

Responses of Chinchilla Inferior Colliculus Neurons to Amplitude-Modulated Tones with Different Envelopes

DONAL G. SINEX,¹ JENNIFER HENDERSON, HONGZHE LI, AND GUANG-DI CHEN²

¹*Department of Speech and Hearing Science, Arizona State University, Tempe, AZ 85287-1908, USA*

²*School of Pharmacy, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73190, USA*

Received: 7 May 2001; Accepted: 20 December 2001; Online publication: 27 February 2002

ABSTRACT

Responses of single neurons in the inferior colliculus of the chinchilla to amplitude-modulated tones were obtained. In one condition, the modulating waveform was a low-frequency sinusoid (SAM tone). In the other, the modulator was a trapezoid with fixed parameters, used to create trains of brief tone bursts presented at various repetition rates (TRAM tone). Modulation frequency (or repetition rate) was varied over the range from 10 to 200 Hz. Many individual neurons exhibited strong selectivity for modulator type. Neurons with pauser discharge patterns to steady-state tones usually exhibited greater responsiveness to SAM tones than to TRAM. In contrast, neurons that responded transiently to steady-state tones usually exhibited greater responsiveness to TRAM tones than to SAM. Neurons with sustained responses to steady-state tones responded strongly to both types of modulated tones. The selectivity for modulator type suggests that transient neurons may play a different functional role in the representation of envelopes than do other types of neurons.

Keywords: inferior colliculus, amplitude modulation, physiology, chinchilla

INTRODUCTION

The temporal envelopes of natural sounds such as speech, animal vocalizations, and environmental sounds convey information that contributes to their detection and identification. For example, human speech is intelligible when it is recoded to remove detailed spectral information but to maintain its temporal envelope in as few as 4–8 frequency bands (Shannon et al. 1995; Loizou et al. 1999). It is also known from psychophysical studies that the auditory system can make use of coherent amplitude modulation in different frequency bands to improve the detection of low-level tones (Hall et al. 1984) and the identification of reduced-cue speech (Carrell and Opie 1992). The ability to process the amplitude envelope in this way likely contributes to the segregation and grouping of spectral components from sounds that have overlapping spectra but must be processed simultaneously. This capability has been referred to as “auditory scene analysis” (Bregman 1990) or auditory image analysis (Yost 1991). These observations suggest that the representation of the amplitude envelope is an important aspect of auditory neural processing. Accordingly, responses of auditory neurons to amplitude-modulated tones have often been studied. At the level of the inferior colliculus, responses to sinusoidally amplitude-modulated (SAM) tones have been reported for several species. However, responses to carriers modulated by waveforms other than sinusoids have been studied less often.

One rationale for measuring responses to SAM tones and for determining the SAM-tone modulation transfer function (MTF) is that it may lead to a better

Correspondence to: Donal G. Sinex, Ph.D. • Department of Speech and Hearing Science • Arizona State University • Box 871908 • Tempe, AZ 85287-1908. Telephone: (480) 727-6153; fax: (480) 965-0965; email: sinex@asu.edu

understanding of the representation of other sounds with nonsinusoidal envelopes. For example, Delgutte et al. (1998) developed a model based in part on neural MTFs obtained with SAM tones that was reasonably successful in predicting responses to the envelope of speech. The MTFs at each modulation frequency included both gain (Rees and Møller 1983, 1987) and phase terms. This allowed the MTF to be treated as a filter that could be applied to predict how envelopes of complex sounds like speech would be represented by neurons. Predictions of the discharge patterns of inferior colliculus (IC) neurons were less accurate than predictions for auditory nerve fibers and cochlear nucleus neurons, however. That result suggests that the SAM-tone MTF does not fully characterize the ability of IC neurons to represent time-varying amplitude information, even in a more complete form that incorporates phase as well as amplitude information.

This article considers the extent to which IC responses to amplitude-modulated waveforms depend upon the nature of the modulation. If the representation of the envelope depends strongly on the type of modulation, the MTF based on responses to sinusoidal modulation will be of limited value for predicting or understanding the representation of real-life sounds whose envelopes may include rapid or abrupt amplitude changes. The major finding of this study was that chinchilla IC neurons do exhibit strong selectivity for different types of modulation. Neurons' selectivities for particular modulation types were strongly associated with their temporal discharge patterns to steady-state tones.

METHODS

Animal preparation

Responses were obtained from single neurons in the IC of the chinchilla. The procedures were similar to those described previously (Nuding et al. 1999; Chen and Sinex 1999; Sinex and Chen 2000) and were approved by the Institutional Animal Care and Use Committee at Arizona State University. Animals were anesthetized by an intraperitoneal (IP) injection of sodium pentobarbital (65 mg/kg) or Dial in urethane (0.5 mL/kg of a solution containing 40 mg/mL of diallylbarbituric acid and 40 mg/mL of urethane). Supplemental IP injections of the original anesthetic were given as required to maintain a surgical level of anesthesia. The selectivity for modulator type described below was observed with both anesthetics. The skull was opened, and a portion of the cerebral cortex was aspirated to expose the dorsal surface of the IC. Tungsten electrodes insulated with Parylene-C (Micro probe, Inc., Potomac, MD) with impedances

greater than 1 M Ω were placed just above the dorsal surface of the IC under visual control, then advanced from outside the sound-attenuating booth (Acoustic Systems, Austin, Tx) with a hydraulic microdrive (Trent-Wells, South Gate, CA).

The electrode trajectory passed through the central nucleus of the IC, in a parasagittal plane. The neurons exhibited secure responses to tone bursts and sharp tuning with characteristic frequencies (CF) that increased with increasing depth. These properties are consistent with those of neurons in the central nucleus of the IC (Merzenich and Reid 1974; Langner and Schreiner 1988). Recording locations were confirmed by histology in 8 animals in which 27 units were studied. In those animals, one or two small electrolytic lesions were made at the location of a recorded unit or at the end of an electrode track, which allowed the locations of other units in the same animal to be determined. In animals for which histology was not obtained, the stereotaxic coordinates of the recording sites were consistent with those for which the recording sites were confirmed and with those of units from previous studies for which the same surgical approach and anatomical reconstruction procedures were used (Nuding et al. 1999; Chen and Sinex 1999). It is likely that nearly all of the neurons included in these analyses were from the central nucleus (ICc), but some may have been recorded at locations dorsal to the ICc.

Data collection and analysis

Stimulus generation, stimulus presentation, and data collection were controlled by a PC-compatible computer. Waveforms were digitally synthesized. The outputs of digital-analog converters were passed through programmable attenuators and anti-aliasing filters, then delivered to a closed acoustic system incorporating two Beyer DT48 earphones, one for each ear. For each experiment, the acoustic system was calibrated with a probe-tube microphone (Etymotic ER-7 or B&K 4134) whose tip was located 2–3 mm from the tympanic membrane.

When a neuron was isolated, estimates of CF and binaural sensitivity were obtained manually. Following this initial characterization of the neuron, a detailed frequency-response map was obtained with an automated procedure that presented tones at multiple frequencies and levels (Nuding et al. 1999). These data were used to make a more precise estimate of CF and threshold at CF. Responses to steady-state CF tones with 5-ms rise times were also obtained and used to classify each neuron according to the shape of its peristimulus time (PST) histogram (Nuding et al. 1999). For this study, neurons with three PST histogram shapes were of greatest interest: transient,

pauser, and sustained. Transient units exhibited a burst of short-latency spikes that ended within 20–25 ms after the onset of a CF tone. Pauser units exhibited a short-latency transient response and a second sustained response, separated by a silent interval of several millisecond duration. Sustained units responded throughout the CF tone, although the PST histogram often included a larger peak at tone onset. A small number of units with longer first-spike latencies (>20 ms) or with temporally complex discharge patterns to pure tones were also encountered.

