Cues for Sound Localization Are Encoded in Multiple Aspects of Spike Trains in the Inferior Colliculus

Steven M. Chase and Eric D. Young

Department of Neurobiology, University of Pittsburgh and Department of Statistics, Carnegie Mellon University, Pittsburgh, Pennsylvania; and Center for Hearing and Balance and Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland

Submitted 11 June 2007; accepted in final form 27 January 2008

Chase SM, Young ED. Cues for sound localization are encoded in multiple aspects of spike trains in the inferior colliculus. J Neurophysiol 99: 1672–1682, 2008. First published January 30, 2008; doi:10.1152/jn.00644.2007. To localize sound, information from three cues—interaural timing differences (ITDs), interaural level differences (ILDs), and spectral notch cues (SNs)—must be properly integrated. The inferior colliculus (IC) receives convergent input from neurons encoding all three cues. Using virtual space stimuli and information theoretic techniques, we investigated the coding of the various localization cues in single neurons of the IC under different recording schemes. Here we focus on the analysis of information encoded by first-spike latency, in comparison to previous results on discharge rate and ongoing spike timing. The results show that the localization cues converge to different degrees in particular neurons. ITD information is conveyed most strongly by spike rate, with small amounts of independent information in latency and ongoing spike timing. ILD information shows a similar pattern, with larger mutual information values for all three cues. For these cues, ongoing spike timing does not typically contribute independent information over that captured by a joint rate/first-spike latency code. SNs are coded by both rate and first-spike latency, but ongoing spike timing significantly enhances their representation in a best frequency–dependent manner, as long as the temporal envelope of the stimulus can be used in the decoder. The differential coding of the localization cues suggests that information about multiple cues could be multiplexed onto the responses of single neurons.

INTRODUCTION

Sensory objects are represented in the brain in multiple ways. These can be different features of the object (loudness, frequency content) or multiple cues for one feature, like the multiple cues for sound localization. When streams of information representing different aspects of a sensory object converge at a neuron, the neuron may discard some of the information and provide a reduced and specialized representation of the object or it may combine information to produce a higher-order representation of the object. At intermediate steps of the latter process it may be useful to preserve separately aspects of more than one of the afferent information streams. How multiple information streams are multiplexed onto neurons’ spike trains is not well understood.

We address the general question of information integration in the context of sound localization coding. The central nucleus of the inferior colliculus (IC) serves as a major integration point of the auditory system, receiving convergent input from nearly every brain stem auditory nucleus (for reviews, see Winer and Schreiner 2005). Among other inputs, neurons in the IC receive information about the three main cues used for sound localization: interaural level differences (ILDs), interaural timing differences (ITDs), and monaural spectral cues (SNs, for spectral notches). IC neurons are usually sensitive to more than one localization cue (Benevento and Coleman 1970; Caird and Klinke 1987; Chase and Young 2005; Delgutte et al. 1995, 1999). The representation of localization cues has usually been analyzed using discharge rate as the response variable (Palmer and Kuwada 2005). However, other features of the response, like first-spike latency and the temporal patterning of spike trains, are also known to encode stimuli in the auditory system (Furukawa and Middlebrooks 2002; Narayan et al. 2006; Nelken et al. 2005).

We have previously used virtual-space stimuli to investigate the simultaneous representation of two localization cues using either a rate code (Chase and Young 2005) or a spike distance metric (SDM) code (Chase and Young 2006; Victor and Purpura 1997). SDM codes are sensitive to spike timing distributed throughout the neural response. However, when given an analysis window consisting of the entire response duration, we found that the SDM code is not very sensitive to changes in the timing of single spikes, such as first-spike latency differences (cf. Fig. 7 of Chase and Young 2006). The present work expands on our previous work by investigating the coding of localization cues in first-spike latency and comparing the relative efficacy of each code in representing the different cues. Latency provides information about all three cues; as is the case for rate responses, the latency information is largest for ILDs, but there is also latency information about ITDs and SNs. Because of the differential coding of the localization cues in different aspects of spike trains, our results show how cues can be multiplexed onto the activity of single neurons.

METHODS

The surgical procedure, recording protocol, and stimulus design were described in Chase and Young (2005), and will be presented only briefly here. The results presented here are based on the same data set as in previous papers (Chase and Young 2005, 2006).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
Surgical procedure

Acute recording experiments were performed on adult cats with clean external ears, obtained from Liberty Labs (Waverly, NY). Animals were anesthetized for surgery with xylazine [1 mg/kg, administered intramuscularly (im)] and ketamine (40 mg/kg, im). The cats were decerebrated by transecting the brain between the superior colliculus and the thalamus. After decerebration, anesthesia was discontinued. Throughout the experiment, the cat’s temperature was maintained between 37.5 and 38.5°C with a feedback-controlled heating pad.

