Statistical Analyses of Temporal Information in Auditory Brainstem Responses to Tones in Noise: Correlation Index and Spike-distance Metric

YAN GAI[1,2] AND LAUREL H. CARNEY[1,2,3,4,5]

[1]*Department of Biomedical and Chemical Engineering, Syracuse University, Syracuse, NY 13244, USA*

[2]*Institute for Sensory Research, Syracuse University, Syracuse, NY 13244, USA*

[3]*Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244, USA*

[4]*Department of Biomedical Engineering, University of Rochester, P.O. Box 603, 601 Elmwood Avenue, Rochester, NY 14642, USA*

[5]*Department of Neurobiology and Anatomy, University of Rochester, P.O. Box 603, 601 Elmwood Avenue, Rochester, NY 14642, USA*

ABSTRACT

Gai and Carney (J Neurophysiol 96:2451–2464, 2006) previously explored the detection of tones in noise based on responses in the anteroventral cochlear nucleus; that study focused on temporal information in discharge reliability and analyses of neural responses related to the fine structure or envelope of the stimulus. Two additional temporal approaches, the correlation index (Joris et al., Hearing Res 216–217:19–30, 2006) and the spike-distance metric (Victor and Purpura, J Neurophysiol 76:1310–1326, 1996; Netw Comput Neural Syst 8:127–164, 1997), are tested in the present study. Trends in the correlation index as a function of stimulus level are similar to those of the synchronization coefficient (also called the vector strength) when the tone is presented alone. However, the present study found that trends in the correlation index did not agree with those of the synchronization coefficient for tones presented with relatively high-level background noise. Instead, trends in the correlation index generally agreed with those of the temporal reliability metric discussed in Gai and Carney (J Neurophysiol 96:2451–2464, 2006); that is, the correlation index decreased with increased tone level in the presence of relatively high-level background noise. The spike-distance metric, which was based on absolute spike times or on interspike intervals, was compared to the temporal measures described above, which were generally based on relative spike times. The results confirm that the spike-distance metric is not an optimal temporal metric. In addition, absolute spike times of primary-like responses generally contained much less temporal information than absolute spike times of chopper response types. The present study highlights the importance of relative spike-timing information as characterized by traditional and novel temporal measures.

Keywords: detection, reliability, synchronization, mutual information, timing

Correspondence to: Laurel H. Carney · Department of Biomedical Engineering · University of Rochester · P.O. Box 603, 601 Elmwood Avenue, Rochester, NY 14642, USA. Telephone: +1-585-2763948; fax: +1-585-7565334; email: laurel.carney@rochester.edu

INTRODUCTION

The detection of signals in noise is a fundamental topic of neurosensory research. The stimulus features that determine detectability and how they are represented and processed by the nervous system are still unknown. Previous studies have examined average discharge rate and synchronization to fine structure at

various stages of the auditory pathway in response to tones in noise (Rhode et al. 1978; Geisler and Sinex 1980; Gibson et al. 1985; Costalupes 1985; Young and Barta 1986; Miller et al. 1987; Rees and Palmer 1988; May and Sachs 1992). Gai and Carney (2006) describe several temporal analyses, based on information related to stimulus fine structure, envelope, and discharge reliability, of tone-in-noise responses in the anteroventral cochlear nucleus (AVCN) in gerbil. Comparisons made between these analyses and with psychophysical detection performance suggest that some representations of temporal information can be more robust for detection than average discharge rate. That study also explored the feasibility of neural mechanisms for processing each type of information.

The present study extends Gai and Carney (2006) by applying two additional temporal approaches, the correlation index and the spike-distance metric, to the same set of tone-in-noise responses recorded in that study. The correlation index (Louage et al. 2005; Joris et al. 2006) measures the occurrence of simultaneous discharges across stimulus repetitions. For pure-tone responses, the correlation index and vector strength have similar trends as a function of tone level (Joris et al. 2006). For responses to sinusoidally amplitude-modulated (SAM) tones, the correlation index measures synchronization to modulation frequency, and for noise responses, synchronization to the noise envelope (Louage et al. 2005; Joris et al. 2006). The correlation index has not previously been applied to tone-in-noise responses to address the question of how the index is influenced by the combination of neural synchronization to tone frequency and to envelope, which can be counteracting cues in the presence of noise (Gai and Carney 2006).

The temporal metrics described above generally rely on relative spike times, that is, the temporal features that determine those measures are "floating" rather than anchored in time. Victor and Purpura (1996, 1997) describe a spike-distance metric that can be applied to absolute spike times or interspike intervals and quantifies the dissimilarity across spike trains over varying time scales. The mutual information quantified by the spike-distance metric indicates the efficiency of spike timing in coding the stimuli at different temporal resolutions. Although Victor and Purpura (1997) state that this approach may not incorporate all types of temporal information as an optimal temporal measure, this approach has been used in recent auditory studies as a general evaluation of the contribution of temporal information compared to that of average rate (Chase and Young 2006; Huetz and Edeline 2004/2005; Huetz et al. 2006). In the present study, the spike-distance metric was compared to other temporal measures for tone-in-noise responses. We hypothesize that the discrepancy between the spike-distance metric based on spike times and other temporal measures is due to differences in the information carried by absolute and relative spike times.

METHODS

Animal preparation, stimulus generation, and recording procedures

This study involved further analysis of responses collected for Gai and Carney (2006) in which the experimental procedures are detailed. Briefly, gerbils were anesthetized with ketamine and xylazine. Single-neuron extracellular recordings in the AVCN were obtained with a single-barrel glass electrode through an opening in the temporal bone. The sound stimulus was a white noise (30 or 40 dB SPL spectrum level from 0.1 to 10 kHz) or the noise plus a simultaneously gated CF tone that varied in level (e.g., 30–80 dB SPL with 10 dB steps). Each stimulus was presented 30 to 50 times, had a duration of 250 ms, and was repeated every 475 ms. The noise waveform was randomly generated and was kept frozen for each dataset (i.e., the same stored digital waveform was used for all stimulus repetitions). For each stimulus, both amplitude and phase were digitally compensated using an acoustic transfer function obtained at the beginning of each experiment with a probe-tube microphone (Etymotic ER7).

Correlation index

The correlation index (CI) is the peak value of the shuffled autocorrelogram (SAC) (Louage et al. 2005). An autocorrelogram based on all-order interspike intervals has been used in various studies of auditory responses, such as modeling studies of the neural representation of pitch (Licklider 1951; Meddis and Hewitt 1991). The SAC is constructed by including the intervals across repetitions but excluding the intervals within each spike train. The CI is a measure of the degree of similarity of discharges in response to a repeated stimulus. The calculation was the same as that described in Joris et al. (2006). For each pair of spike trains, the total number of coincidences of spike times across repetitions was counted. For M repetitions, there were $M(M-1)/2$ different pairs of spike trains. To match the calculation of SAC in Joris et al. (2006) in which the order of the two spike trains matters (e.g., spike train #1 vs. spike train #2 yields different results from spike train #2 vs. spike train #1), $M(M-1)$ pairs of spike trains were counted, so that

half of the CI values were duplicates of the other half. The total number of coincidences across all paired combinations is denoted as N_c. CI was computed as:

$$\text{CI} = \frac{N_c}{M(M-1)r^2\omega D} \tag{1}$$

where r is the average discharge rate (spikes/s), ω is the temporal analysis window (i.e., coincidence window), and D is the stimulus duration (250 ms). The temporal analysis window determines the window during which two spikes are considered coincident. Window sizes of 0.20 and 0.57 ms were used in this study. These are the two smallest temporal resolutions used in Gai and Carney (2006); larger temporal windows were not used in this study because they could result in multiple spikes falling within one window, which is not appropriate for these analyses.

To determine a detection threshold based on the CI, the lowest tone level that yielded a significantly different CI compared to the CI of the noise alone was identified. A bootstrapping algorithm (Moore et al. 2002) was used to test the significance of the difference in CI between responses to noise-alone (M repetitions) and noise-plus-tone (M repetitions) stimuli. The difference was denoted as ΔCI_0. Spike trains in response to both noise-alone and noise-plus-tone stimuli were mixed together and regrouped randomly into two new groups (M spike trains each), and the difference in CI between the two new groups, denoted as ΔCI, was computed. A distribution of ΔCI using 500 random groups was computed. The statistical significance was obtained by comparing ΔCI_0 and ΔCI. Because more than 80% of neurons showed decreased CI with increased tone level for both noise levels (see the "Results" section), only significant decreases in CI derived with the above method were considered. The estimated detection threshold was the lowest tone level that had a significantly lower CI than that for noise-alone for any temporal analysis window, as long as the decrease in CI remained significant for all higher tone levels for that particular window.

Spike-distance metric and mutual information

The present study applied the spike-distance metric based on absolute spike times and on sequences of interspike intervals. The dissimilarity of two spike trains, as measured by the spike-distance metric based on spike times, can be interpreted as how much effort is required or how costly it is to transform one spike train into the other with a certain degree of temporal freedom (Victor and Purpura 1997). More specifically, one spike train can always be created from another spike train by adding (cost=1) and deleting (cost=1) spikes or by shifting the timing of spikes (cost=$q|\Delta t|$; $|\Delta t|$ is the time

difference between the two spikes; the unit of q is s^{-1}). The action is determined by minimizing the total cost of transferring one spike train into the other. When $|\Delta t|<2/q$, the cost of shifting the spike is lower than adding and deleting spikes, and vice versa. Therefore, q determines the temporal resolution of the analysis (temporal resolution is proportional to $1/q$). When $q=0$ s^{-1}, the spike distance is equal to the difference between spike counts of the two spike trains. The minimal spike distance based on spike times was computed using an efficient algorithm presented by Victor and Purpura (1996). The minimal spike distance based on interspike intervals was computed in the same way, but the sequences of interspike intervals replaced the sequences of spike times in the algorithm.

