Functional Specialization in Auditory Cortex: Responses to Frequency-Modulated Stimuli in the Cat’s Posterior Auditory Field

PETER HEIL AND DEXTER R. F. IRVINE

Department of Psychology, Monash University, Clayton, Victoria 3168, Australia

Heil, Peter and Dexter R. F. Irvine. Functional specialization in auditory cortex: responses to frequency-modulated stimuli in the cat’s posterior auditory field. J. Neurophysiol. 79: 3041–3059, 1998. The mammalian auditory cortex contains multiple fields but their functional role is poorly understood. Here we examine the responses of single neurons in the posterior auditory field (P) of barbiturate- and ketamine-anesthetized cats to frequency-modulated (FM) sweeps. FM sweeps traversed the excitatory response area of the neuron under study, and FM direction and the linear rate of change of frequency (RCF) were varied systematically. In some neurons, sweeps of different sound pressure levels (SPLs) also were tested. The response magnitude (number of spikes corrected for spontaneous activity) of nearly all field P neurons varied with RCF. RCF response functions displayed a variety of shapes, but most functions were of low-pass characteristic or peaked at rather low RCFs (<100 kHz/s). Neurons with strong responses to high RCFs (high-pass or nonselective RCF response function characteristics) all displayed spike count—SPL functions to tone burst onsets that were monotonic or weakly nonmonotonic. RCF response functions and best RCFs often changed with SPL. For most neurons, FM directional sensitivity, quantified by a directional sensitivity (DS) index, also varied with RCF and SPL, but the mean and width of the distribution of DS indices across all neurons was independent of RCF. Analysis of response timing revealed that the phasic response of a neuron is triggered when the instantaneous frequency of the sweep reaches a particular value, the effective F. For a given neuron, values of effective F were independent of RCF, but depended on FM direction and SPL and were associated closely with the boundaries of the neuron’s frequency versus amplitude response area. The standard deviation (SD) of the latency of the first spike of the response decreased with RCF. When SD was expressed relative to the rate of change of stimulus frequency, the resulting index of frequency jitter increased with RCF and did so rather uniformly in all neurons and largely independent of SPL. These properties suggest that many FM parameters are represented by, and may be encoded in, orderly temporal patterns across different neurons in addition to the strength of responses. When compared with neurons in primary and anterior auditory fields, field P neurons respond better to relatively slow FMs. Together with previous studies of responses to modulations of amplitude, such as tone onsets, our findings suggest more generally that field P may be best suited for processing signals that vary relatively slowly over time.

INTRODUCTION

The auditory cortex of mammals contains multiple fields distinguished by anatomic and physiological criteria (for reviews, see Aitkin 1990; Clarey et al. 1992). However, it is only in some echolocating bats, and only with respect to the processing of echoes or pulse-echo pairs brought about by reflection of the (rather stereotyped) biosonar pulses emitted by these species, that the role of different fields with respect to sound processing is reasonably well understood (for review see, e.g., Suga 1988). Our lack of understanding of the functional roles of the multiple fields in other species is largely due to the fact that cortical areas outside the primary auditory cortex (AI) have received very little attention.

The posterior field (P), one of several tonotopically organized areas in the cat auditory cortex (Reale and Imig 1980), is one of the better studied areas (e.g., Kitzes and Hollrigr 1996; Orman and Phillips 1984; Phillips and Orman 1984; Phillips et al. 1995, 1996; Schreiner and Urbas 1988). A remarkable feature of field P is its very high percentage of neurons with responses to increments in the sound pressure level (SPL) of tonal stimuli that show nonmonotonic behavior (Kitzes and Hollrigr 1996; Phillips and Orman 1984; Phillips et al. 1995), i.e., with increasing SPL, the number of spikes increases to a maximum, defining a “best SPL” (after Brugge and Merzenich 1973), and then decreases. Many of these cells cease to discharge at high SPLs particularly when tones are shaped with short rise times, leading to circumscribed frequency versus SPL response areas (Heil and Irvine 1998; Phillips and Orman 1984; Phillips et al. 1995). This predominance of nonmonotonicity in field P, and the observation that best SPL was rather stable against variation of tone burst rise time across a range of short rise times, whereas with longer rise times “tuning” to SPL disappeared (Phillips et al. 1995), has led to the proposal that field P might be an area specialized for coding of intensity, particularly of brief sounds, which necessarily have short rise times (Kitzes and Hollrigr 1996; Phillips and Orman 1984; Phillips et al. 1995).

However, a problem for this concept is that neurons in field P responded particularly well to tones with long rise times (see, e.g., Figs. 3–6 in Phillips et al. 1995). We have recently reinvestigated this proposal and have arrived at a different conclusion regarding the possible functional role of field P (Heil and Irvine 1998), viz. that field P may be better suited to process slowly varying envelopes, such as onsets of tones of long rise times, than rapidly varying envelopes, such as onsets of tones of high SPL and short rise time (Heil and Irvine 1998). Our reinvestigation of the proposed specialization of field P for intensity coding was driven by the recognition that neurons in this field, like those in AI, respond preferentially to the onset of a tonal stimulus (Phillips and Orman 1984; Phillips et al. 1995, 1996; Reale and Imig 1980), but that paradoxically these onset responses previously have been analyzed with respect to SPL, a parameter that characterizes the steady-state portion of such stimuli. Our reanalysis of the strength and of the timing of the onset responses of field P neurons therefore was focused on the initial time course of the envelope as it had been in our previous studies of AI (Heil 1997a,b; Heil and Irvine 1996).
In effect, tone burst onsets constitute aperiodic amplitude modulations (AM). The present study was undertaken to examine the hypothesis that field P may be involved, more generally, in the processing of slowly time-varying signals by focusing on the responses of field P neurons to another class of dynamic signals, viz. frequency-modulated (FM) sounds. FMs are almost ubiquitous components of many natural sounds and are likely to be of some significance for cats. FM sounds are produced by cats themselves for intraspecies communication (e.g., Brown et al. 1978; Härtel 1975) and by their major prey, such as birds and rodents. Human speech also contains many FM components and formant transitions (see Pickett 1980). FM sounds can vary along multiple dimensions, such as form of modulation (e.g., linear, exponential, sinusoidal), modulation range or frequency excursion, modulation frequency, ‘speed’ of FM, direction of frequency change, and others. It is therefore nearly impossible to characterize the responses of single neurons along all of those dimensions. To test our hypothesis and to allow for the most straightforward comparison with neuronal responses to FM stimuli obtained from other auditory cortical fields in this species (Heil et al. 1992b,c; Mendelson and Cynader 1985; Mendelson et al. 1993; Phillips et al. 1985; Tian and Rauschecker 1994) and comparison with neuronal responses to aperiodic AM (i.e., to tone burst onsets) (Heil and Irvine 1998), we have selected aperiodic linear FM sweeps of either direction of FM and parametrically varied FM speed, i.e., the rate of change of frequency (RCF).

In most previous studies of FM processing, the emphasis has been on analyzing the number of spikes or discharge rate of neurons as a function of FM parameters, and more-or-less sharply tuned filters or selectivities have been demonstrated for a range of parameters, including RCF and FM direction (for auditory cortex, see, e.g., Heil et al. 1992b,c; Kowalski et al. 1995; Mendelson and Cynader 1985; Mendelson et al. 1993; Shamma et al. 1993; Tian and Rauschecker 1994). Alternatively, or additionally, one could envisage that information about FM parameters is encoded in the temporal sequence of response initiation of different neurons. To shed more light on the feasibility of rate and temporal coding of FM stimuli, we therefore have analyzed aspects of the timing of the responses of field P neurons to FM sweeps in addition to the strength of these responses.

