Frequency-Modulation Encoding in the Primary Auditory Cortex of the Awake Owl Monkey

Craig A. Atencio,1,2,3 David T. Blake,2,3,4 Fabrizio Strata,2,3 Steven W. Cheung,3 Michael M. Merzenich,2,3 and Christoph E. Schreiner1,2,3

1Bioengineering Graduate Group, University of California, San Francisco, University of California, Berkeley; 2W. M. Keck Foundation Center for Integrative Neuroscience; 3Coleman Memorial Laboratory, Department of Otolaryngology–Head and Neck Surgery, University of California, San Francisco, California; and 4Center for Synapses and Cognitive Neuroscience, Medical College of Georgia, Augusta, Georgia

Submitted 5 April 2007; accepted in final form 11 August 2007


The vocalizations of New World monkeys such as the owl monkey, squirrel monkey, and marmoset are complex and contain spectral and temporal features that vary with time (Jürgens 1986; Wang 2000; Winter et al. 1966). A feature of many vocalizations is a change in energy across frequencies, usually as a consequence of muscular tension changes in the vocal tract and/or larynx. The resulting smooth changes in the prominent vocalization element is represented in the auditory cortex if we wish to understand the neural encoding of natural vocalizations. Previous studies investigated responses to these transitions, modeled as isolated FM sweeps, in a number of vertebrates, including birds (Heil and Scheich 1992; Heil et al. 1992a), rats (Mendelson and Ricketts 2001; Poon and Yu 2000; Zhang et al. 2003), ferrets (Kowalski et al. 1995; Nelken and Versnel 2000; Shamma et al. 1993), and cats (Heil et al. 1992b; Mendelson and Grasse 1992; Mendelson et al. 1993; Phillips et al. 1985; Tian and Rauschecker 1994, 1998).

Recently, we examined the topographic organization of responses in the primary auditory cortex (AI) of anesthetized rats and squirrel monkeys to FM sweeps (Godey et al. 2005; Zhang et al. 2003). Both studies used high-density topographic mapping across AI and showed the functional organization of FM sweep responses. Characteristic frequency (CF) could be predicted from FM sweep responses. Direction selectivity was correlated with CF, indicating that sweep direction was topographically represented in AI, with upward FM sweep selective sites present in the low-frequency portion of AI and downward selective sites present in the high-frequency portion.

In this study we extend our experimental technique and examine the responses of neurons in the awake owl monkey to FM sweep stimuli. A main question is whether the coding of dynamic stimuli, such as FM sweeps, is fundamentally different in the awake animal because there has been evidence of potential differences in the response pattern to stationary stimuli in anesthetized and awake preparations (Lu et al. 2001). The owl monkeys used in this study were chronically implanted with a multielectrode recording array using previously established methods (Blake and Merzenich 2002; deCharms et al. 1998, 1999). The use of the implanted awake owl monkey allows auditory function to be studied under more natural, and behaviorally controlled, conditions. Because nonhuman pri-
mates are an attractive model for human perception and neural processing of complex sounds, this preparation is advantageous when studying neural coding in AI.

In addition to FM sweep responses, we computed spectrotemporal receptive fields (STRFs) using random tone-pip stimuli (Blake and Merzenich 2002; deCharms et al. 1998; Rutkowski et al. 2002). Theoretically, for a neuron that processes stimuli in a quasi-linear fashion, the STRF provides a sufficient description of the spectrotemporal preferences of the neuron (Blake and Merzenich 2002; Theunissen et al. 2000). We used the STRFs to predict the preferences of AI neurons for several FM sweep parameters. By predicting direction selectivity and best speed from the STRFs and comparing them to observed FM responses, we determined the adequacy of simple linear prediction in describing the dynamic behavior of cortical neurons (Blake and Merzenich 2002). We conclude this report by discussing our findings in view of FM sweep studies in other species and their significance for the neural processing of vocalizations.

METHODOLOGY

Physiological recordings

Animal welfare was monitored by the IACUC at University of California, San Francisco, and conformed to National Institutes of Health guidelines. All data were obtained using an approved IACUC protocol. We obtained the data from two chronically implanted owl monkeys, Aotus nancymae. All methods were reported previously and are only briefly described here (deCharms et al. 1998, 1999). The AI was targeted 2–3 mm anterior and just lateral to the temporal–frontal fissure in the lateral bank of the lateral sulcus. Transcranial recording through burr holes confirmed AI response characteristics and expected tonotopy (Imig et al. 1977; Recanzone et al. 1999). Recording array implants were then placed over AI. Surgery was performed under aseptic barbiturate anesthesia and histological confirmation of AI was performed in one animal. Recordings were made from chronically implanted arrays of ≤49 individually placed ultrathin parylene-coated iridium microelectrodes (Micro Probe, Potomac, MD). Electrodes had exposed tip lengths of 5 to 7 μm, yielding impedance values of about 2 MΩ and good isolation of signals from individual neurons. These signals were filtered from 0.3 to 8 kHz, sampled at 20 kHz, digitized, stored to disk, and spike-sorted to yield well-resolved single neurons. Stimuli were presented free field and frontally, approximately 24 in. from the center of the head of each monkey, in a double-walled acoustic isolation chamber while the animal sat in a primate chair without head restraint. Recordings were obtained before behavioral sessions. Animals had been trained to wait throughout these data collection periods and were monitored either live or by a closed-circuit camera. Animals were never observed to be drowsy throughout these periods, and this protocol was surprisingly effective at having the animals remain calm and alert. In practice, gaze was within a 15°–20° radius and motion during recordings was minimal. Recordings during substantial head movements were discarded. Spike trains were simultaneously collected from multiple sites in the cortex during the presentation of the stimulus sets. Implant best frequencies spanned the range of frequencies found on the superior temporal gyrus, ranging from 0.1 to 7 kHz (Imig et al. 1977).

