Published in final edited form as:

J Neurophysiol. 2006 November ; 96(5): 2327–2341.

Phase Locking of Auditory-Nerve Fibers to the Envelopes of High-Frequency Sounds: Implications for Sound Localization

Anna Dreyer^{1,4} and Bertrand Delgutte^{1,2,3,4}

¹ Eaton Peabody Laboratory, Massachusetts Eye and Ear Infirmary, Boston

² Department of Otolaryngology, Harvard Medical School, Boston

³ Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts

⁴ Speech and Hearing Bioscience and Technology Program, Harvard–Massachusetts Institute of Technology Division of Health Sciences and Technology, Cambridge, Massachusetts

Abstract

Although listeners are sensitive to interaural time differences (ITDs) in the envelope of high-frequency sounds, both ITD discrimination performance and the extent of lateralization are poorer for high-frequency sinusoidally amplitude-modulated (SAM) tones than for low-frequency pure tones. Psychophysical studies have shown that ITD discrimination at high frequencies can be improved by using novel transposed-tone stimuli, formed by modulating a high-frequency carrier by a half-wave-rectified sinusoid. Transposed tones are designed to produce the same temporal discharge patterns in high-characteristic frequency (CF) neurons as occur in low-CF neurons for pure-tone stimuli. To directly test this hypothesis, we compared responses of auditory-nerve fibers in anesthetized cats to pure tones, SAM tones, and transposed tones. Phase locking was characterized using both the synchronization index and autocorrelograms. With both measures, phase locking was better for transposed tones than for SAM tones, consistent with the rationale for using transposed tones. However, phase locking to transposed tones and that to pure tones were comparable only when all three conditions were met: stimulus levels near thresholds, low modulation frequencies (<250 Hz), and low spontaneous discharge rates. In particular, phase locking to both SAM tones and transposed tones substantially degraded with increasing stimulus level, while remaining more stable for pure tones. These results suggest caution in assuming a close similarity between temporal patterns of peripheral activity produced by transposed tones and pure tones in both psychophysical studies and neurophysiological studies of central neurons.

INTRODUCTION

Interaural time differences (ITDs) serve two important functions in binaural hearing: localization of sound sources and detection of sounds in noisy environments. Although ITDs are the dominant cue for localizing sources containing low-frequency energy (Macpherson and Middlebrooks 2002; Wightman and Kistler 1992), the use of ITDs in localization tasks is not restricted to low frequencies. Although interaural time differences in the fine structure of high-frequency sounds cannot be detected, the auditory system can exploit the ITDs present in the *envelope* of high-frequency stimuli to lateralize sounds (Henning 1974; Leakey et al. 1958; McFadden and Pasanen 1976). However, the perceptual ability to use ITDs in high-frequency envelope cues seems weaker than that for low-frequency fine-structure cues. For instance, the extent of laterality is narrower and ITD discrimination thresholds are poorer for

sinusoidally amplitude-modulated (SAM) tones than for low-frequency pure tones of the same frequency as the modulator (Bernstein and Trahiotis 1985,1994).

Colburn and Esquissaud (1976) indicated that the relatively poor ITD resolution for high-frequency stimuli could be a result of either differences in the temporal properties of the spike trains in the auditory nerves or differences in how the CNS processes low- and high-frequency inputs, or a combination of the two. Van de Par and Kohlrausch (1997) hypothesized that differences in lateralization sensitivity between SAM and pure tones may be attributed to the difference in “the internal representation after transformation in the inner hair cells” for the two stimuli. To a first-order approximation, this transformation can be described as half-wave rectification followed by low-pass filtering. The low-pass filtering is attributed to both the inner hair cell (IHC) membrane time constant (Palmer and Russell 1986) and jitter in transmission at the synapses between the hair cell and primary neurons (Anderson et al. 1971). Half-wave rectification arises in part from the IHC receptor potential versus cilia displacement characteristic.

According to this simplified model for IHCs, a low-frequency pure tone is transformed into a half-wave-rectified sine wave (Fig. 1, *top*), whereas a high-frequency SAM tone is transformed into a sinusoid if the carrier frequency (f_c) is above the cutoff frequency of the low-pass filter (Fig. 1, *middle*). Therefore van de Par and Kohlrausch (1997) suggested a novel stimulus, called a “transposed” tone, designed to produce the same temporal discharge patterns in high-frequency auditory-nerve fibers (ANFs) as a pure tone does in low-frequency fibers. A sinusoidal carrier is multiplied by a half-wave-rectified low-frequency sinusoid, producing a high-frequency transposed tone with the envelope of the low-frequency half-wave-rectified sine wave (Fig. 1, *bottom left*). After low-pass filtering, the response to a transposed tone is expected to be a half-wave-rectified sinusoid resembling the response to a pure tone.

