filters generated under high mean, filters generated under high variance did well at predicting responses to stimuli with the same statistics. The mean correlation coefficient between data and prediction was 0.54 ± 0.17 (SD) under high variance and 0.66 ± 0.18 under low variance. However, predictions made by swapping filters—using high variance filters to predict low variance data and vice versa—were only slightly worse than predictions made with matched filters (Figure 6D). Although an increase in variance caused changes in the shapes of some filters, these changes were smaller, less common, and less systematic than those observed with an increase in mean.

Effects of Mean and Variance on Nonlinearities

Although variance had only small effects on filter shape, it had a significant effect on neural gain. To examine the effects of stimulus mean and variance on the gain of the neural response, we calculated the nonlinear relationship between the filtered stimulus and the instantaneous firing rate under each condition (Experimental Procedures). Filters were normalized such that the variance of their output was equal to the variance of their input (Baccus and Meister, 2002) to ensure that changes in gain were not due to changes in filter shape.

Figures 7A and 7B show the effects of increased mean and variance on the nonlinearities of two cells. An increase in mean (high mean [red] versus low mean [gray]) led to a decrease in the gain of the cell in Figure 7A and an increase in the gain of the cell in Figure 7B. An increase in variance (high variance [green] versus low variance [gray]) led to a decrease in the gain of both cells. These decreases in gain with variance may be considered adaptive, as they allow the cells to maintain a similar range of firing rate fluctuations when the range of stimulus fluctuations increases, and so to take better advantage of their signaling capacity.

The trends that we observed in these examples continued across our population. An increase in stimulus mean led to both increases and decreases in gain,

with no net change in gain across the population. An increase in stimulus variance led to systematic decreases in gain. To quantify changes in gain, we calculated the average slope of each nonlinearity, excluding subthreshold and saturation regions where the slope fell below 5% of its maximum (Experimental Procedures). Figure 7C shows slopes under high mean versus low mean: points are distributed on both sides of the diagonal, indicating no systematic change in gain. Figure 7D shows slopes under high variance versus low variance. Here, most points (27/32) lie below the diagonal, indicating a decrease in gain.

The fly H1 neuron has been shown to reduce its gain in proportion to the increase in the standard deviation of the stimulus (Brenner et al., 2000; Fairhall et al., 2001). If our cells adapted proportionally, the increase in mean should produce no change in gain, but the increase in variance should decrease gain by a factor of 1/3 (blue dashed line in Figure 7D), to compensate for a 3-fold increase in the standard deviation.

To quantify the amount of variance adaptation in each cell, we calculated an adaptation index (see Experimental Procedures). The adaptation index is 0 if there is no change in gain with variance and 1 if the gain under high variance is exactly 1/3 of the gain under low variance. The distribution of adaptation scores is shown in Figure 7E and is highly skewed (median = 0.74): many of the cells show close to proportional adaptation, while some show less than proportional adaptation. A few cells showed no adaptation or adaptation in the opposite direction.

Time Course of Changes in Filter Shape and Gain

Our analysis of field L response properties in different statistical conditions indicates that their encoding changes with both mean and variance. Are these changes due to time-dependent mechanisms, or do they represent fixed nonlinearities, for example differences in the relative strength of excitatory and inhibitory inputs at different intensities? To examine this question, we looked at the time course of changes in filter shape and gain.

To ask how fast filter shapes changed, we calculated filters from different epochs of the response before and after the step change in mean. Figure 8A shows filters derived from three epochs surrounding the switch from low to high mean: the last 500 ms of the low mean response (black), the first 100 ms after the switch to high mean (purple), and the last 500 ms of the high mean response (gray). Although the mean firing rate is still changing during the first 100 ms after the switch, the filter derived from these data (purple) has the narrower shape and larger negative component characteristic of the fully adapted high mean filter (gray); it is quite different from the wide integrator shape of the filter under low mean (black).

To ask whether filters changed this fast across our population, we calculated correlation coefficients between filters from these three epochs (Figure 8B). Correlation coefficients between early (first 100 ms) and late (last 500 ms) filters from the same condition were significantly greater than correlation coefficients between early filters and filters from the end of the previous condition. Correlation coefficients between early and late filters were not significantly different from correlation coefficients between two fully adapted filters taken 4500 ms and 4900 ms after the switch. Together, these data suggest that filters change shape within 100 ms after a change in mean stimulus level.

To look at the time course of changes in gain, we calculated nonlinearities from different epochs before and after a change in variance. Figure 8C shows nonlinearities calculated during three epochs around the switch from low to high variance. The nonlinearity calculated from the last 500 ms of the low variance condition has a high gain (black), while nonlinearities taken from the first 100 ms (purple) and last 500 ms (gray) of the high variance response both show a lower gain.

This rapid change in gain can also be seen in the PSTHs of this cell's response to repeated trials. If the decrease in gain occurred slowly, we would expect firing rate fluctuations under high variance to be initially much larger than firing rate fluctuations under low variance. As shown in Figure 8D, firing rate fluctuations under high and low variance are nearly identical by the time of the first large peak, about 50 ms after the switch, indicating that the gain of the cell has decreased by this time.

To quantify the speed of gain change across the population, we compared the gain of nonlinearities taken from the three epochs of the response. The gain of nonlinearities derived from the first 100 ms of data were significantly different from the gain of nonlinearities under the previous condition ($p = 3.7e^{-5}$ for increased variance, $p = 2.5e^{-5}$ for decreased variance) but were not significantly different from the gain measured later under the same condition ($p = 0.75$ for increased variance, $p = 0.43$ for decreased variance). Together, these data indicate that nonlinear gain changes within 100 ms of a change in stimulus variance.