When there are several sets of spike trains in response to several stimuli, a confusion matrix can be created to quantify the discharge consistency within each stimulus and across different stimuli (Victor and Purpura 1996). The confusion matrix consists of the joint probability $p(s, r)$ for stimulus s and response r. The mutual information (MI) between stimulus set S and response set R is computed as (Cover and Thomas 1991):

$$\text{MI} = \sum_{s \in S} \sum_{r \in R} p(s, r) \log_2\left(\frac{p(s, r)}{p(s)p(r)}\right) \tag{2}$$

where $p(s)$ and $p(r)$ are the marginal probabilities. The mutual information derived from the confusion matrix indicates the efficiency of the responses in coding the stimuli. If absolute spike times are critical for coding a certain set of stimuli, then the spike-distance metric based on spike times should be larger for responses to different stimuli than for responses to different repetitions of the same stimulus. A detailed illustration of this procedure for evaluating the mutual information measured by the spike-distance metric can be found in Chase and Young (2006).

In the present study, a confusion matrix was created from spike distances in the responses to noise-alone and noise-plus-tone stimuli with different tone levels. High mutual information suggested that the tone-in-noise responses were highly "separated" in a metric space, regardless of the monotonicity of the underlying changes. For example, when $q=0$ s^{-1}, the spike-distance metric was equal to the change in average discharge rate between the responses to tone-plus-noise and noise-alone stimuli. High values of mutual information can be achieved, even if the rate change is nonmonotonic with tone level, as long as the rates differ in response to different stimuli. In general, the average rate and several of the temporal metrics presented in Gai and Carney (2006) are monotonic with tone level; however, some nonmonotonicity can

be observed for two adjacent tone levels because of variability in the responses. Therefore, this study only included responses to three well-spaced tone levels when computing the spike-distance metric: noise-alone, noise plus the highest-level tone (70 or 80 dB SPL), and noise plus the third-highest-level tone (50 or 60 dB SPL). These stimuli were chosen to provide potentially large distances between the responses.

Values of q were logarithmically spaced between 10 and 17,783 s^{-1} with four values per decade, which yielded a smooth change in mutual information values based on spike distances for AVCN neural responses; $q=0$ s^{-1} was also included for comparison to average-rate coding. The computations of the confusion matrix and mutual information were the same as described in Chase and Young (2006). Note that a detection threshold was not obtained from the mutual information; rather, the goal was to compare rate coding and spike-timing coding for each neuron. (MI is measured in bits; the upper limit of MI is $\log_2 3 = 1.59$ bits for the three stimuli described above presented with equal probability.)

The mutual information based on the synchronization coefficient was also computed. For synchronization, the following expression is equivalent to Eq. 2 (above) (Cover and Thomas 1991):

$$\text{MI} = \sum_{s \in S} \sum_{r \in R} p(r|s) p(s) \log_2 \left(\frac{p(r|s)}{p(r)} \right) \qquad (3)$$

where r represents the synchronization coefficient of responses for a single trial. The conditional probability, $p(r|s)$, was estimated based on responses to all repetitions of a given stimulus, s.

All the MI values were debiased to correct for the sample size bias (Chase and Young 2005). The amount of bias of the peak SDM was small (the average values across all the neurons were 0.029, 0.025, and 0.037 bits for MI based on spike times, intervals, and synchronization coefficients, respectively).

RESULTS

Tone-in-noise responses of the same 102 AVCN neurons presented in Gai and Carney (2006) were reanalyzed in this study with the correlation index and spike-distance metrics. Neurons were classified into three response types based on Blackburn and Sachs (1989). Primary-like response types have irregular spike intervals and are associated with AVCN bushy cells (Cant and Morest 1979). Chopper response types have regular spike intervals and are associated with AVCN stellate cells (Smith and Rhode 1989). Neurons with responses that do not belong to the first two classes were classified as unusual response types. Results for four representative

neurons (from Figs. 2, 3, 4, and 5 in Gai and Carney 2006; two primary-likes and two choppers) are presented first, followed by population results. The post stimulus time histograms (PSTHs) of these neurons in response to noise-alone stimuli and the noise plus a high-level tone stimuli are shown in Figures 1A, 2A, 3A, and 4A (different temporal-discharge patterns are observed for noise-alone and noise-plus-tone responses). Figures 1, 2, 3, and 4 (B, average discharge rate, C, synchronization to tone frequency, and D, temporal reliability) are replotted from Gai and Carney (2006; their Figs. 2, 3, 4, and 5) for comparison to the newer temporal analyses presented in this study. Filled symbols indicate significant increases and open symbols indicate significant decreases of a particular metric when a tone was added to the noise.

The average discharge rate showed a significant increase at high tone levels for the low-frequency chopper (Fig. 1B) but not for the other three neurons (Figs. 2B, 3B, and 4B). The synchronization coefficient to tone frequency showed a significant increase at high tone levels for three neurons (Figs. 1C, 2C, and 4C). The increase of the synchronization with increasing tone level was due to phase-locking to the pure tone: when the signal-to-noise ratio of the stimulus increased, the synchronization increased (Rhode et al. 1978; Costalupes 1985; Miller et al. 1987). On the other hand, AVCN neurons also show synchronization to stimulus envelope. When a tone is added to a noise, the stimulus envelope flattens, and the flattening of the envelope can be used as a cue for the detection of tones in noise (Richards 1992). Correspondingly, fluctuations in the instantaneous discharge rate of AVCN neurons, caused by locking to the envelope, decreased with increasing tone level (Gai and Carney 2006). In short, the strengths of synchronization to tone frequency and to the envelope change in different directions with increasing tone level. The decrease in temporal reliability with increasing tone level (Figs. 1D, 2D, and 4D) indicated that synchronization to the envelope dominated the temporal reliability.

Correlation index

Figures 1E, 2E, 3E, and 4E show the CI at different tone levels with a 30-dB SPL spectrum level noise and two temporal analysis windows, 0.20 and 0.57 ms. A common trend was the decrease of CI with increasing tone level. Note that the synchronization to the tone increased significantly with tone level for three of four neurons (Figs. 1C, 2C, and 4C). Therefore, the CI did not represent synchronization to tone frequency when background noise was present, as it did for tones in quiet. In fact, the change of CI with tone level was similar to that of the temporal reliability metric
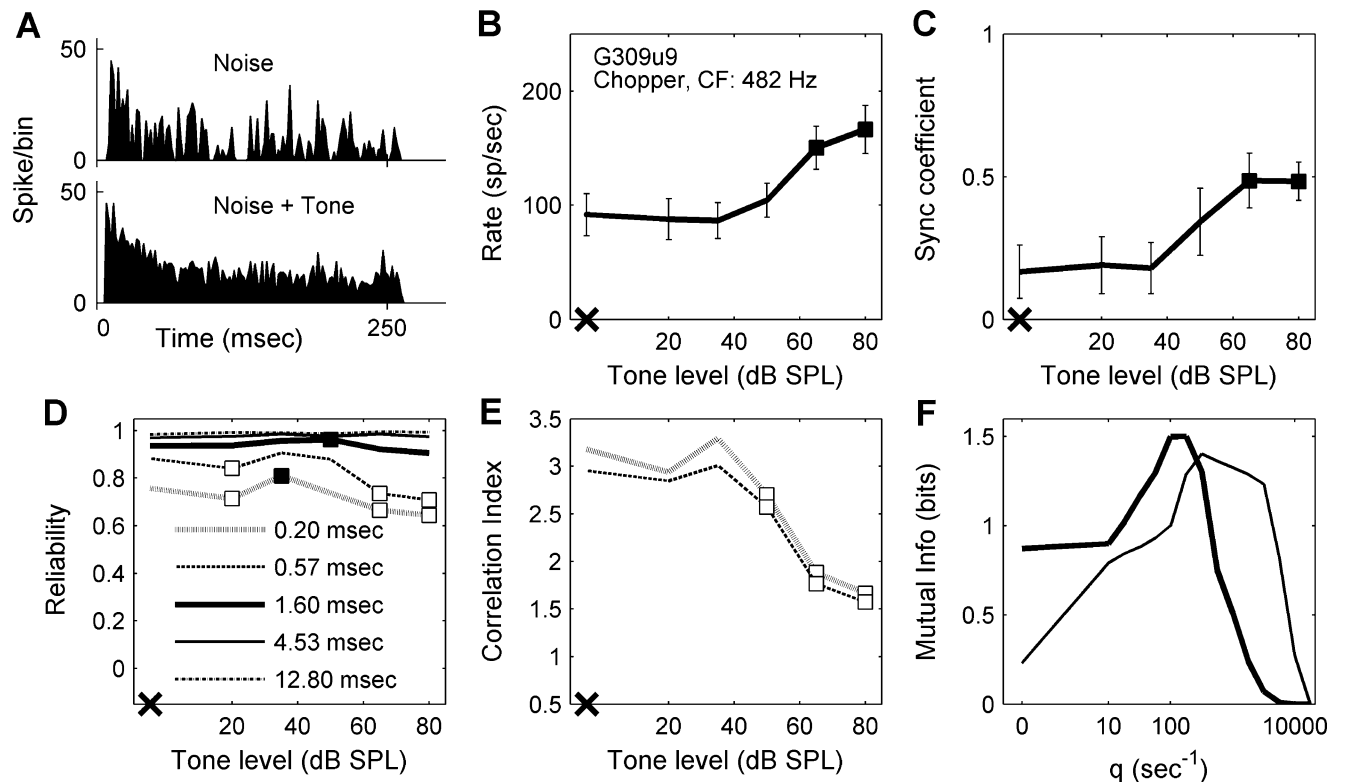
FIG. 1. Responses of a chopper to tones in noise. **A**, **B**, **C**, and **D** are replotted from Gai and Carney ([2006](#), their Fig. 2). **A** PSTHs for a 30-dB SPL spectrum level noise (*top*) and the same noise plus a 65-dB SPL tone. Bin width=2 ms. **B** Masked rate-level function. **C** Synchronization coefficient to the tone. **D** Temporal reliability measured with five temporal analysis windows. **E** CI measured with the two smallest temporal analysis windows (see legend in **D**). **F** Mutual information based on spike-time distances of tone-in-noise responses with varying *q*. *Thick line* mutual information based on all spike times. *Thin line* mutual information based on a fixed number of spikes (see text). All *filled squares* indicate significant increases, and all *open squares* indicate significant decreases ($p<0.05$ or $d' \geq 1$). The *x* on the abscissa indicates responses to noise alone. G309u9: CF= 482 Hz, SR=0 sp/s. Pure-tone rate threshold in quiet=22 dB SPL. Fifty trials per stimulus condition. *Error bars* are standard deviations.