The IC was approached dorsally by aspirating occipital cortex and, when necessary, removing part of the bony tentorium. The ear canals were exposed and fitted with ear tubes for sound delivery, and the bullae on both sides were vented with 30 cm of PE 90 tubing. At the end of the experiment, the cat was killed with an overdose of barbiturate anesthetic. All procedures were performed in accordance with the guidelines of the Institutional Animal Care and Use Committee of the Johns Hopkins University.

Electrodes were advanced dorsoventrally through the IC to sample neurons with various best frequencies (BFs). All neurons used for the analyses had narrowly tuned nonhabituating responses recorded in a dorsoventral BF sequence (as in Ramachandran et al. 1999); these results and the placement of the electrode suggest that the data are from the central nucleus of the IC.

Recording protocol

All recordings were made in a sound-attenuating chamber. Sounds were presented via speakers placed on hollow ear bars inserted into the ear canals. In situ speaker calibrations showed responses that are uniform (±4.6 dB SPL) between 40 Hz and 35 kHz. Platinum/iridium microelectrodes were used for single-neuron recording; neurons were isolated with a Schmitt trigger or a template-matching program (Alpha-Omega Engineering, Nazareth, Israel).

Stimulus design

Two sets of virtual space stimuli were used in this study, based on a single 330-ms token of broadband noise sampled at 100 kHz (interstimulus interval: 1 s). Each set was manipulated to vary independently in two parameters and each parameter was adjusted in five steps, for a total of 25 stimuli per set. To build up statistics sufficient for information theoretic analyses, each stimulus set was repeated multiple times (20–200, depending on how long the neuron was held; median: 90 repetitions). Stimuli were presented with 10-ms linear onset and offset ramps, interlaced in fixed order, such that all 25 stimuli were presented once, then the entire stimulus set was presented again, and so forth.

In the first stimulus set, ITD and ILD were manipulated. The frozen noise token was filtered through a spatially averaged head-related transfer function (HRTF; obtained from the cat data of Rice et al. 1992), which imparts to the stimulus the spectral characteristics of the head and ear canal, independent of spatial location. The stimulus was then split into two streams (one for each ear) that were delayed relative to one another to impart an ITD and attenuated relative to one another to impart an ILD. ITDs and ILDs were chosen to correspond roughly to spatial locations in the horizontal plane of −60, −30, 0, 30, and 60° azimuth, where negative values refer to locations in the ipsilateral hemifield (Kuhn 1977; Rice et al. 1992; Roth et al. 1980).

For ITDs, the five cue values were −160, −80, 0, 80, and 160 ms; for ILDs, they were −13.8, −8.4, 0, 8.4, and 13.8 dB. The ILDs were imparted by setting the levels in the ipsilateral ear to 0, 0, 0, −8.4, and −13.8 dB (re the reference level) and the levels in the contralateral ear to −13.8, −8.4, 0, 0, and 0 dB. The ILDs were arranged in this way to help disambiguate ILD sensitivity from monaural or binaural sound-level sensitivity, as explained previously (Chase and Young 2005).

In the second stimulus set, ILD and SN were varied. The SN cue was imparted by filtering the frozen noise token through one of five midline HRTFs containing a prominent spectral notch, representing elevations ranging from 0 to 30° in 7.5° steps (see Fig. 2D of Chase and Young 2005). The stimulus was then split into two streams and an ILD cue was imparted as in the ITD/ILD stimulus set. Interaural spectral differences were not considered in this work because the same stimulus spectrum was sent to each ear. Before presentation, the stimuli in this set were resampled (resample command in MATLAB, The MathWorks, Natick, MA) such that the five SNs surrounded the BF of the neuron under study with the notch frequency of the third HRTF at BF. A by-product of this resampling is a BF-dependent change in stimulus length. For presentation, stimuli >400 ms were truncated at 400 ms, whereas stimuli <200 ms were repeated (including the 10-ms linear ON–OFF ramps) to be ≥200 ms long. This latter step was necessary only for neurons with BFs >18.5 kHz (n = 10).

Mutual information

Data were analyzed by computing the mutual information (MI) between stimulus (S) and response (R) under the assumption of several different response variables. The MI is defined by (Cover and Thomas 1991)

\[
MI(S; R) = \sum_{s} \sum_{r} p(s, r) \log \frac{p(s, r)}{p(s)p(r)}
\]

where \(p(\cdot)\) denotes a probability.

