Short-Term Adaptation of Auditory Receptive Fields to Dynamic Stimuli

Mark N. Kvale1,2 and Christoph E. Schreiner2
1Department of Physiology and 2Department of Otolaryngology, W. M. Keck Center for Integrative Neuroscience, Sloan-Swartz Center for Theoretical Neurobiology, University of California, San Francisco, California 94143-0444

Submitted 20 May 2003; accepted in final form 9 October 2003

Kvale, Mark N. and Christoph E. Schreiner. Short-term adaptation of auditory receptive fields to dynamic stimuli. J Neurophysiol 91: 604–612, 2004; 10.1152/jn.00484.2003. Short-term adaptation and recovery from adaptation have a strong impact on the processing of dynamic stimuli. Adaptive effects on neuronal activity have been studied most commonly for changes in first-order statistics of stimuli such as stepwise increments or decrements in stimulus amplitude. However, changes in higher moment statistics, such as the variance of the amplitude distribution in visual stimuli, also can invoke pronounced adaptation behavior. We demonstrate here that neurons in the inferior colliculus (ICC) of the cat show adaptation to dynamic auditory stimuli that differ in the variance of their modulation depth distribution. In addition, it is shown that neurons show adaptation to other higher moment statistics (e.g., kurtosis) of the modulation envelope. The time course of adaptation is specific for the altered stimulus property and the direction of parameter change. The use of dynamic stimuli allows an estimate of the effects of the adaptation on the temporal response properties of the neurons. We demonstrate that temporal receptive fields of neurons undergo change during the course of adaptation. We show that adaptation to variance in the ICC has many similarities to that in the retina and suggest that adaptation to variance is a general property of sensory systems that allows them to effectively deal with a nonstationary environment.

INTRODUCTION

It is well known that the auditory system shows short-term adaptation to constant tones and tone bursts at the level of the auditory nerve (e.g., Rhode and Smith 1985). Firing rates from single auditory nerve fibers as well as response magnitudes to compound measures, such as compound action potentials (Abbas and Gorgia 1981; Eggermont and Spoor 1982; Spoor et al. 1976) and frequency following responses (Chimento and Schreiner 1990), show a pronounced transient change after intensity changes of the signal. For amplitude increments, the responses show a large initial increase that declines and asymptotes to a new, higher steady-state level. Typically, this is modeled as an exponential spike rate adaptation with a single time constant (Harris and Dallos 1979; Huang 1981; Kiang et al. 1965; Smith and Zwislocki 1975; Smith 1979). Amplitude decrements result in a rapid reduction of response magnitude with a gradual recovery to a new steady-state level below the previous steady-state response; several investigators have fit the adaptation with two exponential processes (Chimento and Schreiner 1991; Smith and Brachman 1980a,b; Westerman and Smith 1984; Westerman 1985). Adaptation behavior to changes in mean amplitude has also been investigated in a forward masking paradigm (Findlay and Adam 1997) in the superior olivary complex and the central nucleus of the inferior colliculus (ICC) that reflect the peripheral adaptation processes as well as additional local adaptation and network effects. Evidence of adaptation to constant sinusoidal amplitude modulated (AM) stimuli in the ICC has also been seen (e.g., Kvale and Schreiner 1997).

Recently (Smirnakis et al. 1997) have shown that retinal neurons show firing rate adaptation to changes not only in mean amplitude but also to changes in the variance of the amplitude distribution of a visual stimulus. Pronounced effects in the adaptation behavior of retinal neurons were evoked by change in the amplitude variance or modulation depth around a constant mean amplitude. That is, two stimuli with the same average intensity but different contrast distributions resulted in clear changes in the firing rate with a subsequent adaptation process at the transition between the two stimuli. The auditory analog would be to modulate a carrier signal of fixed mean intensity with two Gaussian white noise stimuli of the same mean, but different variance. We describe here the response of single neurons in the ICC to such changes in the statistics of the modulation envelope. In addition, we investigated whether sudden changes in higher moments of the stimulus envelope, such as the kurtosis of the temporal envelope given constant mean amplitude and constant variance, also results in the induction of an adaptation process. We suggest in the discussion that similarities between the ICC and retina are a result of similarities in their processing and that adaptation to variance may be a general feature of sensory systems.