(the correlation coefficient between the PSTHs constructed from even-numbered and odd-numbered stimulus repetitions) of the four neurons (Figs. 1D, 2D, 3D, and 4D). The detection thresholds based on CI and temporal reliability were the same for the two neurons in Figures 3 and 4, but different for the neurons in Figures 1 and 2. For example, the lowest tone level that resulted in a significantly decreased ($p<0.05$, bootstrapping; see the "Methods" section) CI for the neuron in Figure 2E was 60 dB SPL, whereas the lowest tone level that resulted in significantly decreased temporal reliability for the same neuron was 40 dB SPL (Fig. 2C).

Figure 5D shows the detection thresholds based on CI for noise spectrum levels of 30 dB SPL (left) and 40 dB SPL (middle) for all recorded neurons. Thresholds based on the average discharge rate, synchronization to tone frequency, and temporal reliability are replotted in Figure 5A–C for comparison. The thresholds based on CI (Fig. 5D, left and middle) were comparable to the detection thresholds based on temporal reliability (Fig. 5C, left and middle). A small number of neurons showed detection thresholds equal to or lower

than psychophysical detection thresholds in cat (solid lines; Costalupes 1983, corrected for noise duration based on Wier et al. 1977). In the right panel, most neuron thresholds increased with noise level and were close to the psychophysical threshold increments (solid line) but more variable than the threshold increments based on temporal reliability (Fig. 5C, right).

In Joris et al. (2006), 0.05 ms was chosen as a standard temporal analysis window based on the effect of the coincidence window size on the CI. Applying the CI with a 0.05-ms temporal window to AVCN responses (not shown in this study) generally yielded worse detection performance than the results for the two temporal windows presented above (0.20 and 0.57 ms). As stated in the "Methods" section, windows larger than these were not included as no more than one spike was allowed to fall in a window.

Spike-distance metric and mutual information

The spike-distance metric based on spike times will be the focus of this section. The spike-distance metric based on interspike intervals will be only briefly
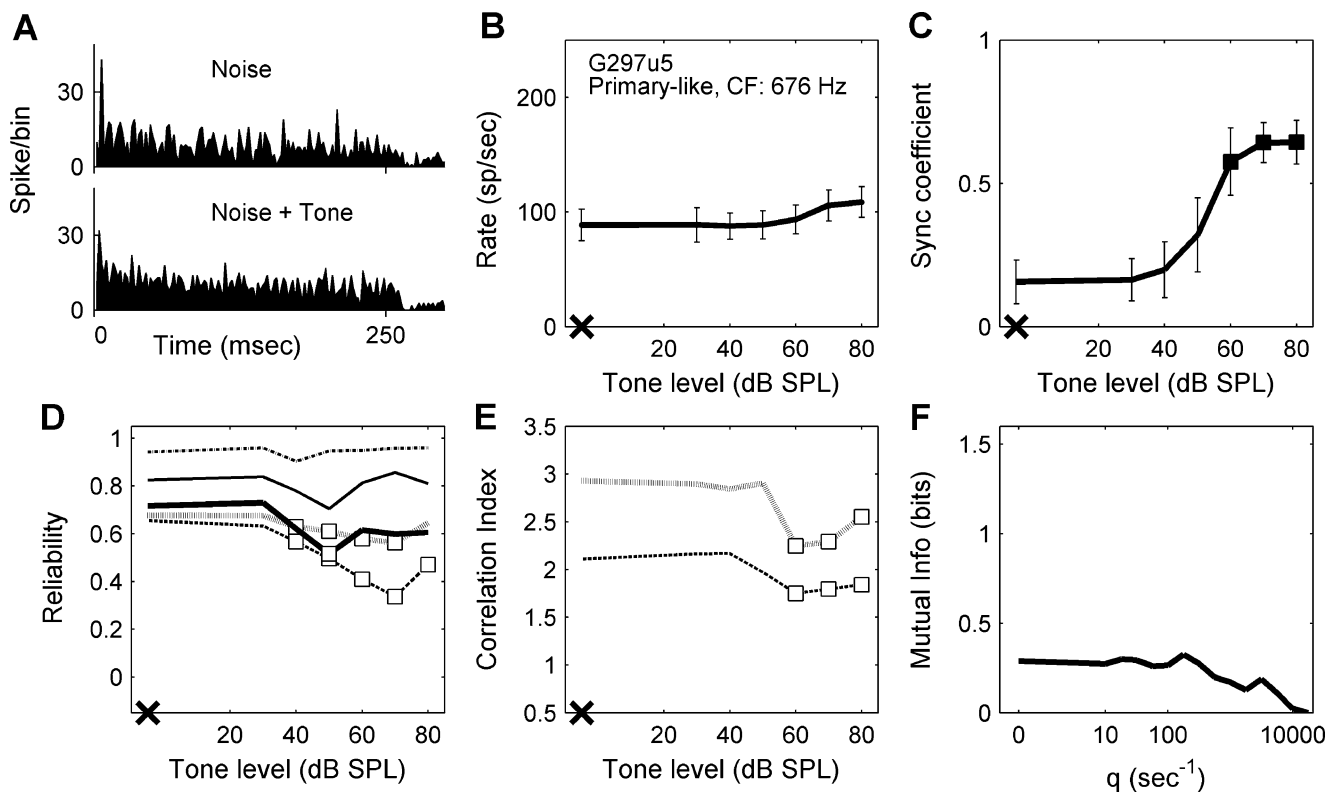
FIG. 2. Responses of a primary-like. Same format as Figure 1. **A**, **B**, **C**, and **D** are replotted from Gai and Carney (2006, their Fig. 3). The legend for **D** and **E** is the same as that in Figure 1. Tone level in **A** (*bottom*) is 70 dB SPL. Fifty tone-in-noise trials. G297u5: CF=676 Hz, SR=49 sp/s, pure-tone rate threshold in quiet=20 dB SPL.
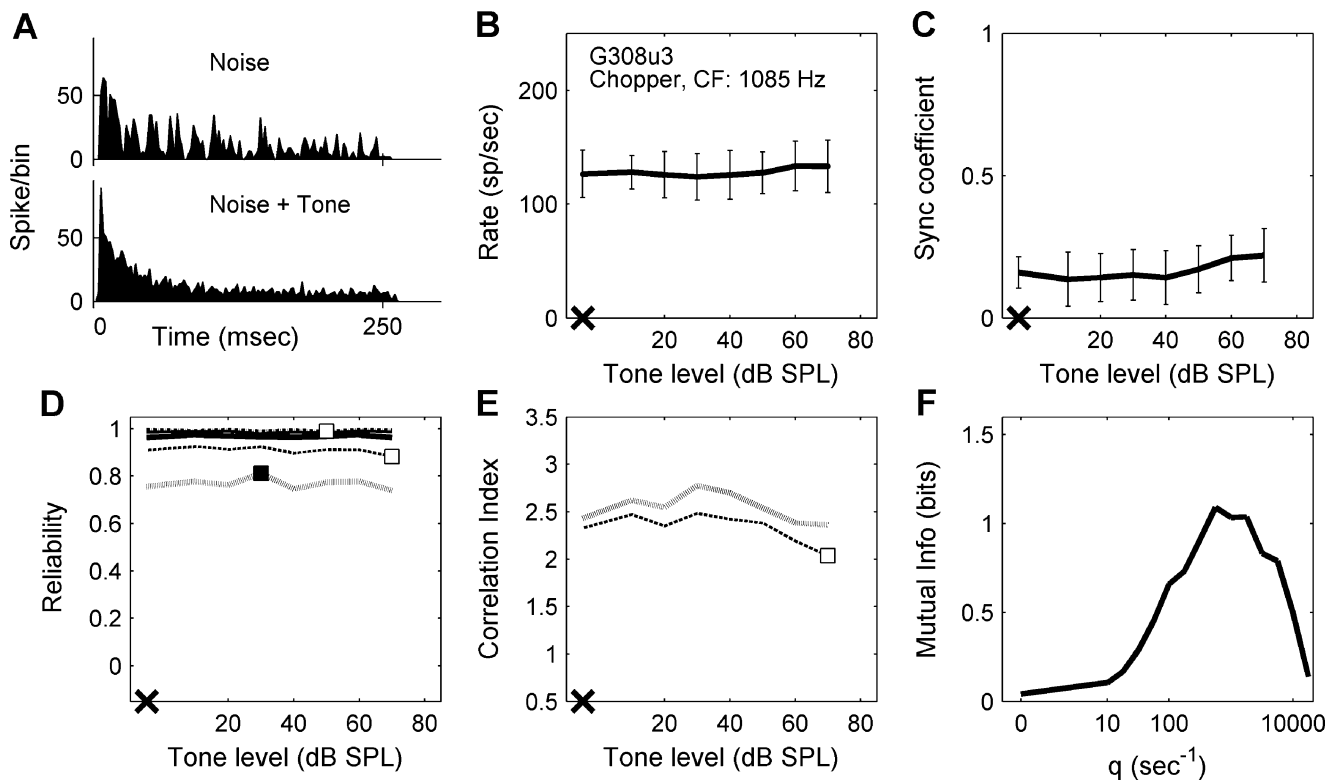