For modulation frequencies <150 Hz, high-frequency transposed tones produce extents of laterality and ITD sensitivity better than those produced by SAM tones and comparable to those produced by pure tones (Bernstein 2001; Bernstein and Trahiotis 2002). These findings suggest that, for low modulation rates, the poorer sensitivity to ITDs at high frequencies compared with low frequencies results at least in part from differences in the temporal discharge patterns produced in the auditory nerve by the two stimuli. However, as modulation frequency increases, lateralization performance in response to transposed tones degrades relative to performance for a comparable pure tone, suggesting that additional factors beyond peripheral representation may also play a role.

The only available neurophysiological study of responses to transposed tones (Griffin et al. 2005) focused on ITD-sensitive neurons in the inferior colliculus (IC), the principal auditory nucleus in the midbrain, using relatively low stimulus levels. Griffin et al. found that ITD sensitivity of IC neurons for transposed tones with low modulation frequencies was comparable to that for pure tones at low and moderate stimulus levels. IC neurons are thought to derive their ITD sensitivity by synaptic inputs from neurons in the lateral and medial superior olives (LSO and MSO), whose function depends critically on phase-locked inputs from the two ears. A quantitative description of phase locking to transposed tones at the earliest stages of processing in the auditory pathway is therefore essential for understanding the central neurophysiological results (Griffin et al. 2005) as well as the psychophysical results (Bernstein 2001; Bernstein and Trahiotis 2002). In this paper, we quantitatively compare the phase locking of ANFs to transposed tones, SAM tones, and pure-tones.

Although the idealized picture of peripheral auditory processing shown in Fig. 1 is a useful first approximation, several arguments suggest that the temporal discharge patterns of ANFs to transposed tones may differ from pure-tone response patterns in significant ways. First, the

relationship between IHC receptor potential and cilia displacement cannot be characterized as simple half-wave rectification because a fraction of the transduction channels are open when the cilia displacement is zero, resulting in sustained current flow into the hair cell and spontaneous release of neurotransmitter at the base of the hair cell. As a result, the instantaneous rate of discharge of ANFs in response to low-frequency pure tones falls below the spontaneous rate (SR) during approximately every other half cycle (Johnson 1980), contrary to the idealized model used for motivating transposed tones. Second, phase locking to SAM tones degrades at a higher sound pressure level (SPL) arising from cochlear compressive nonlinearities (Joris and Yin 1992; Smith and Brachman 1980). Given the amplitude-modulated nature of transposed tones, a similar level dependency of phase locking may be expected. Therefore this paper places particular emphasis on the level dependency of phase locking to transposed tones compared with pure tones. We find that, whereas phase locking to transposed tones with low modulation frequencies (<250 Hz) is similar to that to pure tones at sound levels near neural thresholds, this is not the case for higher levels and higher modulation frequencies. Preliminary reports of these findings were previously presented (Dreyer and Delgutte 2004; Dreyer et al. 2005).

METHODS

Procedure

Healthy adult cats were anesthetized with Dial in urethane (75 mg per kg of body weight) and maintained in an anesthetized state as judged by the toe-pinch reflex (Kiang et al. 1965). The auditory nerve was exposed by opening the posterior fossa and retracting the cerebellum. The tympanic bullae were opened and the round window was exposed. The cat was administered 4.5 ml per kg of lactated Ringer solution, with an additional 4.5 ml booster per day to prevent dehydration, and dexamethasone (0.26 mg/kg) as a prophylactic against brain edema throughout the experiment. Body temperature was maintained at 37°C with a heating blanket.

Single-unit recordings were conducted with the animal placed on a vibration-attenuating table in an electrically shielded, soundproof chamber. A metal electrode near the round window recorded the compound action potential (CAP) in response to click stimuli to assess the sensitivity and stability of the cat's cochlear response.

Stimuli were generated with a 16-bit A/D converter (NIDAC 6052e, National Instruments) using a sampling rate of 50 kHz. An electrodynamic speaker (Realistic 40-1377) delivered sound stimuli near the cat's tympanic membrane through a closed acoustic assembly. Stimuli were digitally compensated for the frequency response of the acoustic system by measuring the magnitude and phase of the sound pressure at the tympanic membrane as a function of frequency and designing digital inverse filters.

Single-unit activity was recorded with glass micropipettes filled with 2 M KCl. The electrode was inserted into the auditory nerve with the aid of a microscope and then advanced with a micropositioner (Kopf 650). The microelectrode signal was band-pass filtered and sent to a custom spike detector and timer that records spike arrival times with a precision of 1 μ s.

A click stimulus at about 55 dB SPL was used as search stimulus while advancing the electrode into the nerve. On locating a responsive fiber, a threshold tuning curve was measured with an automatic tracking procedure (Kiang and Moxon 1974) using 50-ms tone bursts, and the characteristic frequency (CF) and threshold at the CF were noted. The fiber's spontaneous discharge rate (SR) was measured over 20 s.