**Methods**

**Surgery**

The data for this study were obtained from the same 10 adult cats (5 of each sex; weighing between 2.2 and 4.8 kg) that generated the data of our previous report (Heil and Irvine 1998). For detailed accounts of anesthesia and surgery, the reader is referred to that study. Briefly, anesthesia was introduced with 40 mg/kg ip pentobarbitone sodium and maintained at a surgical level throughout the experiment with either intravenous pentobarbitone (cats 96-13, 96-26, 96-27, 96-29, 96-33, and 96-37) or an intramuscular cocktail of ketamine and xylazine (cats 96-39, 96-40, 96-42, and 96-43). The electrocardiogram was monitored continuously, and rectal temperature was held near 38°C. The trachea was cannulated, and a head holder fixed to the skull. A round-window electrode and electrode applied to a length of fine-bore polyethylene tubing, allowing static pressure equalization within the middle ear, were inserted in each bulla, which then were resealed. The skull and the dura overlying field P in the caudal bank of the posterior ectosylvian sulcus (PES) and parts of AI on the middle ectosylvian gyrus were removed. A specially designed Perspex chamber was fixed to the skull surrounding the opening. The chamber was filled with warm saline and sealed with a glass plate, on which a small hydraulic microdrive was mounted and which housed the glass-insulated tungsten microelectrode.

**Acoustic stimulation and recording procedures**

The cat was located in a sound-attenuating chamber. Stimuli were produced digitally (Tucker Davis Technology) and presented via calibrated sealed sound delivery systems. Each system consisted of a STAX SRS-MK3 transducer in a coupler whose sound delivery tube fitted snugly into the meatal stub. The calibration procedure is described and a representative transfer function of the system illustrated in Heil et al. (1992b). The frequency response of the system was flat within ±5 dB up to ~12 kHz, beyond which the output decreased by ~5 dB/octave. Compound action potential thresholds were monitored via the round window electrodes.

Neuronal activity was recorded in field P along electrode penetrations down the caudal bank of PES, for lengths of ~6 mm, at distances of ~0.5–1.5 mm from the center of the sulcus and at various dorsoventral positions. Tone bursts, clicks, and noise bursts were used as search stimuli and were delivered at various SPLs and varying laterality. Once a neuron was well isolated and found to be responsive to tone bursts, initial estimates of its characteristic frequency (CF; frequency of lowest response threshold), its response area in frequency versus SPL coordinates (henceforth termed response area), and its preferred laterality of stimulus presentation (viz., monaural ipsilateral, monaural contralateral, or binaural with identical stimuli at each ear (i.e., diotic)) were obtained using audiovisual criteria. Then quantitative data on the neuron’s response area were obtained, generally with diotic stimulation. All tone bursts were of 200 ms duration and were shaped with cosine-squared rise/fall functions of 8.5 ms (or in some cases 17 ms) rise and fall times. Tone amplitudes ranged from below threshold to ~90 dB SPL (re 20 μPa) generally in steps of 10 dB. In some instances, and over near-threshold ranges, step size was 5 dB. The frequency range and step size on a linear scale were selected so that they covered the neuron’s response area with reasonable resolution (see Figs. 1A and 2A) (Heil and Irvine 1998). Frequency/SPL combinations were presented at 1 Hz in pseudorandom order with ~15 repetitions, each with a different pseudorandom order. Intermixed with the tone burst trials were trials (1 of 20) during which no stimulus was presented and which were used to obtain a measure of the neuron’s spontaneous discharge rate (SR).

If the neuron was still well isolated, its responses either to tone bursts of different rise times, as described in Heil and Irvine (1998), or to FM stimuli, or both, were recorded. The selection of parameters for FM sweeps was based on the neuron’s response area. FM sweeps were shaped with linear rise and fall times of 5 ms, during which the frequency was held constant. The start and end frequencies present during the rise and fall time, respectively, generally fell outside the neuron’s frequency response range. At the signal’s full amplitude the frequency changed linearly from either a low to a high frequency (upward) or vice versa (downward). The rate of change of frequency (RCF) was varied parametrically, in some cases from as low as 0.5 kHz/s to as high as ~4,000 kHz/s, in equal steps on a logarithmic scale for each FM direction. Because start and end frequencies were fixed, FM sweep duration decreased with increasing RCF. FM sweeps longer than 3 s could not be generated by the system. Up to 20 repetitions of FM sweeps of different RCF and direction were presented in pseudorandom order. In several cases, FM sweeps of different SPLs were tested. The times of occurrence of discriminated spikes were stored on disk with 10-μs resolution for off-line analysis.
Data analysis

Response areas were based on the number of spikes discharged during the entire tone burst duration. Contour plots of the response areas (e.g., Figs. 1A and 2A) were generated with commercial software (Surfer, Golden Software). In addition, mean first-spike latency and its standard deviation (SD) were obtained for each frequency/SPL combination. For analysis of FM responses, it was inappropriate to simply count the number of spikes that occurred during the stimulus. Because FM sweep duration was inversely related to RCF, a neuron’s response to short-duration/high-RCF stimuli could occur after the ends of such sweeps (e.g., Fig. 2E).

Therefore, the response analysis windows included a 100-ms period immediately after each sweep. Furthermore, the number of spikes discharged was corrected for the SR of the neuron. This was necessary because otherwise any spontaneous activity would have contributed more spikes to longer-duration/lower-RCF stimuli. SR was measured in the 100-ms interval preceding each sweep, and the average SR was derived from spike counts in all intervals. The number of spontaneous spikes expected during the response analysis window was then extrapolated from this average. The number of spikes per FM sweep, corrected for SR, was the measure of response R and was plotted against RCF to obtain RCF response functions for each FM direction (e.g., Fig. 1, B and C).

As in previous studies of FM processing (see, e.g., Britt and Starr, 1976; Heil et al. 1992a–c; Kowalski et al. 1995; Mendelson and Cynader, 1985; Phillips et al. 1985), a measure of a neuron’s directional sensitivity (DS) was obtained by a comparison of the responses to upward ($R_{\text{up}}$) and downward FM sweeps ($R_{\text{down}}$)

$$DS = (R_{\text{up}} - R_{\text{down}})/(R_{\text{up}} + R_{\text{down}})$$

This index can vary between $+1$ and $-1$, where the sign indicates the preference for upward (+) or downward (–) sweeps, and the magnitude the degree of preference. A disadvantage of this index is that it is affected drastically by small changes in $R$, when the values of $R$ are low. For example, when for a given pair of FM sweeps only a single spike is recorded in any number of repetitions for each FM direction, DS is zero. But, when for one direction no such spike occurs, DS is either 1 (when $R_{\text{down}}$ is 0) or $-1$ (when $R_{\text{up}}$ is 0). Therefore, a DS index was calculated for a given RCF only when the response to at least one FM direction was $>0.5$ spikes per stimulus.

For FM sweeps, the time of occurrence of the first spike in the response analysis window will not accurately reflect the response latency when the neuron is spontaneously active particularly for long-duration sweeps and when the response occurs late during the sweep. Therefore, the poststimulus time histograms (PSTHs) to each individual repetition of a given sweep as well as the composite PSTH to all repetitions of that stimulus were examined off-line. In cases with low spontaneous activity and tightly time-locked responses (e.g., Fig. 1, E and I), inspection of these PSTHs allowed a reasonable judgement of which first spikes, if any, might have been spontaneous ones. In such cases, it was also possible to discern when a neuron had failed to respond to a particular repetition, although a presumed “spontaneous” spike might have occurred much later during that repetition as compared with the others. Therefore, to obtain more reliable estimates of response latency, analysis windows were selected to exclude such presumed spontaneous first spikes, and latency (relative to stimulus onset) and SD of the first spike within those windows were computed. However, in cases of higher spontaneous activity and/or temporally diffuse responses to a FM sweep, this procedure was not practicable, and no analysis of latency was performed in these cases.