FIG. 3. Responses of a chopper. Same format as Figure 1. **A**, **B**, **C**, and **D** are replotted from Gai and Carney (2006, their Fig. 4). The legend for **D** and **E** is the same as that in Figure 1. Tone level in **A** (*bottom*) is 70 dB SPL. Fifty tone-in-noise trials. G308u3: CF=1085 Hz, SR=0 sp/s, pure-tone rate threshold in quiet=21 dB SPL.
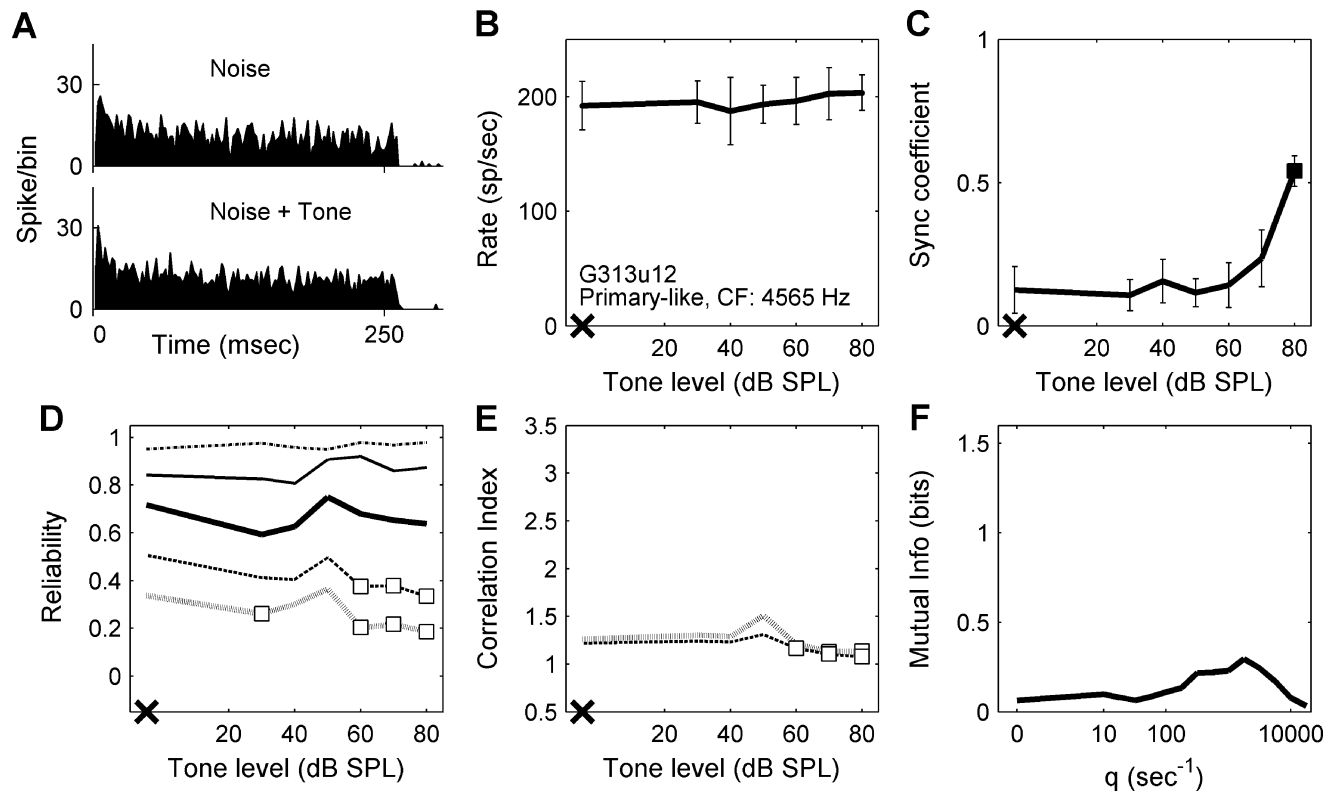
FIG. 4. Responses of a primary-like. Same format as Figure 1. **A**, **B**, **C**, and **D** are replotted from Gai and Carney (2006, their Fig. 5). The legend for **D** and **E** is the same as that in Figure 1. Tone level in **A** (*bottom*) is 70 dB SPL. Thirty tone-in-noise trials. G313u12: CF=4565 Hz, SR=38 sp/s, pure-tone rate threshold in quiet=12 dB SPL.

described at the end of the "Results" section because it generally yielded worse performance than the spike-distance metric based on spike times. Note that the mutual information was evaluated with respect to three fixed tone levels as described in the "Methods" section.

Best temporal resolution. Figures 1F, 2F, 3F, and 4F show the mutual information (MI) based on spike-time distances for the four representative neurons (thick lines). The x-axis is the cost, q, which corresponds to a temporal resolution of $1/q$ (Victor and Purpura 1997); $q=0$ s^{-1} represents average-rate coding. For the two neurons in Figures 3F and 4F, the MI at $q=0$ s^{-1} was close to 0, consistent with the fact that these two neurons had flat masked rate-level functions (Figs. 3B and 4B). The y-axis ranges between zero and the upper limit of mutual information, $\log_2 3=1.59$ bits. For the two neurons shown, the MI peaked at relatively high values of q, indicating that temporal coding for detection based on absolute spike times was most efficient when spike-timing precision was relatively high (peak MI values occurred at $1/q=1.78$ and 0.56 ms for neurons c and d, respectively). Note that although none of the temporal measures shown in Figure 3C–E had detection threshold lower than 70 dB SPL, the detection threshold based on an envelope-related temporal metric (Gai and Carney 2006, their Fig. 4I; not shown in

this study) was 50 dB SPL, which was comparable to the psychophysical detection threshold (approximately 50 dB SPL). The chopper neuron in Figure 1F had a relatively high value of MI at $q=0$ s^{-1}, as its average discharge rate showed a significant change with tone level (Fig. 1B). However, the best coding for detection was achieved at a resolution of 10 ms. This is a relatively large time scale that did not support any of the temporal measures presented in Gai and Carney (2006). The implication of the large time scale will be discussed below with the population results. The primary-like neuron in Figure 2F was even more interesting because including spike-timing information did not improve the detection efficiency over that of average-rate coding, yet all the previously tested temporal metrics outperformed the rate coding for this neuron (Fig. 2B–E). This example suggested that the spike-distance metric was less suitable for the detection task than the other temporal metrics, at least for some neurons.

To describe population results, MI as a function of q was categorized into symmetrical bandpass-shaped (Figs. 3F and 4F), lowpass-shaped (Fig. 2F), and asymmetrical bandpass-shaped (Fig. 1F). The symmetrical bandpass shape described MI functions with low average-rate information. The lowpass shape described MI functions with the maximum information con-
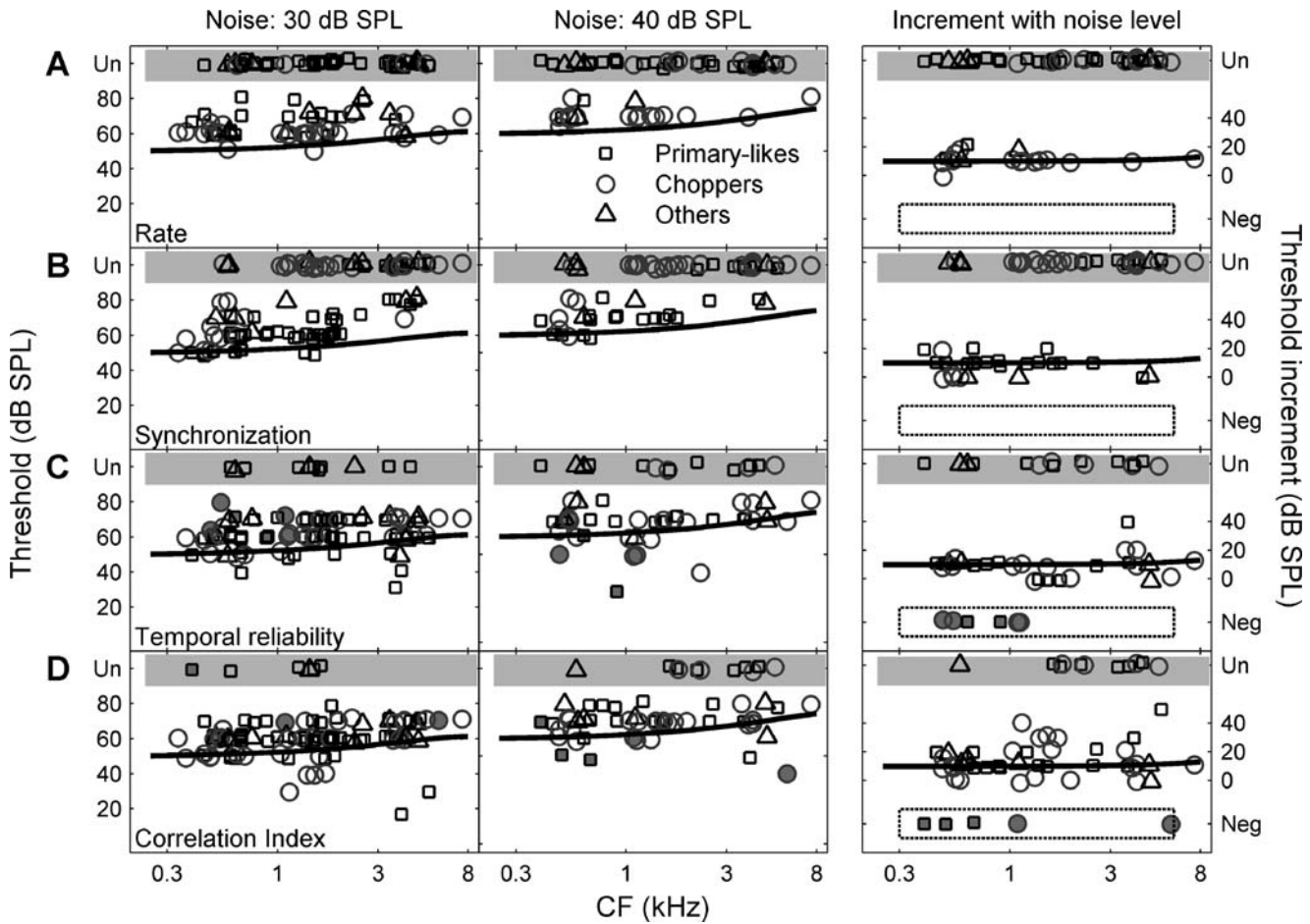
FIG. 5. *Left* and *middle panels* detection thresholds based on average discharge rate of the neuron (**A**), synchronization to the tone frequency (**B**), temporal reliability (**C**), and CI (**D**) for noise spectrum level of 30 dB SPL (*left*) and 40 dB SPL (*middle*). **A**, **B**, and **C** are replotted from Gai and Carney (2006, their Fig. 6). *Solid lines* indicate psychophysical thresholds in cat (Costalupes 1983). *Shaded areas labeled Un* indicate unmeasurable thresholds (no significant change occurred up to the highest tone level tested). *Right panel* shows threshold increments when noise level increased from 30 to 40 dB SPL. The *solid line* indicates psychophysical threshold increments. *Shaded areas labeled Un* indicate unmeasurable thresholds at 40 dB SPL noise level. *Filled symbols in right panel labeled Neg* indicate decreased thresholds with increasing noise level (these neurons are also marked with *filled symbols* in the *left* and *middle panels*).