All procedures were approved by the MIT and MEEI Institutional Animal Care and Use Committees (IACUC).

Stimuli

Low-frequency pure tones and high-frequency SAM and transposed tones were synthesized digitally. Transposed tones were generated as described by van de Par and Kohlrausch (1997) and Bernstein and Trahiotis (2002). All SAM tones were 100% modulated and all transposed tones had half-wave-rectified sinusoidal envelopes. Pure-tone frequencies (f) and the modulation frequencies (f_m) of both SAM and transposed tones were 60, 125, 250, 500, or 1,000 Hz. The frequencies of these pure tones were always within a half-octave of the fiber CF. The carrier frequency (f_c) of SAM and transposed tones was usually set at the CF. In some fibers, f_c was also set slightly below or above CF, such that the threshold at f_c would be about 20 Db above the threshold at CF. These off-CF responses are important because they may convey much of the temporal information at higher sound levels when on-CF responses are saturated. On the average, f_c was 0.16 ± 0.05 ($\pm$ SD) octave above the CF and 0.25 ± 0.12 octave below the CF for these off-CF conditions.

Because only limited amounts of pure-tone data were collected, pure-tone responses previously collected by Tsai and Delgutte (1997) were also used and account for most of the pure-tone responses (250/283). In their experiments, the pure-tone frequency was chosen from a fixed grid of frequencies separated by half-octave steps and was typically within a quarter-octave of the CF, so that it always fell within the tip portion of the tuning curve. We include the Tsai and Delgutte data for frequencies <1 kHz for the majority of our analyses and include their data only for frequencies between 1 and 3 kHz for the analyses shown in Figs. 4 and 11.

Responses to pure tones were collected primarily from low-CF fibers, whereas responses to transposed and SAM tones were recorded from high-CF fibers (>3 kHz) to minimize phase locking to the fine structure. In our experiments, responses were measured as a function of stimulus level from rate threshold to 60 dB above threshold in 10-dB steps using 20 repetitions per level. Sound pressure levels did not exceed 85 dB SPL to avoid injuring the cochlea. For all stimuli, stimulus level was defined based on the root-mean-square amplitude of the acoustic pressure waveform. To facilitate comparisons, the high-frequency stimuli were presented in matched pairs consisting of a 480-ms SAM tone, followed by 520 ms of silence, a 480-ms transposed tone with the same modulation frequency, and another 520 ms of silence. Each presentation of a pure-tone stimulus consisted of a 480-ms tone followed by 520 ms of silence. All stimulus durations include 20-ms rise and fall times using raised-cosine ramps to prevent abrupt transients. In the experiments of Tsai and Delgutte (1997), the pure-tone stimuli were 100 ms in duration, including 2.5-ms (raised-cosine envelope) rise and fall times, and were presented at a rate of 2/s. Responses were measured over an 80-dB range of levels in 2-dB steps, with four repetitions at each SPL. Spikes were then grouped into coarser 10-dB-level bins to increase the number of spikes per level and facilitate comparisons with the data collected in our own experiments.

Analysis

Period histograms, locked either to the modulation frequency for SAM and transposed tones or to waveform frequency for pure tones, were computed from spike times, discarding the first 16 ms after stimulus onset. The strength of phase locking was characterized by the synchronization index (SI), the ratio of the fundamental frequency component of the period histogram to the average firing rate. Also known as *vector strength* (Goldberg and Brown 1969), the SI varies from 0 for a flat histogram with no phase locking to 1 for an impulsive histogram indicating perfect phase locking. Statistical significance of the synchronization index was assessed by the Rayleigh statistic (Mardia and Jupp 2000), using a criterion of $P < 0.001$, and values of SI failing this criterion were discarded.

To obtain an alternate description of temporal discharge patterns more relevant to binaural hearing, phase locking was also characterized using shuffled autocorrelograms (SACs) (Louage et al. 2004). Temporal analysis using SACs compares the timing of spikes between different presentations of the same stimulus. Whereas the synchronization index must be calculated at a predetermined frequency, computation of SACs does not require knowledge of the stimulus. In addition, because the SAC is a cross-correlation between pairs of spike trains, it is particularly relevant to ITD coding because many models of binaural processing are based on interaural correlation (Colburn 1977; Trahiotis and Stern 1995).