Sound presentation

Sound levels at the position of the external ears were calibrated with a Bruel & Kjaer sound level meter. Frequency-modulated sweeps were created using Matlab (The MathWorks, Natick, MA), written to compact disk using standard software, and delivered using a McIntosh audio amplifier. The sweeps traversed the frequency range in the upward (50 to 21,000 Hz) or downward (21,000 to 50 Hz) direction. This range covers the frequency response area normally observed for AI neurons (Recanzone et al. 1999). FM sweeps in New World monkey vocalizations often have bandwidths (BWVs) between 5 and 10 kHz, and thus they effectively cover the frequency response of AI neurons in a manner similar to the synthetic sweeps (Atencio et al. 2005; Bieser 1998; Godey et al. 2005). The sweep speeds were logarithmically spaced: 10, 17, 30, 52, and 90 octaves per second for a total of 10 different stimulus conditions. The sweep speeds were chosen so that they covered the normal range of frequency transitions present in owl monkey vocalizations. All sweeps had 100-ms rise/fall cosine-squared ramps, where either 50 or 21,000 Hz was ramped up or down to the starting amplitude at which point the FM began. The durations of the individual FM sweep stimuli were: 1,071, 713, 490, 368, and 297 ms for the 10, 17, 30, 52, and 90 oct/s sweeps, respectively. The interstimulus time between sweep stimuli was 2 s. Each sweep was nonconsecutively presented 16 times in random order.

Trains of tone pips were used to construct spectrotemporal receptive fields (STRFs) as described in previous reports (Blake and Merzenich 2002; deCharms et al. 1998). The tone-pip trains were synthesized by randomly selecting frequencies from 84 possible values spanning 7 octaves from 110 to 14,080 Hz in 1/12th-octave steps. For all stimuli, every 1/12th-octave frequency band contained an independent Poisson train of tone pips with the same mean rate of tone-pip presentation. Each individual tone pip was 20 ms in duration, with 5-ms cosine-squared onset/offset ramps. The random tone-pip stimulus was presented continuously for 5 min, with 1 s of silence between stimuli. Tone-pip and FM sweep stimuli were presented at the same sound level, between 50 and 70 dB SPL.

Analysis

FM sweep stimuli responses were analyzed in a manner similar to that in previously published work (Godey et al. 2005). Responses to FM sweep stimuli were quantified by first computing the poststimulus time histogram (PSTH) for each stimulus condition using 10-ms bins. For each of the 10 PSTHs a Gaussian fit was to the response in the PSTH. Each Gaussian had the form

$$f(t) = A \exp \left( -\frac{(t - \text{latency})^2}{2SD^2} \right)$$

The Matlab (The MathWorks) 6.5 function “fminsearch” was used to obtain the parameters in the fit that minimized the squared error between the Gaussian and the portion of the PSTH that corresponded to the presence of the stimulus. Thus three parameters were obtained: 1) $A$, the amplitude of the Gaussian; 2) latency, the position along the time axis at which the peak occurs; and 3) $SD$, a measure of the spread of the Gaussian fit. Neurons were further analyzed if fits were significant for four of five PSTHs in both the upward and downward directions ($\chi^2$ test, $P < 0.05$).

Next, we related the parameters from the Gaussian fits to the corresponding pure-tone responses. We pursued this because as an upward FM sweep travels from 50 to 21,000 Hz it will encounter the low-frequency boundary of a neuron’s frequency tuning curve. The same process occurs for downward sweeps, although the high-frequency boundary of the frequency tuning curve is encountered initially in this case. The frequencies at which the sweep encounters the tuning curve are termed trigger frequencies, and may be calculated using the latencies of the PSTH peaks. From these trigger frequencies we can estimate the BW and CF of a particular neuron. The procedure that was used to compute the trigger frequencies, illustrated in Fig. 1, was similar to that in previous studies (Heil 1997; Heil and Irvine 1998; Heil et al. 1992b; Nelken and Versnel 2000). First, latencies
Schematic shows the procedure for upward sweeps. A: Gaussian curves were fitted to each poststimulus time histogram (PSTH) for the 5 sweep speeds (10, 17, 30, 52, and 90 oct/s). Latencies, SD, and areas of each Gaussian curve were obtained. B: latencies (in milliseconds), denoted by $L_{10}, L_{17}, L_{30}, L_{52},$ and $L_{90},$ were plotted against the inverse of the sweep speed (in s/oct). Slope of the best-fitting line between the data points was obtained (in octaves). This slope was used to calculate the upward trigger frequency according to $F_{UP} = 50 \times 2^S$ Hz. For downward sweeps the same procedure was used, with the trigger frequency found using $F_{DN} = 21,000 \times 2^{-S}$ Hz.

Results

In these studies we recorded detailed neural responses from 128 well-isolated neurons from implanted electrode arrays. Awake recordings were made in AI in the left superior temporal gyrus of two adult owl monkeys (OM1 and OM2). AI was histologically verified in one experiment (OM1). AI neurons have characteristic short-latency responses and there is an orderly progression of CF across the surface of AI, allowing for consistent identification of AI (Fitzpatrick and Imig 1980; Imig et al. 1977; Recanzone et al. 1999). The targeted depth range of the electrodes encompassed layers III and IV.
FM sweep poststimulus time histograms

The responses of a representative owl monkey AI neuron to FM sweep stimuli are shown in Fig. 2, where column A shows the responses to upward sweeps and column B the responses to downward sweeps. Each poststimulus time histogram (PSTH) represents the response of the cell to 16 presentations of the same FM sweep stimulus. The bar under each PSTH represents the duration of the FM sweep. As can be seen from the PSTHs, this cell responded most vigorously with a prominent activity peak to upward low-speed sweeps, showing a maximal response to the lowest sweep speed presented, 10 oct/s.