In most cases, modulated tones were presented monaurally to the contralateral ear. Binaural presentation was also used if a neuron exhibited weak responses to contralateral stimulation or if the recording time allowed both monaural and binaural data to be collected. When binaural presentations were made, the stimuli were presented with the interaural time delay or level difference that elicited the largest response to CF tones. In the small number of cases for which both monaural and binaural stimuli were presented, the results were not strongly dependent on the ear or ears of presentation. All of the individual examples presented below were obtained with monaural stimulation of the contralateral ear. For 43 of the 52 units represented in Figure 8, responses were obtained with monaural contralateral stimuli.

Stimuli

The stimuli were tones whose amplitudes were modulated by one of two different low-frequency waveforms. In one condition, the modulating waveform was a low-frequency sinusoid (SAM tone). In the other, the modulator was a trapezoid with fixed parameters, used to create trains of brief tone bursts that were presented at various repetition rates (TRAM tone).

SAM tones were synthesized by multiplying a sine wave carrier by a second sine wave, the modulator. The depth of modulation was always 100%, to be most comparable to the depth of trapezoidal modulation, which could not be varied (described below). The modulator starting phase was delayed by 0.25 cycle, so that the modulation depth was always 100% onset and offset. No additional ramp was applied. Modulator frequency (Fmod) usually varied between 10 and 200 Hz. The long-term spectrum of a SAM tone has a component at the carrier frequency and sidebands at frequencies corresponding to the carrier frequency \pm Fmod. With 100% depth of modulation, the amplitude of each sideband is 6 dB lower than the carrier amplitude.

For TRAM tones, the modulating waveform was a 5-ms trapezoid that resulted in a brief tone burst with a fixed 1-ms linear rise time, a fixed 3-ms steady state, and a 1-ms linear fall time. Trains of these tone bursts

were presented at various repetition rates. For convenience, the term Fmod will be used to refer to the repetition rate of TRAM tones as well as to the modulation frequency of SAM tones. The range of Fmod for TRAM tones was the same as for SAM tones, 10–200 Hz. The long-term spectra of TRAM tones had more complicated shapes and always included more components than did the spectra of SAM tones. The spectrum had its largest component at the carrier frequency and multiple sidebands at the carrier frequency \pm integer multiples of Fmod. The sideband amplitudes were attenuated if they fell near frequencies corresponding to the carrier frequency \pm 0.25 kHz. The spacing of these spectral minima was a consequence of the duration and rise-fall time of the trapezoidal modulating waveform.

Examples of modulated waveforms are shown in Figure 1. For clarity, only the first 100 ms each waveform is shown, although the actual stimuli had durations of 250–300 ms. Figure 1A represents SAM tones with Fmod = 20 and 100 Hz. Each SAM waveform was periodic, repeating after the modulation period, which was 50 or 10 ms for these examples. The rate of amplitude change and the length of time during each cycle over which the stimulus amplitude exceeded an arbitrary criterion necessarily covaried with Fmod. Low-frequency sinusoidal modulation thus produced waveforms with slow amplitude change and a longer “effective duration” within each modulation cycle. “Effective duration” is used as a descriptive term to refer to the length of time during a cycle that the waveform amplitude was large enough to elicit spikes. For higher modulation frequencies, amplitude change was more rapid and the effective duration was shorter.

Figure 1B represents waveforms modulated at the same rates by the 5-ms trapezoid. As was the case with SAM tones, the waveforms repeated after the modulation period. However, with this modulator, the rate of amplitude change and the effective duration did not change with Fmod. At Fmod = 20 Hz, the TRAM waveform was very different from the SAM waveform. At Fmod = 100 Hz, and at higher Fmod in general, SAM and TRAM waveforms were more similar to one another (but never identical).

The carrier frequency was chosen to match the neuron’s characteristic frequency (CF), rounded to the nearest 200 Hz. Fmod usually varied from 10 to 200 Hz. SAM tones with Fmod > 200 Hz were rarely presented. TRAM tones with Fmod > 200 Hz could not be generated since, at those modulating frequencies, the duration of the trapezoid would have exceeded the length of the modulation period. For a given unit, the carrier frequency and the range of Fmod were the same for SAM and TRAM tones (unless SAM tones with Fmod > 200 Hz were presented). The overall duration of the tone was nominally 250 or

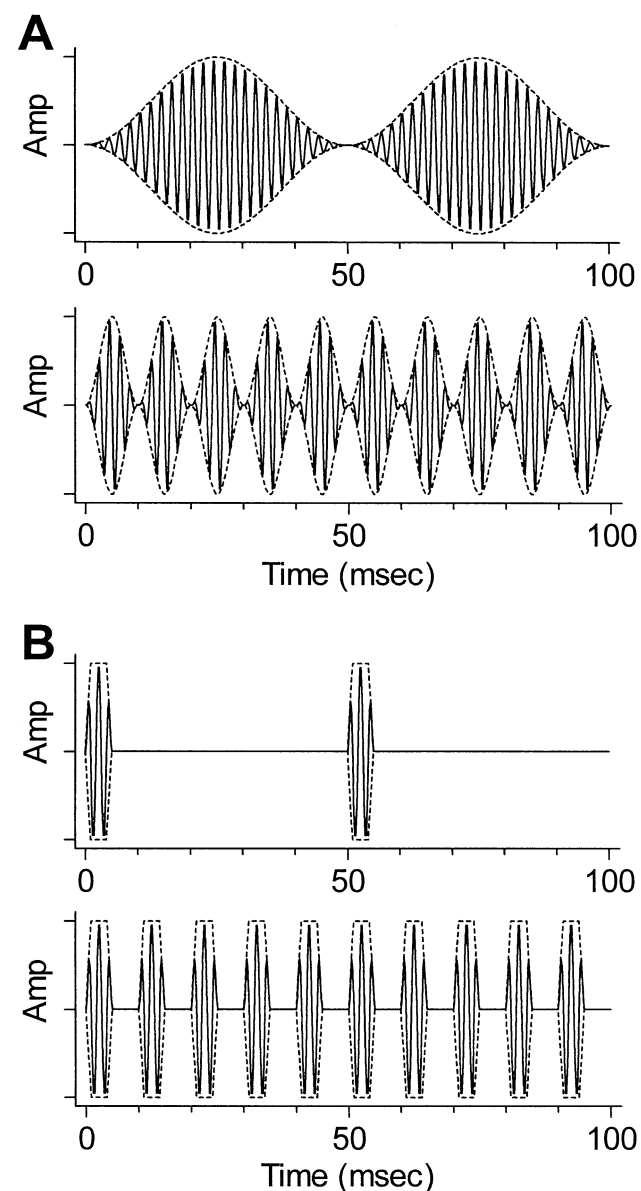


FIG. 1. Examples of modulated waveforms, showing differences in modulator types. In each example, the envelope is marked with a dashed line. The carrier frequency for this illustration was 500 Hz. **A.** SAM tones. *Top:* Fmod = 20 Hz, modulation period = 50 ms. *Bottom:* Fmod = 100 Hz, modulation period = 10 ms. The envelope during each individual modulation cycle rose more abruptly at the higher modulation frequency. **B.** TRAM tones. *Top:* Fmod = 20 Hz, modulation period = 50 ms. *Bottom:* Fmod = 100 Hz, modulation period = 10 ms. The trapezoidally modulated pulse within each 10-ms modulation period was identical and independent of modulation frequency.

300 ms. The actual duration was changed so that the tones always ended after a full modulation cycle. For example, a tone modulated at 25 Hz with a nominal duration of 250 ms would have an actual duration of 280 ms (7 modulation cycles at 40 ms/cycle). Each stimulus waveform was repeated 25–50 times with a silent interval of 300 ms between presentations (plus

a few milliseconds for computer processing that was done between presentations).