The stimulus sets (described earlier) consist of 25 stimuli containing variation in two stimulus parameters. To compute the MI between the neural responses and just one cue, a 5-stimulus set was defined by combining the stimuli across the other parameter. Thus the MI for individual cues quantifies the information carried about one cue in the presence of variation in the other cue. For all codes, data were analyzed over a window starting at stimulus onset and extending for 200 ms.

For spike rate, the MI was computed directly from empirical distributions of spike counts; that is, \(p(s, r)\) was the probability of getting a certain spike count \(r\) for a stimulus \(s\). Bias was estimated (and removed) using a bootstrap procedure (Efron and Tibshirani 1998). This method has been fully explained, including the debiasing techniques, in a previous publication (Chase and Young 2005). The data processing here was slightly different from that in our previous papers. Here we analyzed spikes from a fixed interval of 0–200 ms, regardless of the duration of the stimulus, whereas previously we used the full stimulus duration. Also here we analyzed all neurons, regardless of response type, whereas previously we analyzed only response types V, I, and O in the scheme of Ramachandran and colleagues (1999). These changes result in small differences between comparable figures (Fig. 5, D–F of this paper vs. Fig. 5, A, B, and D in Chase and Young 2006).

For the SDM code, the response variable is an estimate of the localization cues depending on both spike rate and spike timing (Victor and Purpura 1997). The application of this technique to these stimulus sets is described in detail in a previous publication (Chase and Young 2006). Briefly, the distance between two spike trains is defined as the sum of the costs of the elementary steps it takes to transform one spike train into the other, where the allowed steps are spike deletion (cost of 1), spike addition (cost of 1), and spike shift (cost of \(q|\Delta t|\)). Here, \(\Delta t\) is the time difference between a spike in one train and the nearest spike in the other train and \(q\) is a variable cost parameter with units of s\(^{-1}\) that effectively sets the temporal resolution of the code. For a given \(q\), there exists a minimum cost solution for the distance between any two spike trains. The distances are used to construct a confusion matrix for the responses as follows. To
estimate the localization cues most likely to have elicited a particular spike train, the average distance between that train and the spike trains elicited by each of the 25 stimuli is calculated. The estimated stimulus is the stimulus eliciting spike trains with the minimum average distance to the spike train under consideration. The $M_{\text{ISDM}}$ is computed directly from the confusion matrix by estimating the probabilities in Eq. 1 from the confusion matrix; for this calculation, $s$ is the actual stimulus and $r$ is the estimated stimulus. This calculation was repeated at several values of cost $q$; in addition to $q = 0$, the cost was sampled logarithmically from 10 to 15,850 s$^{-1}$ at 5 costs per decade. For each neuron, the maximum value of $M_{\text{ISDM}}$ across all costs was used (corresponding to median costs of 39.8 s$^{-1}$ for ILD cues, 100 s$^{-1}$ for SN cues, and 0 s$^{-1}$ for ITD cues; see Chase and Young 2006).

Although the SDM method extracts information in spike timing beyond that conveyed by spike rate, it is relatively insensitive to small changes in temporal patterns (see Fig. 7 of Chase and Young 2006). Because latency variations usually occur only in the first ten or so milliseconds of a neuron’s response to a stimulus (e.g., Fig. 2A), the SDM decoder can underestimate the information carried by first-spike latency.

First-spike latency is actually a vector of two response variables. The first element characterizes the likelihood that a particular stimulus will elicit at least one spike; the second element characterizes the timing of that spike, should it occur (Nelken et al. 2005). Both of these elements can carry information. If $M_{\text{ISP}}$ is the information between the stimuli and the likelihood of eliciting at least one spike, and $M_{\text{ITM}}$ is the information between the stimuli and the timing of the first spike, when it occurs, then the first-spike latency information $M_{\text{IFSL}}$ is related to its components through

$$M_{\text{IFSL}} = M_{\text{ISP}} + p_{\text{spike}}M_{\text{ITM}}$$

where $p_{\text{spike}}$ is the average probability, over all stimuli, of the stimulus set eliciting at least one spike. A proof of this decomposition can be found in APPENDIX III of Nelken et al. (2005). $M_{\text{ISP}}$ was computed by directly counting the numbers of times each stimulus elicited at least one spike. $M_{\text{ITM}}$ was computed using Victor’s binless method (Victor 2002), which is based on nearest-neighbor distances. Essentially, the probability of observing a spike at time $t$ is estimated from the density of samples at time $t$. Specifically, we use Eq. 12 from that work

$$M_{\text{ITM}} = \frac{1}{N} \sum_{j=1}^{N} \log_2 \left( \frac{\lambda_j}{\hat{\lambda}_j} \right) - \frac{1}{N} \sum_{k=1}^{N} \log_2 \left( \frac{N - 1}{N - 1} \right)$$

Here, $j$ indexes the first spike samples and $k$ indexes the stimuli. The nearest-neighbor distance computation comes from the first term on the right-hand side, where $\lambda_j$ is the minimum Euclidean distance (here, the difference in spike times) between spike $j$ and any other first spike, regardless of stimulus, and $\hat{\lambda}_j$ is the minimum Euclidean distance between spike $j$ and any other first spike elicited by the same stimulus. The number of stimuli $S = 25$ for the full stimulus set or 5 when considering an individual localization cue. $N_j$ is the number of times the stimulus $k$ elicited a spike, whereas $N$ is the total number of first-spikes observed across all stimuli.