tained in average rate. The asymmetrical bandpass shape describes MI functions with substantial rate information, but the maximum information was obtained when spike-timing information was included. Figure 6 shows the MI based on spike-time distance for all recorded neurons grouped by response types (primary-like, chopper, and unusual).

A common trend in the best temporal resolution was that neurons with symmetrical bandpass-shaped functions had consistently higher values of best q than did neurons with asymmetrical bandpass-shaped functions (Fig. 6). However, the best temporal resolution derived in this manner can be hard to interpret for the following two reasons. First, as q increases, the number of spikes required for sampling the space also increases (Victor and Purpura 1997). Because of limited data samples, however, the same number of spikes was used for different q values; therefore, the best q can be

underestimated. Second, it is likely that the presence of average-rate information at $q=0$ s^{-1} biased the best q toward lower values for the asymmetrical bandpass-shaped functions. To examine this possibility, the MI with varying q values for the first representative neuron was recomputed with artificially reduced rate information by including only the first 23 spikes (the mean discharge count for noise alone) from the stimulus onset for each noise-alone or noise-plus-tone response. The resulting MI was plotted in Figure 1F (thin line). The curve shifted to the right (toward higher q values) of the original MI that was computed with all recorded spikes (Fig. 1E, thick line). The same result was obtained by including only the last 23 spikes of each response (not shown); therefore, the shift of the MI curve was not due to a difference in the information carried by onset and offset responses. The shift of the MI curve was also not due to a generally shorter
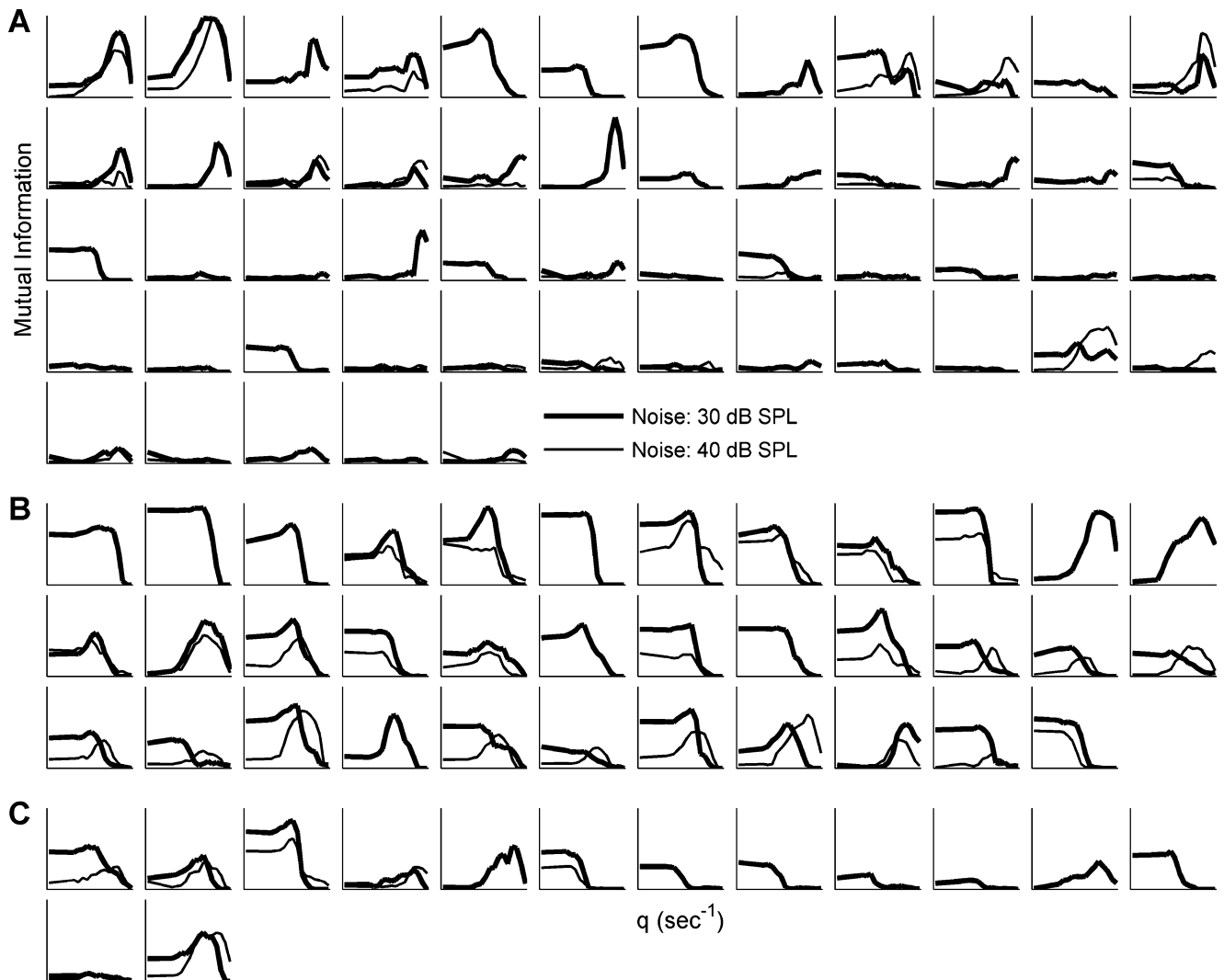
FIG. 6. Mutual information based on spike-time distances of tone-in-noise responses with varying q for all primary-like (**A**), chopper (**B**), and unusual (**C**) response types. In each group, neurons are arranged with increasing CF from *left* to *right, top* to *bottom*. Thick *lines* 30 dB SPL spectrum level noise, *thin lines* 40 dB SPL spectrum level noise. The x- and y-axis scales are the same as in Figure 1D (the y-axis ranges from zero to the upper limit of mutual information, $\log_2 3 = 1.59$ bits). All spikes were used in this analysis.

response duration for the truncated recordings used above; simulations using responses to only the first 150 ms (the entire stimulus duration was 250 ms) did not produce a shift of best q.

The bias toward smaller q values for neurons with changes in average discharge rate can also cause changes in best q at different noise levels. Figure 7A replots the MI of a chopper at 30 and 40 dB SPL noise spectrum level from the population results (Fig. 6B, third row, tenth column). The best temporal resolution was higher at 40 dB SPL (thin line) than at 30 dB SPL (thick line) noise spectrum level. However, when the same strategy was used to reduce the rate information, as shown in Figure 7B, the best temporal resolution did not vary across noise level. Therefore, the best temporal resolution will not be further explored in this study.

Maximum mutual information. In Figure 6 (population results), neurons are arranged with increasing CF for each response type. In general, the maximum MI decreased substantially with CF for primary-likes (Fig. 6A), but only slightly for choppers (Fig. 6B). This trend can be better visualized in Figure 8 where the peak MI is plotted as a function of CF. The maximum MI for choppers was generally higher than that of primary-likes and unusual neurons at all measured CFs.