Each modulated tone was presented at a series of levels, from just above the threshold for a steady-state CF tone to 70 or 80 dB SPL in 10- or 20-dB steps. The levels of modulated tones are reported as the SPLs of steady-state tones with the same peak SPL. Thus, a SAM tone presented at 70 dB SPL had the same peak sound pressure as a TRAM tone presented at 70 dB SPL. The RMS sound pressure of SAM tones presented at a fixed peak SPL is independent of Fmod and will always be 4.3 dB lower than the RMS pressure of the steady-state tone with the same peak SPL. The RMS sound pressure of TRAM tones presented at a fixed peak SPL was also lower than the RMS pressure of a steady-state tone, and it increased with Fmod: approximately 13 dB lower than a steady-state tone when Fmod = 10 Hz, but only 3 dB lower when Fmod = 100 Hz.

Data analysis

Spike times were measured with 1- μ s resolution and displayed as dot rasters or peristimulus time (PST) histograms. Average discharge rate in spikes/second and vector strength (Goldberg and Brown 1969) were calculated. The statistical significance of vector strength was evaluated by calculating the Rayleigh coefficient (Rhode 1976). Only values that differed from chance at the 0.001 level were considered to be significant. MTFs were generated by plotting average discharge rate or vector strength as a function of Fmod; MTFs were generated separately for each modulator type.

RESULTS

Data were obtained from 78 single units from 30 chinchillas (the average number of units per animal was low because in most of those animals some units were studied with other stimulus configurations not reported here). The units' CFs ranged from 0.17 to 16.6 kHz, which includes nearly all of the chinchilla's audible range (Miller 1970). On the basis of their discharge patterns in response to steady-state CF tones, 33 units were classified as pausers, 18 as transient, and 12 as sustained. Fifteen additional units had long first-spike latencies or unusual discharge patterns. These units responded poorly to both types of modulated tones and are not included in the analyses below.

Effects of modulation type and modulation frequency on discharge patterns

The major finding of this study was that neurons classified by their discharge patterns to steady-state

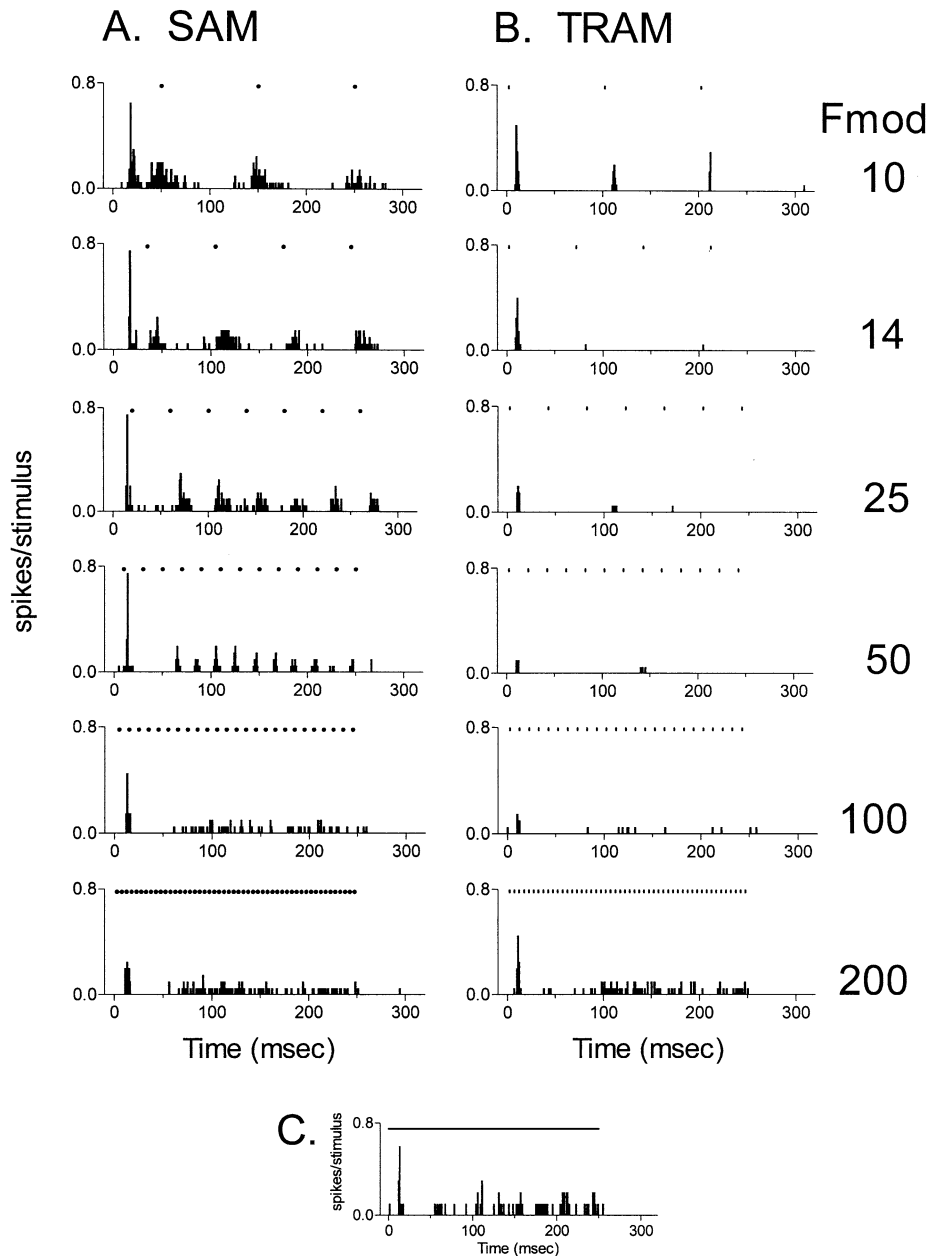


FIG. 2. Discharge patterns of a representative pauser unit to modulated tones. Modulation frequency increases from top to bottom, as shown by the labels at the far right. For all modulated tones, the carrier frequency was 2.0 kHz and the presentation level was 60 dB SPL. Unit 9908-3, CF = 20 kHz. **A.** Responses to SAM tones. Dots above the PST histograms indicate the locations of peaks in the SAM-tone envelope. **B.** Responses to TRAM tones. Vertical symbols above the PST histograms indicate the locations of the centers of the individual trapezoids in the TRAM-tone envelope. These points occurred with the same intervals as peaks in the corresponding SAM tones, although not at the same times relative to the onset of the tone. **C.** Discharge pattern in response to a steady-state 2-kHz tone at 60 dB SPL. The horizontal line above the PST histogram indicates the tone duration of 250 ms.

CF tones responded differentially to SAM and TRAM tones. The differences will be illustrated with discharge patterns from three representative neurons exhibiting different PST patterns, followed by summary analyses of the larger sample of neurons.

Examples of PST histograms obtained for SAM and TRAM tones at different Fmod are shown in Figures 2–4. Figure 2 illustrates the responses of a representative pauser neuron. Responses to SAM tones are shown in the left column. For modulation frequencies up to 50 Hz, the PST histograms exhibited a pattern of modulation that generally followed the modulation in the stimulus waveform. The pattern of peaks in the histograms followed waveform peaks (indicated by symbols), although for the lowest values

of Fmod, spikes could occur before the envelope peak. Exceptions occurred at stimulus onset, which tended to elicit a response peak in every case; the latency of this peak was unrelated to the waveform envelope. In addition, for Fmod = 50 Hz and higher, the PST histograms exhibited a pause or absence of spikes between 20 and 55 ms, similar to that seen in response to CF tones.