To help control for errors due to small sample sizes, if a stimulus did not elicit a spike on at least eight repetitions, its contribution to the information was set to zero by deleting the corresponding terms in the second summation in Eq. 3. Spikes were recorded with 10-μs precision. To avoid zero distances, a random number uniformly distributed over $[-5, 5]$ μs was added to each spike time.

In practice, we find the nearest-neighbor method of $M_{\text{IFSL}}$ computation of Victor (2002) to be nearly bias free. To test this, we took neurons from which $\geq 100$ repetitions of each stimulus were collected and computed the information contained in randomly selected subsets of the spike trains. Figure 1A shows an example of this calculation for three randomly selected neurons from which 100 stimulus repetitions were collected. Here $M_{\text{ITM}}$ is plotted as a function of $N_{\text{reps}}$, the size of the subset in repetitions. Each calculation was repeated four times, with a different random subset each time. Although there is a fair amount of scatter in the data points for the smaller subsets, for $\geq 20$ repetitions the mean values are relatively stable. The variation in the results at $N_{\text{reps}} = 100$ is due to the random jitter added to the spike times. Figure 1B shows the average MIs measured using 100 repetitions of each stimulus plotted against the average MIs measured using only 20 randomly selected repetitions of each stimulus. There is good agreement between the two measures and no tendency for the 20-repetition estimates to be either smaller or larger than the 100-repetition estimates ($P = 0.69$, signed-rank test). For this reason, we did not attempt to debias these MI measurements.

The first-spike latency and the discharge rate of responses are correlated in most cases. To evaluate the degree of redundancy in rate and latency codes, we computed $M_{\text{JOIN}}$, the information in a joint latency-rate spike code, as follows

$$M_{\text{JOIN}} = M_{\text{ITR}} + \frac{1}{N} \sum_{r} p(r)\log_2 \left( \frac{N_{\text{ITR}}}{N_{\text{ITR}}} \right)$$

This equation is derived from the chain rule in a way similar to Eq. 3. The first term on the right side of the equations is the usual MI for rate (Chase and Young 2005). The second term is the additional information in latency, beyond that in rate. Note that this term corresponds only to the timing information, because $M_{\text{ISP}}$ is zero (and $p_{\text{spike}}$ is one) for spike trains with a fixed (nonzero) number of spikes.

**RESULTS**

Electrophysiological recordings were made from well-isolated single neurons in the IC of decerebrate cats. Once a
neuron was isolated and its BF determined, one of the two virtual-space stimulus sets was presented. In all, 101 neurons were recorded, 35 with the ITD/ILD stimulus set and 83 with the SN/ILD stimulus set (some neurons were held long enough to record both sets of stimuli; in these cases, both of the measured ILD MI values are included in the plots).

Figure 2 shows two examples of spike rasters recorded in response to the SN/ILD stimulus set. Each box represents the responses to multiple presentations of one stimulus; ILD parameters are given at the left, SN parameters at the right. With the use of frozen noise, spike times are very consistent across multiple repetitions of the same stimulus. A common feature of the responses of IC neurons to these frozen noise stimuli is the appearance of “clusters” of spikes in the rasters. These clusters correspond to the alignment of single spikes across multiple repetitions of the same stimulus with some small jitter (and not to bursts of spikes in response to one stimulus). In Fig. 2A, there is a clear ILD-dependent change in the latency of these spike clusters that disappears within 10 ms. This is typical of the responses of IC neurons in this data set; the timing of the first cluster of spikes is often more strongly dependent on the localization cues than the timing of subsequent spike clusters (Chase and Young 2007). The strong variation in first-spike timing exhibited by this neuron makes first-spike latency an obvious candidate for a neural code. Figure 2 shows an example in which spike rate and first-spike latency show correlated changes to both ILD and SN.

We computed the first-spike latency information ($MI_{rsp}$) between the localization cues and the first-spike latency variations exemplified by Fig. 2. Again, with this stimulus design, information can be computed about the entire set of 25 stimulus cues or, by combining rows of the probability distributions, the information between responses and the individual cues can be computed.