It is not appropriate to compare the spike-distance metric with other temporal measures because the spike-distance metric is quantified by information and other temporal measures are quantified by detection threshold. Nevertheless, if any temporal metric is correlated to the spike-distance metric, it would most likely be the synchronization to tone frequency because this was the
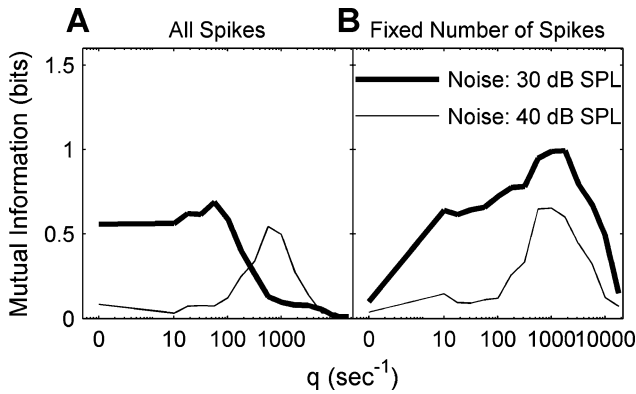
FIG. 7. Mutual information based on spike-time distances of noise-plus-tone responses with varying q at different noise spectrum levels. **A** Mutual information based on all spikes recorded. **B** Mutual information with fixed number of spikes (the number of spikes in each repetition for noise-alone stimuli and noise-plus-tone stimuli was no more than the mean number of spikes for noise-alone stimuli). *Thick lines* 30 dB SPL spectrum level noise, *thin lines* 40 dB SPL spectrum level noise. G315u20: CF=1579 Hz chopper.

only metric that showed decreased detection performance with CF for primary-likes (Fig. 5B). To compare the spike-distance metric and the synchronization to tone frequency, the MI of the synchronization coefficient was computed (see the "Methods" section).
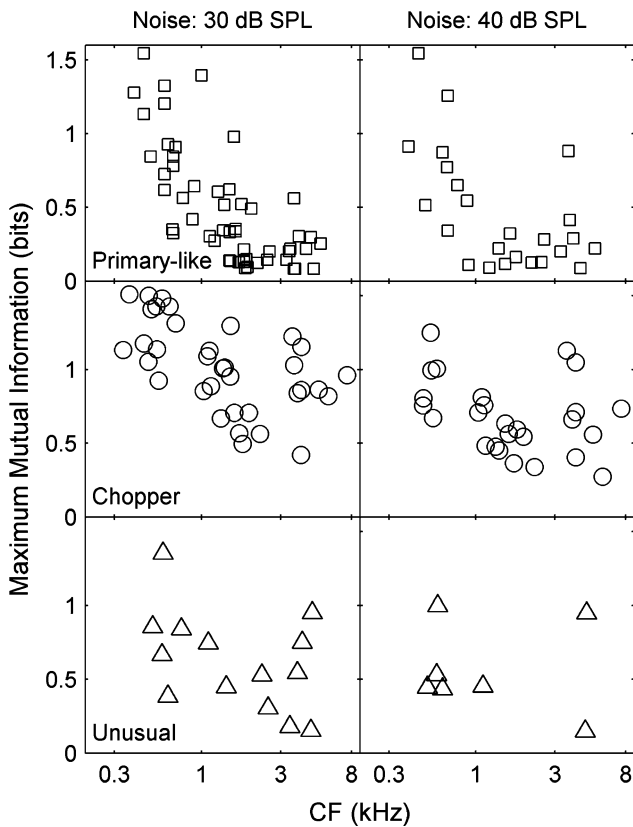

FIG. 8. Maximum mutual information based on spike-time distances of noise-plus-tone responses vs. CF. *Left column* 30 dB SPL spectrum level noise, *right column* 40 dB SPL spectrum level noise. All spikes were used in this analysis.

Figure 9 shows the MI based on synchronization to tone frequency vs. CF. The decreased MI based on synchronization seemed to resemble the maximum MI based on spike-time distance. However, when individual neurons were considered, these two measures were not significantly correlated ($r=0.04$ and 0.12 for 30 and 40 dB SPL noise, respectively). Figure 10 shows the maximum MI plotted against the MI based on synchronization. In fact, the MI based on synchronization was higher than the maximum MI based on spike-time distance for 72% and 63% of primary-likes at 30 and 40 dB SPL noise spectrum level, respectively. (The maximum MI was sometimes achieved with average-rate information; if only timing information was considered, the maximum MI would be even lower.)

In summary, absolute spike times of choppers carried more useful information for the detection of tones in noise than those of primary-likes, especially at mid and high CFs. In addition, the spike-distance metric was found to be generally unrelated to the previously studied temporal metrics.
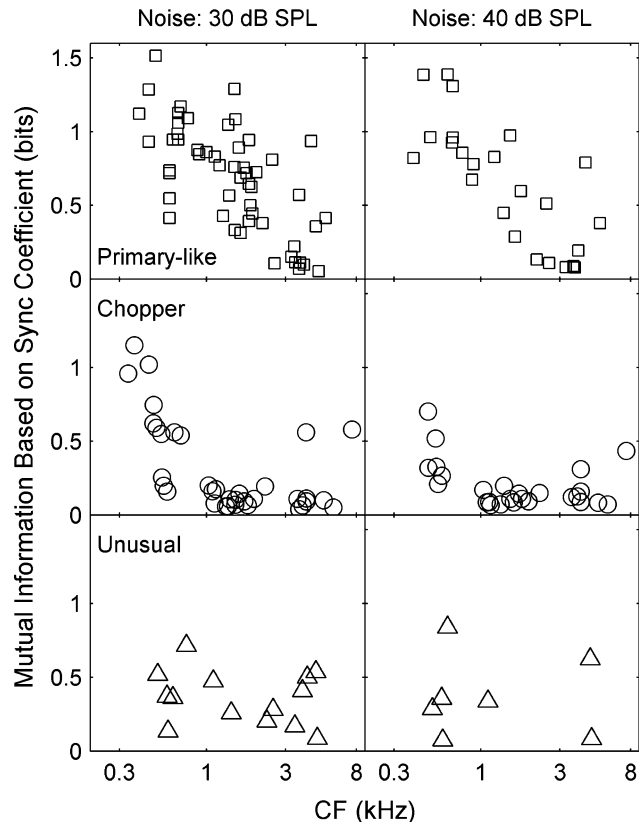

FIG. 9. Mutual information based on synchronization to tone frequency vs. CF. To match the computation of mutual information based on spike-time distance, the synchronization was also computed based on three sets of neural responses (to noise-alone stimuli, the third highest tone level in noise-plus-tone stimuli, and the highest tone level in noise-plus-tone stimuli). *Left column* 30 dB SPL spectrum level noise, *right column* 40 dB SPL spectrum level noise. All spikes were used in this analysis.
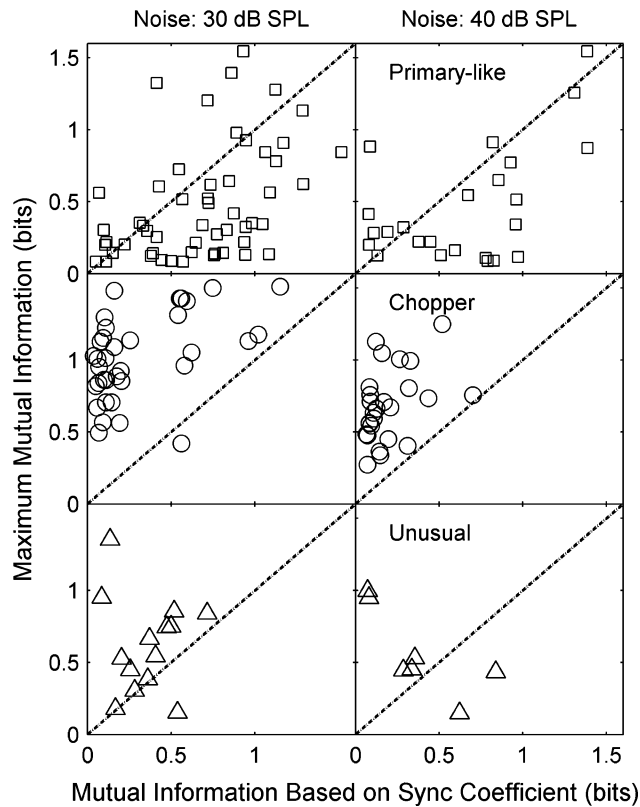
FIG. 10. Maximum mutual information based on spike-time distances vs. mutual information based on synchronization to tone frequency. *Left column* 30 dB SPL spectrum level noise, *right column* 40 dB SPL spectrum level noise. All spikes were used in this analysis.

information when the noise spectrum level increased. This observation suggested that when the average rate was saturated by high-level noise, detection information was more likely to be carried by absolute spike times.

Spike distance based on spike times vs. spike distance based on interspike intervals. The above results for the spike-distance metric were determined by the exact spike times with respect to stimulus onset. Figure 12 compares the spike-distance metric based on spike times (solid lines) with the spike-distance metric based on interspike intervals (dashed lines) for the four representative neurons. The maximum MI based on intervals was clearly lower than the maximum MI based on spike times in Figure 12A, C, and D. In Figure 12B, the two estimates of MI were comparable, and both were low.