Responses to TRAM tones are shown in the right column in Figures 2–4. For the lowest Fmod, 10 Hz, the neuron's PST histogram exhibited a peak that corresponded to each modulation cycle. At higher values of Fmod, up to 100 Hz, responses other than a spike at onset were absent. This is in contrast to the response to SAM tones, for which waveforms modu-

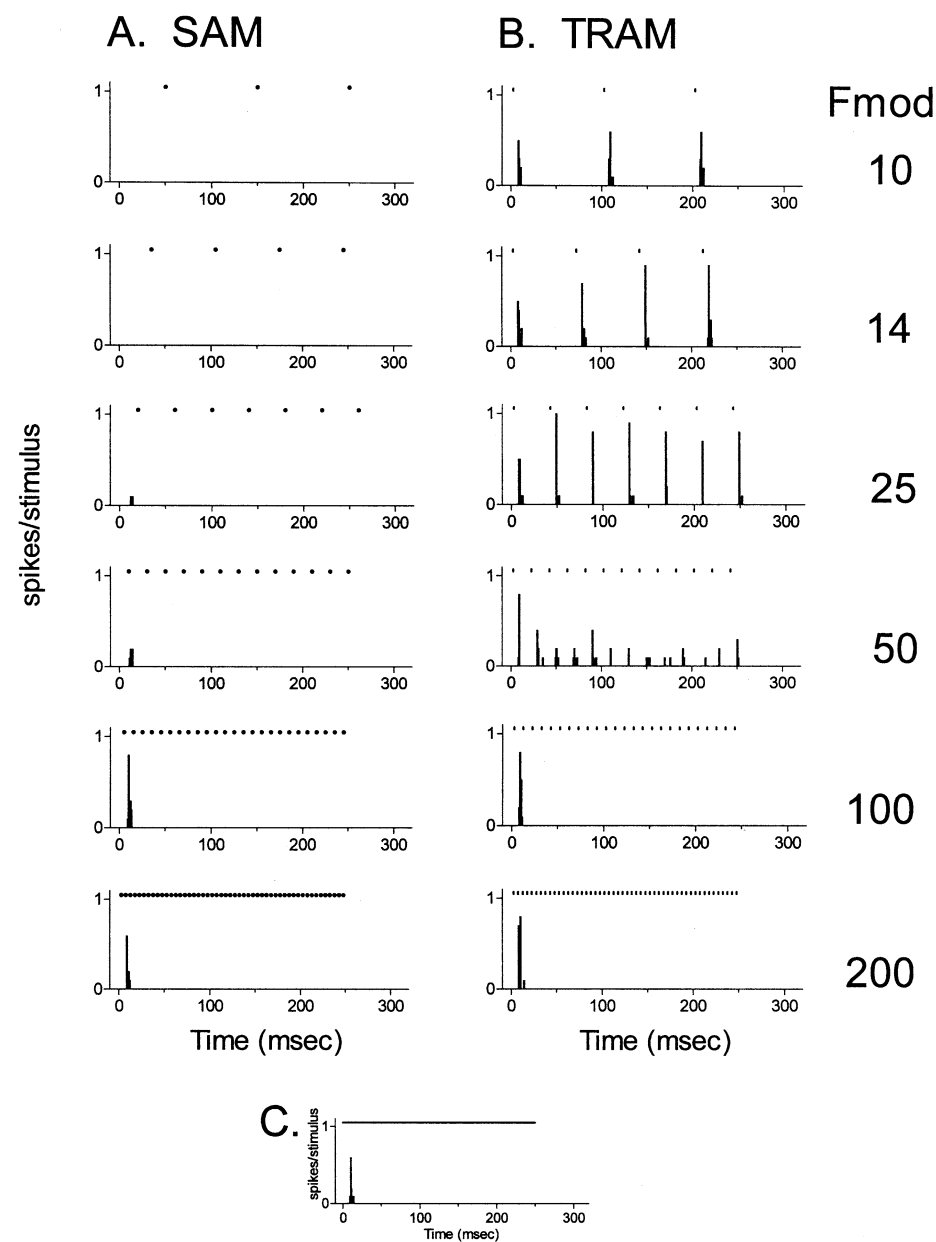


FIG. 3. Discharge patterns of a representative transient unit to modulated tones. See Figure 2 for description of the layout. The unit was recorded in the same animal as the unit in Figure 2. For all modulated tones, the carrier frequency was 4.8 kHz and the presentation level was 60 dB SPL. Unit 9908-2, CF = 4.8 kHz. The response to a 4.8-kHz steady-state tone at 60 dB SPL is shown in C.

lated at 14–50 Hz were effective in driving the unit with periodicities that matched the envelope of the stimulus.

Figure 3 presents comparable data for a transient neuron from the same animal; in this example, the trends in the responses were reversed. TRAM tones elicited strong responses that followed the envelope of the stimulus up to 50 Hz. At 100 and 200 Hz, the neuron responded at onset but did not follow the envelope; the temporal discharge patterns at these modulation frequencies were similar to the response to steady-state CF tones and thus provided no information about the waveform envelope. SAM tones, on the other hand, were almost completely ineffective, except when Fmod was 100–200 Hz. At those modu-

lation frequencies, the neuron responded at the onset of the stimulus, similar to the response to TRAM tones and steady-state tones, providing no information about the envelope of the stimulus.

Units with sustained responses to steady-state tones exhibited strong responses that could follow the envelopes of modulated waveforms up to the highest tested Fmod; they were also less likely to exhibit preferences for one type of modulation. Examples are shown in Figure 4. For SAM tones presented at low Fmod, spikes could occur throughout about one-half of the modulation cycle, and the overall discharge rate for SAM tones was greater than the rate in response to TRAM tones. For Fmod = 33 Hz and higher, the responses to the two types of modulation