Figure 3A shows the MI between first-spike latency and the full set of stimulus cues for the population of neurons studied, plotted as a function of BF. On average, first-spike latency carries $0.68 \pm 0.35$ (SD) bits of information about the stimuli, and there are no obvious trends with BF. Although it is not shown, there are also no trends with the IC neuron response classes, as defined by Ramachandran et al. (1999). The components of the first-spike latency information are shown in Fig. 3, B and C. A substantial number of neurons (48/118) have no information in $MI_{rsp}$ because each repetition of each stimulus elicited at least one spike. Of those neurons with $MI_{rsp} > 0$, roughly 23% of $MI_{FSL}$ stems from $MI_{rsp}$, with the remainder coming from actual changes in spice timing. There are no obvious differences in $MI_{rsp}$ or $MI_{tim}$ across the two stimulus sets, with the possible exception that for very low frequency neurons (BFs <1 kHz), $MI_{tim}$ for the SN/ILD stimulus set is larger than for the ITD/ILD stimulus set.

There are differences in the codes that are effective in representing the individual cues. Figure 4 shows the $MI_{FSL}$ of the individual localization cues plotted as a function of BF. First-spike latency carries, on average, $0.43 \pm 0.28$ bits of information about ILD, $0.23 \pm 0.26$ bits of information about SN, and only $0.09 \pm 0.12$ bits of information about ITD. There are no obvious differences across BF, except for the decrease in ITD information at high BFs, which is also seen in a spike rate code (Fig. 6; Chase and Young 2005). Of course no ITD information is expected at high BFs, based on the duplex theory of sound localization (Dreyer and Delgutte 2006; Macpherson and Middlebrooks 2002; Rayleigh 1907; Simon et al. 2002).

![Figure 2](http://jn.physiology.org/)

**FIG. 2.** Spike raster plots of the first 50 ms of the responses of 2 example neurons to the spectral notch/interaural level difference (SN/ILD) stimulus set. Thick black lines divide different SN parameters, shown to the right of each plot; thin black lines divide different ILD parameters, shown at left. A: responses of a 1.3-kHz neuron, from which 100 repetitions of each stimulus were collected. B: responses of a 13.2-kHz neuron, from which 55 repetitions of each stimulus were collected.
Although the average $M_{FSL}$ for the ITD cue is quite small, this alone does not indicate that first-spike latency is not the code for ITD. For example, it could be that the range of ITD cues used in these stimuli probe these neurons less effectively than the range of ILD cues. One way to investigate this possibility is to compare the representation of the various cues under different codes. Because the first-spike latency code and the rate code often covary (as in Fig. 2B; see also Nelken et al. 2005), it is useful to calculate the information conveyed by first-spike latency that is independent of the information conveyed by rate. To do this, it is necessary to compare the information conveyed by a joint rate/first-spike latency code ($M_{JOINT}$, Eq. 4) to the information conveyed by rate alone ($M_{RAT}$). The difference between the two is the amount of additional information contained in first-spike latency.

Figure 5, A–C shows $M_{JOINT}$ plotted against $M_{RAT}$ for the three localization cues. Vertical deviations from the diagonal represent cases where estimation noise makes $M_{RAT}$ appear greater than the other cue. For the ITD cue (Fig. 5A), the points cluster close to the diagonal (mean deviation: 0.06 ± 0.10 bits); there is little information conveyed by first-spike latency about ITD that is not conveyed by spike rate. Both ILD and SN show larger deviations from the diagonal (ILD: 0.17 ± 0.17 bits; SN: 0.16 ± 0.13 bits), indicating extra information available in first-spike latency over that available in the rate. The independent information conveyed by first-spike latency is significantly smaller for ITD cues than for either ILD or SN cues ($P < 10^{-5}$ in both cases, rank sum). This significant difference remains if the range of $M_{RAT}$ is constrained (to be 0–0.5 or 0–1 bits) so that the comparison is done for cases with comparable $M_{RAT}$.

Figure 5, D–F shows a similar analysis for the information coded in the timing of spikes throughout the response, as computed using the SDM. The SDM method incorporates information from both spike rate and spike timing, so it provides a lower bound to the total information available in the spike train. The vertical distance between data points and the diagonal lines in Fig. 5, D–F is the additional information conveyed by spike timing over that available in discharge rate.