RESULTS

Database

Only a minority of $\sim 200$ well-isolated neurons in field P responded to the tone burst stimuli tested (see also Heil and Irvine 1998), and response areas, which were used to select FM parameters, were obtained from 80 neurons. Sixty-six of those neurons were studied with FM sweeps over a range of RCFs for both FM directions at least one SPL. The CFs of these neurons ranged from 0.4 to 18 kHz, with a median of 3.4 kHz. Eighteen neurons were studied at two (12 neurons) or more (6 neurons) SPLs. The data obtained from neurons 96-42/07 and 96-27/06 are among the most extensive and are illustrated in Figs. 1 and 2. These figures will be used throughout the results section to outline analytic procedures and some of the basic findings.

Sensitivity to the rate of change of frequency

Start and end frequencies of FM sweeps fell $\sim 1–2$ kHz outside the boundaries of a given neuron’s response area (Figs. 1A and 2A). The responses of nearly all field P neurons to such FM sweeps were affected by their RCF. A total of 93 RCF response functions for either FM direction were obtained from the 66 neurons studied at one or multiple SPLs. These RCF response functions displayed a variety of shapes. For convenience, these functions were classified into six categories: band-pass (BP), low-pass (LP), high-pass (HP), complex (CX), nonselective (NS), or nonresponsive (NR). A function was classified as BP if responses to RCFs higher and lower than the RCF eliciting the maximum response (best RCF) were $<50\%$ of maximum. If only RCFs higher or lower than the best RCF were $<50\%$ of maximum, the function was classified as LP or HP, respectively. If the responses to all RCFs tested were within $50\%$ of maximum, the function was classified as NS. Functions that exhibited multiple distinct maxima were classified as CX. If none of the RCFs tested elicited $>0.5$ spikes/stimulus, the function was classified NR.

The largest proportion of functions recorded exhibited LP characteristics (viz., 35 and 36%, for upward and downward FM sweeps, respectively), followed by BP (26 and 24%), CX (13 and 17%), NR (15 and 14%), HP (8 and 9%), and NS (3 and 0%). When the RCF response functions to the two FM directions were compared for a given neuron and latency when the neuron is spontaneously active particularly for these functions were classed into six categories. For analysis of latency was performed in these cases.

The SPL of FM sweeps in a given direction often had pronounced effects on RCF function category. For upward FM sweeps, this procedure was not practicable, and no analysis of latency was performed in these cases.

The SPL of FM sweeps in a given direction often had pronounced effects on RCF response function characteristics, such as function category or best RCF. In only 4 of 18 neurons, 2 of which were classified as NR at both SPLs tested, were these characteristics invariant with changes in SPL. Drastic effects of SPL on RCF function category are seen for upward FM sweeps in neuron 96-42/07 (Fig. 1B).
FIG. 1. Responses of neuron 96-42/07 to tones and to FM sweeps. A: response area, obtained with diotically presented tone bursts shaped with 8.5-ms cosine-squared rise and fall times. Amplitude and frequency steps were 10 dB and 0.5 kHz, respectively. Measure of response is the total number of spikes per stimulus recorded during each tone (200 ms). Isoresponse areas are shown by different shades of gray as identified by the scale (right). Plots were generated with commercial software (Surfer, Golden Software), using the kriging interpolation algorithm and a ratio of search radii reflecting the relative spacing of frequency and amplitude steps. Four horizontal lines represent the range (3–8 kHz) and SPLs (20, 30, 55, and 80 dB SPL) of the FM stimuli. Positions of the arrowheads on these lines pointing to the right and left represent the effective F0 to upward and downward FM sweeps, respectively, as determined from analysis of response latency to FM sweeps and as described in RESULTS. B and C: responses to upward and downward FM sweeps plotted as a function of rate of change of frequency (RCF). Different symbols identify different sound pressure levels (SPLs; see key in B). Negative values (below broken line) arise when the number of spikes recorded is less than that expected from spontaneous activity. D: directional sensitivity as a function of RCF and SPL (symbols as in B). E and F: poststimulus time histograms (PSTHs) of the responses of this neuron to 15 repetitions of upward and downward FM sweeps of 20 dB SPL and different RCF. Each PSTH starts at stimulus onset and ends 100 ms after stimulus offset. Stimulus offset is identified by a small marker below the baseline of each PSTH. Duration of the FM portion of the stimulus, given by ΔF/RCF, decreases with increasing RCF. Because in this case ΔF = 5 kHz (see A), and each stimulus has rise and fall times of 5 ms each, the time scales of the PSTHs decrease from 2,610 ms for a RCF of 2 kHz/s (top PSTHs) to 113 ms for a RCF of 1,970 kHz/s (bottom PSTHs). G and H: response latency for upward and downward FM sweeps plotted as a function of FM sweep duration (i.e., stimulus duration minus rise and fall times) and SPL (symbols as in B). I and J: index of frequency jitter for upward and downward FM sweeps plotted as a function of RCF and SPL (symbols as in B).
FIG. 2. Responses of neuron 96-27/06 to tones and to FM sweeps. A: response area, obtained with diotically presented tone bursts shaped with 1.7-ms cosine-squared rise and fall times. Amplitude and frequency steps were 10 dB and 1 kHz, respectively. Other conventions as in Fig. 1A. B–D: responses to upward (B) and downward (C) FM sweeps, and directional sensitivity (D) plotted as a function of RCF. Different symbols identify different SPLs (see key in B). E: PSTHs of the responses to 10 repetitions of 20 dB SPL downward FM sweeps from 14 to 5 kHz at different RCFs. Other conventions as in Fig. 1, E and I. F: response latency for downward FM sweeps plotted as a function of sweep duration and SPL (symbols as in B). G and H: SD of response latency and corresponding frequency jitter for downward FM sweeps plotted as a function of RCF and SPL (symbols as in B).
FIG. 3. A: responses of neuron 96-40/13 to upward and downward FM sweeps of 2–10 kHz and 15 dB SPL as a function of RCF. B: resulting directional sensitivity as a function of RCF. Response latency for upward and downward FM sweeps (see key in A) as a function of sweep duration. D: index of frequency jitter for upward and downward FM sweeps (see key in A) as a function of RCF.

The function obtained at 20 dB SPL was classified as LP, those obtained at 30 and at 55 dB SPL as BP, and that obtained at 80 dB SPL as NR. The corresponding functions for downward FM sweeps were all classified as BP, but best RCF varied considerably and nonuniformly. Best RCF was highest at 55 dB SPL, but decreased for lower and higher SPLs (Fig. 1C). Somewhat less severe effects of SPL are seen for neurons 96-42/14 (Fig. 4A) and 96-27/06 (Fig. 2, B and C), although for the latter neuron RCF response function category to upward FM sweeps is affected and there are changes in best RCF for both FM directions. In different neurons, best RCF could increase, decrease, or change nonuniformly with SPL. There were no obvious systematic relationships of the SPL-related changes in best RCF or in the maximum number of spikes per FM sweep with characteristics of the neuron’s response area or of its spike count—level function for CF tones. However, because the number of different SPLs that could be tested before the cell was lost was limited, systematic relationships cannot be ruled out. Best RCF for upward sweeps was not significantly correlated with the width of the response area at the corresponding SPL or with CF, while for downward sweeps there were weak correlations of the logarithm of best RCF with those of response area width and CF (r = 0.310 and r = 0.234, respectively; n = 79; P < 0.05).

In any event, the above observations reveal that the best RCF measured at a single SPL may not be representative for that neuron. Figure 7, which shows the distributions of best RCFs recorded from field P neurons the RCF response functions of which fell into BP, LP, HP, and CX categories, therefore includes the best RCFs at all SPLs tested in a given neuron. Side maxima, obtained from CX functions, are also included and identified by different shading. The histograms show that, while the distributions of best RCFs in field P are wide, covering at least three orders of magnitude, low best RCFs predominate. For many LP functions, the best RCF assigned was identical with the lowest RCF that could be tested (e.g., neuron 96-40/13; Fig. 3A). Therefore, the histograms in Fig. 7 likely underestimate the relative frequency of very low best RCFs.