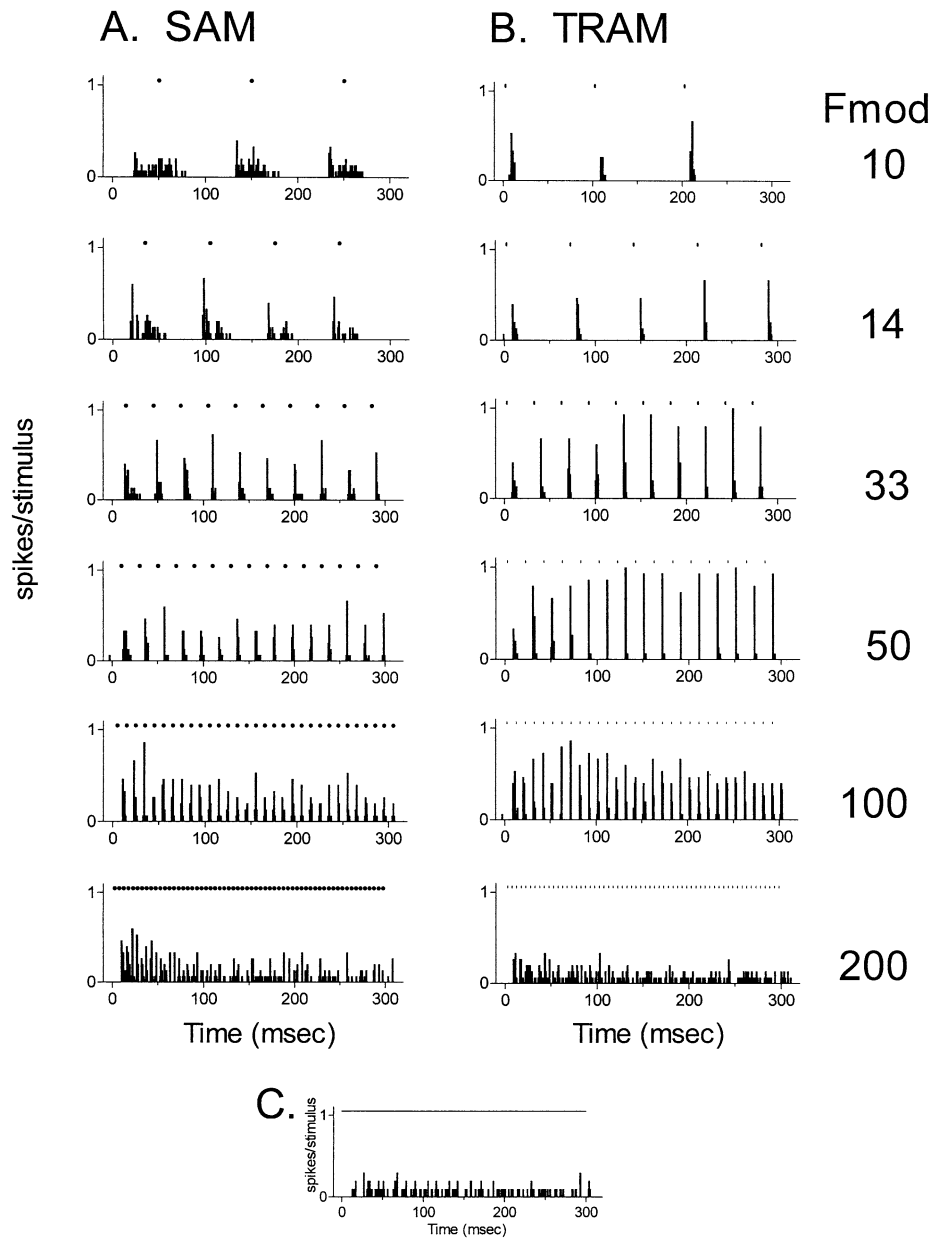


Fig. 4. Discharge patterns of a representative sustained unit to modulated tones. See Figure 2 for the description of the layout. For all modulated tones, the carrier frequency was 6.0 kHz and the presentation level was 50 dB SPL. Unit 0008-3, CR = 6 kHz. The response to a 6-kHz steady-state tone at 50 dB SPL is shown in C.

were similar in the appearance of the PST histograms and in discharge rate. This similarity was in contrast to the tendency for pauser and transient units to respond preferentially to one type of modulation.

MTFs for individual neurons

MTFs based on average discharge rate for the same three units used as examples in the previous section are shown in Figure 5. For the pauser neuron (Fig. 5A), rates in response to SAM tones were high at every Fmod. For TRAM tones, discharge rates were lower at every Fmod except 200 Hz.

Comparable MTFs for the transient unit are shown in Figure 5B. As noted previously, TRAM tones eli-

cited higher discharge rates than SAM tones for this unit. The MTF for TRAM tones was bandpass in shape, with the highest rates observed for tones modulated at 25 Hz. For SAM tones, rates were quite low for Fmod < 100 Hz. At 100 and 200 Hz, rates to TRAM and SAM tones were similar but low compared with the strongest responses elicited by TRAM tones.

The MTF for the sustained unit is shown in Figure 5C. For both SAM and TRAM tones, discharge rate increased with Fmod up to 100 Hz. For TRAM tones, the unit produced approximately 1 spike per modulation cycle for all Fmod between 10 and 100 Hz. Thus, the average discharge rate was nearly equal to Fmod, and the rate increase largely reflected the increase in the number of modulation cycles presented.

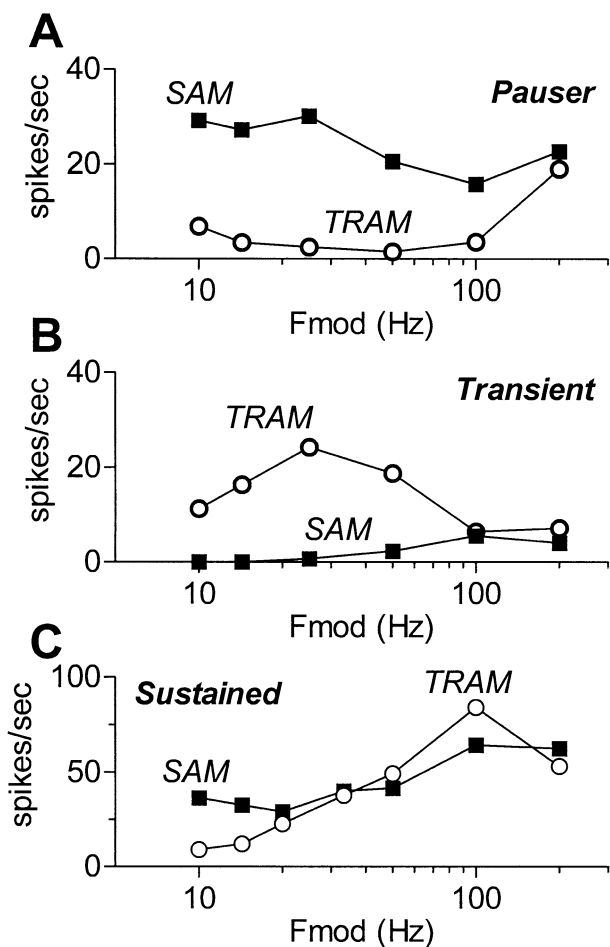


FIG. 5. MTFs based on average discharge rate for the units shown in figure 2-4.

Discharge rate decreased at 200 Hz, producing a bandpass MTF. Rates for SAM tones were also influenced by the number of modulation cycles, although in that case individual cycles at the lowest Fmod were more effective in eliciting spikes than were individual cycles at higher modulation frequencies. As noted above, the sustained unit did not exhibit strong selectivity for modulator type.

MTFs based on vector strength are shown in Figure 6 for the same units as in Figure 5. Selectivity for modulator type was also apparent in the temporal discharge patterns, especially in the range of Fmod over which significantly synchronized responses were observed. For the pauser unit, shown in Figure 6A, significant phase locking was observed for Fmod up to 50 Hz for SAM tones, but only to 14 Hz for TRAM tones. TRAM tones elicited higher values of vector strength; however, it is most likely a consequence of the lower spike counts elicited by TRAM tones. For the transient unit (Fig. 6B), statistically significant phase locking was observed for TRAM tones at all tested Fmod up to 100 Hz. For SAM tones, in con-

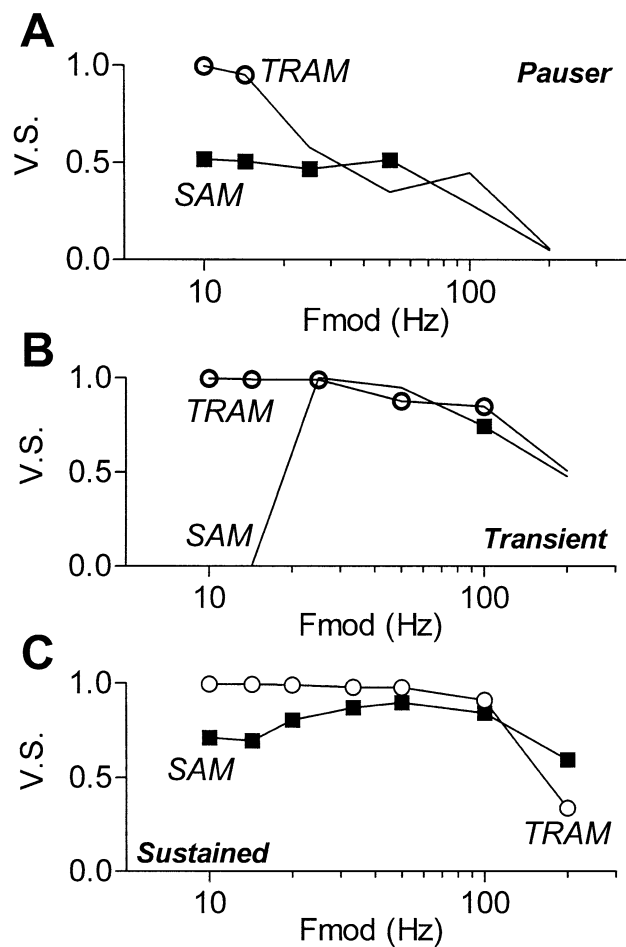


FIG. 6. MTFs based on vector strength for the units shown in Figures 2-4. All values are plotted, but only those that reached statistical significance as described in the Methods Section are marked with symbols.

trast, reliable phase locking was observed at Fmod = 100 Hz but not at any other modulation frequency. Figure 6C shows the vector-strength MTF for the sustained unit. Significant synchrony was observed at every tested Fmod, including 200 Hz, and differences related to modulator type were smaller.