Our analyses indicate that changes in filter shape and nonlinear gain occur quickly, well before the mean firing rate has finished adapting. This suggests that gradual firing rate adaptation corresponds to a change in the threshold or set-point of firing, rather than to a change in gain or in temporal feature selectivity. To ask whether firing rate adaptation is correlated with a change in threshold, we plotted nonlinearities from several epochs around the switch from low to high mean. Figure 8E shows these nonlinearities for a single cell. Following the increase in mean, the nonlinearity undergoes a rapid shift upward and to the left (purple versus thin black), as well as a small change in slope. The shift occurs immediately and is probably due to the change in filter shape rather than to a change in neural sensitivity. Over the next 500 ms, the nonlinearity moves gradually back toward the center (violet and light blue), but does not change slope. The slow decay in firing rate thus seems to determine the shifting position of the nonlinearity.

To confirm that this trend was consistent across our population, we calculated the correlation between the mean firing rate during each 100 ms long epoch and the y intercept of the nonlinearity measured during that epoch. The average correlation coefficient across cells was high—0.82 ± 0.02 (SE)—suggesting that firing rate adaptation is best considered as a change in the position or set-point of the nonlinearity.

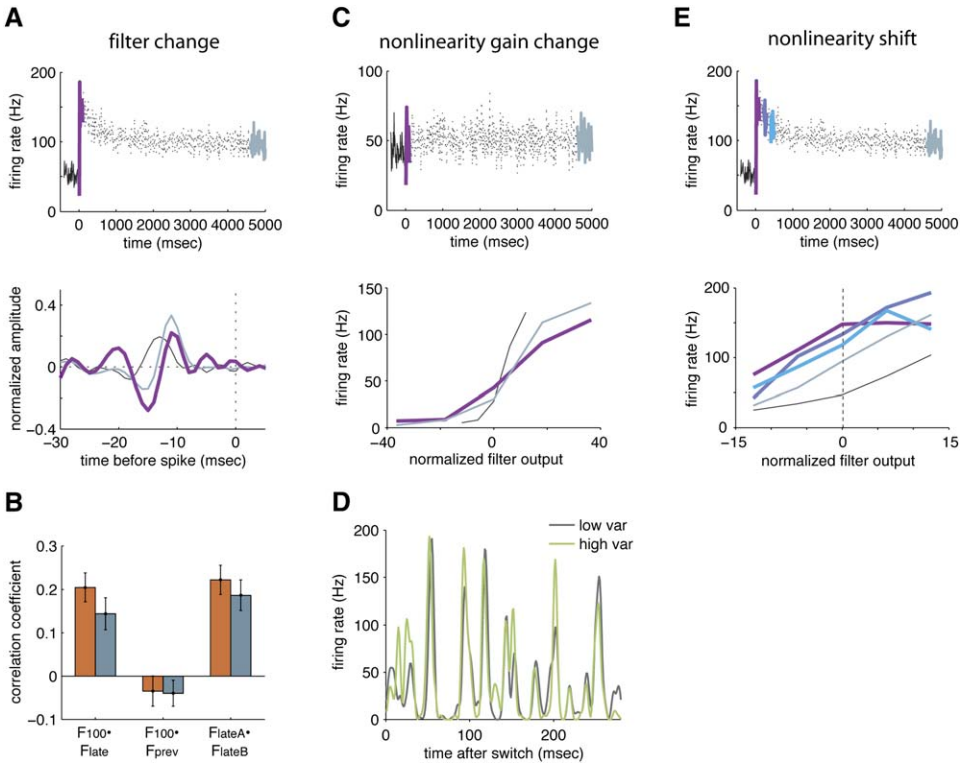

Figure 8. Changes in Filter Shape and Gain Both Occur within 100 ms of a Change in Stimulus Statistics

(A) An example of rapid filter change. (Top) PSTHs showing the mean firing rate of a single cell around the switch from low to high mean. Black, last 500 ms of low mean; purple, first 100 ms of high mean; gray, last 500 ms of the high mean. (Bottom) Filters derived from each of the three epochs shown above. The filter derived from the first 100 ms of high mean (purple) resembles the filter from the last 500 ms of high mean (gray) much more than it does the filter from the last 500 ms of low mean (black).

(B) Comparisons of filter similarity across the population. First column: average correlation coefficient between filters derived from the first 100 ms and last 500 ms of the same condition (red: high mean, 0.20 ± 0.03, SE; gray: low mean, 0.14 ± 0.04). Second column: average correlation coefficient between filters derived from the first 100 ms of one condition and from the last 500 ms of the previous condition (red: low mean to high mean transitions, −0.03 ± 0.03; gray: high mean to low mean transitions, −0.04 ± 0.03). Correlation coefficients between early and previous filters were significantly smaller than those between early and late filters ($p = 3.0e^{-5}$ for high mean, $p = 1.3e^{-3}$ for low mean). Third column: average correlation coefficient between filters derived from two 100 ms epochs at the end of the same condition (red: high mean, 0.22 ± 0.03; gray: low mean, 0.19 ± 0.04). Correlation coefficients between early and late filters were not significantly different from those between two late filters ($p = 0.33$ for high mean, $p = 0.15$ for low mean). These data indicate that filters from the first 100 ms after the switch were significantly different from filters under the previous condition, but were no more different from the fully adapted filters than two fully adapted filters were from each other.

(C) An example of rapid change in nonlinear gain. (Top) PSTH showing the mean firing rate of a single cell around the switch from low variance to high variance. Black, last 500 ms of low variance; purple, first 100 ms of high variance; gray, last 500 ms of high variance. (Bottom) Nonlinearities derived from each epoch after the switch from low variance to high variance. The nonlinearity derived from the first 100 ms (purple) and last 500 ms (gray) of high variance both have a lower gain than the low variance nonlinearity (black).

(D) Rapid gain adaptation is evident in responses to repeated trials. PSTH of the response of the neuron in (C) to the first 280 ms after the switch from low to high variance (green), and after the switch from high to low variance (gray). The peaks are approximately the same within 50 ms after the switch, indicating that the cell is already compensating for the difference in the magnitude of stimulus fluctuations.