The predominance of neuronal responses in field P to relatively low RCFs also is emphasized by Fig. 8, which shows for both FM directions the average number of spikes per sweep per neuron as a function of RCF. Data from all neurons and SPLs are included, except data from NR response functions. The apparent peaks of these average functions near 2 and 6 kHz/s for downward and upward FM sweeps, respectively, possibly result from the fact that, due to the upper stimulus duration limit of 3 s (see METHODS), very low RCFs could not be tested in some neurons, among them neurons with particularly strong responses and RCF response functions with LP characteristics (e.g., neuron 96-40/13; Fig. 3A). Thus we cannot exclude the possibility that, had it been possible to test very low RCFs for all
neurons, the average functions would peak at still lower RCFs or would even be of LP shape.

**Sensitivity to FM direction**

The directional sensitivity of field P neurons was assessed by a DS index, which was calculated for certain RCFs at all SPLs tested (see METHODS). For a given neuron, this index generally varied as a function of RCF and SPL. For neuron 96-42/07, for example, and at low SPLs most DS indices were positive, whereas at 55 dB SPL, the index changed systematically from positive to negative values with increasing RCF and at 80 dB SPL, all indices were negative (Fig. 1D). In some cases, the DS magnitude appeared to change systematically with RCF, though without obviously reversing sign (e.g., neurons 96-42/14 and 96-39/04; Figs. 1B and 5B). In other cases, DS indices appeared to scatter unsystematically around some mean, with or without reversing sign (e.g., Figs. 2B, 3B, and 6B). The considerable variation of DS with RCF for most neurons is emphasized by Fig. 9A, which shows all calculated DS indices plotted as a function of RCF. Data obtained from the same neuron and with the same SPL are connected. Note that hardly any lines run horizontally but rather course in all directions. Nevertheless, at any given RCF, DS indices span the entire range without obvious shifts in the distribution. The latter point is more clearly illustrated by Fig. 9B, which plots the mean (●) ± 1 SD of DS indices averaged within RCF octave bands as a function of RCF. Note that the distributions of DS values are rather symmetrical around zero (−−−−) and of similar width along the entire RCF range, except at the extreme ends, where only very few data points were collected.

For a better comparison with other studies, the DS index at best RCF and the mean DS, obtained by averaging the DS indices across different RCFs for a given neuron and SPL, were extracted, even though for some neurons neither measure provides a representative estimate of their directional sensitivity. Figure 9C provides a scatterplot of the mean DS against the DS index at best RCF. The two measures are correlated significantly (r = 0.835), yet data points scatter considerably around the line of unity slope (−−−−). The rectangular box marks DS magnitudes of 0.3, a boundary often chosen as the criterion for direction selectivity (e.g., Heil et al. 1992a–c; Mendelson and Cynader 1985; Tian and Rauschecker 1994). Using this criterion, ~58% of points would qualify as reflecting direction selectivity when based on the response at best RCF and ~53% when based on mean DS. Nevertheless, highest directional sensitivities were not necessarily associated with the best RCF. The data illustrated in Figs. 4 and 5 represent examples where the best RCFs are in fact correlated with the lowest directional sensitivities (cf. Figs. 4, A with B, and 5, A with B).

Figured 9D illustrates the effects of SPL on mean DS. Data from a given neuron are represented by the same symbol and connected by lines. Most lines are fairly horizontal, indicating little effect of SPL on mean DS, at least over the
limited SPL ranges studied. Nevertheless, some lines deviate markedly from horizontal, such as those for neuron 96-42/07.

Timing of responses

For RCF response functions of LP characteristic, one could argue that the neuron’s response is simply proportional to the time spent by an FM sweep within the neuron’s excitatory response area. However, this was not the case. The responses of field P neurons to FM sweeps were generally rather phasic in nature (Figs. 1, E and I, and 2E) and, at least for long-duration/low-RCF sweeps, were shorter than the time spent by the sweep within the neuron’s response area. This is evident, for example, from consideration of Fig. 1 A and E or I. At 20 dB SPL, ~20–25% of the duration of the FM sweeps, which cover the 3–8 kHz range, is spent traversing the response area of neuron 96-42/07. However, the duration of that neuron’s response to low RCFs is only some 10% of the duration of the stimulus.

Laten
cy and effective $F_i$. The PSTHs illustrated in Figs. 1, E and I, and 2E show that the timing of the phasic response relative to the duration of the FM stimulus shifts systematically as a function of RCF. For example, for upward FM sweeps and for an RCF of 2 kHz/s, the response of neuron 96-42/07 occurs during the first half of the stimulus, whereas with increasing RCF, the response shifts more and more toward the end of the stimulus or even beyond (Fig. 1E). Such a “phase” shift also was observed routinely in the responses of neurons in AI (Heil et al. 1992b). There it was argued that such a shift would occur if the responses of a given neuron to FM sweeps in a given direction were in fact triggered at, or evoked by, the same instantaneous frequency (termed effective $F_i$), independent of RCF, and if the responses occurred with some fixed delay $L_0$ after the occurrence of that frequency. The response latency $L_{RCF}$ then would be given by

$$L_{RCF} = kD_{RCF} + L_0 + RT$$

where RT is the rise time of the FM stimulus (viz., 5 ms; during which the frequency was constant) and $D_{RCF}$ the duration of the FM portion of the stimulus, which is inversely proportional to RCF. Thus $L_{RCF}$ should increase linearly with $D_{RCF}$, and the effective $F_i$ can be calculated from the slope $k$ of that relationship. For upward and downward FM sweeps, effective $F_i$ would be given by the terms ($F_L + k*\Delta F$) and ($F_H - k*\Delta F$), respectively, where $F_L$ and $F_H$ are the lowest and highest instantaneous frequencies of the sweep and $\Delta F$ is the modulation depth, given by $\Delta F = (F_H - F_L)$.

Figure 1, F and J, shows, for neuron 96-42/07 and for upward and downward FM sweeps of different SPL, plots of the response latency against the FM duration $D_{RCF}$. The data for the two or three longest sweeps at 20 dB SPL are not shown to provide a higher resolution for the data for shorter sweeps but follow the same relationship as those of shorter sweeps. For most of these conditions, the growth of latency with FM duration is very well approximated by a linear function. Only for upward sweeps of 55 dB SPL (●) might there

---

**Fig. 5.** A–D: responses of neuron 96-39/05 to upward and downward FM sweeps of 0.1–6 kHz and 50 dB SPL. Organization of the figure is analogous to that of Fig. 3.
be systematic deviations from linearity. Linear functions also provided good approximations of the data for other neurons (see Figs. 2F and 3C–6C), although in one case, there was clearly a systematic deviation from linearity (not shown). Only 3 of 96 correlation coefficients were <0.95. The medians of the correlation coefficients for upward and downward sweeps were 0.996 and 0.997, respectively. The value of the delay $L_0$, obtained from the y intercepts of the linear regressions, was correlated with the near-minimum latency of the neuron recorded with tone bursts ($r = 0.523; n = 97; P < 0.0001$). The slope and origin of that relationship were not different from 1 and 0, respectively (not shown).

**Effective $F_i$ and Response Area.** For neuron 96-42/07, the values of effective $F_i$ obtained from the slopes $k$ of the relationships between latency and FM duration are identified in Fig. 1A by arrowheads on four horizontal lines. The lines represent the range and SPLs of the FM stimuli and the positions of the arrowheads pointing to the right and left represent the effective $F_i$ to upward and downward FM sweeps, respectively. Note that, in general, effective $F_i$ and CF or best frequency do not coincide. Rather, effective $F_i$ for upward sweeps is always lower than that for downward sweeps. In other words, the response to FM sweeps is triggered when the frequency is modulated toward the best frequency. Effective $F_i$ values change with SPL and roughly follow the neuron’s response area boundary, although they do not coincide exactly with that boundary. Very similar data were obtained for neuron 96-27/06 (Fig. 2A).