Several features in Fig. 2, A and B are noteworthy. First, the duration of the response, as shown by the width of the peaks in the PSTHs, is not uniform across different sweep speeds and directions. The lowest sweep speeds often evoked the broadest responses, whereas the faster sweep speeds invariably evoked the narrowest responses. Second, the onset time of the response changed in a systematic fashion, as expected, due to differences in the duration of the FM sweep stimuli and the frequency selectivity of the neuron.

To quantify the nature of the response profiles, Gaussian curves were fit to each of the 10 PSTHs obtained for each FM sweep stimulus. In Fig. 2 these are seen as the black outlines superimposed on each PSTH. Two main measures were used to quantify the FM response. The first was the latency of the Gaussian curve, which provided an estimate of the time to the peak response, as shown in each PSTH. The second was the area under each Gaussian fit, a metric for response strength. The results of these computations for the responses in Fig. 2, A and B are shown in Fig. 2, C and D. In Fig. 2C plots of latency versus the inverse of FM sweep speed are shown for upward and downward directions. These data contain a significant nonzero correlation (Upward: \( r = 0.99, P < 0.001 \); Downward: \( r = 0.99, P < 0.001 \), t-test). Ordinary least-squares regression on this data gave the slope of the lines and allowed, using a simple formula (see METHODS), calculation of the frequency at which the neuron began to respond to the FM sweep (Bain and Engelhardt 1992; Draper and Smith 1981).

For the cell in Fig. 2 the frequency calculated from the upward sweeps is 1,016 Hz and from the downward sweeps this value is 1,201 Hz. Because these two frequencies represent the trigger frequencies of the cell to sweep stimuli, each may be taken as an estimate of the effective low- and high-frequency boundary edges of the cell’s frequency–response area. From these values CF is estimated as the average of the low and high frequencies, or 1,109 Hz, which is in excellent agreement with the pure-tone CF estimate of 1,100 Hz. The bandwidth of the cell is estimated as the difference between the two trigger frequencies, which in this case is 185 Hz, whereas the response bandwidth calculated from pure tones was 1,453 Hz.

Plots of normalized area versus sweep speed are shown in Fig. 2D. These plots show response strength versus FM sweep speed. The normalized areas are calculated by dividing all responses, in both sweep directions, by the largest area value, determined from both directions. These curves show a cell’s selectivity for a particular sweep speed in either the upward or the downward direction. For a highly selective cell, preferring only upward sweeps at 10 oct/s, it would be expected that the upward area–speed curves would contain a strong peak at 10

![Graphs showing FM sweep poststimulus time histograms](http://jn.physiology.org/)

**FIG. 2.** PSTH of responses for unit om1-b24-bl32-20 to FM sweep stimuli. Responses to upward FM sweeps (A) and responses to downward sweeps (B) at the 5 different sweep speeds. Horizontal bar below the time axis represents the duration of each sweep. Superimposed on each PSTH in black is the best-fitting Gaussian curve for each sweep response. C: plot of latency of the Gaussian fits shown in A and B vs. inverse sweep speed. Curves for upward and downward sweeps have significantly nonzero slopes (Upward: \( r = 0.99, P < 0.001 \); Downward: \( r = 0.99, P < 0.001 \), t-test). D: plot of areas of Gaussian fits vs. sweep speeds for upward and downward directions. Both area curves are normalized by the same maximum area value. Response is greater for upward sweeps [direction selectivity index (DSI) = 0.14], although note the response is greatest for slow sweep speeds in both directions (best speed = 27 oct/s; speed tuning = 0.57), as well as the similar shapes of the 2 curves (correlation coefficient = 0.98).
oct/s and have values near zero elsewhere, as well as a flat curve near zero for the downward direction response. The responses in Fig. 2D show that this neuron preferred FM sweeps at low sweep speeds in both directions. These curves show that for this neuron slower speeds elicit the strongest response and that in both directions a low-pass area–speed characteristic is present. Also, in comparing the normalized area–speed curves in both directions it is clear that the response in the upward direction is stronger, although the shapes of the two curves are similar ($r = 0.98, P < 0.01, t$-test). This was a general trend because the correlation between the upward and downward area–speed curves was normally quite high (see following text). The direction selectivity index (DSI) for this unit was calculated to be 0.14, in accord with the greater responses this neuron showed to upward sweeps. Best speed was 27 oct/s and the speed tuning was 0.57, in accord with this site’s higher selectivity for sweeps at slower sweep speeds.

Another representative example of responses to FM sweep stimuli is shown in Fig. 3. The greatest response occurs at 30 oct/s for sweeps in the upward direction. In the downward direction the neuron, although not selective for one speed, did respond to the whole range of sweep speeds. Thus this example denotes a neuron that is most selective to upward sweeps and one sweep speed in particular while being responsive, though nonselective, to sweeps in the downward direction.

Figure 3, C and D again shows the same analysis procedure outlined in Fig. 2, C and D. As in Fig. 2C straight-line fits were made to the latency–speed$^{-1}$ data, where a significant nonzero correlation was found (Upward: $r = 0.99, P < 0.001$; Downward: $r = 0.99, P < 0.001, t$-test). As shown in the area–speed curves of Fig. 2D this neuron was maximally responsive to sweep speeds near 30 oct/s (best speed: 41 oct/s) and had a band-pass area–speed curve in the upward direction (DSI = 0.21). The speed tuning was 0.29, reflecting only a moderate selectivity for specific upward sweep speeds. The downward direction area–speed curve was flat, showing that in case upward and downward sweeps were processed differentially ($r = 0.64, P < 0.05, t$-test).

The responses of a downward sweep selective neuron are shown in Fig. 4. The greatest response occurs at 10 oct/s in the downward direction, although responses in the upward direction are comparable. Thus this example shows a neuron that is most selective to slower sweep speeds, with responses that show a similar selectivity for FM sweep speed in the upward and downward directions.