(E) Slow decays in firing rate represent shifts in the nonlinearity. Nonlinearities from different epochs after the switch from low to high mean. Black, 500 ms preceding the switch; purple, 0–100 ms after the switch; violet, 200–300 ms after the switch; light blue, 400–500 ms after the switch; gray, last 500 ms of the high mean response. Immediately after the switch from low mean to high mean, the nonlinearity shifts up and to the left, then moves slowly back down and to the right.

Discussion

The goal of our study was to investigate how coding of complex amplitude modulations in a primary auditory area depends on the distribution of stimulus intensities. To characterize coding, we developed a stimulus that mimics several features of natural song stimuli while still permitting a randomized and systematic search of possible stimuli. From responses to these stimuli, we were able to generate linear-nonlinear models that successfully predict many features of the neurons' responses to novel stimuli.

The filters we extracted from our neurons revealed sensitivity to a range of temporal features and timescales. Peak temporal frequency sensitivities varied from 10 to 150 Hz. Neurons showed sensitivity to many combinations of onsets, offsets, and continuous stimulation. This range of sensitivities may be important for tracking the complex contours of zebra finch syllables. Previous studies in this area found a lower range of temporal frequency sensitivities and a preponderance of onset units (Sen et al., 2001; Woolley et al., 2005). Our results support findings in other species showing a broader range of temporal frequency sensitivities

and response types in unanesthetized animals (Liang et al., 2002; Wang et al., 2005).

To examine how coding changes with the distribution of stimulus amplitudes, we altered the mean and the variance of amplitude modulations in our stimulus. Generating linear-nonlinear models under three different statistical conditions revealed that changes in mean and variance led to distinct types of adaptive changes in coding.

The first adaptive change that we described was a remodeling of filter shape that depends on the mean amplitude of the stimulus. Under low mean conditions, neurons act more like low-pass integrators. As the stimulus mean increases, negative components of the neural filter grow stronger and decrease in latency, causing neurons to behave more like band-pass differentiators. This transition could help prevent firing rate saturation at high sound levels. It could also allow cells to signal effectively over a wider range of stimulus levels. The amplitudes of natural sounds tend to be similar at nearby points in time (Attias and Schreiner, 1998; Singh and Theunissen, 2003), so responding only to changes in amplitude—which are more rare—may be efficient at high signal-to-noise levels. At low signal-to-noise levels, however, integrating over a longer time can improve the chance of detecting a quiet sound.

Several groups (Woolley et al., 2005; Narayan et al., 2005) have shown that analyzing songs with band-pass or differentiating filters allows for better theoretical discrimination of different individuals' songs. However, these studies only analyzed songs under high signal-to-noise conditions. It would be interesting to ask whether low-pass filters permit better song discrimination when songs are soft or noise is prevalent.

Changes in filter properties similar to those we describe have also been observed in retinal ganglion cells. At low light intensities, ganglion cells lose their receptive field surrounds and increase their integration times (Enroth-Cugell and Lennie, 1975). Atick (1992) used information theory to show how these changes in receptive field structure can maximize information transmission in bright versus dim conditions. Similar principles should apply to the changes we observe in auditory receptive field structure when the volume of an auditory signal changes.

The second adaptive change we described was a reduction in gain when the stimulus variance increased. This reduction in sensitivity allows the cell to match its range of outputs more closely to its range of inputs. Similar properties have been described in the auditory midbrain (Dean et al., 2005) and in several areas of the visual system, including the retina (Chander and Chichilnisky, 2001; Baccus and Meister, 2002), the LGN (Mante et al., 2005), and the fly H1 neuron (Brenner et al., 2000; Fairhall et al., 2001). In the fly, gain adaptation is proportional to the increase in the standard deviation of a velocity stimulus. In our system, neurons showed a distribution of variance adaptation. Many cells showed close to proportional gain adaptation, while some showed weaker adaptation, and a few showed no adaptation or adaptation in the opposite direction. The difference between the fly motion-detection system and the songbird auditory system may be related to the number of neurons involved. In the fly, a single neuron must encode visual motion under all behavioral conditions, while in the songbird, a population of neurons may allow some specialization for different statistical regimes.

If changes in filter shape and gain arose through some time-dependent process, we would expect them to occur slowly after a step change in stimulus statistics. Instead, we found that both these changes occur within 100 ms of the switch in stimulus statistics. In a few cells (data not shown), we attempted to estimate filters from even shorter time intervals and found that they changed essentially within the timescale of the filter—as fast as we could measure them. Fast changes in sensitivity have also been described in retinal filters (Baccus and Meister, 2002) and in the fly H1 neuron (Brenner et al., 2000; Fairhall et al., 2001). This does not rule out the possibility that these changes are due to adaptive processes operating within a few tens of milliseconds—such as fast adaptation measured in cortical cells (Nowak et al., 2003; McCormick et al., 1985) or fast adaptive processes acting in the cochlea (LeMasurier and Gillespie, 2005). However, it suggests that the changes we observed may best be described and modeled as "adaptive nonlinearities" in processing rather than as traditional time-dependent adaptations (Borst et al., 2005).

The fact that these changes occur so quickly can constrain models of how these properties might arise. One hypothesis is that the positive and negative peaks in our filters reflect excitatory and inhibitory inputs, whose latency and relative strength depend differently on the volume of the stimulus. Dramatic changes in filter shape with mean stimulus amplitude will arise if inhibitory inputs have a higher threshold than excitatory inputs and if their gain rises more steeply as a function of stimulus amplitude. This arrangement of inhibition has been described in several parts of the ascending auditory system, particularly in type IV units of the dorsal cochlear nucleus (Nelken and Young, 1994; Yu and Young, 2000).