To understand the coding of sounds in central auditory stations, such as the ICC, one must be able to decipher the code, i.e., to infer which stimulus produced a given spike train. The best method we have for inferring stimuli is to compute receptive fields (RFs). Some ICC RF aspects have been calculated for modulation response in steady-state regimes (Krishna and Semple 2000; Langner and Schreiner 1988; Rees and Moller 1983), with responses typically averaged over minutes. Since the processing of natural stimulus sequences does not occur in a steady-state mode and since the incoming sequences are often interrupted or have abrupt changes in their spectral and temporal statistics, the steady-state RFs may not be suitable to predict responses to non-steady-state stimuli. To address the question of how much and in which manner RFs are modified during the transition into a steady state, or during the time course of adaptation, we obtained temporal RFs at different instants during the studied adaptation process and compared their properties. This was possible because the stimulus envelopes that were used for the induction of adaptation were
also suitable for a reverse correlation analysis and extraction of the temporal RF of the neurons.

METHODS

Surgical approach

The responses of single units in the ICC of cats were studied. A dorsal approach to the ICC was made by removing the overlying cerebrum (sparring the auditory cortex) and parts of the bony tentorium. Electrode trajectories were tilted at 45° relative to the sagittal plane, which approximately compensated for the tilt of the “frequency-band lamina” (Oliver and Morest 1984). If necessary, electrode position in the inferior colliculus (IC) was stabilized by applying a 2% solution of agarose in Ringer solution to the surface of the IC. The initial surgical procedure was conducted under pentobarbitol sodium (25–29 mg/kg iv to effect). The experiment was conducted under a surgical level of anesthesia maintained by a continuous infusion of ketamine hydrochloride (approximately 1–2 mg/kg/h) and diazepam (approximately 1–2 mg/kg/h) in lactated Ringer solution. The infusion rate was adjusted according to several criteria, which included ECG, respiration, body temperature, and the status of periodically checked reflexes. Additionally, the animals received periodic injections of dexamethasone (0.14 mg/kg sc) to prevent brain edema and atropine (0.04 mg/kg) to reduce salivation. All procedures were in agreement with institutional and national guidelines for animal care.

Stimulus generation

Pure-tone stimuli for the measurement of frequency response areas were digitally generated by a fast microprocessor (TMS 32010), converted into an analog signal using two 16-bit deglitched digital to analog converters (DACs) (60- to 120-kHz sampling rate), and subjected to appropriate antialiasing filtering. Adaptation stimuli (described below) were generated on an Alpha workstation (Compaq) and converted to an analog signal using two shielded electret speakers (STAX analog signal using a Sound Blaster board (Creative Labs) at a 44.1-kHz sampling rate. Sounds were delivered with a Sokolich-type sound system (Sokolich 1981) employing two shielded electret speakers (STAX 54). The system has a nearly flat frequency–response characteristic (approximately 8 dB) ±7 kHz and shows a roll-off of 10 dB per octave for higher frequencies. The calibration of the speaker system was performed with a Bruel and Kjaer 2209 sound level meter and a General Radio Waveform Analyzer 1521-B. Experiments were conducted in a sound-shielded room (IAC).

Single-unit responses were recorded with parylene-coated tungsten electrodes (1–2 MΩ at 1 kHz). Action potentials were amplified and stored on DAT tape (Cygnus) and were discriminated off-line using a custom-built spike sorter that implemented a Bayesian inference engine to cluster similar spike shapes and decompose spike overlaps (Lewicki 1994).

To find characteristic frequencies (CF) of the units studied, tone bursts of 50-ms duration including a rise/fall time of 3 ms were delivered. Frequency–response curves to the tone bursts were constructed from responses to 675 different frequency/level combinations that were presented in a pseudorandom order over a frequency range of 3–5 octaves and an intensity range of 70 dB (Schreiner et al. 1992). All of the stimuli used in these experiments consist of amplitude modulations of a sinusoid carrier at the characteristic frequency of the unit

\[ s(t) = \frac{1}{2} [1 + g(t)] \sin(\omega_c t), \]  

with \( g(t) \) the modulation function. Two types of modulation functions were used: Gaussian white noise and binary m-sequences.

The Gaussian white noise is a sequence of independent and identically distributed (i.i.d.) values picked from a Gaussian distribution

\[ N(\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left( -\frac{x^2}{2\sigma^2} \right). \]  

Because the values are i.i.d., the power spectrum of a finite length of Gaussian white noise is approximately white, i.e., the spectrum has approximately constant power up to a frequency \( \omega_m \) given by half the sampling rate and then falls off beyond that.

A binary m-sequence \( m(t) \) is a two-level pseudorandom sequence of +1s and −1s. The sequence length is \( L = 2^n - 1 \), where \( n \) is the order of the sequence. M-sequences have two important properties (Golomb and Kjaer 2002): they have nearly zero means and the autocorrelation functions take on the impulse-like form

\[ S_{\text{aut}}(\tau) = \delta(\tau) \]  

\[ = \begin{cases} 1 & \text{if } \tau = 0 \\ 0 & \text{otherwise} \end{cases} \]  

The power spectrum of an m-sequence (the Fourier transform of its autocorrelation function) is approximately white up to the cutoff frequency imposed by discretization. Thus both m-sequences and Gaussian white noise have white noise spectra but differ greatly in the shape of their distributions.