The latency–speed$^{-1}$ data for this neuron show a significant nonzero correlation (Upward: $r = 0.99, P < 0.001$; Downward: $r = 0.99, P < 0.001, t$-test). The estimated trigger frequency in the upward direction was 641 Hz and in the downward direction was 1,128 Hz, giving a CF estimate of 885 Hz (pure-tone CF of 831 Hz). The bandwidth estimated from the trigger frequencies was 487 Hz, showing a considerable discrepancy between FM sweep and pure-tone bandwidth estimate (903 Hz). As shown in the area–speed curves of Fig. 4D the responses to downward sweeps were greater than those to upward sweeps, reflected in a DSI value of −0.14. This neuron also responded to sweeps in either direction in a similar manner ($r = 0.91, P < 0.01, t$-test). The neuron was responsive over a range of speeds, with the best speed estimated as 33

![Graphs](http://jn.physiology.org/)

**Fig. 3.** A and B: PSTH of responses at unit om2-b25-bl32-43 to FM sweep stimuli. Format of the figure is identical to Fig. 2. Responses for this site are greatest to upward sweeps at 30 oct/s, whereas in the downward direction the responses are relatively uniform and show little speed selectivity. C: latency of the Gaussian fits shown in A and B vs. inverse. Form of plot is exactly as in Fig. 2C. Curves for the upward and downward sweeps have significant nonzero slopes (Upward: $r = 0.99, P < 0.001$; Downward: $r = 0.99, P < 0.001, t$-test). D: plot of area vs. sweep speed for upward and downward sweeps. This neuron responded best to upward sweeps of 30 oct/s. Correlation coefficient between the 2 curves is −0.089. For this site DSI = 0.21, best speed = 41 oct/s, and speed tuning = 0.29.
oct/s and the speed tuning as 0.39, reflecting a moderate selectivity for specific sweep speeds.

Comparison of pure-tone and FM-derived response properties

CF estimates from the FM sweep data and from the STRF are in close register (Fig. 5A). The FM estimate of CF was calculated as the average of the upward and downward trigger frequencies and compared with the CF estimated from the STRF (Linden et al. 2003; Nelken and Versnel 2000; Sen et al. 2001). The data shown in Fig. 5A are from animal OM1 because sufficient STRFs were available only for that subject.

FIG. 5. Correlations between pure tone estimates and FM sweep responses. A: characteristic frequency (CF) derived from FM sweeps vs. CF from pure tones ($r = 0.86$, $P < 0.001$, $t$-test). B: bandwidth (BW) derived from FM sweeps vs. pure tone BW ($r = 0.64$, $P < 0.001$, $t$-test). Diagonal lines represent unity relationships. Spectrotemporal receptive field (STRF) pure-tone data were obtained from subject OM1 only.

Tone-pip CFs plotted against the CF estimate from FM sweep stimuli reveal a significant correlation (OM1: $r = 0.86$, $P < 0.001$, $n = 68$, $t$-test). Thus FM sweep data accurately predict the CF from pure-tone responses (Heil and Irvine 1998; Heil et al. 1992b; Tian and Rauschecker 1994).

We also estimated the frequency response bandwidth of tone pips and compared it to that estimated from FM sweeps. FM sweep bandwidth estimates were obtained by subtracting the downward from the upward trigger frequencies. In some cases the estimate was negative, and these data points were excluded from this analysis (Nelken and Versnel 2000). Figure 5B shows that the two bandwidth estimates are significantly correlated (OM1: $r = 0.64$, $P < 0.001$, $t$-test). In this comparison it is implicitly assumed that when the FM sweep stimulus first excites a neuron, the energy in the sweep is located at either the low-frequency (upward sweep) or high-frequency (downward sweep) edge of the cell’s frequency response area. For a simple response pattern the BW estimate should yield reasonable estimates of the excitatory BW obtained from tone pips. The majority of FM sweep–derived bandwidths underestimate the tonal response bandwidth (Fig. 5B; diagonal line represents unity relation). This underestimation, as well as the occurrence of negative bandwidth estimates, may be due either to nonlinearities in the responses or to asymmetric strength differences of the inhibitory influences on the low- or high-frequency sides of the receptive field that significantly affect the response latency of the excitatory activity (Heil 1997; Heil et al. 1992a; Kowalski et al. 1995; Shamma et al. 1993; Zhang et al. 2003).

In many species the direction selectivity index is significantly correlated with CF. In this study we did not find a significant correlation ($r = -0.04$, $P = 0.38$, $t$-test). The lack
of a significant negative correlation between DSI and CF is not in accord with results from other studies (Godey et al. 2005; Mendelson and Ricketts 2001; Zhang et al. 2003). Because CF is mapped topographically in owl monkey AI (Imig et al. 1977), this leaves an open question as to whether the DSI–CF correlation is present in the owl monkey, or whether our negative finding is a consequence of incomplete sampling of the full frequency range.

**Population distributions of FM sweep preferences**

Response distributions for the four FM sweep response parameters for both cases (OM1 and OM2) are shown in Fig. 6. The two columns represent the response distributions for each monkey. **Direction selectivity**: The median DSI across all neurons is 0.04, and that for each individual subject is 0.01 and 0.10, reflecting that, as a population, the neurons are equally responsive to both upward and downward sweeps. **Best speed**: The overall median is 36 oct/s, with 35 and 37 oct/s the median values for the individual populations. This corresponds to the middle of the range of FM sweeps found in New World monkey twitter vocalizations (Bieser 1998; Nagarajan et al. 2002). **Sweep speed tuning**: The two monkeys show similar distributions, with the median of the distributions at 0.57 and 0.45, leading to an overall value 0.51 that reflects moderate selectivity for specific sweep speeds. Thus whereas most neurons prefer sweeps that are between 25 and 45 oct/s, they are responsive to a broader range of sweep speeds.