Summary of discharge rates across units and conditions

To evaluate the generality of the trends shown in Figures 2-4, scatterplots comparing discharge rates elicited by SAM and TRAM tones are shown in Figures 7 and 8. Each point in Figure 7 compares the average number of spikes elicited by a SAM tone presented at a particular Fmod and SPL to the average number of spikes elicited by a TRAM tone presented at the same Fmod and SPL. Spikes were counted over the full duration of the tone. All the responses were obtained at 60 dB SPL, at three different modulation frequencies.

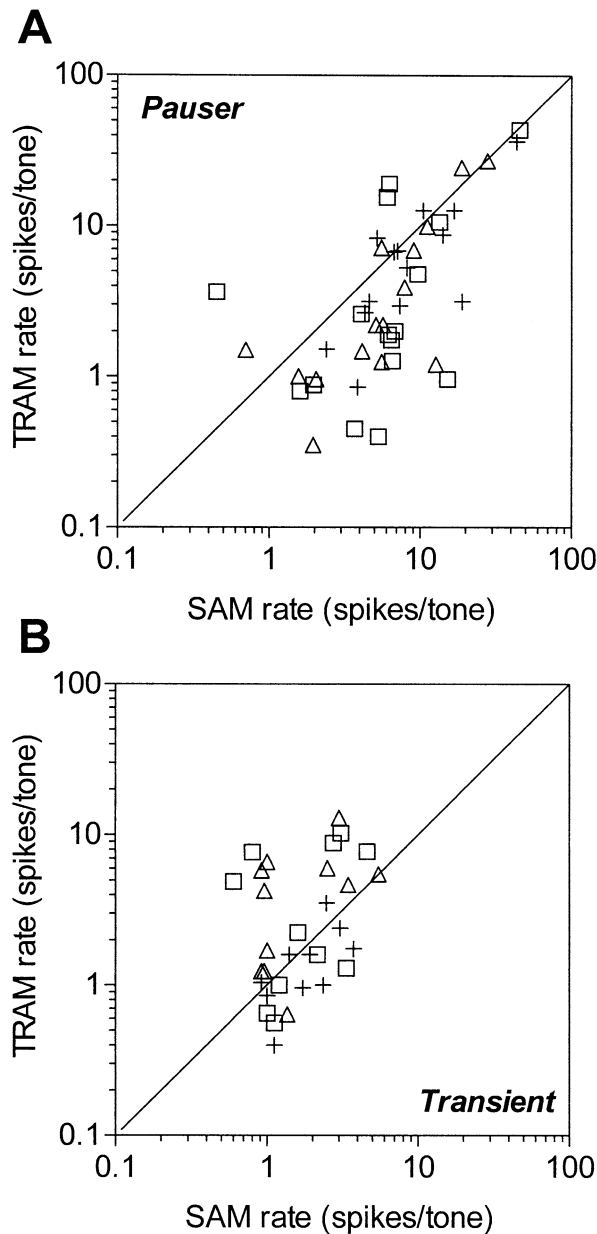


FIG. 7. Comparison of discharge rates elicited by SAM and TRAM tones. Discharge rate is given as the average number of spikes per tone. All tones were presented at 60 dB SPL, and rates obtained for three different Fmod are shown separately. Within each panel a discharge rate elicited by a SAM tone is compared with the same unit's rate in response to a TRAM tone presented at the same Fmod and peak SPL. **A.** Responses from 17 pauser units: Fmod = 20 Hz (Δ), 50 Hz (\square), or 100 Hz (+). **B.** Responses from transient units: Symbols are the same as in **A.**

Data from pauser units are shown in Figure 7A. The majority of points fell below the diagonal line, indicating higher discharge rates in response to SAM tones. The difference in rates was greatest for units with moderate discharge rates, up to about 20 spikes per SAM tone, and absent when discharge rate reached 40–50 spikes per tone. The pattern was similar for each Fmod included in the figure. In contrast, for transient

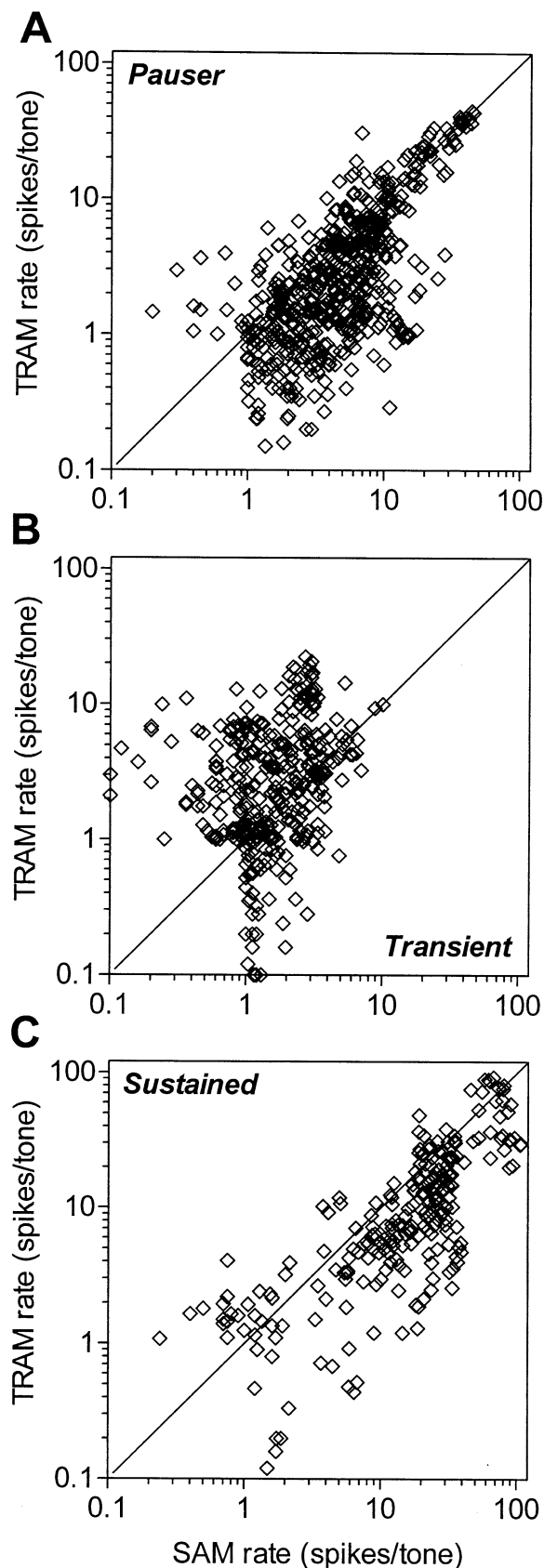
units (Fig. 7B), the majority of points fell above the diagonal line, indicating higher discharge rates in response to TRAM tones. The distribution of rates was similar for Fmod of 20 and 50 Hz. However, at 100 Hz, rates were low for both modulator types; transient units typically did not follow high modulation frequencies as well as pauser and sustained units did (see Fig. 5).

Figure 8 is similar to Figure 7, except that it includes every case for which equivalent stimuli (differing only in modulator type) were presented; it also presents responses of the sustained units. It excludes conditions for which neither stimulus elicited an average response of at least 1 spike per presentation, typically low-SPL or high-Fmod conditions for which the lack of responsiveness was presumably unrelated to characteristics of the modulator. Responses of pauser units are shown in Figure 8A. Although there are many more points, the pattern was consistent with that shown in Figure 7A; rates elicited by SAM tones were generally greater than rates elicited by TRAM tones. The difference in discharge rate was statistically significant (Wilcoxon signed rank sum test, 2-tailed, $p < 0.0001$). Responses of transient units are shown in Figure 8B. As in Figure 7B, rates to TRAM tones tended to be higher than rates to SAM tones. This rate difference was also statistically significant (Wilcoxon signed rank sum test, 2-tailed, $p < 0.0001$). Figure 8C compares rates elicited from sustained units. As was the case for pauser units, most points fell below the diagonal. However, the distribution of points did not resemble the distribution for pauser units, for which discharge rates in response to TRAM tones could be quite low. For sustained units, both SAM and TRAM tones were effective, but SAM tones elicited more spikes per tone because of the longer “effective duration” of these signals within individual modulation cycles. The rate difference was statistically significant (Wilcoxon signed rank sum test, 2-tailed, $p < 0.0001$). Overall, the data in Figures 7 and 8 indicate that the selectivity illustrated in Figures. 2–4 was not limited to isolated cases or to certain Fmod or presentation levels.