Each SAC was computed from the N (usually 20) stimulus presentations at each stimulus level. Spike trains from each stimulus presentation were paired, or shuffled, with the $N - 1$ others to create $N(N - 1)/2$ distinct pairs of spike trains. For each pair, the forward and backward time intervals between all spikes from the first spike train and the spikes from the second spike train were calculated. All intervals were tallied into a histogram using a $50\text{-}\mu\text{s}$ bin width as suggested by Louage et al. (2004). The histograms were then normalized by the square of the average firing rate r_0 , the bin width b , the recording duration (minus discarded times) D , and number of stimulus presentations N by dividing by $N(N - 1)r_0^2 bD$. The ordinate of the normalized SAC measures the extent to which a temporal structure related to the stimulus exists in the spike trains, with unity indicating randomly distributed spikes (Louage et al. 2004). For periodic stimuli, this temporal structure primarily reflects phase locking to the stimulus.

Two measures of phase locking were derived from normalized SACs (Louage et al. 2004): the maximum peak height and the peak half-width. The *peak height* measures the extent to which spikes occur at the same intervals on repeated presentations of the same stimulus relative to randomly distributed intervals. The *half-width* measures the temporal precision, or spread, of the intervals around the peak. Precise phase locking is indicated by a large peak height and/or a small half-width. To compute these two metrics, we first collapsed the SAC extending over a span of $\pm$ two stimulus cycles onto a folded histogram extending over one cycle.¹ This procedure is theoretically justified because the SAC for a periodic stimulus should be periodic in the limit of a large amount of data.

The SI and SAC analyses were applied to all responses to SAM tones, transposed tones, and pure tones. These metrics were then used to characterize how phase locking depends on sound pressure level, modulation frequency, and fiber characteristics.

RESULTS

The results are based on recordings from 182 auditory nerve fibers in nine cats. Responses to SAM and transposed tones were collected from 39 high-CF fibers in four cats (CF range: 2.9–13.5 kHz, median 5.7 kHz). These data include 119 level series for matched pairs (same f_c and f_m) of SAM and transposed tones. Responses to low-frequency ($f \leq 1,000$ Hz) pure-tones were recorded from 70 low-CF fibers from these four and five additional cats (CF range: 170–1,600 Hz, median 740 Hz). Overall, the data include 5.1% low-SR, 37.2% medium-SR, and 57.7% high-SR responses. Responses to pure tones $>1,000$ Hz were also recorded from an additional 73 higher-CF fibers (CF range: 1–3.3 kHz). These higher-frequency data are used only in Figs. 4 and 11.

¹This folding might appear to contradict our statement that computation of the SAC does not require knowledge of the stimulus period. Folding improved the signal-to-noise ratio of the SAC and was particularly useful for low discharge rates or when phase locking was weak, but it is not theoretically necessary to compute the metrics. Measuring the peak height and half-width does not require knowledge of the stimulus period when signal-to-noise ratio is not an issue.

In the following sections, we first compare phase locking to pure tones, SAM tones, and transposed tones presented at the fiber's CF using the synchronization index (SI) measure. Then, we compare phase locking for on- and off-CF presentations of SAM and transposed tones. Last, we examine alternate, autocorrelation-based metrics to quantify phase locking to each of the three stimuli.

Phase locking to pure, SAM, and transposed tones at CF

A set of period histograms measured as a function of stimulus level in response to a pure tone, a transposed tone, and a SAM tone (Fig. 2A) are shown in Fig. 2B. Responses to the pure tone are recorded from a low-CF fiber (385 Hz), whereas responses to the SAM and transposed tones are from a high-CF fiber (7,000 Hz). The modulation frequency of the transposed and SAM tones (250 Hz) roughly matches the frequency of the pure tone (354 Hz).

At low stimulus levels, responses to both SAM and transposed tones are restricted to approximately half of the stimulus cycle, indicating a high degree of phase locking to the envelope. As the level increases, spike discharges gradually spread to the other half-cycle, indicating a decrease in phase locking, consistent with previous reports for SAM tones (Cooper et al. 1993; Javel 1980; Joris and Yin 1992; Smith and Brachman 1980). However, for each level, responses to the transposed tone are restricted to a smaller part of the stimulus cycle than responses to the SAM tone. Although the firing rates in response to both SAM and transposed tones saturate at 50 dB SPL (Fig. 2C), phase locking reaches a maximum at lower levels, consistent with previous observations.

For the low-frequency pure tone, the fiber's response occupies virtually the same portion of the stimulus cycle from threshold (20 dB SPL) $\leq$ 60 dB above threshold (80 dB SPL), well beyond the point where the rate begins to saturate (Fig. 2C). Thus consistent with previous observations (Johnson 1980), this medium-SR fiber phase locks to the pure tone over a wider range of levels than the rate-based dynamic range.

These observations are reflected in the synchrony level functions derived from the period histograms (Fig. 2B). For the transposed and SAM tones, the SI is highest near threshold and falls dramatically with increasing stimulus level. For each stimulus level, SI is larger for the transposed tone than for the SAM tone. Synchrony to the pure tone remains fairly stable with level. Near threshold, the synchrony values are similar for the pure tone and the transposed tone.