The last row in Fig. 6 quantifies the direction specificity of speed tuning. Shown in these plots are the distributions of correlation coefficient values derived between the upward and downward area–speed curves. The majority of the values are >0.5, reflecting similar shapes of the area–speed curves in both directions. Although this metric does not account for the magnitude of the response, it does imply that most sites in AI have similar speed tuning properties to upward and downward sweeps.

**Spectrotemporal receptive fields**

For many recording sites \(n = 68\) randomly presented tone pips were presented and STRFs were constructed. An STRF can be interpreted as the average stimulus that precedes an action potential. Although additional characterization is required to determine the specificity of the response to the stimulus (Escabi et al. 2005), the STRF is, in principle, suited to assess the time-dependent nature of a neuron’s responses to complex stimulation patterns, including FM components. Examples of representative STRFs are shown in Fig. 7. Figure 7A shows a neuron’s STRF with a CF of about 500 Hz. This STRF shows clear excitatory and suppressive regions. Of note is the slanted shape of the suppressive region, which is consistent with a neuron that prefers upward direction FM sweeps because a stimulus that begins at low frequencies and transitions to high frequencies would interact minimally with the suppressive region while interacting maximally with the excitatory portion of the receptive field (Shamma et al. 1993; Suga 1965a,b, 1968). These STRF features are consistent with the neuron’s responses to FM sweep stimuli. The experimentally determined DSI for this neuron was +0.17, indicating upward sweep preferences.

![Fig. 6. Population distributions of FM sweep response parameters. Each column represents data from each subject. Each row represents a different response parameter. First row: DSI values, where each distribution is centered about 0. Second row: best speed values, which are centered about 36 oct/s. Third row: speed tuning values are centered about 0.51; most sites in primary auditory cortex (AI) show moderate tuning for sweep speed. Last row: correlation coefficient values between the area–speed curves in the upward vs. downward directions. Majority of values are >0.5, implying that sweep speed is processed with similar selectivity in the upward and downward directions.](http://jn.physiology.org/)

An STRF with different characteristics is shown in Fig. 7B. The neuron has a CF of about 800 Hz and clearly defined excitatory and suppressive regions. The STRF for this neuron differs from the STRF in Fig. 7A in that the suppressive region is symmetrically positioned relative to the CF. From the symmetry of these regions this neuron would react to upward and downward FM sweeps in a similar fashion. Thus qualitatively the STRF would be indicative of a relatively nondirectional selective neuron, and is consistent with the observed small DSI (0.05) for this neuron. In the cases shown in Fig. 7, A and
Other STRF shapes are also possible, with further examples of FM sweep direction selectivity (deCharms et al. 1998). B, the STRFs appear to give an adequate qualitative indication that suppression is visible in the STRFs in Fig. 7, F (DSI = 0.14; FM sweep data for this neuron are shown in Fig. 2). Two STRF examples for neurons that responded best to downward sweeps are shown in Fig. 7D (DSI = 0.14; FM sweep data for this neuron are shown in Fig. 2). Two STRF examples for neurons that responded best to upward sweeps are shown in Fig. 7E (DSI = −0.14; FM sweep data for this neuron are shown in Fig. 4) and in Fig. 7F (DSI = −0.09). Little high-frequency suppression is more apparent for the neuron in Fig. 7F versus Fig. 7E, although the direction selectivity for FM sweep stimuli was greatest for the neuron in Fig. 7E. Thus in some cases the suitability of STRFs to encode FM sweeps is not immediately apparent.

To quantitatively assess how well STRFs can capture FM sweep responses, we predicted the responses to the sweep stimuli using approaches that were similar to those in previous reports (Andoni et al. 2007; Brimijoin and O’Neill 2005). The predicted response can be obtained by following the procedure illustrated in Fig. 8A. First, the convolution between the stimulus with the STRF, in the time domain, was calculated. This result was then summed across frequency. Last, the response was half-wave rectified to approximate the nonlinearity present in extracellular recordings (Andoni et al. 2007). The resulting predicted FM response was then analyzed in the same manner as the observed FM PSTHs (shown in Figs. 2–4). This process was performed for each FM sweep stimulus (n = 10), and then used to calculate DSI and best speed values.

When this procedure was carried out for the example STRFs shown in Fig. 7 the observed and predicted DSIs were often similar. The observed/predicted DSIs for Fig. 7A (0.17/0.16), Fig. 7B (0.05/0.03), Fig. 7D (0.14/0.10), and Fig. 7F (−0.09/−0.03) were in relative agreement. For Fig. 7C (0.33/0.04) and Fig. 7E (−0.14/0.03) the values were not congruent. Over the population of neurons, the STRF-based DSI predictions were significantly correlated with DSI values determined from FM sweep stimuli (Fig. 8B; r = 0.64, P < 0.0001, t-test). Over the DSI range of −0.25 to 0.25 the predictions and data were well matched; i.e., in this range the processing of FM sweep stimuli is well approximated by the quasi-linear prediction method. The greatest deviations from a unity relation between actual and predicted DSI values (solid line) are for highly direction selective units (DSI values >0.25 or less than −0.25). In this regime the STRF underestimates direction selectivity preferences, indicating an increased degree of nonlinear processing relative to neurons that are relatively nondirection selective. For the majority of neurons, however, DSIs are well predicted from the STRF.

FM sweep responses derived best speed preferences and the corresponding STRF predictions are shown in Fig. 8C. Again, a significant correlation is seen between observed and predicted results (r = 0.67, P < 0.0001, t-test). However, the predicted speed range differs from the observed range of best speeds present in the FM data. The observed best speeds range from 10 to 60 oct/s, whereas the range of predictions is limited to 35 to 50 oct/s. This bias toward higher speeds in the STRF...
predictions may reflect systematic influences of the stimulus ensemble used to derive STRFs. For example, the duration of the tone pips was fairly short compared with the duration of the slow FM stimuli. Also, longer-duration adaptation and/or suppression effects might be underestimated by the STRF.