To compute the dynamic kernels, i.e., a system-theoretical descriptor of the temporal behavior of the neurons (Marmarelis and Marmarelis 1978), we model the ICC response with a system \( F \) through which a scalar stimulus \( s(t) \) is passed to give a response \( r(t) \)

\[ r(t) = F[s(t)]. \]  

\( F \) is taken to be a linear functional plus a DC component. In real experiments, the input and output signals are sampled into discrete sequences with time \( t \) becoming an integer indexing the sequence. The system then can be written as the discrete convolution

\[ r(t) = h_0 + \sum_{n=0}^{L-1} h(t_n)s(t - t_n) \]  

with kernels \( h_0 \) and \( h(t_n) \) to be determined. We assume that the system has a finite memory of \( M \) time steps so that at most \( M \) of the \( h(t_n) \) coefficients are nonzero. Given that both Gaussian and m-sequence stimuli have an approximate δ-function autocorrelation, \( h(t_n) \) can be computed by using the Lee–Schetzen technique of cross-correlation (Marmarelis and Marmarelis 1978).
higher order moment properties of the distributions. For instance, the shape of their distributions. This leads to differences in the even the same average sound spectrum. But they strongly differ in the same mean and variance and, as indicated previously, they induce contributions of the two noises is shown in Fig. 2. Both distributions have $0.25 \text{ m}$ with a variance of $0.0625$. A sketch of the amplitude distribution of intensity steps causes no change in the bandwidth of the signal induced by spectral splatter. So both components will have the same kurtosis, so the only change across a transition is in the bandwidth of a pure tone, a step change in the variance of the Gaussian distribution. Note that, unlike a step change in intensity of the signal, a step change in the variance of the Gaussian distribution causes no change in the bandwidth of the signal induced by spectral splatter. So both components will induce a narrow-band noise of bandwidth $\omega_n$ around a carrier frequency $\omega_c$, but the average shape of the spectrum will not change across the transition.

A continuous stimulus set consisted of 30 repetitions of $A$ followed by $B$ yielding the alternating sequence $ABAB \ldots AB$. This stimulus set was played four to six times to the unit. To create a dynamic kernel that characterizes the temporal receptive field, one needs to generate an approximately ergodic distribution of stimuli within each time slice, i.e., the noise distribution must be well sampled to gain tight bounds on the cross-correlation. To do this, each $A$ and $B$ contained a different random sample of Gaussian white noise. Thus the stimulus set became $A_1 B_1 A_2 B_2 \ldots A_n B_n$.

In the adaptation to higher order moments experiment, the stimulus also had two components. $A$ was a Gaussian white noise of zero mean with a variance of 0.0625 and $B$ was the $m$-sequence stimulus $0.25\text{m}(t)$ with a variance of 0.0625. A sketch of the amplitude distributions of the two noises is shown in Fig. 2. Both distributions have the same mean and variance and, as indicated previously, they induce the same average sound spectrum. But they strongly differ in the shape of their distributions. This leads to differences in the even higher order moment properties of the distributions. For instance, the Fisher kurtosis

$$\kappa = \left( \frac{1}{N} \sum_{i=1}^{N} \frac{(x_i - \bar{x})^4}{\sigma^4} \right) - 3$$

is a fourth-order correlation. For the Gaussian process the kurtosis is $\kappa = 0$, whereas the kurtosis of an $m$-sequence is a minimal $\kappa = -2$.

Thus this stimulus tested sensitivity to changes in shape of the amplitude distribution of the noise.

The statistical procedure for calculating response kernels and firing rate dynamics is a standard one (Marmarelis and Marmarelis 1978). Depending on the length of the stimulus, several to many trials of the same stimulus were played to the animal. Reverse correlation kernels for each stimulus set were calculated using the Lee–Schetzen cross-correlation method. The set of individual kernels were averaged together and used to compute a standard error to determine the noise. Fitting the firing rate dynamics in the adaptation experiments to derive time constants is a problem in nonlinear regression and the Levenberg–Marquardt method (Press et al 1992) was used for its superior numerical stability and speed.

**RESULTS**

**Variance adaptation**

We first studied the effects of the adaptive variance stimuli in 31 single units in the ICC (20 penetrations in 2 cats). The characteristic frequencies of the units measured ranged from 6 to 10 kHz.