DISCUSSION

Units classified according to temporal discharge patterns to steady-state tones represent envelopes differently

Responses of IC neurons to SAM tones have been the focus of many studies in several different species (cat: Langner and Schreiner 1988; Schreiner and Langner 1988; rat: Rees and Møller 1983 1987; Palombi et al 2001; guinea pig: Rees and Palmer 1989; gerbil: Hell et al. 1995; Krishna and Semple 2000; rabbit: Batra et al. 1989; monkey: Müller-Preuss et al. 1994; bat:



Condon et al. 1996). Few studies have directly compared responses to different modulators or systematically related the responses to modulated tones to properties of neurons, such as the shape of the PST histogram in response to unmodulated tones. The study most similar to the present one was reported by Condon et al. (1996), who obtained responses to SAM tones from neurons in the IC of the little brown bat. They found that neurons with transient responses to steady-state tones responded poorly to SAM. In contrast, neurons that responded tonically to steady-state tones responded strongly to SAM tones. Finally, one transient unit that responded weakly to SAM tones exhibited strong responses to trains of short tone bursts comparable to TRAM tones (pulsatile amplitude modulation, in their terminology). TRAM tones apparently were not presented to every unit in their sample, and only one example a transient unit was shown. However, the results that were presented were consistent in every respect with those described here. Cooler and Feng (1992) reported similar findings from a study of units in the midbrain of the frog.

Several studies have mentioned, usually without extensive discussion, that a small proportion of transient units exhibited weak responses to SAM tones that resembled those reported here. The most detailed report is from Palombi et al. (2001) for IC neurons in young and old Fischer 344 rats. Approximately 40% of their units were judged unresponsive to SAM tones, and, of these, 81% exhibited only onset responses to steady-state tones. Overall, only about half of the onset units were responsive to SAM tones. Responses to waveforms comparable to TRAM tones were not reported.

In most studies, smaller numbers of unresponsive transient units were observed. Rees and Møller (1983) described one unit that was unresponsive to SAM tones; they found that the unit could be driven with tones modulated by a square wave. Rees and Palmer (1989) noted that some transient units in the guinea pig IC failed to respond to SAM tones. However, they did not provide a detailed quantitative description of these units' responses, or mention whether these units could be driven with other types of modulated tones. Krishna and Semple (2000) reported that for three transient units in their sample of gerbil IC units, SAM

←

FIG. 8. Comparison of discharge rates elicited by SAM and TRAM tones. Similar to Figure 7 except that many different Fmod and presentation levels are represented and one unit may contribute many points to the figure. **A.** Pauser units: 604 pairs of points from 28 units are plotted. **B.** Transient units: 388 pairs of points from 16 units are plotted. **C.** Sustained units: 252 pairs of points from 8 units are plotted

tones were ineffective at all tested Fmod. They also noted that transient units exhibited a slightly different pattern of bandpass tuning to SAM tones than did units of other PST classes; the MTFs of transient units lacked “regions of suppression,” in which increasing modulation depth produced rate decreases. Yang and Pollak (1997) observed that sustained units in the dorsal nucleus of the lateral lemniscus (DNLL) of the mustache bat followed SAM tones up to Fmod as high as 800 Hz; onset units, on the other hand, exhibited cutoff frequencies on the order of 300 Hz or less. The relatively restricted bandwidth of the onset units studied by Yang and Pollak (1997) was consistent with the pattern reported here. Overall, the data that are available suggest that it may be a general property of the IC that neurons categorized on the basis of the temporal pattern of responses to steady-state tones can exhibit selectivity for particular modulating waveforms. The degree of selectivity may be species-specific.

Although the results presented here demonstrate that SAM and TRAM tones elicit different discharge patterns from different types of IC units, they do not permit a strong conclusion about the particular stimulus properties that underlie the difference. That is because SAM and TRAM tones differ on more than one stimulus dimension. For example, rise time, which was always rapid for TRAM tones but could be quite slow for SAM tones, might account for the lack of effectiveness of SAM tones presented to transient units. However, duty cycle also varied, being constant across Fmod for SAM tones but variable for TRAM tones, and that variation might contribute to the observed selectivity. Additional experiments would be required to determine the extent to which these or other stimulus properties affect transient and pauser units differently.

Functional significance of this difference

To evaluate the potential significance of the observed selectivity for modulator type, it is helpful to consider the contributions made by the envelope to auditory perception. For this discussion, the term envelope will be used to refer to the pattern of amplitude fluctuation over time within one or more narrow bands within a sound's spectrum rather than the overall envelope. These envelopes have been shown to carry information about speech-sound identity (Shannon et al. 1995; Loizou et al. 1999). They also contribute to “auditory scene analysis” (Bregman 1990; Yost 1991), the ability to isolate a source of sound from a background of competing sounds. The presence of coherent amplitude modulation in different spectral regions is one of the acoustic cues that identifies components arising from the same source; it can be exploited by the auditory system for scene analysis. Onset synchrony is another temporal cue

that contributes to scene analysis. Little is known about the neural processing that allows these acoustic cues to be used, however.

Real-life sounds, including human speech, animal vocalizations, and environmental sounds, have modulated envelopes, and the modulators used in this study were intended to be representative of the range of envelope shapes and modulation frequencies that might be encountered. For example, in human speech, alternating consonants and vowels produce envelopes with fluctuations in the range of 2–50 Hz (Plomp 1983; Rosen 1992). These envelopes will have complex shapes, but they are likely to be more similar to SAM tones than to TRAM tones in that the rise and fall of amplitude is usually gradual (one exception to this generalization is discussed below). Recent psychophysical studies have established that the envelope in as few as four narrow frequency bands can convey enough information to make speech intelligible in quiet conditions (Shannon et al. 1995; Loizou et al. 1999).

Neurons that can represent the envelope up to a modulation frequency of about 50 Hz presumably would play an important role in the representation of speech and environmental sounds and in the decomposition of the auditory scene into a set of isolated sound sources. In the present results, pauser and sustained units were able to represent those frequencies, at least when the modulation was sinusoidal. Most transient units were not as effective in representing sinusoidal modulation over the same range, although they were capable of following nonsinusoidal modulation. Units that respond only at the onset of a steady-state tone are common throughout the central auditory system, including the IC (Nuding et al. 2000), and they presumably play an important role in auditory processing. The present results suggest that their role might be different than that played by sustained and pauser units. One possibility is that transient units and units with pauser or sustained responses extract different sets of temporal cues in parallel. For example, transient units are well-suited to identify onsets, especially abrupt onsets, while pauser and sustained units are better suited for representing ongoing slow modulation. As noted above, both cues contribute to sound source identification.

Representation of speech envelopes in the IC

Studies of the representation in the IC of consonant–vowel syllables differing in voice-onset time (VOT) previously reported by this laboratory (Chen et al. 1996; Chen and Sinex 1999; Sinex and Chen 2000) provide examples in which units with different temporal discharge patterns responded selectively to different components of VOT syllables. The onset of

voicing in these syllables typically produces an abrupt increase in amplitude in certain frequency regions; the rapid amplitude increase more closely resembled the pattern of amplitude change in TRAM tones. The onset of voicing was followed by a segment modulated at the fundamental frequency of voicing (114 Hz for the syllables used in those experiments), which resembles SAM tones more than TRAM tones. At the level of the IC, transient neurons often responded at the abrupt onset of voicing, but responded at low rates (or not at all) during the following vowel. Sustained and pauser units were more likely to respond throughout the vowel, often in synchrony to the 114-Hz fundamental frequency. The result was that slightly different populations of neurons responded at the onset of voicing and during the steady-state vowel. In each case, the neurons' responses tended to be consistent with the patterns observed in the present study.