Properties related to those we describe have also been seen in the ventral cochlear nucleus. The temporal modulation transfer functions of VCN chopper units show a transition from low-pass to band-pass temporal frequency sensitivity with increased stimulus volume (Frisina et al., 1990). In the inferior colliculus (IC), a similar transition from low-pass to band-pass sensitivity has been observed (Rees and Moller, 1987; Krishna and Semple, 2000). This transition could arise in the cochlear nucleus and be filtered to lower frequencies as it is propagated through the ascending auditory system. Alternatively, it could arise anew at each stage of processing due to inhibitory interactions like those we hypothesize.

Two other studies in the mammalian IC described properties related to those seen here. Dean et al. (2005) showed that IC neurons change their rate-level functions depending on the local distribution of stimulus amplitudes. However, because that study calculated rate-level functions by directly comparing stimulus amplitude and firing rate, it could not discriminate changes in neural sensitivity (gain) from changes in feature selectivity (filter shape). Kvale and Schreiner (2004) also saw small changes in neural gain and filter shape of IC units when the stimulus variance changed, but in contrast to our results, these changes occurred gradually during adaptation. We think it is likely that the adaptive

properties we describe first arise early in the ascending auditory system but may be augmented by processing at many stages.

In addition to the filter and gain changes, we observed a range of slow decays in firing rate following a change in the mean amplitude of the stimulus, similar to the wide range of decay times described in auditory cortex by Ulanovsky et al. (2004) and by Bartlett and Wang (2005). The rate of decay was not related to the width of each cell's filter, nor did this decay arise from a change in filter shape or gain—because both of those changes occurred very quickly. Instead, slow decays appear to be related to a shift in the setpoint or threshold of firing: as time progresses, a fixed amount is subtracted from the neuron's response to the same feature.

A striking feature of our results is their similarity to findings in the visual system (e.g., Enroth-Cugell and Lennie, 1975; Baccus and Meister, 2002), particularly in the retina, where techniques similar to ours have been most employed. In both systems, neural filters become more differentiating as stimulus intensity increases. In both systems, neural gain decreases with increased stimulus contrast. And in both systems these changes occur rapidly, perhaps within the timescale of filtering, and independent of slower decays in firing rate. Although the neurons we studied lie at a very different stage in the sensory hierarchy from these retinal cells, they seem to share common computational strategies. These may represent solutions to common problems faced by all sensory systems that must represent a wide dynamic range of signals.

The changes in coding that we describe may help neurons to accurately encode stimulus characteristics in very different regimes, but they raise problems for current models of encoding and decoding natural stimuli. For instance, several existing models of primary auditory cortex assume that spectro-temporal features are encoded by a bank of linear filters (Chi et al., 2005). Models of complex song-selective neurons in higher areas of the avian brain also assume that the inputs to these cells are static and largely linear (Margoliash, 1983; Lewicki and Konishi, 1995; Drew and Abbott, 2003). In contrast, our study suggests that the inputs to higher-order auditory areas are highly nonlinear but change in systematic ways that could be effectively modeled. A model that incorporates the nonlinearities we describe should not only make better predictions about how primary neurons respond to complex stimuli but will also be crucial to understanding how the brain constructs high-level feature detectors such as song- or face-selective neurons.

Experimental Procedures

Chronic Recording and Electrophysiology

We used chronically implanted microdrives (Hessler and Doupe, 1999) to record single units from five adult male zebra finches. A detailed description of microdrive construction and implantation is given in that paper. Electrodes (two to three tungsten electrodes, 4–5 MOhms, MicroProbe Inc, Gaithersburg, MD) were implanted 1.5 μm lateral (left) and 1.5 μm rostral to the posterior border of the branch point of the central sinus, at an initial depth of 400 μm.

During recording, the bird was attached to a commutator by a flexible lead and op-amp. Electrical traces were digitized, amplified (1000×), filtered (300–5000 Hz), and recorded using TDT System 3 hardware (Tucker-Davis Technologies, Alachua, FL) interfaced with custom-written Matlab software. The electrode bundle was advanced manually in small steps (40–160 μm). Putative single units were identified on the oscilloscope by their stable spike waveform and clear refractory period. All spikes were resorted offline using a custom-written software window discriminator (Matlab) based on the similarity of overlaid spike waveforms and on clustering of waveform projections in a two-dimensional principal component space. Neural recordings were considered single units if they contained fewer than one violation of 1 ms refractoriness per thousand spikes after sorting (Figure S1). Units that responded to auditory stimuli were found at depths of 1000–2500 μm. Between each recording session, the electrodes were retracted to a position above where the first auditory units were found.

After the final recordings, histological sections were prepared to confirm that electrode tracks, and in some cases marker lesions, were located in field L. Birds were lethally anesthetized and perfused with saline followed by 4% paraformaldehyde. Alternate 40 μm sections of fixed brain tissue were Nissl-stained and labeled for enkephalin (mouse anti-leucine enkephalin monoclonal antibody, Accurate Chemical & Scientific Corporation), a marker for the nucleus interface (Nif), which abuts the anterior end of field L2a. All sites were identified to be in field L layers L1, L2, and very occasionally L3. Although the preferred frequency and linearity of cells varied across these areas as described in Sen et al. (2001), we saw no systematic differences in the sensitivity to stimulus statistics across layers.

Stimulus Presentation

During recording, the bird was placed inside a small cage (20 cm × 20 cm floor area) within a sound-attenuating chamber (Acoustic Systems). The chamber lights were kept off to minimize movement, and birds were monitored using an infra-red camera. Birds generally sat in one corner of the cage for the duration of the experiment although the commutator permitted free movement within the cage. Auditory stimuli were presented free-field from a small speaker (Bose) located 50 cm from the center of the bird's cage. Using a calibrated microphone (B&K) we verified that 250 Hz to 10 kHz tones designed to play at 80 dB appeared at 79.6 ± 2.3 dB, and that the highest harmonic distortion peak observed was less than 23.3 dB (65.9 ± 8.7 dB SNR).