The majority of units (25/31) in the ICC showed initially a strong and rapid change in firing rate after each variance transition followed by a gradual change in the firing rate to a steady state. An example of the firing rate following transitions from low-to-high and high-to-low variance is shown in Fig. 3. The stimulus was played at a mean intensity of 20–30 dB.
above response threshold to the stimulus, where the response threshold is defined as a firing rate 50% greater than that of the spontaneous firing rate. In this case 120 AB trials were presented. The variances for the two stimulus conditions were 0.0156 and 0.0625 for all studied neurons. As has been reported for the visual case (Smirnakis et al. 1997), the low variance to high variance transition produced an initial transient increase in the firing rate, which decayed with time to a steady-state level. Conversely, a high variance to low variance transition produced an initial suppression of the firing rate with an increase in time to the steady-state level. Fitting the firing rates to a sum of steady-state and exponential contributions, we found that the high-to-low transition had a decay time of $\tau_{hl} = 0.30 \pm 0.08$ s and the low-to-high transition had a decay time of $\tau_{lh} = 0.14 \pm 0.05$ s, which resulted in a decay time asymmetry ratio of $\tau_{lh}/\tau_{hl} = 2.1$. In the population studied, 16 (52%) of the units showed a transient increase in the firing rate for a low-to-high transition and a transient firing rate decrease for a high-to-low transition, i.e., these units showed a monotonic transient-rate/variance behavior. By contrast, 9 (29%) of the units showed a transient increase in the firing rate for a low-to-high transition but little transient decrease in the firing rate for a high-to-low transition, resulting in a nonmonotonic transient-rate/variance behavior. Finally, 6 (19%) of the units showed little adaptation. Among these three behavior groups, there was no statistically significant difference in average firing rates in response to a stationary Gaussian white noise modulated carrier at the unit’s CF. Pure tone tuning curves and temporal kernels generated from the noise stimulus also showed no clustering tendencies as a function of adaptation behavior.

In the monotonic transient-rate/variance group, temporal asymmetries in the decay time were found in all the cells, with an average asymmetry ratio of $1.7 \pm 0.4$. The asymmetry for the second, nonmonotonic group was less expressed, with a ratio of $1.5 \pm 0.5$. To investigate whether strongly adapting units had a different time course from weakly adapting or nonadapting units, we plotted amplitude and decay time of the two transition types, shown in Fig. 4. The plots show large scatter in decay time for amplitude ratios near 1.0, reflecting the numerical difficulty of estimating decay times for shallow exponentials in noise. For amplitude ratios away from 1.0, there appeared no consistent trend of decay time with amplitude ratio, indicating that overall time scale of the adaptation is independent of the strength of the response.

Adaptation to changes in higher moments

One can regard the Gaussian white noise modulated carrier as generating a Gaussian distribution of intensities of the carrier. In this sense, the units are adapting to a stationary intensity distribution of the signal. One may ask whether the units will respond only to a change in the variance component of the distribution or to a change in the higher order components as well. To investigate this, we used the combination Gaussian white noise and m-sequence stimulus described above. The Gaussian white noise and m-sequence distributions have the same power spectrum, but different amplitude distributions. As seen under METHODS, it is possible to set the amplitude distributions of the two stimuli so that they have the same mean and variance, but have large differences in the Fisher kurtosis: $\kappa = 0$ for a Gaussian process and a minimal

![Fig. 4. Correlation of amplitude ratio versus decay time. A: population scatter plot for 31 units of the amplitude ratio versus the decay time to steady state for the low-to-high transition. The amplitude ratio of the transition is defined as $(r_i/r_{ss})/\sqrt{\kappa}$, where $r_i$ is the firing rate immediately after transition and $r_{ss}$ is the steady-state firing rate. B: corresponding scatter plot for the high-to-low transition.](http://jn.physiology.org/)
To investigate the relationship between adaptation to changes in higher order moments and adaptation to variance, we played both stimuli (interleaved) to a subset of the units. In five units we were able to gather sufficient data to compare the adaptive response directly. Of the five, four units showed adaptation to both low-to-high and high-to-low variance transitions, and one unit showed adaptation to only low-to-high transitions. In Fig. 6 it is shown that there is little correlation between variance and kurtosis decay times for the low-to-high transition. A linear fit yields $\tau_{\text{var}} = 0.54 \tau_{\text{kur}} + 0.45$, with a correlation $r^2 = 0.68$ (corrected for small $n$) (Sokal and Rohlf 1981) and significance $P < 0.21$ for $H_0: r^2 = 0$. The decay times for higher order adaptation are significantly longer than those for variance adaptation ($P < 0.000003$ for $H_0: \mu_{\text{hho}} = \mu_{\text{ss}}$, paired $t$-test). The amplitude ratio between pre- and posttransition firing rates shows a linear correspondence between variance and higher order moments responses for the low-to-high transition: $\tau_{\text{var}} = 0.59 \tau_{\text{kur}} + 0.40$, with a correlation $r^2 = 0.96$ (corrected for small $n$) (Sokal and Rohlf 1981) and significance $P < 0.008$ for $H_0: r^2 = 0$. Units that respond strongly to changes in the variance also do so to changes in the higher order moments. In general, dynamics for higher order moment adaptation is almost twice as slow as that for variance adaptation.