Figure 13 shows the maximum MI of the two metrics for all recorded neurons. Overall, the maximum MI based on intervals was comparable to or lower than the maximum MI based on spike times with only a few exceptions for which the maximum MI based on intervals was slightly higher. No clear differ-
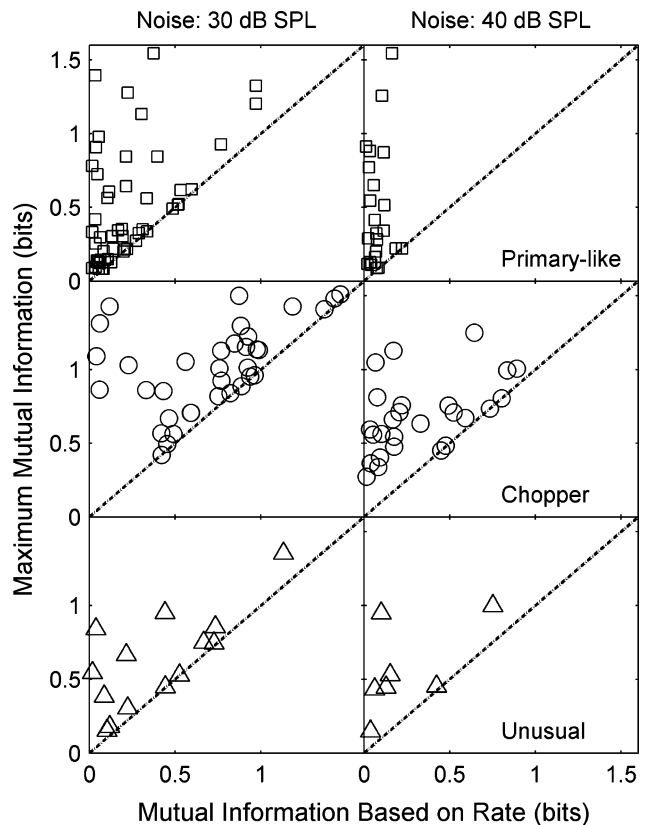
Average rate vs. timing. Figure 6 shows that for each type of AVCN neuron, some were dominated by average-rate information (lowpass-shaped), some by spike-timing information (symmetrical bandpass-shaped), and some by both (asymmetrical bandpass-shaped). Note that when the noise spectrum level increased from 30 dB SPL (thick lines) to 40 dB SPL (thin lines), 36% of neurons showed decreased rate performance and improved spike-timing performance as the MI changed from lowpass or asymmetrical bandpass shape into symmetrical bandpass shape. Figure 11 plots the maximum MI against the MI based on average rate. Symbols on the diagonals indicate cells for which the best coding for detection of tones in noise was average-rate coding, and symbols above the diagonals indicate cells for which the best coding for detection of tones in noise was spike-timing coding. When the noise level increased, the symbols generally moved away from the diagonals for primary-likes (Fig. 11, top). Although this trend was harder to visualize for the other response types (Fig. 11, middle and bottom), Figure 6B and C show that the decrease of mutual information contained in average rate was generally more severe than the decrease of temporal

FIG. 11. Maximum mutual information based on spike-time distances vs. mutual information for average-rate coding (mutual information at $q=0$). *Left column* 30 dB SPL spectrum level noise, *right column* 40 dB SPL spectrum level noise. All spikes were used in this analysis.
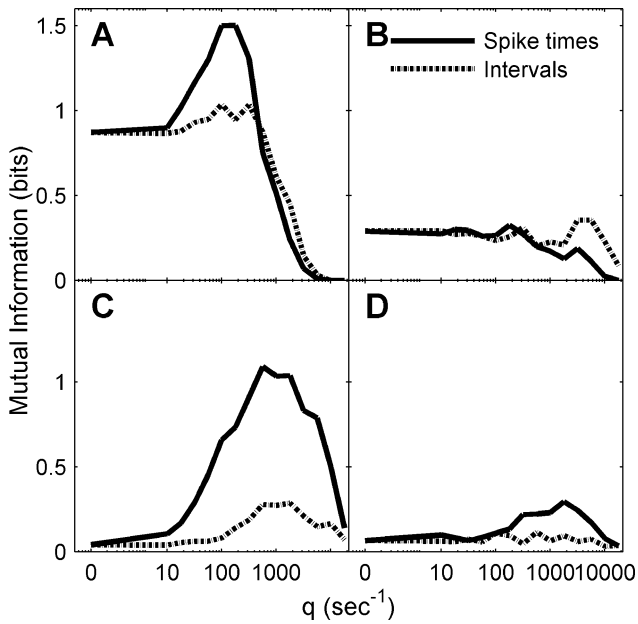
FIG. 12. Mutual information based on spike times (*solid lines*) vs. mutual information based on intervals (*dashed lines*). **A**, **B**, **C**, and **D** correspond to the four representative neurons shown in Figures 1, 2, 3, and 4.

ences among different neuron types were observed in this comparison.

DISCUSSION

Correlation index

Phase-locking is a response feature that contains important temporal information related to the sound stimulus; however, no physiological mechanism has been found to provide an accurate measure of the phase-locking, particularly for signals in the presence of background noise. The computation of vector strength requires knowledge of the duration of the tone cycle, which is not available to the brain. The shuffled autocorrelogram (SAC) is a relatively simple method to examine fine structure or envelope-related information for both periodic and aperiodic stimuli (Louage et al. 2004, 2005), yet a mechanism for computing the interspike intervals that it requires is still unclear.

Joris et al. (2006) shows that the correlation index, which can be implemented by relatively simple neural mechanisms such as coincidence detection (assuming the recorded responses of a single neuron across trials were similar to the simultaneous responses of a population of neurons), represents the synchronization to tone frequency for neural responses to pure tones, and it does not have the drawback of compressing high synchrony values. However, the present study showed that in the presence of broadband noise, the

correlation index decreased with tone level for the majority of AVCN neurons. For low-CF neurons that partially synchronize to both fine structure and envelope, increasing the level of a tone added to noise increases synchronization to tone frequency and decreases fluctuations in time-varying rate due to the flattening of the stimulus envelope (Gai and Carney 2006). Therefore, the correlation index is affected by two counteracting cues and does not represent either fine structure or envelope cues alone. As shown in Gai and Carney (2006), the dominance of envelope over fine structure caused a decrease in the temporal reliability with increasing tone level.

The dominance of envelope over fine structure in the correlation index is also implied by the finding of Louage et al. (2004) that a larger temporal consistency is observed in auditory-nerve responses to noise than to tones. The present study showed that the correlation index changed with tone level in a manner similar to the temporal reliability metric. It should also be noted that the temporal analysis windows used in this study (0.20 and 0.57 ms) were relatively large compared to that used in Joris et al. (2006; 0.05 ms); both envelope and fine structure play
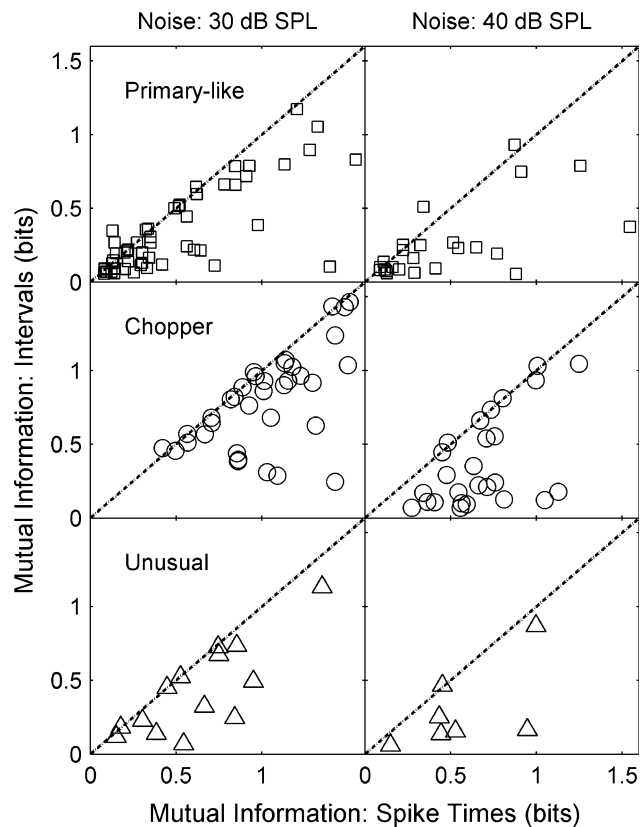

FIG. 13. Maximum mutual information based on spike-time distances vs. maximum mutual information based on interspike intervals. *Left column* 30 dB SPL spectrum level noise, *right column* 40 dB SPL spectrum level noise. All spikes were used in this analysis.

a role in determining the correlation index, especially for the larger windows.

The difference in detection performance between the correlation index and temporal reliability was not surprising, given the different computational methods and statistical tests used for these two metrics. The correlation index is physiologically more feasible than temporal reliability, as coincidence detection is a simpler neural mechanism than correlation of response patterns. Nevertheless, it is still unclear what mechanism could be used to normalize the coincidence count by average discharge rate (Eq. 1). This normalization is critical for the detection of tones in noise because the unnormalized correlation index can either increase or decrease with tone level, depending on the presence of increased average rate or decreased spike-timing coincidence. When both of these effects were present, simulations without normalization showed that coincidence detection was more sensitive to changes in average rate than to changes in spike-timing coincidence (Gai and Carney 2006).

The correlation index focuses on single-channel performance (stimulus tone frequencies were set equal to neuron CFs), as was true for the metrics described in Gai and Carney (2006). Studies that combine information across frequency channels using simple coincidence detection or cross correlation have shown certain advantages in information processing (Carney et al. 2002; Deng and Geisler 1987). Cross-frequency analyses may benefit by adopting the more sophisticated temporal metrics explored in this study; to test this hypothesis, future studies using off-CF tones in noise are required.

Spike-distance metric vs. other temporal metrics

Temporal information carried by spike times can be studied in different forms. The present study examined the effect of temporal feature sequence on information related to the detection of tones in noise. The spike-distance metric based on spike times measures information carried by absolute spike times. Note that the phrase "absolute spike time" does not necessarily refer to "precise" spike time; different temporal resolutions can be achieved by temporal smoothing, binning, or varying the cost factor for moving spikes. Rather, absolute spike timing refers to spike timing with respect to stimulus onset. On the other hand, synchronization to tone frequency and PSTH fluctuation (the latter was hypothesized to be dominated by locking to the stimulus envelope; Gai and Carney 2006) reflect response features that are not directly related to spike timing with respect to stimulus onset, but are related to temporal features that are "floating" in time. For example, two spike trains locked to different phases of the same stimulus

would not have the same absolute spike timing but could have the same synchronization coefficient. More generally, two different spike trains in response to two different stimuli can have the same synchronization coefficient. In other words, the synchronization and PSTH fluctuation metrics are influenced by temporal features regardless of the sequence of their occurrence, that is, whether they occur early or late in the response.