Stimulus Construction

Although it is likely that field L neurons respond to other paramerters of acoustic stimuli (Elhilali et al., 2004), we designed our stimulus to isolate responses to modulations of the log amplitude envelope. Recent studies in field L (Gill et al., 2006) and in mammalian inferior colliculus (Escabi et al., 2003) suggest that modulations of the log envelope drive cells better and are more linearly encoded than modulations of the linear envelope. Our stimulus consisted of two parts: a slowly varying envelope with fixed statistics that was repeated exactly in every experiment, and a rapidly varying carrier that could be adjusted for the frequency preference of each cell.

The log envelope, or "modulation signal" [$n(t)$], consisted of Gaussian noise filtered to have an exponential power spectrum:

$$P(f) = e^{-f/50Hz} \quad (1)$$

After filtering, the modulation signal was normalized to have unit standard deviation and zero mean. To generate a linear voltage envelope [$E(t)$] from the logarithmic modulation signal [$n(t)$], we exponentiated it according to:

$$E(t) = 1e^{-5} \times 10^{[\mu + \sigma \times n(t)]/20} \quad (2)$$

where μ is the mean amplitude of the stimulus in dB, and σ is the standard deviation. When multiplied by a noise carrier with unit standard deviation, this produces a sound whose local amplitude in dB is given by

$$A(t) = \mu + \sigma \times n(t) \quad (3)$$

and whose overall RMS amplitude in dB is given by

$$RMS = \mu + \frac{\log(10)}{20}\sigma^2 \quad (4)$$

We presented stimuli with three different statistics: low mean/low variance (30 ± 6 dB SD, RMS = 34 dB), low mean/high variance

(30 ± 18 dB, RMS = 67 dB), and high mean/low variance (63 ± 6 dB, RMS = 67 dB). The low mean/high variance and high mean/low variance stimuli were designed to have the same RMS power. Stimulus statistics changed abruptly every 5 s, with the low mean/low variance condition appearing in between each presentation of low mean/high variance or high mean/low variance. A single trial consisted of 400 continuous 5 s segments, half repeats of the same segment, and half unique. Repeated and unique segments were randomly distributed throughout the sequence.

The carrier was a Gaussian noise stimulus, digitally created and filtered online. Upon encountering a cell, we first determined its frequency preference using broad- (500–8000 Hz), and narrowband (500 Hz wide, 750–7750 Hz center frequencies) noise bursts, presented at 60 dB SPL. If broadband noise drove the cell robustly, a broadband carrier was used (n = 21/36), otherwise we chose the narrowband carrier that best drove the cell. At several sites (n = 8) we repeated the experiment using both a narrowband and a broadband carrier. Changing the carrier altered the precise shape of the filter but did not affect our basic findings about changes with mean and variance or the distribution of filter widths. Because the carrier was varied throughout the experiment, responses to repeated trials represent the response to the modulation signal, averaged over many carriers.

We used a variation of this stimulus to characterize filters at 30, 40, 50, and 60 dB. This stimulus consisted of unique segments only. All segments had a standard deviation of 5 dB, and the order of statistical conditions was completely randomized.

Extraction of Filters and Nonlinearities

Filters were calculated in the Fourier domain according to:

$$F(\omega) = \frac{\langle s^*(\omega) r(\omega) \rangle}{\langle s^*(\omega) s(\omega) \rangle} \quad (5)$$

where $F(\omega)$ is the Fourier transform of the filter, $s(\omega)$ is the Fourier transform of the normalized log stimulus envelope $n(t)$, and $r(\omega)$ is the Fourier transform of the spike train. The * indicates the complex conjugate. The numerator in this equation is equal to the spike-triggered average, while the denominator is equal to the power spectrum of the log stimulus envelope, which by design is exponential (Equation 1). We verified that deviations of the full stimulus power spectrum from this predicted power spectrum did not significantly impact our calculations (Figure S2).

In practice, division or decorrelation by the power spectrum results in a noisy estimate of the filter, because it boosts power in high frequencies that are poorly sampled in $n(t)$. To recover meaningful filters, we placed an exponential cutoff on $F(\omega)$, given by

$$\begin{aligned} c(\omega) &= 1 \text{ for } |\omega| < \text{cutoff} \\ c(\omega) &= e^{-|\omega - \text{cutoff}|/10} \text{ for } |\omega| \geq \text{cutoff} \end{aligned} \quad (6)$$

The cutoff frequency for each cell was placed where the power spectrum of the raw spike-triggered average fell below two standard deviations of the power spectrum of an average of random spike times. The highest cutoff across conditions was used for all calculations involving a single cell. Cutoffs ranged from 23 to 167 Hz.

Filters were considered significant if at least 10 ms of the filter lay outside three standard deviations of the random-triggered average. Nonsignificant filters arose exclusively when the firing rate during a particular condition was extremely low. All but one cell recorded produced significant filters under at least two conditions. To estimate the error in our filter calculations, we divided our spikes into five random pools and calculated separate filters from each. The standard deviation of these five estimates are shown as error bars in the figures.

The nonlinearity describes the probability of spiking given a value of filter output, $P(spike|F \otimes stim)$, and was calculated using a Bayesian formula described in Brenner et al., (2000).

$$P(spike|F \otimes stim) = \frac{P(F \otimes stim|spike) P(spike)}{P(F \otimes stim)} \quad (7)$$

where $P(spike)$ is the mean firing rate over the stimulus condition. $P(F \otimes stim)$ is the distribution of the mean-subtracted filtered stimulus, and $P(F \otimes stim|spike)$ is the distribution of spike-triggered mean-subtracted filtered stimulus segments.