We conclude this study by noting that our result, which showed that spectral bandwidth cannot be adequately estimated from the responses to FM, is similar to that found in the anesthetized squirrel monkey (Atencio et al. 2005; Godey et al. 2005). A recognizable difference between awake and anesthetized monkeys, as illustrated in Figs. 2–4, is that awake monkeys have PSTHs that are robust and qualitatively more tonic (Atencio et al. 2005; Godey et al. 2005). To examine these response differences in further detail we compared the response durations, across all neurons, in the awake owl monkey in the present study, to those in the anesthetized squirrel monkey (Atencio et al. 2005; Godey and colleagues 2005). Figure 9A shows the proportion of neuronal sites in the awake owl monkey and anesthetized squirrel monkey as a function of the response duration to 10 oct/s upward and downward sweeps. Upward and downward sweeps were combined because they were not statistically different in either preparation (P > 0.10, rank-sum test). The slowest experimental FM speeds were chosen because their responses are the most likely to reveal tonic components compared with very fast FM sweeps. The figure shows that a higher proportion of sites in the anesthetized squirrel monkey have shorter response durations (Owl Monkey median: 16.9 ms; Squirrel Monkey median: 7.6 ms, P < 0.0001, rank-sum test), although a considerable range of responses is evident. Figure 9B shows the complete population responses. In Fig. 9C the mean response duration of the population data is shown for both preparations. The mean response duration in this plot is taken as the SD of the Gaussian fits to the PSTHs. The data show that at every speed the awake owl monkey responses are longer than those in the anesthetized case (Owl Monkey median: 12.0 ms; Squirrel Monkey median: 7.6 ms, P < 0.0001, rank-sum test), although a considerable range of responses is evident. Figure 9, A and B shows the complete population responses. In Fig. 9C the mean response duration of the population data is shown for both preparations. The mean response duration in this plot is taken as the SD of the Gaussian fits to the PSTHs. The data show that at every speed the awake owl monkey responses are longer than those in the anesthetized squirrel monkey (P < 0.001, t-test for each sweep speed). Thus the earlier description of qualitative differences between the two preparations is indeed statistically significant. In the awake owl monkey the higher proportion of longer-duration responses, which vary with changing FM sweep duration, may indicate a higher proportion of tonic, compared with phasic, responses in this preparation.

**DISCUSSION**

A goal of this study was to systematically characterize responses of single neurons to FM sweeps in AI of the awake owl monkey, a New World primate with many vocalizations (Wright 1994). An implanted electrode array was used in single animals to achieve this goal, allowing recordings to take place over repeated sessions. Our main findings were that sites in the awake owl monkey AI respond in a temporally precise manner to FM sweeps and that the best responses are well matched to the
natural frequency modulations present in vocalizations. Neurons were also classified according to FM sweep speed tuning and to how similar sweeps in the upward versus downward direction were processed. We found that in the awake owl monkey the STRF was a good descriptor of FM sweep processing. The STRF accurately predicted FM sweep responses over a broad range, although for units that were exceptionally direction selective the STRF prediction procedure did not capture this subpopulation. CFs estimated from the STRF and FM sweeps were highly correlated, as were predicted and observed direction selectivity and best FM sweep speed. In the following discussion we place our findings in the context of earlier work, discuss the use of the STRF in descriptions of AI neural processing, and then compare results obtained from awake and anesthetized New World monkey FM sweep experiments.

**Comparison to other species and studies**

Three key findings have been consistently reported in population studies of FM sweep responses in AI. These are the population distributions of direction selectivity and best speed and the correlation between DSI and CF. The findings are often confounded by the different methodologies used, such as different FM sweep stimuli, different species, and different anesthetics. Despite these variations in published accounts, comparisons may still be made across studies, giving a general picture of FM processing in AI.

In anesthetized squirrel monkey, the global distributions of DSI and best speed in AI are similar to the current values in the awake owl monkey, suggesting that the impact of anesthesia on the coding of dynamic attributes may be much less compared with coding for static stimuli (Godey et al. 2005; Lu et al. 2001; Wang et al. 2005). Distributions of DSI in ferrets are also similar to those in the awake owl monkey, although speed preferences are higher. The discrepancy between these best speed values may be due to species-specific differences in environment or processing, or because some FM sweeps in the ferret study were very fast (>150 oct/s) and thus resembled broadband transients (Kowalski et al. 1995; Shamma et al. 1993).

Rat and squirrel monkey AI, unlike that of the owl monkey, show a significant correlation between CF and DSI (Godey et al. 2005; Mendelson and Ricketts 2001; Zhang et al. 2003). FM sweep responses from AI in the cat using logarithmic (Mendelson et al. 1993) and linear (Heil et al. 1992b) sweeps revealed a DSI population distribution that was centered about 0, as in the present study. A weak correlation between direction selectivity and CF was also found (Mendelson et al. 1993). The lack of a CF and DSI correlation in the current study is puzzling but may be a consequence of limited CF sampling with the fixed recording array. Also, in contrast to a previous report in the awake cat, we did not find any case in which a neuron responded exclusively to one sweep direction, or evidence of two temporally separated phasic responses during the presentation of sweep stimuli (Whitfield and Evans 1965).

In the bat auditory cortex, direction selectivity is more pronounced (Razak and Fuzessery 2006). The majority of DSI values are >0.2, which implies significantly more selectivity, most likely as a consequence of the increased ethological significance of FM in bat echolocation. Although the selectivity is higher, the underlying mechanisms in auditory cortex, and in other stations, may be similar (Razak and Fuzessery 2006). The timing between excitatory and inhibitory inputs may explain intra- and extracellular FM selectivity results (Gordon and O’Neill 1998; Zhang et al. 2003). Duration tuning may also contribute (Razak and Fuzessery 2006), although this aspect requires a more thorough evaluation in New World monkeys than is presently available.