**Adaptation effects on temporal RFs**

So far we have only considered adaptation effects on the response amplitude. Does adaptation affect other response properties of neurons? To answer this, we will look at response to the noise stimuli.

By cross-correlating the modulation envelope with a spike train we computed the temporal modulation response kernel for a unit. Since we used different segments of Gaussian white noise as the modulation signal in each AB trial, we in effect presented an ergodic set of stimuli in each small interval of time, which allows one to calculate a dynamic temporal modulation kernel. The dynamic behavior of the temporal modulation kernels of a single unit showing variance adaptation is shown in Fig. 7. The figure depicts kernels at consecutive 200-ms intervals, averaged over 200 ms, during the adaptation process.

![Graph](http://example.org/graph.png)

**FIG. 5.** Adaptation to higher order moments. Firing rates as a function of time are shown for a single unit adapting to a stimulus which alternates between Gaussian white noise and m-sequence modulation envelopes. The error bars represent ±1 SE. GWN, Gaussian white noise; MS, m-sequence.

**FIG. 6.** Variance versus higher order moments. A: scatterplot of decay times for adaptation to variance in the low-to-high transition versus adaptation to higher order moments in the m-sequence to Gaussian transition. All points lie well above the line of equal decay times. B: scatterplot of amplitude ratios for adaptation to variance in the low-to-high transition versus adaptation to higher order moments in the m-sequence to Gaussian transition. As before, the amplitude ratio is defined as $(r_f - r_{ss})/r_{ss}$, where $r_f$ is the firing rate immediately after transition and $r_{ss}$ is the steady-state firing rate.

---

$k = -2$ for an m-sequence. The stimuli with identical mean and variance but different kurtosis were played to 27 single units in the ICC. Seventeen of the units showed statistically significant changes in firing rate. The responses for a typical unit are shown in Fig. 5. Similar to the adaptation to changes in variance, a clear adaptation process was observed. Increase in kurtosis resulted in a rapid increase of the firing rate followed by a gradual transition to a steady state. Decrease in kurtosis was accompanied by an initial decrease of firing rate followed by short-term adaptation. In this case, the magnitude of the initial rate change is much smaller than that of the variance adaptation. However, the time constant of the adaptation process appears to be markedly longer than that of the variance adaptation: the decay time after the transition from Gaussian to m-sequence distributions ($\tau_{\text{gm}} = 0.57 \pm 0.11$ s, $n = 16$ vs. the high-to-low variance decay time $\tau_{\text{hl}} = 0.30 \pm 0.08$ s, $n = 17$; $t$-test, $P < 0.0000001$). The adaptation time constants are again asymmetrical, with Gaussian to m-sequence decay time being $\tau_{\text{gm}} = 0.57 \pm 0.11$ s and m-sequence to Gaussian decay time being $\tau_{\text{mg}} = 0.26 \pm 0.09$ s, for an asymmetry of $\tau_{\text{gm}}/\tau_{\text{mg}} = 2.2$. 

---

By cross-correlating the modulation envelope with a spike train we computed the temporal modulation response kernel for a unit. Since we used different segments of Gaussian white noise as the modulation signal in each AB trial, we in effect presented an ergodic set of stimuli in each small interval of time, which allows one to calculate a dynamic temporal modulation kernel. The dynamic behavior of the temporal modulation kernels of a single unit showing variance adaptation is shown in Fig. 7. The figure depicts kernels at consecutive 200-ms intervals, averaged over 200 ms, during the adaptation process.
process immediately after low-to-high and high-to-low variance transitions. After the low-to-high transition, overall amplitude of the kernel decreased in size as time increases, in conjunction with the firing rate. Latency of the peak of the kernel also increased with adaptation by approximately 3 ms. In addition, the shape of the kernel broadened, with the width at half-maximum increasing from 4.6 to 5.8 ms. For the high-to-low transition, the kernels were initially suppressed, with an increase in amplitude and small decrease in latency as adaptation occurred. The modulation transfer functions (MTFs) in Fig. 7 were calculated from a smoothed Fourier transform of the temporal kernels; the best modulation frequency (BMF) is the location of maximum amplitude of the MTF. The MTFs for later time intervals show a shift to lower frequencies relative to the MTF for 0–200 ms. The corresponding MTFs for the high-to-low transition show only a slight increase in BMF.