The slope of the nonlinearity depends on the scale of the filter F: if the amplitude of F is increased, the width of the distribution $P(F \otimes stim)$ will also increase, decreasing the slope of the nonlinearity. We normalized each filter so that the variance of its output was equal to the variance of its input (Baccus and Meister, 2002). This ensured that changes we measured in slope were not due to changes in filter shape. The mean was subtracted from each stimulus prior to calculating the nonlinearity to minimize a leftward shift due to the decrease in the integral of the filter with increased mean. To ensure that the nonlinearity was well sampled, we restricted our analysis to a region from two standard deviations below the mean of the filtered stimulus distribution, to two standard deviations above. To measure the reliability of the nonlinearity, we performed a jackknife operation, where 1/5 of the spikes were excluded in each of five estimates of the nonlinearity. Error bars on nonlinearities shown in figures represent the standard deviation of the jackknife estimate (Sen et al., 2001):

$$std = \sqrt{\frac{n-1}{n} \sum_j (x_j - \langle x_j \rangle)^2} \quad (8)$$

where n is the number of jackknife estimates, and x_j is the jth estimate of the nonlinearity. To calculate nonlinearities as a function of time for a single condition, we used the filter obtained by pooling all data from that condition, as our analysis showed that filters do not change over this time period.

To predict responses to the repeated modulation signal, we convolved it with the neural filter, quantized the result, then transformed the quantized signal according to the nonlinearity. PSTHs were obtained by smoothing the mean spike count per 1 ms bin with an 8 ms wide hanning window. When comparing actual and predicted PSTHs, we omitted the first 500 ms when mean firing rates were strongly adapting, as our filters were not designed to capture this response feature. The quality of the prediction depended strongly on the amount of data recorded.

Analysis of Filters and Nonlinearities

Parameters of filters were defined as follows: the 50% width was the duration in milliseconds over which the absolute value of the filter was greater than or equal to half its maximum absolute value. The best modulation frequency (BMF) of a filter was the frequency at which its power spectrum was maximum. To minimize noise, power spectra were calculated from a segment of each filter, from 25 ms before the absolute value of the filter reached 25% of its maximum, to the time of the spike. The ratio of positive to negative filter components (Pos/Neg) was obtained by dividing the sum of all positive parts of the filter by the absolute value of the sum of all negative parts of the filter. Each parameter was measured on five estimates of the filter to obtain error bars.

To eliminate subthreshold and saturation regions from our estimates of nonlinear gain, values of the nonlinearity slope less than 5% of the maximum gain under any condition were excluded prior to averaging. To compare gain across conditions, we compared the log ratio of gain under the two conditions to zero.

Analysis of Decay Times

To estimate the decay time following a change in mean, we fit a single exponential of the form

$$r(t) = A + \Delta R \times e^{(-t/\tau)} \quad (9A)$$

to the unsmoothed PSTH. In this formula, $r(t)$ is the PSTH, ΔR is the magnitude of the decay, τ is the decay time constant, and A is the steady-state firing rate. Although several decays exhibited complex transients, we focused our analysis on the final slow decay. To avoid mistakes in fitting introduced by transients, we fit only a portion of the response:

$$r(t_{lat} : end) = A + \Delta R \times e^{(-t/\tau)} \quad (9B)$$

where t_{lat} was a latency parameter from 1 to 500 ms. Fits for latency values that did not converge were excluded. Of the remaining fits, we chose the one that resulted in the smallest mean squared error

between the actual PSTH and the exponential model. A few PSTHs were fit with only a subset of latency values (n = 4). Decay time constants were considered significant only if they plateaued within our 5 s trial (all cells but one) and if they decayed over a range of greater than 5 Hz (32/35 cells under high mean, 31/35 cells under low mean).

Supplemental Data

The Supplemental Data for this article can be found online at http://www.neuron.org/cgi/content/full/51/6/845/DC1/.

Acknowledgments

This work was supported by grants from HHMI (to K.I.N.) and NIH (NS34835 and MH55987 to A.J.D.). We would like to thank T. Sharpee, D. Schoppik, and W. Bialek for very helpful suggestions related to experimental design and analysis; A. Arteseros for performing the histology; and M. Kao, B.D. Wright, and M.P. Stryker for comments on the manuscript.

Received: April 12, 2006
Revised: June 27, 2006
Accepted: August 10, 2006
Published: September 20, 2006

References

Atick, J.J. (1992). Could information-theory provide an ecological theory of sensory processing. Network: Computation in Neural Systems *3*, 213–251.

Attias, H., and Schreiner, C.E. (1998). Coding of naturalistic stimuli by auditory midbrain neurons. In Advances in Neural Information Processing Systems 10, M.I. Jordan, M.J. Kearns, and S.A. Solla, eds. (Cambridge, MA: MIT Press).

Baccus, S.A., and Meister, M. (2002). Fast and slow contrast adaptation in retinal circuitry. Neuron *36*, 909–919.

Barbour, D.L., and Wang, X. (2003). Auditory cortical responses elicited in awake primates by random spectrum stimuli. J. Neurosci. *23*, 7194–7206.

Bartlett, E.L., and Wang, X. (2005). Long-lasting modulation by stimulus context in primate auditory cortex. J. Neurophysiol. *94*, 83–104.

Borst, A., Flanagin, V.L., and Sompolisnky, H. (2005). Adaptation without parameter change: Dynamic gain control in motion detection. Proc. Natl. Acad. Sci. USA *102*, 6172–6176.

Brenner, N., Bialek, W., and de Ruyter van Steveninck, R. (2000). Adaptive rescaling maximizes information transmission. Neuron *26*, 695–702.

Calhoun, B.M., and Schreiner, C.E. (1998). Spectral envelope coding in cat primary auditory cortex: linear and non-linear effects of stimulus characteristics. Eur. J. Neurosci. *10*, 926–940.

Chander, D., and Chichilnisky, E.J. (2001). Adaptation to temporal contrast in primate and salamander retina. J. Neurosci. *21*, 9904–9916.