Figure 3 shows synchrony-level functions for pure, transposed, and SAM tones for the entire data set. Responses to SAM and transposed tones with all modulation frequencies are included, as well as pure-tone responses for all frequencies $< 1,000$ Hz. Sound level is expressed in decibels with respect to rate threshold to facilitate comparison between fibers. For the most part, responses are well separated based on the stimulus type, even though data from different frequencies and modulation frequencies are pooled: low-frequency fibers phase lock with the greatest precision to pure tones, then high-frequency fibers to transposed tones, followed by high-frequency fibers for SAM tones. Near threshold, fibers tend to show similarly high phase locking to transposed tones and pure tones, although synchrony falls more rapidly with level for transposed tones than for pure tones. Synchrony to SAM tones is poorer near threshold and degrades faster with level than for either transposed or pure tones.

With a few exceptions, the sound pressure level of the maximum synchrony occurred between rate threshold and rate saturation for all stimuli, with an average ($\pm$ SD) of 15.9 ± 10.3 , 5.2 ± 8.7 , and 2.5 ± 7.1 dB above threshold for pure, transposed, and SAM tones, respectively. Our average for SAM tones is lower than the value of 10 ± 5.0 dB with respect to threshold given by Joris and Yin (1992), although the rather coarse (10-dB) spacing between adjacent stimulus

levels in our experiments makes a precise assessment difficult. Further, the difference could be partially accounted for by clarifying how sound pressure level of the SAM tone is defined: root-mean-square of the waveform versus the peak of the waveform envelope, as well as possibility of different definitions of rate threshold. Consistent with previous observations for SAM tones (Joris and Yin 1992), the SI to SAM and transposed tones decreased monotonically above the level of maximum synchrony and sometimes leveled out at higher levels.

To quantify the fall of synchrony with level for each stimulus, a regression line was fit to the decreasing portion of each SI-level function above the maximum. Only responses containing at least three statistically significant ($P < 0.001$) SI values based on the Rayleigh statistic were used in this analysis. The upward sloping portions of the SI level functions were not fit because they typically contained fewer than three data points. The mean downward slopes ($\pm$ SD) were -0.0036 ± 0.0043 /dB for pure tones, -0.0064 ± 0.0028 /dB for transposed tones, and -0.0103 ± 0.0026 /dB for SAM tones. Our slope estimates for SAM tones are similar to those (-0.012 ± 0.003 /dB) given by Joris and Yin (1992). The thick lines in Fig. 3 show the mean slopes for the three stimulus types. The slope relating phase locking to level is almost two times shallower for pure tones than for transposed tones. In addition, phase locking to transposed tones is always greater than phase locking to SAM tones for both the average line and the majority of the data. ANOVA revealed a highly significant effect of stimulus type on the slopes [$F(2,202) = 71.5$, $P < 0.001$]. Post hoc multiple comparisons with Bonferroni adjustment indicates that slopes for each stimulus are significantly different from slopes for the other two stimuli ($P < 0.05$).

Although SI-level curves for the three stimuli are fairly well separated (Fig. 3), synchrony to SAM and transposed tones is expected to vary with modulation frequency. As modulation frequency increases, the sidebands are increasingly attenuated by the cochlear filter centered at the CF, thereby decreasing the modulation depth of the mechanical input to the hair cells. In addition, there appears to be a temporal limitation in the ability of AN fibers to phase lock to high modulation frequencies (Greenwood and Joris 1996; Joris and Yin 1992). We therefore examined the dependency of phase locking on modulation frequency for each stimulus type (Fig. 4A). Following Johnson (1980), we use the maximum SI ($SI_{\max}$) over all stimulus levels as a measure for comparison. For low modulation frequencies ($f_m \leq 250$ Hz) of the transposed tone, $SI_{\max}$ is comparable to that for low-frequency pure tones at 250 Hz. However, synchrony to transposed tones begins to degrade for modulation frequencies between 250 and 500 Hz, whereas phase locking to pure tones does not start to degrade until 1,000 Hz. $SI_{\max}$ is always lower for SAM tones than for transposed tones, although the rate of decrease at higher modulation frequencies is similar for both stimuli.

Our finding that $SI_{\max}$ is greater for pure tones than for transposed tones >250 Hz might be confounded by the fact the tone-burst stimuli used by Tsai and Delgutte (1997) (on which most of our pure-tone data are based) had shorter durations and higher repetition rates than those of our SAM and transposed tones. However, this is unlikely because a two-way ANOVA comparing $SI_{\max}$ for short tone bursts (100 ms; Tsai and Delgutte, 1997) and long-duration tones (15–30 s; Johnson 1980) revealed neither a significant main effect of duration [$F(1,442) = 0.19$, $P = 0.66$] nor a significant interaction between duration and frequency [$F(6,442) = 1.38$; $P = 0.22$].