For a single time slice 0–200 ms after transition, BMF and latency were found to display a general inverse relationship across the population of cells ($r = -0.764$ with significance $P < 5.74 \times 10^{-4}$ for $H_0: r^2 = 0$). Changes in the BMF and latency as the units adapt after the low-to-high variance transition is shown in Fig. 8, for all neurons from the first 200 to 800 ms after the transition. Most units showed an increase in latency ($\Delta t = 1.3 \pm 0.4$ ms) and a slight corresponding decrease in BMF ($\Delta BMF = -5.6 \pm 4.1$ Hz) with adaptation. It is concluded that temporal response properties are dynamically adjusted during the course of the adaptation process, resulting in poorer temporal resolution in the adapted state than in the unadapted state.
DISCUSSION

The main objective of this investigation was to determine whether there exist components in the ICC that respond to higher order nonstationarities in a stimulus and, if so, to characterize those responses. The results above show that such adaptive units exist and that they display several typical behaviors.

The first is that units exposed to sudden changes in the variance of a Gaussian white noise modulation undergo a short-term adaptation of firing rate with a relaxation time on the order of hundreds of milliseconds. The firing rate experiences a transient increase with an increase in variance and a transient suppression with a decrease in variance. Furthermore, the steady-state firing rate for the high variance case is larger than the steady-state firing rate for the low variance case, even though both stimuli have the same average intensity. One possible explanation for the difference in steady-state firing rates is a static nonlinearity in the response to stimuli. For example, if the response varies as the stimulus squared, then large variance stimuli, with proportionately more large events, will generate a greater firing rate than the small variance stimuli. If one further hypothesizes that adaptation is due to a mechanism dependent on firing rate, such as synaptic adaptation mechanisms (Abbott et al. 1997), then the transient increase with a positive change in variance and a transient suppression with a negative change in variance become possible.

Our results, however, cannot be easily explained by a pure firing rate induced adaptation. We saw in the results that there was no correlation between firing rate and decay time. Furthermore, the decay time for adaptation to higher order moments is nearly twice the decay time for adaptation to variance, despite the fact that firing rate changes for higher order moments were comparable to some of the variance firing rate changes. Note, however, that this argument is based on a single layer of synaptic adaptation; there are many levels of adaptive processing prior to the ICC that contribute to the observed effects, and the circuit level explanation for the data remains to be seen.

More generally, neural adaptation refers to a neural response that is dependent on the history of the stimulus, the history of the animal, or both. These histories lead to a context sensitivity of the response. In this work, the context is the history of the variance or higher order moments of the modulation envelope and the state of the animal, where the stochastic aspects of the modulation envelope and state of the animal have been averaged over trials. The work of Malone and Semple (2001) investigates the context dependence of response to monaural frequency transitions and finds evidence of unit response dependent on both stimulus context and discharge history of the unit. They hypothesize that, among many possible mechanisms, differential adaptation of excitatory and inhibitory afferents may play a role in unit adaptation in the IC. The work of Ulanovsky et al. (2003) also considers the context sensitivity of response to frequency shifts in an “oddball” paradigm and demonstrate results consistent with adaptation to pure tones.

Adaptation to the mean intensity for pure tones in single fibers of the cat auditory nerve has been studied (e.g., Chimetto and Schreiner 1991; Smith and Zwischenk 1975; Westerman and Smith 1984) and it is useful to compare their results to ours. A somewhat different stimulus paradigm was used, alternating high and low intensity pure tones with variable periods of silence between them. Both “rapid” and “short” adaptation time scales were found. The rapid time scales were 3–8 ms (low-to-high) and 13–32 ms (high-to-low). The short time scales were 65–125 ms (low-to-high) and 80–280 ms (high-to-low). The ICC neurons do not show change on the rapid time scale; most ICC temporal kernels have a width greater than 5 ms. Compared with the auditory nerve, the time scale for variance adaptation in the ICC is 40% greater than the short time scale for the low-to-high transition and 60% greater for the high-to-low transition. The variance adaptation in the ICC thus takes place on a longer time scale than the auditory nerve short time scale mean adaptation. The large difference in time scales shows that variance adaptation cannot be simply explained by latency effects or a static nonlinear transformation of the mean adaptation mechanism. There are further computations being performed between the auditory nerve and ICC to detect and adapt changes in the variance. Because of the long time scale and range of time scales seen for adaptation rates, the computation being performed most likely involves a network of neurons involving both feedforward and feedback connections.