Temporal resolution of IC neurons

Sinex and Chen (2000) questioned whether IC neurons can be described as having a general ability to represent temporal information, that is, an ability that is independent of the stimulus used to measure it. In auditory neurophysiology, it is often implicitly assumed that a general temporal ability is reflected in the pattern of responses elicited by SAM tones, and the SAM-tone MTF has often been taken as the appropriate measure of the temporal resolution (Langner and Schreiner 1988; Schreiner and Langer 1988; Delgutte et al. 1998). It is intuitively appealing to assume that temporal resolution is a general property of auditory neurons that contributes to, among other things, the representation of speech sounds. Although it is certainly true that there must be a limit to temporal resolution, the present results indicate that the pattern of responses to SAM tones does not completely specify what that limit is. Clearly, if only SAM tones had been studied, the conclusion would have been that transient units in the chinchilla IC possess poor temporal resolution, even though the same units represent the temporal information in TRAM tones very well. While there must be a relation between temporal response properties of auditory neurons and their ability to represent complex envelopes such as those in speech, the present results and those of Delgutte et al. (1998), Eggermont (1995), and Sinex and Chen (2000) indicate that the relation is not always the simplest or most obvious one.

ACKNOWLEDGMENTS

Preliminary results have been presented as abstracts (Sinex and Chen 1998; Sinex et al. 2000). Supported by Grant DC341 from NIDCD to D.G.S.

REFERENCES

- BATRA R, KUWADA S, STANFORD TR. Temporal coding of envelopes and their interaural delays in the inferior colliculus of the unanesthetized rabbit. *J. Neurophysiol.* 61:257–268, 1989.
- BREGMAN AS. *Auditory Scene Analysis*. Cambridge, MA, MIT Press, 1990.
- CARRELL TD, OPIE JM. The effect of amplitude comodulation on auditory object formation in sentence perception. *Percept. Psychophys.* 52:437–445, 1992.
- CHEN GD, NUDING SC, NARAYAN SS, SINEX DG. Responses of single neurons in the chinchilla inferior colliculus to consonant–vowel syllables differing in voice onset time. *Aud. Neurosci.* 3:179–198, 1996.
- CHEN GD, SINEX DG. Effects of interaural time differences on the responses of chinchilla inferior colliculus neurons to consonant–vowel syllables. *Hear. Res.* 138:29–44, 1999.
- CONDON CJ, WHITE KR, FENG AS. Neurons with different temporal firing patterns in the inferior colliculus of the little brown bat differentially process sinusoidal amplitude-modulated signals. *J. Comp. Physiol. A* 178:147–157, 1996.
- DELGUTTE B, HAMMOND BM, CARIANI PA. Neural coding of the temporal envelope of speech: relation to modulation transfer functions. In: Palmer AR, Rees A, Summerfield Q, Meddis R (eds) *Psychophysical and physiological advances in hearing*. London, Whurr Publishers Ltd., pp 595–603, 1998.
- EGGERMONT JJ. Representation of a voice onset time continuum in primary auditory cortex of the cat. *J. Acoust. Soc. Am.* 98:911–920, 1995.
- GOLDBERG JM, BROWN PB. Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: Some physiological mechanisms of sound localization. *J. Neurophysiol.* 32:613–636, 1969.
- GOOLER DM, FENG AS. Temporal coding in the frog auditory midbrain: The influence of duration and rise–fall time on the processing of complex amplitude-modulated stimuli. *J. Neurophysiol.* 67:1–22, 1992.
- HALL JW III, HAGGARD MP, FERNANDES MA. Detection in noise by spectro-temporal pattern analysis. *J. Acoust. Soc. Am.* 76:50–56, 1984.
- HEIL P, SCHULZE H, LANGNER G. Ontogenetic development of periodicity coding in the inferior colliculus of the Mongolian gerbil. *Aud. Neurosci.* 1:363–383, 1995.
- KRISHNA BS, SEMPLE MN. Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus. *J. Neurophysiol.* 84:255–273, 2000.
- LANGNER G, SCHREINER CE. Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *J. Neurophysiol.* 60:1799–1821, 1988.
- LOIZOU PC, DORMAN M, TU Z. On the number of channels needed to understand speech. *J. Acoust. Soc. Am.* 106:2097–2103, 1999.
- MERZENICH MM, REID MD. Representation of the cochlea within the inferior colliculus of the cat. *Brain Res.* 77:397–415, 1974.
- MILLER JD. Audibility curve of the chinchilla. *J. Acoust. Soc. Am.* 48:513–523, 1970.
- MÖLLER AR, REES A. Dynamic properties of the responses of single neurons in the inferior colliculus of the rat. *Hear. Res.* 24:203–215, 1986.
- MÜLLER-PREUSS P, FLACHSKAMM C, BIESER A. Neural encoding of amplitude modulation within the auditory midbrain of squirrel monkeys. *Hear. Res.* 80:197–208, 1994.
- NUDING SC, CHEN GD, SINEX DG. Monaural response properties of single neurons in the chinchilla inferior colliculus. *Hear. Res.* 131:89–106, 1999.

- PALOMBI PS, BACKOFF PM, CASPARY DM. Responses of young and aged rat inferior colliculus neurons to sinusoidally amplitude modulated stimuli. *Hear. Res.* 153:174–180, 2001.
- PLOMP R. The role of modulation in hearing. In: Klinke R, Hartmann R (eds) *Hearing—Physiological bases and psychophysics*. New York, Springer-Verlag, pp 270–275, 1983.
- REES A, MØLLER AR. Responses of neurons in the inferior colliculus of the rat to AM and FM tones. *Hear. Res.* 10:301–330, 1983.
- REES A, MØLLER AR. Stimulus properties influencing the responses of inferior colliculus neurons to amplitude-modulated sounds. *Hear. Res.* 27:129–143, 1987.
- REES A, PALMER AR. Neuronal responses to amplitude-modulated and pure-tone stimuli in the guinea pig inferior colliculus, and their modification by broadband noise. *J. Acoust. Soc. Am.* 85:1978–1994, 1989.
- RHODE WS. A test for the significance of the mean direction and the concentration parameter of a circular distribution. Technical Report, University of Wisconsin, Department of Neurophysiology, 1976.
- ROSEN S. Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 336:367–373, 1992.
- SCHREINER CE, LANGNER G. Periodicity coding in the inferior colliculus of the cat. II. Topographical organization. *J. Neurophysiol.* 60:1823–1839, 1988.
- SHANNON RV, ZENG F-G, KAMATH V, WYGONSKI J, EKELID M. Speech recognition with primarily temporal cues. *Science* 270:303–304, 1995.
- SINEX DG, CHEN GD. Responses of chinchilla inferior colliculus neurons to different types of amplitude-modulated tones. *Abstr. Assoc. Res. Otolaryngol.* 21:116, 1998.
- SINEX DG, CHEN GD. Neural responses to the onset of voicing are unrelated to other measures of temporal resolution. *J. Acoust. Soc. Am.* 107:486–495, 2000.
- SINEX DG, LI H, HENDERSON J. Modulation transfer functions of chinchilla inferior colliculus neurons. *Abstr. Soc. Neurosci.* 30:283, 2000.
- YANG L, POLLAK GD. Differential response properties to amplitude modulated signals in the dorsal nucleus of the lateral lemniscus of the mustache bat and the roles of GABAergic inhibition. *J. Neurophysiol.* 77:324–340, 1997.
- YOST WA. Auditory image perception and analysis: The basis for hearing. *Hear. Res.* 56:8–18, 1991.