Chi, T., Ru, P., and Shamma, S.A. (2005). Multiresolution spectrotemporal analysis of complex sounds. J. Acoust. Soc. Am. *118*, 887–906.

Dean, I., Harper, N.S., and McAlpine, D. (2005). Neural population coding of sound level adapts to stimulus statistics. Nat. Neurosci. *8*, 1684–1689.

Depireux, D.A., Simon, J.Z., Klein, D.J., and Shamma, S.A. (2001). Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. J. Neurophysiol. *85*, 1220–1234.

Drew, P.J., and Abbott, L.F. (2003). Model of song selectivity and sequence generation in area HVc of the songbird. J. Neurophysiol. *89*, 2697–2706.

Eggermont, J.J. (1993). Wiener and Volterra analyses applied to the auditory system. Hear. Res. *66*, 177–201.

Eggermont, J.J., Johannesma, P.M., and Aertsen, A.M. (1983). Reverse-correlation methods in auditory research. Q. Rev. Biophys. *16*, 341–414.

Elhilali, M., Fritz, J.B., Klein, D.J., Simon, J.Z., and Shamma, S.A. (2004). Dynamics of precise spike timing in primary auditory cortex. J. Neurosci. *24*, 1159–1172.

Enroth-Cugell, C., and Lennie, P. (1975). The control of retinal ganglion cell discharge by receptive field surrounds. J. Physiol. *247*, 551–578.

Epping, W.J., and Eggermont, J.J. (1986). Sensitivity of neurons in the auditory midbrain of the grassfrog to temporal characteristics of sound. II. Stimulation with amplitude modulated sound. Hear. Res. *24*, 55–72.

Escabi, M.A., and Schreiner, C.E. (2002). Nonlinear spectrotemporal sound analysis by neurons in the auditory midbrain. J. Neurosci. *22*, 4114–4131.

Escabi, M.A., Miller, L.M., Read, H.L., and Schreiner, C.E. (2003). Naturalistic auditory contrast improves spectrotemporal coding in the cat inferior colliculus. J. Neurosci. *23*, 11489–11504.

Fairhall, A.L., Lewen, G.D., Bialek, W., and de Ruyter van Steveninck, R. (2001). Efficiency and ambiguity in an adaptive neural code. Nature *412*, 787–792.

Fortune, E.S., and Margoliash, D. (1992). Cytoarchitectonic organization and morphology of cells of the field L complex in male zebra finches (Taenopygia guttata). J. Comp. Neurol. *325*, 388–404.

Fortune, E.S., and Margoliash, D. (1995). Parallel pathways and convergence onto HVc and adjacent neostriatum of adult zebra finches (Taeniopygia guttata). J. Comp. Neurol. *360*, 413–441.

Frisina, R.D., Smith, R.L., and Chamberlain, S.C. (1990). Encoding of amplitude modulation in the gerbil cochlear nucleus: I. A hierarchy of enhancement. Hear. Res. *44*, 99–122.

Gehr, D.D., Capsius, B., Grabner, P., Gahr, M., and Leppelsack, H.J. (1999). Functional organisation of the field-L-complex of adult male zebra finches. Neuroreport *10*, 375–380.

Gentner, T.Q., and Margoliash, D. (2003). Neuronal populations and single cells representing learned auditory objects. Nature *424*, 669–674.

Gill, P., Zhang, J., Woolley, S.M., Fremouw, T., and Theunissen, F.E. (2006). Sound representation methods for spectro-temporal receptive field estimation. J. Comput. Neurosci. *21*, 5–20. Published online April 22, 2006. 10.1007/s10827-006-7059-4.

Grace, J.A., Amin, N., Singh, N.C., and Theunissen, F.E. (2003). Selectivity for conspecific song in the zebra finch auditory forebrain. J. Neurophysiol. *89*, 472–487.

Hausberger, M., Leppelsack, E., Richard, J., and Leppelsack, H.J. (2000). Neuronal bases of categorization in starling song. Behav. Brain Res. *114*, 89–95.

Heil, P., and Scheich, H. (1985). Quantitative analysis and two-dimensional reconstruction of the tonotopic organization of the auditory field L in the chick from 2-deoxyglucose data. Exp. Brain Res. *58*, 532–543.

Hessler, N.A., and Doupe, A.J. (1999). Singing-related neural activity in a dorsal forebrain-basal ganglia circuit of adult zebra finches. J. Neurosci. *19*, 10461–10481.

Hose, B., Langner, G., and Scheich, H. (1987). Topographic representation of periodicities in the forebrain of the mynah bird: one map for pitch and rhythm? Brain Res. *422*, 367–373.

Kim, K.J., and Rieke, F. (2001). Temporal contrast adaptation in the input and output signals of salamander retinal ganglion cells. J. Neurosci. *21*, 287–299.

Kim, P.J., and Young, E.D. (1994). Comparative analysis of spectrotemporal receptive fields, reverse correlation functions, and frequency tuning curves of auditory-nerve fibers. J. Acoust. Soc. Am. *95*, 410–422.

Kowalski, N., Depireux, D.A., and Shamma, S.A. (1996a). Analysis of dynamic spectra in ferret primary auditory cortex. I. Characteristics of single-unit responses to moving ripple spectra. J. Neurophysiol. *76*, 3503–3523.

Kowalski, N., Depireux, D.A., and Shamma, S.A. (1996b). Analysis of dynamic spectra in ferret primary auditory cortex. II. Prediction of unit responses to arbitrary dynamic spectra. J. Neurophysiol. *76*, 3524–3534.

Krishna, B.S., and Semple, M.N. (2000). Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus. J. Neurophysiol. *84*, 255–273.

Kvale, M.N., and Schreiner, C.E. (2004). Short-term adaptation of auditory receptive fields to dynamic stimuli. J. Neurophysiol. *91*, 604–612.