For ITD coding (Fig. 5D), there is little extra information conveyed by the SDM code, as was the case for latency (Fig. 5A) (mean SDM gain: 0.02 ± 0.06 bits). This result again suggests that ITD information is primarily carried by spike rate. For ILD coding there is an approximately equivalent amount of extra information available in the SDM code (0.12 ± 0.18 bits; Fig. 5E) as in the latency code (Fig. 5B). SN information is also provided by both latency (Fig. 5C) and spike timing throughout the response (0.31 ± 0.41 bits; Fig. 5F), although in this case there is more information in the SDM code than there is in the joint rate/first-spike latency code ($P = 0.02$, signed-rank test).

To further dissect the coding of the localization cues, Fig. 6 shows the information conveyed about each cue as a function of BF. For reference, the top row shows the information contained in spike rate. The middle row shows the FSL gain, or the extra information contained in first-spike latency that is not available in rate. This corresponds to $M_{JOINT} - M_{RAT}$, the
Fig. 5. Comparison of the information contained in different codes. In each plot, information in a temporal code is plotted against the information contained in a rate code alone. Top row: the plots show information contained in a joint rate/first-spike latency code. Bottom plots: information recovered by a spike distance metric (SDM) code. In both cases, vertical displacement from the diagonal represents information available in the temporal code that is not available in rate. A: ITD information in a joint rate/first-spike latency code compared with the ITD information in rate alone. B: same as A, for ILD. C: same, for SN. D: ITD information recovered by an SDM code, compared with the information contained in rate alone. E: same as D, for ILD. F: same, for SN.

vertical deviations of the points in Fig. 5, A–C. Unlike the rate information in the top row, the FSL gain does not show any obvious BF dependence for any of the cues. The SDM gain, corresponding to the vertical deviations shown in Fig. 5, D–F, is shown on the bottom row. Here there are differences among the three localization cues. The most striking feature of Fig. 6 is the difference in the BF dependence of the SN information carried by the three possible codes (right column). Whereas the rate information is highest in the mid to high BF range, the FSL gain is relatively independent of BF, and the SDM gain is highest at low BFs.

In discussing Fig. 5, we emphasized that the timing gains for ITD are smaller than for ILD and SN. Figure 6 shows this effect clearly, in that the scatter of points is more restricted near a timing gain of 0 for ITD than for ILD or SN. However, it is also true that MI values in general are smaller for ITD than for ILD or SN (Fig. 4 and the top row of Fig. 6); this begs the question of whether the timing gains as a fraction of \( M_{IRATE} \) are smaller for ITDs. Unfortunately, the scatter in the data do not allow a firm answer to this question because most of the timing gains for ITD are near the values of the bias and scatter in the data.

It has been shown in experiments on the auditory cortex of anesthetized cats that a joint code consisting of spike count and first-spike latency (Furukawa and Middlebrooks 2002) or mean spike latency (Nelken et al. 2005) is sufficient to capture the majority of information transferred by auditory neurons. In Fig. 7, A–C, we compare the information contained in a joint rate/first-spike latency code to the maximum information contained in an SDM code by plotting \( M_{JOIN} \) against \( M_{SDM} \). Points that lie above the diagonal indicate instances where the SDM code underestimates the total information available in the spike trains. These are cases in which latency is particularly informative, but the SDM analysis does not capture that information. Points below the diagonal, however, indicate information available in spike timing beyond that available in the first spike.

For the ITD cues (Fig. 7A), the points lie close to the diagonal, although they have a (nonsignificant) tendency to cluster above it (22 of 35, \( P = 0.16 \), signed-rank test). ILD cues (Fig. 7B) show a significant tendency to lie above the diagonal (79 of 118, \( P = 0.002 \), signed-rank test). Thus a joint rate/first-spike latency code usually captures slightly more ILD information than the SDM code can recover. For SN cues (Fig. 7C), many points fall below the diagonal, indicating that spike timing carries extra information about SNs not captured by the joint code. Previously we showed that this information is in the form of responses to envelope fluctuations over the duration of the stimulus (Chase and Young 2006). The difference between \( M_{ISDM} \) and \( M_{JOIN} \) is plotted as a function of BF for the SN cues in Fig. 7D. Low BF units are almost entirely dominated by SDM coding. In the physiological range of SN cues (6 to 20 kHz), there are neurons where \( M_{ISDM} \) dominates \( M_{JOIN} \) and neurons where the opposite holds true. In the highest BF neurons, \( M_{JOIN} \) and \( M_{ISDM} \) are roughly equivalent.