A less pronounced adaptation to the higher order statistical properties of a stimulus, such as the kurtosis, was also shown to exist. The time scales are nearly twice as long as those for the variance adaptation, indicating a different dynamics for this higher order adaptation process. For a subset of neurons in which we were able to record both variance and higher order moment adaptive properties, there was a rough correspondence between the two time scales, which indicates that the local environment may be important in determining adaptation dynamics. The fact that adaptation to the mean, the variance, and to higher order properties all exist and have different dynamical properties lends evidence to the view that adaptation in the auditory system is more than an automatic gain control and has important dynamical effects on neural coding in the higher stations.

Further evidence for nontrivial adaptation dynamics is provided by changes in the temporal RF as a function of time after the variance transition. For a low-to-high transition, latencies are initially short and increase with time, and the width of the kernel is initially narrow and increases with time. Both changes are independent of changes in firing rate and show not only that probability of firing changes as the system adapts but that the details of the unit response changes as well. A wider kernel with increasing time may indicate an increasing lack of synchronization of the response with the stimuli in the adapted state.

The time-dependent kernels found here complement results found in our previous work in the ICC. In a previous study (Kvale and Schreiner 1997) we studied marginal temporal response kernels in the presence of a dominant stimulus and perturbative narrowband noise and found kernels with dominant initial peaks and an overall biphasic or triphasic shape, as we do in this study. In another study (Escabi and Schreiner 2002) we looked at the spectrotemporal response to both wideband coherent and wideband noise ripple stimuli. The units we look at here correspond to the units sensitive to the phase of the modulation envelope seen in the former study and temporal slices through the spectro-temporal receptive fields (STRFs) at CF again show a prominent initial peak with possible multiphasic tail.
The previous studies investigating RFs averaged spike trains over many minutes to generate steady-state temporal or spectrotemporal receptive fields. Previous studies showed linear responses to stimuli, but never considered adaptation of units to the stimuli employed. In contrast, the current study uses narrowband noise to probe adaptation of response as reflected in time-dependent and context-dependent RFs. Thus the current study explores a new direction in response space, modeling single units in the ICC as time-dependent linear filters.

The modulation behavior captured by the cross-correlation analysis matches the range of previous AM findings of the ICC using noise approaches (Escabi and Schreiner 2002; Kvale and Schreiner 1997; Rees and Møller 1983) or a sinusoidal AM approach (Krishna and Semple 2000; Langner and Schreiner 1988). The fact that both temporal RFs and modulation transfer functions have consistent properties across multiple studies and stimulus types indicates that there is a population of ICC units that is phase sensitive and largely linear in its response to modulation envelopes.

As adaptation progresses after a low to high variance transition, the best modulation frequency decreases by 5–10%. This decrease results from the increasing spike jitter as a unit adapts. The modulation frequency at which the time scale for spike jitter becomes comparable to the modulation period sets a high-frequency cutoff, because, when spikes may occur at any phase, the vector average of the phase tends to zero as jitter increases with adaptation time, the cutoff modulation period increases and the cutoff modulation frequency decreases.

In part, this work was inspired by the investigation of variance adaptation in the retina by Smirnakis et al. (1997). Underlying this is the hypothesis that the ICC and the retina both process their respective stimuli in a similar fashion, and consequently their adaptive responses to phenomena such as sudden changes in the variance may follow similar time courses.

The similarity in processing of the ICC and the retina is as follows. Both sound and light are wave-like excitations, with sound frequencies typically on the order of $10^3$ Hz and light frequencies on the order of $10^{14}$ Hz. The value $10^{14}$ Hz corresponds to a time scale of femtoseconds, which is much faster than any neural time scale; thus rods and cones cannot detect the phase of a light wave, but instead detect the modulation envelope of the incoming light wave. The auditory system is fast enough to detect the phase of at least low frequency sound waves, but in the ICC, units with CFS above approximately 500 Hz largely ignore the phase and respond primarily to the modulation envelope. Changes in the modulation envelope with time are called temporal contrast in the visual field and modulation depth in the auditory field, but they amount to the same thing. In both the visual and auditory domains, sudden changes in modulation properties impart information about the world and it may be supposed that both systems will respond usefully to these events, i.e., pass the information on to higher stations. The results above show that, at least for adaptation to changes in variance, the ICC and retina have the same qualitative time course (although different time scales) of adaptation in firing rate. Our results then point to adaptation to variance as a more general scheme for neural systems to deal with nonstationary stimuli than first supposed.

How are the stimuli used in this work related to natural sounds and auditory scenes? The stimuli are fully synthetic and yet retain some “naturalistic” features of natural sounds. There exist many naturally occurring sources of narrow-band noise, such as wind noise and animal vocalizations. Changes in variance and higher order moments of the modulation envelope may occur during changes in background sounds and in inherently nonstationary phenomena, such as vocalizations, speech, and sound generated from wind and water turbulence. The wide variety of phenomena with varying amplitude distributions suggests that sensitivity to these changes is, at the level of the ICC, part of a general processing strategy rather than specific feature detection.