The general validity of these findings is illustrated in Fig. 10, which shows scatterplots of effective $F_i$ for upward and downward sweeps against CF (Fig. 10A), for upward sweeps against the low-frequency border (Fig. 10B), and for downward sweeps against the high-frequency border (Fig. 10C) of the response area. The borders were defined by a response of 10% of the maximum response. Note that for upward sweeps, effective $F_i$ values are mostly lower and for downward sweeps mostly higher than CF, emphasizing that for FM sweeps starting outside and traversing a neuron’s response area the response is generally triggered when the frequency is modulated toward CF. Effective $F_i$ values are associated closely with the boundary of the response area on the side of the starting frequency (near-frequency boundary), but, in most instances and at least for downward sweeps, the frequency has to be modulated beyond that boundary for the response to be triggered.

In summary, these data are consistent with the idea that the response of a field P neuron to FM sweeps traversing its response area is triggered whenever the instantaneous frequency reaches a particular value, the effective $F_i$, that is roughly independent of the speed but dependent on the direction and SPL of FM.

**Precision of Spike Timing and Frequency Jitter.** The measure of response latency was the mean latency of the first spike within the analysis window relative to stimulus onset. The jitter around the mean on individual presentations was quantified by the standard deviation (SD) of the latency of the first spike. As shown for neurons 96-42/07 and 96-27/06 (Figs. 1, G and K, and 2G), SD generally decreased
with increasing RCF, though the slope of such functions became shallower with increasing RCF, even when plotted on double-logarithmic axes. SPL appeared to have little effect on SD.

Although in absolute terms the precision of spike timing increases with increasing RCF, it may be more appropriate to consider the precision relative to the rapidity of changes in the signal, which of course increases with RCF. Variability in response latency corresponds to variation in the frequency at the instant when the response occurs or, taking the delay $L_0$ into account, when the response is triggered. The SD of a neuron’s response latency therefore can be converted into an index of this frequency variability or jitter (in kilohertz). Because for all FM sweeps used here the frequency changed linearly as a function of time, the conversion is achieved by simply multiplying SD with RCF. When the SD decreases with increasing RCF with a slope of 0 in a log-log plot, this index of frequency jitter remains constant. Conversely, if a neuron’s SD to FM sweeps were constant, i.e., independent of RCF, the frequency jitter would increase as a function of RCF with a slope of 1 in a log-log plot (Fig. 11A). Note that this conversion is independent of whether or not the neuron’s response is triggered at the same instantaneous frequency at different RCFs.

For neuron 96-42/07, the index of frequency jitter obtained in this way is plotted as a function of RCF, separately for upward and downward sweeps and for different SPLs (Fig. 1, H and L). Note that for both FM directions and for all SPLs, with the possible exception of downward sweeps of 20 dB SPL, the frequency jitter increases systematically, and with increasing slope, as RCF increases. A similar be-
FIG. 9. Directional sensitivity of field P neurons. A: directional sensitivity (DS) indices plotted as a function of RCF. Data obtained from the same neuron and with FM sweeps of the same SPL are connected. B: mean (●) ± 1 SD (○) of DS indices in A averaged within octave bands of RCF and plotted as a function of RCF. C: scatterplot of mean DS (see RESULTS) against DS at best RCF. — — — , unity slope and 0,0 origin. Box demarcates DS magnitudes of 0.3, often arbitrarily selected as the lower limit for directional selectivity. D: mean DS plotted as a function of the SPL of FM sweeps. Data from the same neuron are identified by the same symbol and connected by lines. — — — in A, B, and D mark the boundary between preferences for upward and downward sweeps (positive and negative DS indices, respectively).

behavior was also seen in other neurons (Figs. 2H and 4D–6D). In agreement with the lack of effects of the SPL of FM sweeps on SD, SPL appeared to have only small, if any, effects on frequency jitter. Note that the frequency jitter versus RCF functions obtained from a given neuron for different SPLs are in rather close register and do not reveal obvious SPL-dependent differences (Figs. 1, H and L, 2H, and 4D).

Figure 11 plots, for all neurons and SPLs for which this analysis was possible, frequency jitter as a function of RCF, separately for upward (A) and downward (B) sweeps. Data obtained from the same neuron and with the same SPL are connected. Although there is crossing of lines, the general increase of frequency jitter with increasing RCF is obvious. Note that in general the slopes for the neuronal data are <1 at low RCFs but, at least for some neurons (see further), approach 1 as RCF increases. The overall increase of frequency jitter with increasing RCF is more clearly illustrated in Fig. 11C, which plots the mean frequency jitter of values within octave bands of RCF as a function of RCF, separately for upward and downward sweeps.

Also note that, on average, the frequency jitter for downward sweeps is slightly larger than for upward sweeps (Fig. 11C). A similar result emerges when the frequency jitter obtained from a given neuron and for a given RCF and SPL to downward sweeps is plotted against that to the corresponding upward sweeps (Fig. 11D). In 219 of 341 such comparisons, frequency jitter to downward sweeps was larger than to upward sweeps (points above line of unity slope in Fig. 11D). This difference was highly significant (P < 0.001 and P < 0.000001 before and after log-transformation of the data; 2-tailed paired Student’s t-test).

RELATIONSHIP OF TEMPORAL PRECISION AND RCF RESPONSE FUNCTION CHARACTERISTICS. Figure 11, A and B, also illustrates that the range of frequency jitter (on a logarithmic scale) covered by the neuronal population is rather wide at low RCFs but narrows as RCF increases. This appears due to fact that those neurons for which the index of frequency jitter was large at low RCFs did not respond at high RCFs. For neuron 96-40/13, for example, and for either FM direction, the frequency jitter at low RCFs is relatively large but narrows as RCF increases. In contrast, the frequency jitter of neuron 96-42/14 at low RCFs is relatively large but narrows as RCF increases.
half of the data points at low RCFs but only a small proportion at high RCFs.

These observations indicate a relationship between SD, or its corresponding frequency jitter, and RCF response function category or best RCF. However, such a relationship is not easy to quantify. As SD depends on RCF, one approach could be to measure the SD to one selected RCF for all neurons. However, there was no RCF to which all neurons responded reasonably well. In a given neuron, SD decreases with increasing RCF (Figs. 1, G and K, and 2G), but in many neurons SD did not reach, or come close to, an asymptotic minimum. As an alternative measure of precision of response timing, we therefore have used the near-minimum SD, obtained from the neuron’s responses to tone burst onsets (Heil and Irvine 1998). Use of this measure is exploratory, as we have insufficient data to establish the relationship between minimum SD for FM stimuli and tone bursts.

Figure 12A shows a plot of the RCF response function category for both FM directions against tone-burst minimum SD. Neurons with RCF response functions that were classified as HP or NS, i.e., neurons that responded well to high RCFs, all had rather small minimum SDs (i.e., <5 ms and even <2.1 ms for all but 1 neuron). In contrast, most neurons with rather large minimum SD (in excess of 6 ms) either did not respond to any of the FM sweeps tested (NR) or showed LP characteristics, i.e., preferred low RCFs and responded poorly, if at all, to high RCFs. The RCF response functions of those neurons with large minimum SDs that fell into the BP or CX categories peaked at rather low RCFs (see numbers in kHz/s at data points in Fig. 12A). Figure 12B, which provides a scatterplot of best RCFs against minimum SD, shows that neurons of small minimum SD can be associated with the entire range of best RCFs, whereas those of large minimum SD all prefer slow RCFs. These data provide additional support for the relationship between precision of spike timing and RCF response function category and best RCF that was indicated in Fig. 11.