Both the maximum synchrony and the slope of the synchrony-level function above the maximum may depend on modulation frequency. SI-level functions for transposed and SAM tones are shown in Fig. 5 separately for each modulation frequency. Data at $f_m = 1,000$ Hz are not shown because few fibers were studied at this modulation frequency. Figure 5 suggests that the shapes of SI-level functions depend on both the stimulus type and the modulation frequency and that the two variables interact. For low modulation frequencies (≤ 125 Hz), synchrony falls at a faster rate with level for SAM tones than for transposed tones, as was the

case for the pooled data in Fig. 3. However, as modulation frequency increases to 250 Hz and, especially, 500 Hz, the synchrony-level functions for the two stimuli become more parallel. Although the slopes at 500 Hz are similar, the two set of curves remain well separated. These results are confirmed in Fig. 4B, which shows the mean slopes of regression lines fit to the SI-level functions as a function of modulation frequency for all three stimuli. Excluding the limited data at 1,000 Hz, the mean slope for SAM tones tends to increase with f_m , whereas the mean slope for transposed tones decreases with f_m , so that the two curves converge at 500 Hz. A two-way ANOVA of slopes with stimulus type (SAM and transposed tones) and f_m as factors indicates that their interaction is highly significant [$F(4,128) = 10.87$, $P < 0.001$]. The mean slopes for transposed tones are always more negative than those for their pure-tone counterparts, except again for $f_m = 1,000$ Hz, where few data points are available.

For both pure and SAM tones, it has been reported that low-SR fibers tend to phase lock with greater precision than medium-SR fibers, which, in turn, phase lock with more precision than high-SR fibers (pure tones: Johnson 1980; SAM tones: Joris and Yin 1992; Joris et al. 1994). Spontaneous activity adds a DC offset to the period histogram by contributing spikes that are not locked to the stimulus and is thus expected to decrease synchrony. To examine whether phase locking to transposed tones also depends on SR, Fig. 6A shows $SI_{\max}$ as a function of f_m separated according to two SR groups (Lieberman 1978): high SR ($>18/s$) and low/medium SR ($<18/s$). We did not have data from enough low-SR fibers to analyze them separately. For all three stimuli, there is a clear trend for responses from low/medium-SR fibers to have greater $SI_{\max}$ than responses from high-SR fibers. This observation is confirmed by two-way ANOVAs, with f_m and SR group as factors, which reveal highly significant main effects of SR group on $SI_{\max}$ for all three stimuli [pure tones: $F(1,117) = 32.4$, $P < 0.001$; transposed tones: $F(1,58) = 37.9$, $P < 0.001$; SAM tones: $F(1,58) = 31.7$, $P < 0.001$].

Whereas the instantaneous firing rate in response to a low-frequency pure tone falls below the SR during one half of each stimulus cycle (Johnson 1980), this is not expected to be the case for high-frequency modulated sounds such as SAM and transposed tones, neglecting possible effects of adaptation (see DISCUSSION). Responses to transposed tones and pure tones would therefore be expected to become less similar with higher spontaneous activity. To further analyze the interaction between SR and stimulus type, we ran a two-way ANOVA on $SI_{\max}$ with SR group and stimulus type (pure or transposed tone) as factors. To minimize the confounding effect of frequency, the analysis included pure-tone data only for frequencies <800 Hz and transposed-tone data for $f_m \leq 250$ Hz. As expected, the analysis showed significant main effects of SR group and stimulus type (both $P < 0.001$). Post hoc multiple comparisons using Bonferroni corrections show that $SI_{\max}$ to pure tones is significantly greater than $SI_{\max}$ to transposed tones for the high-SR group ($P < 0.001$), but not for the low/medium-SR group ($P > 0.05$). Thus temporal discharge patterns to transposed tones are similar to pure-tone responses only for low/medium-SR fibers, and only at low stimulus levels and low modulation frequencies where the phase locking to transposed tones is maximum.

Off-CF responses to SAM and transposed tones

Thus far, we have described phase locking of ANF responses to SAM and transposed tones when the carrier frequency was at the fiber's CF. However, these stimuli will excite fibers over a range of CFs, particularly at higher stimulus levels where the cochlear excitation patterns become broad. If most of the fibers with CFs near the carrier frequency are saturated, so that phase locking to the envelope becomes weak, temporal information about the modulation frequency may be conveyed primarily by fibers with CFs far from the frequency band where the stimulus has the most energy, i.e., far from the carrier frequency. Therefore it is important to characterize phase locking when f_c is presented off the CF.