**Discussion**

The main finding of this work is that there is a diversity of coding mechanisms for the three different localization cues. Although the relative information provided about different cues varies from neuron to neuron, there are some consistent features of the representations. At the level of the IC, ITD information is encoded most strongly by spike rate, with small
contributions from latency and spike timing (Figs. 5, A and D and 7A). Similarly, ILD information is mainly encoded by rate, with lesser information in first-spike latency and the least in ongoing spike timing (Fig. 5, B and E), and the joint rate/FSL code appears to capture most of the information available (Fig. 7B). The coding of SNs varies with BF (Figs. 6, right column and 7D). At high BFs, the majority of neurons code SN by spike rate and, to a lesser extent, by timing and latency. Low BF neurons, however, have little SN information in spike rate and seem instead to represent SN by spike timing distributed throughout the duration of their response.

Temporal representation of SN cues

In a previous study (Chase and Young 2006), responses to the frozen noise SN/ILD stimuli were compared with responses to a random noise set filtered in the same manner. Under these conditions, it was found that although FSL variations were still evident, the extra information recovered by the SDM decoder disappeared. Thus the information carried in temporal patterns has to do with the reproducible nature of the temporal envelope. This fact limits the utility of the information identified by SDM in natural conditions where the stimulus envelope is unknown and not repeatable.

Nevertheless, there are situations in which the temporal coding of SN might prove useful, as for moving sound sources, for comparing the locations of two sources, or for dealing with self-motion of the head or pinnae. As an example, take a case where the task is to judge whether a sound source is moving laterally (such that the distance of the source does not change, and therefore there is no overall frequency blue or red shift). In this case, slight movement of the source will cause subtle differences in the spectrum of the sound reaching the ear due to pinna filtering. These differences will be translated into differences in the temporal responses of some neurons (examples are shown in Fig. 8, discussed in the following text). Whether the brain can use this information is a matter of speculation at present because it seems to require a serial comparison of responses to successive stimuli. However, those changes in the temporal arrangements of spikes do not have to be processed strictly as temporal patterns. Temporal patterns

![Image of diagrams showing MI for different codes as a function of BF. Left column: ITD. Middle column: ILD. Right column: SN. Top row: MI_RATE. Middle row: FSL gain (MI_JOINT - MI_RATE). Bottom row: SDM gain (MI_SD - MI_RATE)].
can be converted to rate changes, as occurs in responses to amplitude modulation between cochlear nucleus, where there is little rate coding of modulation frequency, and IC (where there is a rate code; Ioris et al. 2004). As a test of this possibility, one can predict that the just-noticeable difference in the location of a frozen noise sound source will be smaller than that of a random noise source.

The coding of SNs in low BF neurons is an interesting case. This information has little to do with sound localization because SN cues do not exist in the low-frequency region of cat head-related transfer functions (Musicant et al. 1990; Rice et al. 1992). The response here has more to do with the coding of temporal envelopes. Although these are static stimuli, cochlear filtering of these stimuli will produce temporal envelope modulations in auditory nerve fibers. As the SN is shifted across BF, the temporal modulations change position, and low BF neurons can follow this change. The fact that subtle changes in spectrum can result in large changes in temporal spike patterns has implications for speech processing (Bandyopadhyay and Young 2004).

Multiplexing of information within a single spike train

It should be stressed that the three coding mechanisms investigated here—rate, first-spike latency, and ongoing temporal pattern changes—can occur simultaneously in the same neuron. Figure 8 shows the responses of one neuron to both the ITD/ILD and SN/ILD stimulus sets. In Fig. 8, A and B, the spike raster responses of the neuron to the two stimulus sets are shown. The most striking difference between the responses to the two stimulus sets is the variability from stimulus to stimulus of the rasters in Fig. 8A compared with the regularity of the rasters in Fig. 8B: the locations of the spike clusters in this neuron change as a function of SN, but not ITD. Using the SDM method, the extra information uncovered by spike timing in this neuron is 0.84 bits for SN, compared with 0.00 bits for ITD and 0.21 bits for ILD (averaged across the two stimulus sets).

Although ILD does not cause the global spike cluster rearrangement of SN, there is a clear ILD-dependent latency change in Fig. 8, both A and B. The first-spike latency information is 0.54 bits for ILD, compared with 0.13 bits for ITD and 0.57 bits for SN. The rate responses of this neuron are shown in the surface plots of Fig. 8, C and D. It is clear that all three cue causes changes in mean firing rates, with the greatest extent of information about ITD: 0.4 bits, compared with 0.28 bits for SN and 0.28 bits for ILD.

The fact that rate, latency, and temporal spike patterns all contain information about different cues makes it possible for neurons to multiplex multiple aspects of stimuli in their spike trains. When only discharge rate codes are considered (Chase and Young 2005), there is substantial confounded information, meaning that the information about a full 25-stimulus set is less than the sum of the mutual information about the two stimulus cues that vary within the set, considered separately. Of course this is expected when one response variable is used to represent two independent aspects of the stimulus. When rate and temporal spike pattern (SDM) are considered simultaneously, it is possible to have full multiplexing of the two stimulus cues, with no confounded information (see Figs. 8 and 9 in Chase and Young 2006). Here we show that latency provides additional information that should further improve the multiplexing of localization cues on spike trains in the IC.