Why might adaptation to variance and higher order moments be useful to neural systems? One reason may be to allow the neural system to achieve a homeostasis or firing rate set-point in the presence of a highly variable stimulus. For both the high-to-low and low-to-high transitions, the firing rate relaxes away from the extremes to a middle range of firing rates. Thus adaptation tends to reduce the dynamic range of steady-state response (for a given intensity) to a fairly small and presumably quiescent value.

A second possible use of adaptation to variance and higher order moments is to detect and to emphasize changes in the stimulus. Both the high-to-low and low-to-high transitions show large changes in firing rate that mark the event, but relax to a steady state in the presence of the relatively small changes in the stimulus at constant variance. One may consider adaptation as the system “learning” the current modulation depth distribution in preparation to detect changes in it.

It is yet unresolved how these adaptation processes depend on the magnitude of the transition and relate to other response properties and neural classes. Analysis of the pure tone tuning curves did not reveal any clear link to adaptation behavior. However, all studied neurons showed phase locked response to AM stimuli or had a temporal modulation kernel in the reverse correlation analysis. Therefore none of the studied neurons belonged to the “nonphase locking” class of neurons (Escabi and Schreiner 2002).

Related to this second use is the puzzle of the asymmetry in the relaxation time constants for the two variance transitions. It is possible to achieve an asymmetry in the synaptic adaptation scheme, but this raises the question of why the asymmetry is robust and in the same direction for both the retinal (Smirnakis et al. 1997) and auditory midbrain systems.

We hypothesize that the temporal asymmetry in both these systems is not an accidental consequence of other neural processing, but is rather a consequence of the system trying to detect changes in the statistical properties of the stimulus. It can be shown with Bayesian techniques (DeWeese and Zador 1998) that an ideal observer who tries to detect a change in the variance as soon as possible will take longer to detect the high-to-low transition than the low-to-high transition. Under this hypothesis, asymmetry in decay time is due to the asymmetry in the Gaussian probability distribution function. Changes in the statistics of a stimulus are relevant in many sensory domains and the adaptive responses seen in the visual and auditory realms may be indicative of a universal strategy of biological sensory systems coping with a nonstationary world. We would predict, for example, that somatosensory response would also show adaptation to the statistical properties of a vibrational stimulus and would show temporal asymmetry in the adaptation rates.
CONCLUSIONS

We have demonstrated the existence of units in the ICC that adapt to changes in a nonstationary signal. In these units we have found three novel effects of adaptation to time varying acoustic stimuli. The first effect is that units exposed to discontinuous changes in the variance of a Gaussian white noise modulation undergo a short-term adaptation with a relaxation time on the order of hundreds of milliseconds. Second, enhancement or suppression of the average firing rate and relaxation time of the adaptation depended strongly on the gradient of the variance. For sudden increases in the variance, the response will overshoot and relax the steady-state firing rate fairly quickly and the dynamic kernels will increase in latency. For sudden decreases in variance, the response will be suppressed and rise to a steady-state value more slowly. Third, even when the mean and variance are held constant but higher order statistics of the stimulus are changed, there is adaptation to the change. In our example, the transition between a high kurtosis stimulus, m-sequences, and a low kurtosis stimulus, Gaussian white noise, induced an adaptive change in firing rates and an asymmetry in relaxation times for the two types of transition. This means that sounds that have modulation envelope distributions that change in time can be detected at the level of the ICC and this information is readily available for downstream processing.

These three effects provide evidence that there exist units in the ICC that not only adapt to overall stimulus levels, but adapt to higher order statistical properties of the stimulus as well. Stationary stimuli are characterized by constant statistical quantities such as mean, variance, and higher order moments. Thus response and adaptation to time varying statistical properties allows these units in the ICC to detect nonstationarities in acoustic stimuli. It is well known that many natural stimuli in the acoustic realm have nonstationary properties and one may speculate that the adaptation effects shown in this study are useful for detecting changes in behaviorally relevant features of the acoustic biotope.

ACKNOWLEDGMENTS

We thank Drs. B. Bialek, M. DeWeese, D. Mackay, and A. Zador for useful conversations and Dr. L. Miller and Dr. B. Bonham for assistance in the experiments. We also thank Dr. M. Lewicki for making available his source code for the Bayesian inference engine used in the spike sorting.

GRANTS

We gratefully acknowledge support by The Sloan-Swartz Foundation, National Institute on Deafness and Other Communication Disorders Grant DC-00354–01A1 to M. Kvale, NDS Grant 34835, and DC-02260.

REFERENCES