RELATIONSHIP BETWEEN CHARACTERISTICS OF TONE BURST ONSET RESPONSES AND RCF RESPONSE FUNCTIONS. The responses of neurons in auditory cortex, including field P, to tone bursts are dominated by spikes triggered by tone burst onsets. In essence, the onsets constitute aperiodic unidirectional modulations of SPL (peak pressure), and it is therefore of interest to see whether there are correspondences between a neuron’s response to such modulations of amplitude and to those of frequency. When the tone burst rise time is fixed and its SPL is varied, the rate of change of peak pressure is covaried. Although the magnitude of a neuron’s onset response is not strictly a function of that rate, characteristics of its level functions are informative: when for a fixed rise time the onset response increases nonmonotonically with SPL, it also increases monotonically with rate of change of peak pressure; when the response changes nonmonotonically with SPL, it changes nonmonotonically with rate of change of peak pressure. For the purpose of a crude comparison with RCF response function categories, we have classified the level functions of field P neurons, obtained at CF and with short rise times (≈8.5 ms; cosine-squared rise function), as strongly nonmonotonic, if the neuron ceased to discharge at high SPLs (≈90 dB SPL), or as monotonic/weakly non-

![Figure 10](http://jn.physiology.org/)

**Figure 10.** Effective $F_i$, obtained from the slope of the linear relationship between response latency and FM duration, for upward and downward FM sweeps (see key in A) plotted as a function of CF (A) and of the near-frequency border of the response area at the equivalent SPL (B and C).
monotonic, if the neuron had not ceased to discharge at 90 dB SPL. Figure 13 shows the distribution of RCF response function categories (pooled for the 2 FM directions) for those two classes of level functions. Neurons with RCF response functions that were classified as NS and as HP, i.e., that responded well or particularly well to high RCFs, were associated exclusively with the monotonic/weakly nonmonotonic category, i.e., they also responded well to rather high rates of change of peak pressure. Similarly, neurons with RCF response functions that were classified as LP and as BP, i.e., that responded well to low and intermediate, but not to high RCFs, were associated most often with the strongly nonmonotonic category, i.e., they also did not respond well to high rates of change of peak pressure. Overall, however, the correspondence between the shapes of a neuron’s level function for CF tones and those of its RCF functions at some SPL is limited. The RCF response function category cannot be reliably predicted on the basis of the shape of the level function.

**DISCUSSION**

We have shown here that FM sweeps are effective stimuli for neurons in field P of cat auditory cortex. Because FM sweeps were not used as search stimuli and only those neurons that also responded to tones were tested with FM sweeps, we cannot address the issue of whether there might be neurons in field P that respond selectively to FM stimuli. In the following sections, we will first discuss our basic findings and compare them, where possible, with those of other studies. On a more speculative basis, we then will consider the possible implications of our data for rate and temporal coding of FM stimuli and for functional specializations of different fields within auditory cortex.

**Sensitivity to the rate of change of frequency**

We have demonstrated here that field P neurons are generally quite sensitive to RCF. RCF response functions displayed a variety of shapes, with LP functions being most prominent. Best RCFs assigned to the various RCF response functions were distributed over at least three orders of magnitude, but low best RCF predominated. Sensitivity to RCF is not a unique feature of neurons in field P but has been demonstrated in other auditory cortical fields and at more peripheral levels and in a variety of species (e.g., Brit and Starr 1976; Felsheim and Ostwald 1996; Heil et al. 1992a–c; Kowalski et al. 1995; Mendelson and Cynader 1985; Men-
delson et al. 1993; Møller 1971, 1972; Sinex and Geisler 1981; Suga 1973; Tian and Rauschecker 1994). A direct comparison of most of these data with those of the present study is not straightforward, given the differences in, for example, stimulus paradigms, data analysis, response measures, or temporal response patterns of the neurons recorded (e.g., phasic or sustained). A few studies, however, have employed stimuli similar to those used in our study, viz., linear FM sweeps of different RCF that completely traversed the neuron’s excitatory response area, and have explored the (phasic) responses of single or multiunits in other fields of cat auditory cortex, viz. in AI (Heil et al. 1992b,c; Mendelson and Cynader 1985) and in the anterior auditory field (AAF) (Tian and Rauschecker 1994). Most neurons in these fields responded best to rather high RCFs, most often displaying high-pass RCF response functions or band-pass functions with peaks at relatively high RCFs. The distributions of best RCFs in AI and AAF were such that the number of units increased rapidly with increasing best RCF (for AI, see Fig. 5 in Heil et al. 1992b; for AAF, see Fig. 11 in Tian and Rauschecker 1994). When these histograms are compared with those for field P (Fig. 7), it appears that on average best RCFs in field P are considerably lower than in either AI or AAF. However, to judge the magnitude of this difference, it has to be kept in mind that the data in the different studies were obtained from neuronal populations of different CF and that best RCF may increase with CF (Heil et al. 1992a; Tian and Rauschecker 1994; present study).

The relationship between best RCF and CF may be caused by two factors. First, neurons of higher CF tend to have

FIG. 12. Relationship of RCF response function characteristics and temporal precision. A: RCF response function category for upward and downward FM sweeps (see key) plotted against the minimum SD obtained from the neuron’s responses to tone burst onsets. Numbers at some of the symbols for BP and CX type functions identify “best RCFs” (in kHz/s) and, for CX, also the positions of side maxima. See RESULTS for further description. B: best RCFs for upward and downward FM sweeps plotted against minimum SD.

FIG. 13. Relationship of the degree of nonmonotonicity of level functions to tone bursts and RCF response function category. Level functions were classified as either strongly nonmonotonic or as monotonic/weakly nonmonotonic. See RESULTS for further description.
broader frequency response ranges on a linear scale than neurons of lower CF. Because the response time of linear filters is inversely proportional to bandwidth, neurons with broader bandwidths would be expected to have shorter integration times and hence should respond better to faster RCFs. Such a relationship was found here for downward FM sweeps and also was observed in the chick field L (Heil et al. 1992a) and in cat AI (Mendelson et al. 1993). Consistent with this interpretation, the latter authors also observed that best RCF increased with SPL, as does bandwidth. Second, the frequency representation in the cochlea [see, e.g., Liberman (1982) for cat and Manley et al. (1987) for chick] and in central auditory nuclei [see, e.g., Merzenich et al. (1975); Rajan et al. (1993) for cat AI; Heil and Scheich (1985, 1992) for chick field L] is better described by a logarithmic than by a linear relationship so that, at least to a first approximation, frequency octaves tend to occupy similar distances along the basilar papilla or along the tonotopic axis of the central neuronal structures. Therefore, and as was pointed out by Mendelson et al. (1993) and by Felsheim and Ostwald (1996), exponential, but not linear, FM sweeps produce a close to uniform motion of excitation along the basilar membrane and the tonotopic axis of central neuronal structures. A linear FM sweep of any given RCF would thus correspond to a higher rate in octaves per time when the sweep covers a low-frequency range as compared with a high-frequency range.

Determination of whether RCFs in field P are in fact lower than those in AI and AAF therefore requires comparison of values for samples matched for CF. Our AI data (Heil et al. 1992b,c) unfortunately do not allow this comparison, as they were obtained from neurons along high-frequency (15–22 kHz) isofrequency domains, and consequently there was almost no overlap in the CF distribution with that of our field P sample (see Database). The only study of cat auditory cortex from which data on the CF-RCF relationship for linear FM sweeps can be extracted is that of Tian and Rauschecker (1994). Their Fig. 12A indicates, for upward FM sweeps, median RCFs of ~85 and 130 kHz/s for AAF neurons with CFs ~5 kHz and between 5 and 10 kHz, respectively. The corresponding values from our field P sample are 9 and 12 kHz/s. Although this comparison is qualified by various differences in methodology between the two studies, it strongly supports the view that RCFs in field P are lower than those in AAF at least.