This type of information multiplexing has also been found in the visual system. Sakai et al. (1997) found that red and green inputs could be simultaneously encoded in ganglion cell responses. Gawne (2000) and Gawne et al. (1996) found that first-spike latency and spike rate differentially code for the...
contrast and orientation of visual line stimuli, respectively, whereas Victor and Purpura (1996) found that visual contrast is coded at a higher temporal resolution than visual texture. Meanwhile, Oram (2005) used a model of visual processing based on the differential coding of contrast and orientation to successfully predict reaction time distributions in visual recognition experiments. It is possible that the multiplexing of information in different codes within single neuron spike trains is a general mechanism of increasing information capacity in the brain.

**One versus two stimulus variables**

In this work we used stimuli that varied simultaneously in two parameters. The question arises of how this stimulus design affects the question considered in this paper: how multiple stimulus parameters can be represented by multiple aspects of a spike train. In particular, would it make a difference to our results if we had used stimuli that vary in only one parameter? Insight into this question can be gained using the idea of confounded information, introduced previously (Chase and Young 2005). Define the mutual information between a stimulus parameter X and the response R as $MI(X; R)$ as it was computed here by combining responses across all values of the second stimulus parameter Y. Mutual information could also be computed as $MI(X; R|Y = y)$, by varying only one stimulus parameter (X) and keeping the second one (Y) constant at some reasonable value y. Usually there is not a clear “best” choice for y, so this calculation could be averaged across several values of y, which yields $MI(X; R|Y)$. These two $MI$ values are different and are related as

$$MI(X; R|Y) = MI(X; R) + MI(X; Y|R)$$  \hspace{1cm} (5)

This equation follows from Eqs. A3 and A6 in Chase and Young (2005) and makes the assumption that $MI(X; Y) = 0$, as is the case here. The term $MI(X; Y|R)$ is the confounded information, mentioned earlier. Thus the $MI$ estimated by averaging across cases where only a single stimulus variable changes is larger than the $MI$ when several things are changing at once, which is perhaps not surprising. In real-world conditions, of course, many aspects of a stimulus change at the same time, so it seems reasonable to investigate stimulus coding in a way that allows these effects to be observed.

Note also that if the response $R$ is multidimensional, as discussed in the previous section, then the confounded infor-
information can be small. In this case the stimulus design matters less, i.e., the result will be the same if one stimulus parameter is changed versus two. For the results shown herein, this situation would hold to varying degrees for stimuli varying in ITD and either of the other two parameters, as long as temporal codes are considered. It might hold in cases for simultaneous variation of ILD or SN, but only if envelope locking could be used.

Implications for auditory processing

The differential coding of localization cues is useful only if downstream neurons are able to decode these cues independently. To compute the position of a sound source, information from the three cues must eventually be integrated together. However, the accuracy with which each cue can contribute to the overall location percept changes depending on sound conditions, such as the frequency content of the stimulus, temporal characteristics of the stimulus, or the reverberant qualities of the environment (Faller and Merimaa 2004; MacPherson and Middlebrooks 2002; Wightman and Kistler 1992). In the cat, there is the additional problem that the spatial maps of ITD, ILD, and SN can vary differently as the ears are moved (Young et al. 1996). Thus it is useful for the auditory system to preserve information about the cues separately so that the relative weight with which each cue contributes to the location percept can be changed depending on sound conditions.

In a recent study, Bizley et al. (2007) reported the existence of visually responsive neurons within auditory cortical areas. They found that many neurons responded to both auditory and visual cues, often with unique temporal response components. In a subset of neurons, they studied the visual spatial receptive fields of these cells and noted a significant modulation of this visual spatial receptive field during auditory stimulation. In light of these results, it is possible not only for information from multiple auditory spatial cues to be optimally weighted for the computation of sound source location, but for information from different sensory modalities to be combined, as well. Although this type of spatial location processing is largely speculative at this time, investigations of sensory integration for localization processing may prove fruitful.

Regardless of how the final localization percept is calculated, the results reported here support the hypothesis that the analysis of sound location information above the inferior colliculus must allow for the possibility of temporal codes and simultaneous rate and temporal codes (e.g., Furukawa and Middlebrooks 2002). The question of the means by which temporal codes such as latency could be used by neurons has been considered elsewhere (e.g., Chase and Young 2007; Stecker and Middlebrooks 2003), but remains a challenge for studies of neural coding.

REFERENCES