The FM sweep processing results just described appear to have behavioral consequences. In a study of the European starling, linear FM sweeps covered either high or low frequencies, but not the entire frequency range tested (Klump 1991; Phillips et al. 1985). Upward FM sweeps were most detectable when low frequencies were traversed and downward FM sweeps were most easily discriminated when the sweep traversed high frequencies, consistent with the correlation between FM direction selectivity and characteristic frequency. In
the primate, similarly structured studies have yet to be initiated, although from preliminary work similar results would be expected because frequency transitions appear to have behavioral significance for the processing of Japanese macaque coo vocalizations (May et al. 1988, 1989). Any future primate work that incorporates these behavioral aspects with physiological recording techniques would be a significant advance.

The methodologies commonly used in FM studies were compared in AI of the barbiturate-anesthetized ferret (Nelken and Versnel 2000). In the paradigm that most resembles the one used in our report there was a correlation of approximately +1 between the DSI values obtained from logarithmic and linear sweeps. Best speed preferences were also similar, in that neurons that preferred fast/slow linear sweeps also preferred fast/slow logarithmic sweeps. Thus despite different experimental paradigms and anesthetic differences meaningful conclusions may be drawn.

These remarks are qualified because our recordings were predominantly from laminae IIIb and IV. These laminae are the main targets of thalamic input to cortex and therefore display strong responses with latencies shorter than those found in other layers. Although these layers have been used almost exclusively in mapping studies in the auditory cortex very little work has shown how responses in different layers might lead to different functional maps. Few studies have attempted to relate FM sweep responses across different cortical depths (Mendelson and Cynader 1985; Mendelson et al. 1993; Shamma et al. 1993). These studies concluded that best FM speed and direction selectivity are fairly constant in orthogonal cortical penetrations, although they were limited in the number of neurons sampled and in the consistency of penetration depths. FM response properties may vary as a function of cortical depth due to lamina-specific cortico-cortical connections and general receptive field differences (Code and Winer 1985, 1986; Winguth and Winer 1986), although further clarification is needed.

Last, our analysis procedure implicitly uses spike count as the metric for response strength. Other response metrics, such as spike timing, may also convey information. An examination of this issue is beyond the scope of this work, although it points to an additional means that AI neurons may use to transmit FM information.

Similarities and differences of FM sweep response distributions across animals

The results displayed in Fig. 6 show that similar population response patterns are present across individual animals. Our analysis revealed that selectivity for upward or downward sweeps—the speed of these sweeps that gives the best response—and the correlation between upward and downward area–speed curves are similar across animals. For example, it is perhaps surprising that the DSI distributions are centered near zero in each animal. One might hypothesize a bimodal distribution for directional selectivity because owl monkeys must process either upward or downward sweeps that are present in their social calls, although we did not find this result. DSI values were approximately Gaussian distributed and centered at zero. For best speed the distributions showed consistent trends, with maximal responses at speeds between 20 and 60 oct/s, the normal range of FM sweep speeds found in New World monkey twitter vocalizations. These consistent response distributions imply that there is a general range of FM sweep parameters represented in AI. These response distributions do not permit inference to how these sweep parameters are organized spatially across animals in AI, yet they do point to constraints on the characteristics of populations of neurons that are needed for New World monkeys to process complex communication sounds.

**STRF implications and linearity of responses**

For a significant fraction of neurons in one awake owl monkey we computed STRFs. Although the applicability of the STRF metric to awake recordings has been debated (Barbour and Wang 2003b), our results are consistent with those from other sensory systems, where this approach has been used successfully in primary cortical areas of awake monkeys (David et al. 2004; deCharms et al. 1998; DiCarlo et al. 1998; Reich et al. 2000; Rust et al. 2005). A significant finding of our study was that the STRF accurately predicted FM sweep direction selectivity and best speed in the majority of cases. An STRF is a linear filter for a neuron optimized for the stimulus conditions used to create it (Blake and Merzenich 2002). It predicts responses less well to stimuli that are less similar to the STRF-generating stimulus (Blake and Merzenich 2002). We found that the directional selectivity index and speed tuning of neurons were significantly related to those predicted by the STRF in the awake monkey, indicating that spectro-temporal processing, in addition to first-order spectral and temporal processing (Linden et al. 2003; Miller et al. 2002; Sen et al. 2001), may be estimated with some success from tone-pip–derived STRFs in AI.

STRF predictions were less successful for high speed preferences and high direction selectivity. This degradation of the predictions may result from the nature of the stimulus used. By their nature FM sweeps are spectrotemporally correlated stimuli and are fundamentally different from uncorrelated noise-like stimuli. If auditory neurons are activated in nonlinear ways by correlated stimuli then they may respond more robustly to correlated than to uncorrelated sounds (Escabi and Schreiner 2002). This is expected from cortical neurons because natural sounds contain significant correlations (Attias and Schreiner 1997; Brillinger and Irizarry 1998; Voss and Clarke 1975). Thus a possible conclusion is that highly direction selective neurons perform more nonlinear processing on acoustic sounds. Simply using linear prediction, then, may not be adequate, and further extensions of the model may need to be considered in the future. In fact, this realization has been a motivating factor behind recent calculations of higher-order filters in sensory systems (Sharpee et al. 2004; Touryan et al. 2002, 2005; Yamada and Lewis 1999).

In summary, the STRF procedure can be useful in estimating many receptive field properties, although limitations to the method must be kept in mind. For instance, even though the STRF is a useful descriptor of cortical processing and is well defined mathematically, only linear predictions may be deduced, although the inclusion of nonlinearities may improve its accuracy (Eggermont 1993; Marmarelis and Marmarelis 1978). Also, the construction of the STRF is dependent on the responsiveness of the neuron and the stimulus, so care must be taken when choosing a stimulus set. Given these caveats, the STRF
remains a robust, multidimensional metric of cortical processing that should be useful in future awake animal preparations (Elhilali et al. 2004; Fritz et al. 2005).