Neurons that responded well to high RCFs, i.e., that had RCF response functions of NS or HP characteristic, had monotonic or only weakly nonmonotonic level functions for CF tones (Fig. 13), i.e., also responded well to rapid aperiodic AM. Recall that the onset of a tone, shaped with some rise function of fixed rise time, constitutes an AM whose rapidity increases with SPL. Similarly, most of the neurons that did not respond well to high but did to low and intermediate RCFs, i.e., with RCF response functions that were classified as LP and as BP, did not respond well to rapid AM, i.e., their level functions were strongly nonmonotonic. In this context, it is also worth noting that for cat AI, Schreiner et al. (1988) reported a positive correlation between best RCF and threshold. Likewise, in our study of the topographic organization of various response properties along the isofrequency domain of cat AI (Heil et al. 1992c), we observed that patches of low thresholds to CF tones and patches of low best RCFs were congruous. Low thresholds, in turn, were correlated with strongly nonmonotonic level functions (Heil et al. 1994; Schreiner et al. 1992). Hence, low-threshold nonmonotonic functions to CF tone bursts reflect sensitivity to slow, and suppression/inhibition of responses to rapid, AM. Although the RCF response function cannot be predicted reliably from the shape of the level function, these correspondences and correlations suggest some common mechanisms. In particular, they suggest that the inhibition that causes level functions to be nonmonotonic (e.g., Fuzessery and Hall 1996) may be one of the factors contributing to a neuron’s RCF sensitivity.

Directional sensitivity

We have shown here that the responses of at least some field P neurons are also sensitive to the FM direction. Again, FM directional sensitivity is not unique to field P but can be found at nearly all levels of the auditory pathway as peripheral as the cochlear nucleus (Britt and Starr 1976; Erulkar et al. 1968). Only for auditory nerve fibers and some cochlear nucleus neurons are the PSTHs recorded with FM sweeps of opposite direction roughly mirror-images of one another (Britt and Starr 1976; Sinex and Geisler 1981), such that the numbers of spikes elicited by FM sweeps of opposite direction are roughly the same but the temporal patterns are reversed. The present study was not designed to explore the mechanisms behind the FM directional sensitivity of field P neurons, which must await further study. Basic mechanisms discussed for the generation of FM directional sensitivity include differential effects of spatiotemporal summation of excitatory postsynaptic potentials (EPSPs) caused by orderly (i.e., tonotopic) arrangements of inputs along dendrites (e.g., Erulkar et al. 1968; Fuzessery 1994; Heil et al. 1992a); ideas prompted by the theoretical work on the electrical properties of dendritic trees of Rall (1964; see also review by Segev 1992), and differential effects of temporal summation of EPSPs and inhibitory postsynaptic potentials caused by asymmetric lateral inhibitory areas (e.g., Fuzessery 1994; Heil et al. 1992a; Kowalski et al. 1995; Shamma et al. 1993; Shannon-Hartman et al. 1992; Suga 1965, 1973; for review, see Fuzessery 1994).

When attempting a comparison of directional sensitivity across studies, it has to be kept in mind that a neuron’s directional sensitivity depends on the stimulus configuration. For example, Phillips et al. (1985) used linear FM sweeps of 2-kHz excursion, embedded in constant-frequency segments corresponding to the start and end frequencies of the sweep, and varied the center frequency of the sweep relative to the neuron’s response area. They found, in cat AI, that a neuron responded best when the frequency was modulated toward its CF. Consequently, a neuron preferred upward sweeps when the modulation range was below its CF and downward sweeps when that range was above its CF and showed no directional preference when the range was around CF. This behavior, which appears to be closely related to our finding (see further) that a neuron’s response is triggered when the frequency is modulated toward the CF, was independent of the directional sensitivity revealed with FM sweeps that traversed the neuron’s entire frequency response range (Phillips et al. 1985). Versnel and Nelken (1994) found, in ferret AI, that directional sensitivities obtained with linear versus logarithmic FM sweeps were uncorrelated.
Response triggering by a particular instantaneous frequency also has been demonstrated at other levels of the auditory pathway, in different species and with other forms of FM (Felsheim and Ostwald 1996; Heil et al. 1992a; Schulze et al. 1997; Whitfield and Evans 1965). These studies also have shown that, when the FM starts outside and then enters a neuron’s excitatory response area, the responses are triggered when the frequency is modulated toward the CF (see also Tian and Rauschecker 1994). Effective $F_I$ appears to be more or less closely associated with the near-frequency border of the neuron’s response area (Figs. 1A, 2A, and 10) (Felsheim and Ostwald 1996; Heil 1997c; Heil et al. 1992a,b; Schulze et al. 1997), and thus depends not only on FM direction but also on SPL. In AI, it was found that effective $F_I$ was correlated very closely with the frequency where the slope of the neuron’s response function to tones of different frequencies but identical amplitudes (frequency response function, FRF) was steepest, i.e., where the magnitude of the first derivative of the FRF was maximal (Heil 1997c; Schulze et al. 1997). That point along the FRF identifies the frequency around which the neuron is most sensitive to a change in frequency of a tone burst. The tight correlation between that frequency and the effective $F_I$ seen with FM sweeps may therefore not be coincidental. In any event, the match of effective $F_I$ with the borders of the response area or the frequency of maximum FRF slope emphasizes their significance, as opposed to that of the best frequency, for processing of FM sounds, at least of those that start outside and traverse a neuron’s response area, as discussed previously (Heil 1997c). Because the effective $F_I$ for FM sweeps of a given direction and SPL is independent of the speed of FM for linear (Heil 1997c; Heil et al. 1992b; Schulze et al. 1997; present study) and for exponential (Felsheim and Ostwald 1996) sweeps and independent of the modulation frequency for sinusoidal FM (Heil et al. 1992a; Whitfield and Evans 1965), it is probably fair to say that the effective $F_I$ is independent of the time course of the modulation. The mechanisms that underlie this phenomenon are unknown, but, as will be outlined below, it may have important implications for the neural representation of FM stimuli.

The SD of the response latency to linear FM sweeps decreased rather systematically as a function of RCF in the neurons studied. To our knowledge, this characteristic of a neuron’s response has not received attention with respect to FM stimuli. It was proposed here that SD should be considered relative to the rapidity of the changes in stimulus frequency to obtain a measure of how precisely a neuron’s response occurs at, or is triggered by, a particular instantaneous frequency on different presentations of the same FM stimulus. This measure (i.e., the index of frequency jitter) is independent of whether or not that frequency changes with RCF or other stimulus variables. Although the measure was based on the timing of only the first spike of the response, the responses were generally quite phasic and the total number of spikes evoked, even by optimal FM stimuli, was often rather low (e.g., Figs. 1, B and C, 2, B and C, 5A, and 6A; see also Fig. 8). Unlike response magnitudes and DS indices, frequency jitter showed a more uniform behavior in all neurons, increasing systematically with increasing RCF, and SPL appeared to have little, if any, effect on these functions. On average, the frequency jitter was
somewhat smaller for upward than for downward sweeps (Fig. 11, C and D). It is conceivable that this difference may have its origin in cochlear mechanics because upward sweeps would cancel, or partially cancel, travel time along the cochlear partition, whereas downward sweeps would augment travel time and because these differences are reflected in the latency of cochlear nucleus neurons to FM sweeps (Shore et al. 1987). At any given RCF, there was a range of frequency jitter, but neurons with precision of spike timing at low RCFs that was poor often did not respond to high RCFs (Fig. 11, A and B). A similar relationship emerged when the precision of spike timing to tone burst onsets was compared with the neuron’s RCF response functions (Fig. 12). The possible implications of these findings for the neural representation of FM stimuli will be discussed further.

Implications for FM coding

RATE CODING. With respect to the processing or representation of FM sounds in the auditory system, the emphasis has traditionally been on the analysis of the response magnitudes of the neurons (such as firing rate or other measures of response). The finding that a neuron’s response depends, for example, on the direction or on the speed of FM, with some neurons being direction selective by some criterion (e.g., DS > 0.3) or apparently tuned to RCF, has implied that FM direction and RCF are in fact encoded in that response. The response properties often are compared with the corresponding parameters occurring in those sounds in the animal’s environment that are deemed of particular importance for the animal, such as vocalizations of conspecifics or of prey or predators (Felsheim and Ostwald 1996; Fuzessery 1994; Gaese and Ostwald 1995; Heil 1992a,b; Poon et al. 1991; Tian and Rauschecker 1994). Yet, it is obvious that such sounds, although important, constitute only a very small proportion of the diverse range of sounds in the environment that we, and many animals, are exposed to and that we can discriminate and identify. Indeed, the correspondence between the apparent tuning properties of neurons in species such as rats or cats and the equivalent acoustic parameters in such vocalizations is rather poor (Felsheim and Ostwald 1996; Gaese and Ostwald 1995; Heil 1992a,b; Poon et al. 1991; Tian and Rauschecker 1994).