In general, responses to SAM and transposed tones with carrier frequencies both above and below CF broadly resemble on-CF responses. However, some noticeable differences do exist. Figure 7A shows response patterns of a fiber as a function of level to transposed tones with carrier frequencies at the CF (6,920 Hz), below the CF (5,603 Hz), and above the CF (7,490 Hz). For all three carrier frequencies, the period histograms show the same trend in that the fraction of the stimulus cycle occupied by spike discharges increases with stimulus level, implying a degradation in phase locking. Average discharge rates for this low-SR fiber do not saturate completely, even at high levels (Fig. 7C). As expected, phase locking begins to deteriorate at lower levels for the on-CF stimulus than for the off-CF stimuli because the threshold is lower when f_c is at the CF. The shapes of the period histograms for stimuli below CF resemble those at the CF. However, the period histograms for stimuli above CF clearly show two modes within each stimulus cycle at lower stimulus levels, unlike the on-CF and below-CF histograms. Such bimodal period histograms may result from the sharp transients occurring twice in each stimulus cycle at times when the envelope of the transposed tone shows a slope discontinuity (Fig. 1), which would be expected to induce ringing in the cochlear filters. Despite these complex histogram shapes, Fig. 7B shows that the synchronization index decreases monotonically with the level beyond a maximum for all three carrier frequencies, and the slopes are similar. However, the maximum synchrony is somewhat greater on the CF than either above or below the CF.

Period histograms showing two or more peaks per stimulus cycle were automatically identified from the presence of multiple zero crossings in the numerical derivative of the autocorrelation of the period histogram, which is considerably smoother than the raw period histogram. Multimodal period histograms were seen in response to transposed tones in six of 22 fibers (27%) for carrier frequencies below the CF and in five of 27 fibers (19%) for carrier frequencies above the CF, but were never observed at the CF. Multiple peaks in responses to SAM tones were less prevalent and occurred in only one of 22 fibers (5%) for carrier frequencies below the CF and in four of 27 fibers (15%) for carriers above the CF. These observations are consistent with oscillations seen at the onset and, sometimes, offsets in response to tone bursts with abrupt onsets (Kiang et al. 1979; Rhode and Smith 1985). Such oscillations, which are observed only when the stimulus is presented in the steep gradients of the tuning curve above and below the CF, have been attributed to the ringing of the cochlear filters. Although ringing of the cochlear filters is characterized by oscillations at the characteristic frequency in low-CF fibers, only the envelope of the ringing can be seen in the present study because the CFs of the neurons studied with SAM and transposed tones were always above the limit of phase locking. According to this interpretation, multimodal discharge patterns are more prevalent for transposed tones than for SAM tones because the transposed tones have more abrupt transients.

Figure 8A shows synchrony-level functions for SAM and transposed tones with carrier frequencies at, above, and below the CF for the entire data set. To facilitate comparisons, stimulus level is expressed relative to the level of maximum synchrony at each carrier frequency. For all three ranges of carrier frequencies, the SI-level curves generally decrease monotonically for levels above the maximum and the slopes are similar. Precise trends, however, are hard to identify in Fig. 8A because of the large variability between fibers. To quantitatively compare the level dependency of SI on and off CF, for each fiber, we computed the difference between each off-CF SI-level curve and the corresponding on-CF curve. The resulting change in SI (Δ SI) level functions (Fig. 8B) are generally flat or just slightly increasing for both SAM and transposed tones. Flat Δ SI-level functions indicate that on- and off-CF SI-level functions have the same slopes, whereas sloping Δ SI-level functions indicate that SI decreases with level at different rates on and off the CF.

These effects were quantitatively analyzed by fitting a regression line to the entire set of Δ SI-level functions separately for each stimulus type (SAM and transposed) and each carrier

frequency range (above and below the CF). The slopes and intercepts of the four best-fitting lines are given in Table 1. Because the ΔSI -level functions are plotted with respect to the level maximum, the intercept gives an estimate of the difference in $SI_{\max}$ off and on CF. Statistical analyses (*t*-test) show that the intercepts for SAM tones above CF and transposed tones below CF were significantly negative ($P < 0.05$), indicating a lower $SI_{\max}$ for these two conditions compared with when the carrier is at the CF. On the other hand, the intercepts did not significantly differ from zero for SAM tones below CF and for transposed tones above CF, indicating similar $SI_{\max}$ on and off CF in these conditions. The slopes of the best-fitting lines to the ΔSI -level functions did not, in general, differ from zero (Table 1), indicating that the rates of decrease of synchrony with level are similar on and off CF. The one exception is for SAM tones above the CF, where the slope was significantly positive, indicating that synchrony decreases more rapidly with level for SAM tones above CF than for SAM tones at CF. Despite these statistically significant differences, the overall effects are small and the general pattern of results suggests that the dependency of synchrony on level does not greatly differ on and off CF for either stimulus.