Comparison between and awake and anesthetized New World monkey responses

The same analysis approach used in our previous anesthetized New World monkey study worked without modification in examining the results from the current study (Atencio et al. 2005; Godey et al. 2005). For neurons that responded to the FM sweep stimuli the Gaussian fits were well matched to the PSTHs. Additionally, the range of distributions for direction selectivity, best speed, speed tuning, and the correlation coefficient were similar to those found in the anesthetized squirrel monkey (see Table 1). It appears that the general response properties of New World monkeys to these dynamic sounds are broadly similar in spite of anesthetic state, although this does not exclude the possibility that differences may be present when behavioral tasks are used.

On more detailed examination differences emerged. One is that, unlike in the squirrel monkey, we did not find a significant negative correlation between direction selectivity and CF. A possible explanation may be that in the present study we did not systematically map AI (due to the limitations in array placement) and the identification of specific clustering of direction selective sites may have been missed. Additional weight must be given to this because the correlation between DSsI and CF was significant, although weak, in the anesthetized squirrel monkey (Godey et al. 2005). Also, in the present study we analyzed single-unit responses, whereas our previous study analyzed multiunit responses, which may have averaged out local individual neuron differences.

Best speed and speed tuning parameters also showed subtle differences from our anesthetized study. Median best speed in the present study was about 35.9 oct/s, a slightly lower value than that found in the anesthetized squirrel monkey (41.4 oct/s), although it is closer to the speeds of upward FM sweeps found in natural New World monkey twitter calls, which are typically between 30 and 35 oct/s. Likewise, speed tuning values were higher in the present study than those in the anesthetized squirrel monkey (0.52 vs. 0.43), indicating more selective processing of FM sweep speeds in the awake owl monkey. These differences suggest that further studies need to be completed to understand the factors surrounding the parameter differences that may be due to actual anesthetic effects, species differences, to single- versus multiunit responses, or to inherent variability in the response. Preliminary analysis does show that FM response parameters are different in the squirrel monkey under different anesthetic states (MA Heiser, personal communication).

Comparison of the onset duration in awake owl monkey and anesthetized squirrel monkeys did reveal a tendency toward longer response durations in the awake animals. This is in line with the notion that more sustained response patterns are a common feature in the awake preparation (Lu et al. 2001). It is intriguing, however, that the global differences in FM processing between the anesthetized squirrel monkey and the awake owl monkey remain quite small, suggesting that dynamic stimulus aspects, as captured by the phasic response components, are less affected by anesthesia than tonic components.

In conclusion, most studies implicitly assume that the responses to FM sweeps are related to spectral receptive fields (sRFs). The rationale is that as an FM sweep traverses different frequencies it will reach the boundary of the receptive field and produce a response. Thus the sRF boundaries and FM sweep trigger frequencies should be related (Heil 1997). The use of trigger frequencies allowed us to obtain accurate estimates of the CFs compared with those obtained from STRFs. However, bandwidth estimates were accurate in only one quarter of the sites. They were quite poor in the remaining sites, most likely because of inhibitory surround effects (Shamma et al. 1993). A recent intracellular study suggests that other frequency-dependent asymmetries in sRFs, such as response magnitude and the timing of excitation and inhibition, may also be responsible for the expression of FM preferences (Zhang et al. 2003).

Because the most significant difference between simple tones and FM sweeps is the stationary nature of tones, this suggests that the inaccuracy in our bandwidth estimates is due to the nonstationary nature of FM sweeps, as well as the partial picture of acoustic processing given by STRFs. Indeed, even though FM sweeps are parameterized sounds, and are less complex than natural vocalizations, they still contain frequency correlations. These correlations, as shown in a previous study in the midbrain, may cause auditory neurons to process sounds in ways that are not predictable from the processing of simple, stationary sounds (Escabi and Schreiner 2002). It follows then that most auditory neurons process spectrally dynamic sounds and spectrally simple sounds in different ways (Schreiner 1998).

Our results show systematic relationships between tone-evoked receptive field response parameters and FM sweep response parameters. In many cases STRFs accurately predict direction selectivity and best speed preferences of the neurons under study, although significant deviations were present. Therefore responses in AI to different acoustic properties are not completely described by the STRF prediction procedure, implying that underlying higher-order nonlinear interactions are responsible for their increased direction selectivity. Studies that use a stimulus that dynamically challenges cortical neurons, and can be used to calculate higher-order response properties, are clearly needed for a more complete picture of auditory processing in AI (Barbour and Wang 2003a; deCharms et al. 1998; Escabi and Schreiner 2002; Miller et al. 2002; Sharpee et al. 2004).

| TABLE 1. Comparison between population FM sweep response parameters for the awake owl monkey and anesthetized squirrel monkey |
|---------------------------------|-----------------|-----------------|
|                                 | Awake Owl Monkey | Anesthetized Squirrel Monkey |
| Direction selectivity           | 0.0641 (0.220)   | 0.0349 (0.236)   |
| Best speed, oct/s               | 35.9 (9.23)      | 41.5 (9.22)      |
| Speed tuning                    | 0.521 (0.184)    | 0.430 (0.181)    |
| Correlation coefficient         | 0.372 (0.544)    | 0.369 (0.559)    |

FM sweep response parameters for the awake owl monkey and the anesthetized squirrel monkey. Values represent the population mean, with SD in parentheses. Owl monkey data are from the present study. Squirrel monkey data are from previous work (Godey et al. 2005). The distributions are similar in the two preparations.
GRANTS
The study was supported by National Institutes of Health Research Grants DC-02260 to C. E. Schreiner, NS-10414 to M. M. Merzenich, and NS-34835 and MH-077970 to C. E. Schreiner and M. M. Merzenich; the Coleman Memorial Fund; and Hearing Research Inc.

REFERENCES