Temporal reliability and the correlation index measure the similarity of responses to different repetitions of the same stimulus. These two metrics are related to absolute spike times in response to the same stimulus. How absolute spike times change across different stimuli is not included in these metrics. If responses to different repetitions of the same stimulus are thought of as simultaneous responses from a population of statistically identical and independent neurons to the same stimulus presentation, then these two metrics can differ from the absolute spike coding because no memory of discharge times (across different stimuli) is required.

The spike-distance metric based on interspike intervals is not strictly dependent on absolute spike times, and therefore, can be thought of as a measure of relative spike timing. However, the *sequence* of intervals is important for this distance measure, distinguishing this metric from other temporal metrics based on relative spike times. This study showed that the sequence of interspike intervals generally carried the same or less information than the sequence of spike times.

No direct relationship was found between the results for the spike-distance metric and for other metrics based on relative spike times. Different types of neurons can carry different amounts of temporal information based on absolute spike times. This study showed that information carried by absolute spike times and average rate decreased dramatically with increasing CF for primary-like response types, but only slightly for other types of neurons. This decreasing trend with CF for primary-likes was not observed for any of the other temporal metrics (Fig. 5 in this study and Fig. 6 in Gai and Carney 2006) except for synchronization to tone frequency. A closer examination of the relationship between spike-time distance and synchronization to tone frequency showed no significant correlation for individual neurons. In fact, detection performance was better for most primary-likes when using synchronization to tone frequency than when using the spike-distance metric. Primary-likes apparently carry more information in the form of floating temporal information. Future studies should examine other auditory tasks to test whether the absolute spike times of choppers always carry more information than those of primary-likes.

Note that other types of spike-timing distances (e.g., van Rossum 2001) were not tested in this study. As long as the metric is based on absolute spike times, the general results for related spike-distance measures should be similar to those described in this study.

Best temporal resolution

In previous auditory studies (Chase and Young 2006; Huetz et al. 2006), the temporal resolution that yields the maximum mutual information has been identified as the time scale most likely used by a certain auditory stage when processing temporal information. The present study showed that the best temporal resolution, determined by q, can be biased toward low values by variations in average rate. Intuitively, when two spike trains have different numbers of spikes, the comparison of timing information is more ambiguous. This bias was reduced by basing the analysis on a constant number of spikes beginning at stimulus onset. However, this approach resulted in a loss of information carried by the removed spikes. More importantly, the spike trains were effective responses to stimuli that had different durations. The alternative strategy of equalizing rates by fixing stimulus duration and selectively removing spikes throughout the duration of the response would impair the temporal information. Therefore, to study the best temporal resolution based on absolute spike times, a better strategy to normalize the rate change must be designed. Other studies have also grappled with the issue posed by the effects of different numbers of spikes on the quantification of temporal information (Hung et al. 2002).

Spike-distance metric vs. template approaches

Templates with particular temporal features have been used in previous psychophysical detection models (Dau et al. 1996; Breebaart et al. 2001). In these models, decisions such as whether a tone is present or not in a background noise are based on comparisons of a template to a temporal representation derived from neural responses for a single trial. Models involving comparisons to templates typically rely on absolute stimulus features or absolute spike times (based on instantaneous rate functions). Results based on the spike-distance metric in this study suggested that the use of a template method for a certain task may be appropriate for some but not all types of neurons. For example, a strategy for the detection of tones in noise using a template is unlikely to yield a satisfactory result based on responses of primary-like neurons in the AVCN.

In summary, the spike-distance metric based on spike-timing precision was not an optimal metric for evaluating temporal information. For AVCN primary-like neurons, information useful for detection was more likely to be carried by features that are not anchored in time or by other forms of information that were less dependent on absolute spike times (i.e., temporal reliability and correlation index). The absolute spike times of choppers carry more information than for primary-likes, but there was no evidence concerning the relative amount of information carried by absolute spike times or floating features for choppers. Future studies should examine these issues in neurons at other auditory levels and for other types of auditory tasks. The same analytical and modeling approaches are also applicable to studies of masked detection in other sensory systems.

REFERENCES

BLACKBURN CC, SACHS MB. Classification of unit types in the anteroventral cochlear nucleus: PST histograms and regularity analysis. J. Neurophysiol. 62:1303–1329, 1989.

BREEBAART J, VAN DE PAR S, KOHLRAUSCH A. Binaural processing model based on contralateral inhibition. I Model structure. J. Acoust. Soc. Am. 110:1074–1088, 2001.

CANT NB, MOREST DK. The bushy cells in the anteroventral cochlear nucleus of the cat. A study with the electron microscope. Neuroscience. 4:1925–1945, 1979.

CARNEY LH, HEINZ MG, EVILSIZER ME, GILKEY RH, COLBURN HS. Auditory phase opponency: A temporal model for masked detection at low frequencies. Acta. Acust. Acust. 88:334–347, 2002.

CHASE SM, YOUNG ED. Limited segregation of different types of sound localization information among classes of units in the inferior colliculus. J. Neurosci. 25:7575–7585, 2005.

CHASE SM, YOUNG ED. Spike-timing codes enhance the representation of multiple simultaneous sound-localization cues in the inferior colliculus. J. Neurosci. 26:3889–3898, 2006.

COSTALUPES JA. Broadband masking noise and behavioral pure tone thresholds in cats. I. Comparison with detection thresholds. J. Acoust. Soc. Am. 74:758–764, 1983.

COSTALUPES JA. Representation of tones in noise in the responses of auditory nerve fibers in cats. I. Comparison with detection thresholds. J. Neurosci. 5:3261–3269, 1985.

COVER TM, THOMAS JA. Elements of Information Theory. New York, Wiley, 1991.

DAU T, PUSCHEL D, KOHLRAUSCH A. A quantitative model of the "effective" signal processing in the auditory system. II. Simulations and measurements. J. Acoust. Soc. Am. 99:3623–3631, 1996.

DENG L, GEISLER D. A composite auditory model for processing speech sounds. J. Acoust. Soc. Am. 82:2001–2012, 1987.

GAI Y, CARNEY LH. Temporal measures and neural strategies for detection of tones in noise based on responses in anteroventral cochlear nucleus. J. Neurophysiol. 96:2451–2464, 2006.

Geisler CD, Sinex DG. Responses of primary auditory fibers to combined noise and tonal stimuli. Hear. Res. 3:317–334, 1980.

Gibson DJ, Young ED, Costalupes JA. Similarity of dynamic range adjustment in auditory nerve and cochlear nuclei. J. Neurophysiol. 53:940–958, 1985.

Huetz C, Edeline JM. From receptive field dynamics to the rate of transmitted information: Some facets of the thalamocortical auditory system. Neuroembryol. Aging. 3:230–238, 2004/2005.

Huetz C, Del Negro C, Lebas N, Tarroux P, Edeline JM. Contribution of spike timing to the information transmitted by HVC neurons. Eur. J. Neurosci. 24:1091–1108, 2006.

Hung CP, Ramsden BM, Roe AW. Weakly modulated spike trains: Significance, precision, and correction for sample size. J. Neurophysiol. 87:2542–2554, 2002.

Joris PX, Louage DHG, Cardoen L, van der Heijden M. Correlation index: A new metric to quantify temporal coding. Hear. Res 216–217:19–30, 2006.

Licklider JCR. A duplex theory of pitch perception. Experimentia. 7:128–134, 1951.

Louage DHG, van der Heijden M, Joris PX. Temporal properties of responses to broadband noise in the auditory nerve. J. Neurophysiol. 91:2051–2065, 2004.

Louage DHG, van der Heijden M, Joris PX. Enhanced temporal response properties of anteroventral cochlear nucleus neurons to broadband noise. J. Neurosci. 25:1560–1570, 2005.

May BJ, Sachs MB. Dynamic range of neural rate responses in the ventral cochlear nucleus of awake cats. J. Neurophysiol. 68:1589–1602, 1992.

Meddis R, Hewitt MJ. Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification. J. Acoust. Soc. Am. 89:2866–2882, 1991.

Miller MI, Barta PE, Sachs MB. Strategies for the representation of a tone in background noise in the temporal aspects of the discharge patterns of auditory-nerve fibers. J. Acoust. Soc. Am. 81:665–679, 1987.

Moore DS, McCabe GP, Duckworth WM, Sclove SL. The Practice of Business Statistics: Using Data for Decisions. New York, Freeman, pp. 18–4–18-65, 2002.

Rees A, Palmer AR. Rate-intensity functions and their modification by broadband noise for neurons in the guinea pig inferior colliculus. J. Acoust. Soc. Am. 83:1488–1498, 1988.

Rhode WS, Geisler CD, Kennedy DT. Auditory nerve fiber responses to wide-band noise and tone combinations. J. Neurophysiol. 41:692–704, 1978.

Richards VM. The detectability of a tone added to narrow bands of equal-energy noise. J. Acoust. Soc. Am. 91:3424–3435, 1992.

Smith PH, Rhode WS. Structural and functional properties distinguish two types of stellate cells in the ventral cochlear nucleus. J. Comp. Neurol. 282:596–616, 1989.

van Rossum MC. A novel spike distance. Neural Comput. 13:751–763, 2001.

Victor JD, Purpura KP. Nature and precision of temporal coding in visual cortex: A metric-space analysis. J. Neurophysiol. 76:1310–1326, 1996.

Victor JD, Purpura KP. Metric-space analysis of spike trains: Theory, algorithms and application. Netw. Comput. Neural Syst. 8:127–164, 1997.

Wier CC, Green DM, Hafter ER, Burkhardt S. Detection of a tone burst in continuous- and gated-noise maskers; defects of signal frequency, duration, and masker level. J. Acoust. Soc. Am. 61:1298–1300, 1977.

Young ED, Barta PE. Rate responses of auditory nerve fibers to tones in noise near masked threshold. J. Acoust. Soc. Am. 79:426–442, 1986.