LeMasurier, M., and Gillespie, P.G. (2005). Hair-cell mechanotransduction and cochlear amplification. Neuron *48*, 403–415.

Lewicki, M.S., and Arthur, B.J. (1996). Hierarchical organization of auditory temporal context sensitivity. J. Neurosci. *16*, 6987–6998.

Lewicki, M.S., and Konishi, M. (1995). Mechanisms underlying the sensitivity of songbird forebrain neurons to temporal order. Proc. Natl. Acad. Sci. USA *92*, 5582–5586.

Liang, L., Lu, T., and Wang, X. (2002). Neural representations of sinusoidal amplitude and frequency modulations in the primary auditory cortex of awake primates. J. Neurophysiol. *87*, 2237–2261.

Lohr, B., Wright, T.F., and Dooling, R.J. (2003). Detection and discrimination of natural calls in masking noise by birds: estimating the active space of a signal. Anim. Behav. *65*, 763–777.

Mante, V., Frazor, R.A., Bonin, V., Geisler, W.S., and Carandini, M. (2005). Independence of luminance and contrast in natural scenes and in the early visual system. Nat. Neurosci. *8*, 1690–1697.

Margoliash, D. (1983). Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. J. Neurosci. *3*, 1039–1057.

Margoliash, D., and Fortune, E.S. (1992). Temporal and harmonic combination-sensitive neurons in the zebra finch's HVc. J. Neurosci. *12*, 4309–4326.

McCormick, D.A., Connors, B.W., Lighthall, J.W., and Prince, D.A. (1985). Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. J. Neurophysiol. *54*, 782–806.

Miller, L.M., Escabi, M.A., Read, H.L., and Schreiner, C.E. (2002). Spectrotemporal receptive fields in the lemniscal auditory thalamus and cortex. J. Neurophysiol. *87*, 516–527.

Muller, C.M., and Leppelsack, H.J. (1985). Feature extraction and tonotopic organization in the avian auditory forebrain. Exp. Brain Res. *59*, 587–599.

Narayan, R., Ergun, A., and Sen, K. (2005). Delayed inhibition in cortical receptive fields and the discrimination of complex stimuli. J. Neurophysiol. *94*, 2970–2975.

Nelken, I., and Young, E.D. (1994). Two separate inhibitory mechanisms shape the responses of dorsal cochlear nucleus type IV units to narrowband and wideband stimuli. J. Neurophysiol. *71*, 2446–2462.

Nelken, I., Kim, P.J., and Young, E.D. (1997). Linear and nonlinear spectral integration in type IV neurons of the dorsal cochlear nucleus. II. Predicting responses with the use of nonlinear models. J. Neurophysiol. *78*, 800–811.

Nowak, L.G., Azouz, R., Sanchez-Vives, M.V., Gray, C.M., and McCormick, D.A. (2003). Electrophysiological classes of cat primary visual cortical neurons in vivo as revealed by quantitative analyses. J. Neurophysiol. *89*, 1541–1566.

Phillips, D.P., and Hall, S.E. (1987). Responses of single neurons in cat auditory cortex to time-varying stimuli: linear amplitude modulations. Exp. Brain Res. *67*, 479–492.

Phillips, D.P., Semple, M.N., Calford, M.B., and Kitzes, L.M. (1994). Level-dependent representation of stimulus frequency in cat primary auditory cortex. Exp. Brain Res. *102*, 210–226.

Rees, A., and Moller, A.R. (1987). Stimulus properties influencing the responses of inferior colliculus neurons to amplitude-modulated sounds. Hear. Res. *27*, 129–143.

Sachs, M.B., and Young, E.D. (1980). Effects of nonlinearities on speech encoding in the auditory nerve. J. Acoust. Soc. Am. *68*, 858–875.

Scheich, H., Bonke, B.A., Bonke, D., and Langner, G. (1979). Functional organization of some auditory nuclei in the guinea fowl demonstrated by the 2-deoxyglucose technique. Cell Tissue Res. *204*, 17–27.

Sen, K., Theunissen, F.E., and Doupe, A.J. (2001). Feature analysis of natural sounds in the songbird auditory forebrain. J. Neurophysiol. *86*, 1445–1458.

Shannon, R.V., Zeng, F.G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). Speech recognition with primarily temporal cues. Science *270*, 303–304.

Singh, N.C., and Theunissen, F.E. (2003). Modulation spectra of natural sounds and ethological theories of auditory processing. J. Acoust. Soc. Am. *114*, 3394–3411.

Theunissen, F.E., and Doupe, A.J. (1998). Temporal and spectral sensitivity of complex auditory neurons in the nucleus HVc of male zebra finches. J. Neurosci. *18*, 3786–3802.

Theunissen, F.E., Sen, K., and Doupe, A.J. (2000). Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. J. Neurosci. *20*, 2315–2331.

Ulanovsky, N., Las, L., Farkas, D., and Nelken, I. (2004). Multiple time scales of adaptation in auditory cortex neurons. J. Neurosci. *24*, 10440–10453.

Wang, X., Lu, T., and Snider, R.K. (2005). Sustained firing in auditory cortex evoked by preferred stimuli. Nature *435*, 341–346.

Wild, J.M., Karten, H.J., and Forst, B.J. (1993). Connections of the auditory forebrain in the pigeon (Columba livia). J. Comp. Neurol. *337*, 32–62.

Woolley, S.M., Fremouw, T.E., Hsu, A., and Theunissen, F.E. (2005). Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. Nat. Neurosci. *8*, 1371–1379.

Young, E.D., and Voigt, H.F. (1982). Response properties of type II and type III units in dorsal cochlear nucleus. Hear. Res. *6*, 153–169.

Yu, J.J., and Young, E.D. (2000). Linear and nonlinear pathways of spectral information transmission in the cochlear nucleus. Proc. Natl. Acad. Sci. USA *97*, 11780–11786.