As shown here for linear FM sweeps, a neuron’s response does not only depend on RCF and FM direction but also on the sweep’s SPL. At a given RCF, a neuron’s response can increase, decrease, or change nonmonotonically with SPL. The response also depends on the modulation range. This was shown for rat IC’ (Felsheim and Ostwald 1996) and, as detailed earlier, for cat AI (Phillips et al. 1985). These results demonstrate that a neuron’s response to FM sweeps is some joint function of several parameters, including at least RCF, direction, SPL, and modulation range, but possibly also other parameters, such as form of modulation (see Versnel and Nelken 1994). Clearly, the responses of several neurons would have to be compared to resolve the ambiguities with regard to these signal properties that are inherent in the response of a single neuron.

TEMPORAL CODING. Most of these ambiguities also could be resolved by comparing the relative timing rather than, or in addition to, the magnitudes of responses from different neurons. As was outlined above, a neuron’s response to FM sweeps is triggered when the instantaneous frequency reaches a particular value, the effective \( F_t \). This frequency appears to be independent of the speed, and hence of the time course, of FM. Effective \( F_t \) is associated, more or less closely in different neurons, with the border of the response area, i.e., depends on the width of the response area and on CF. An FM sweep therefore would trigger responses from a population of neurons in a particular temporal sequence (and spatial order, given a tonotopic map), which depends on the neurons’ effective \( F_t \) and on the time course of the stimulus frequency. In a population of neurons with similar delay \( L_o \), the triggering sequence corresponds to the sequence in which the responses actually occur. In effect, the response triggering sequence tracks the signal’s changing frequency. Consider an array of five neurons (numbered 1–5) with increasing effective \( F_t \) from 1 to 5. Any response triggering sequence in ascending numerical order then could represent an upward sweep, a sequence in descending order a downward sweep. The range of modulation could be represented by the range of effective \( F_t \); a wider frequency excursion might activate neurons 1–5, whereas a more narrow excursion might activate only neurons 2–4. The intervals between response triggering in different neurons could represent the speed of FM, with intervals being short for a high RCF and vice versa. Thus the rate at which responses from different neurons are triggered increases with the rate of change frequency. In the same way, the intervals could represent the form of the modulation. For example, in a map where the effective \( F_t \) gradient is logarithmic, exponential sweeps would produce identical intervals while for linear sweeps intervals are the longer the higher the instantaneous frequency.

Note that in this ensemble coding scheme all neurons that respond to a particular FM stimulus would be involved in its coding, irrespective of whether or not they are tuned to, or selective for, a particular parameter, such as RCF or FM direction. Furthermore, it seems favorable for such temporal coding that the responses are rather phasic, a finding routinely made in auditory cortex (Heil et al. 1992a,b; Mendelson et al. 1993; Phillips et al. 1985; Tian and Rauschecker 1994; present study) but also at more peripheral levels (e.g., Britt and Starr 1976; Felsheim and Ostwald 1996; Poon et al. 1991). Through the involvement of different neurons, it seems possible to track FM sounds in which the frequency changes on time scales too rapid to be encoded in changes in the firing rate of single neurons, a problem emphasized by Bialek and Rieke (1992). Such latency-place codes may not only operate for FM signals but could be involved in other sensory processes (for review, see Cariani 1995).

In the context of this idea of temporal tracking of FM stimuli, a functional explanation of the fact that neurons with large minimum SD do not respond to high RCFs (Fig. 12) arises. The orderly sequence in which responses from neurons with different effective \( F_t \) are triggered and that would track the changes in the signal’s instantaneous frequency is increasingly jeopardized when the speed of modulation increases, the more so the lower the neurons’ temporal precision. Hence, tracking of fast FMs is best performed by neurons with relatively small frequency jitter, whereas the responses of neurons with large jitter would interfere with the tracking of such stimuli. The situation
is analogous to temporal coding by the spike train of a single neuron where the transmittable information depends on the number of distinguishable spike sequences, which in turn of course increases with the timing precision (Bialek and Rieke 1992).

If the direction, range, speed, and form of FM were coded in the way described above, i.e., by latency-place mechanisms, the response magnitude of different neurons might be used to represent other stimulus properties. A likely candidate is SPL, which might be changing as a function of frequency. Neurons with responses to a given FM stimulus that are triggered at the same instantaneous frequency might differ in their response magnitude, so that the signal’s SPL at that point in time might be represented by the ratio of the responses of these neurons. The same argument was made for envelope transients (tone onsets) where subpopulations of neurons with responses that are triggered at the same time after stimulus onset, hence at the same instantaneous amplitude, differ in response magnitude (Heil 1997b; Heil and Irvine 1998).

Another interesting aspect of our data that emerged from the analysis of the timing of responses was the finding that the index of frequency jitter increased with increasing RCF (Fig. 11) and did so in all neurons and in much the same way. Furthermore, this behavior appeared independent of SPL, at least in the few neurons and over the suprathreshold ranges studied. It is therefore tempting to speculate that this increasing frequency jitter, rather than DS values, might underlie the psychophysical results reported by Tian and Rauschecker (1994), viz. that discriminability of FM direction decreases with increasing RCF. Assuming that the responses of neurons in the human auditory cortex to FM stimuli show a similar behavior as those reported here for cat field P and that humans might rely on a temporal tracking mechanism, one would expect that, for linear FM sweeps of a given modulation range, direction discriminability should deteriorate with increasing RCF. That discriminability also might be expected to be largely independent of SPL and for a given RCF, should improve with increasing modulation range. In contrast, if the discriminability relied on directional sensitivities in the neuronal responses, the distribution of DS indices (Fig. 9, A and B) would suggest that performance should be independent of RCF.

Areal specializations in auditory cortex

Compared with AI, many neurons in field P have nonmonotonic level functions, i.e., respond well to tones of relatively low and intermediate, but not of high SPL, particularly when tones are shaped with short rise times (Heil and Irvine 1998; Kitzes and Hollrigel 1996; Phillips and Orman 1984; Phillips et al. 1995). However, field P neurons, like those of AI, respond predominantly to tone onset, i.e., that part of the stimulus that constitutes an aperiodic AM. Nonmonotonic level functions therefore reflect good responses to relatively slow or intermediate AM and poor responses to relatively fast AM (Heil and Irvine 1998). On average then, and compared with AI, field P responds better to slowly varying envelopes. Similarly, Schreiner and Urban (1988) found that neuron clusters in field P preferred lower modulation frequencies of sinusoidally amplitude-modulated tones than clusters in AI and AAF. Note that for such tones the envelope varies more rapidly than the higher modulation frequency. The present study has revealed that, on average, field P neurons also respond well to slow, not to fast, FMs. Taken together, these neuronal properties therefore suggest that field P may be best suited or specialized for processing relatively slowly time-varying signals.

This proposed time-domain specialization appears to be linked to the precision of spike timing. As outlined above for FM stimuli and for tone burst onsets elsewhere (Heil 1997b; Heil and Irvine 1997, 1998), the resolution of the proposed temporal tracking mechanism for changes in stimulus frequency or amplitude depends on the precision of spike timing. As the precision is relatively poor in field P (Heil and Irvine 1998; Phillips et al. 1995), the field’s resolution of details in the stimulus time course would be rather limited.

We are grateful to J. F. Cassell, E. Hartley, V. N. Park, D. Sketchley, R. R. Williams, and members of the general departmental staff for excellent technical support and to Dr. R. Rajan for assistance in preliminary experiments.

The study was supported by the National Health and Medical Research Council of Australia.

Address reprint requests to P. Heil.

Received 5 August 1997; accepted in final form 10 February 1998.

REFERENCES