Phase-locking analysis with autocorrelation

Although the synchronization index is an appropriate measure of phase locking to a periodic stimulus when the period histogram is unimodal (as was largely the case in our data), it is a poor model of computations performed by the CNS because it necessitates a priori knowledge of the stimulus period. Recently, Louage et al. (2004) proposed alternative, autocorrelation-based measures of phase locking that capture the intrinsic periodicities in the spike responses without requiring prior knowledge of the stimulus. These measures make particular sense in the present context because interaural cross-correlation is a key component of most models for ITD processing. This section compares phase locking, as measured by SI to autocorrelation-based measures, and examines the extent to which these phase-locking metrics are correlated and can be used interchangeably.

Figure 9 shows the shuffled autocorrelogram (SACs; see METHODS) as a function of stimulus level for the same data that were analyzed with period histograms in Fig. 2. For all three stimuli, the SACs show multiple modes separated by the stimulus period, indicating phase locking to the stimulus. In general, each cycle of the SAC is more symmetric around its maximum than the corresponding period histogram in Fig. 2. For SAM and transposed tones, the increasing width of each SAC mode as level increases parallels the spread of the period histograms to an increasingly large fraction of the stimulus cycle. SACs for the pure tone do not broaden as appreciably with level as SACs for SAM and transposed tones. To quantify phase locking, we first use the height of the largest mode in the normalized SAC, called “peak height” in short (Louage et al. 2004). The peak height varies from 1 for randomly distributed spikes (no phase locking) to arbitrarily large values if phase locking is very precise. Peak-height versus level functions (Fig. 9B) show similarly strong phase locking to the pure tone and the transposed tone at low levels, but a faster decrease in peak height above the maximum for the transposed tone than for the pure tone. In addition, for each level, the peak height for the SAM tone is always lower than that for the transposed tone. These trends parallel those seen with the synchronization index measure in Fig. 2B. Phase locking was also quantified by the “half-width,” the width of the SAC at 50% of the peak height (Louage et al. 2004). The half-width is normalized by the stimulus period to facilitate comparison of responses to stimuli with different frequencies. The normalized half-width varies in the opposite direction of the peak height, from 0 for ideally precise phase locking to 1 for randomly distributed spikes. Half-widths are similar for the pure tone and the transposed tone at low levels in this example (Fig. 9C). At higher levels, half-widths for the SAM tone and the transposed tone increase considerably, reaching 1 for the SAM tone, but stabilize near 0.3 for the pure tone. These trends

parallel the observations made with the peak height and the synchronization index. Thus for these examples, the three measures of phase locking are highly correlated.

Figure 10 shows a scatterplot of synchronization index against SAC peak height for the entire data set. Responses to SAM and transposed tones at all modulation frequencies are included, as well as pure-tone responses for frequencies <1 kHz. Despite including data from different frequencies, there is a clear, monotonic relationship between SAC peak height and synchronization index for each stimulus type. The data were fit using a hyperbolic function

$$PH = 1 + \frac{aSI}{1 - SI} \quad (1)$$

where SI is the synchronization index, PH is the SAC peak height, and a is the only free parameter. This equation has the simplest form that gives both $PH = 1$ when $SI = 0$, and $PH = \infty$ when $SI = 1$. Three different versions of the model were tested, in which either the same curve (i.e., the same value of the parameter a) was fit to all the data, or a curve (a value of a) was fit separately for each stimulus type, or the SAM and transposed tone data were grouped together, while keeping the pure-tone responses separate. The model that minimized the variance of the residuals while using the smallest number of parameters was the last one, in which the transposed and SAM data share the same value of a , but not the pure tone data. This model performed significantly better than the single-curve model [$F(1,1088) = 976.0$, $P < 0.001$], while not performing significantly worse than the separate-curves model [$F(1,1087) = 3.079$, $P < 0.080$]. This best model (thick lines in Fig. 10) accounted for >98% of the variance in the data, confirming the tight relationship between peak height and synchronization index. The best-fitting parameter a was 0.316 for pure tones and 0.460 for SAM and transposed tones. Because the relationship between peak height and synchronization index is the same for SAM tones and transposed tones, all conclusions based on the SI remain valid if we use the peak height instead for these two stimuli.

To verify that similar trends hold whether the SAC peak height or the SI is used as the phase-locking metric, Fig. 11 shows the level-maximum peak height ($PH_{\max}$) as a function of frequency or f_m , which can be compared with $SI_{\max}$ in Fig. 4A. We do not show slope data comparable to Fig. 4B for the peak height because the level dependency of peak height could not be fit by a straight line. As observed with the synchronization index metric (Fig. 4A), $PH_{\max}$ for transposed tones at modulation frequencies <250 Hz is similar to pure-tone $PH_{\max}$ at 250 Hz and is higher than $PH_{\max}$ for SAM tones. However, the two metrics differ in that $PH_{\max}$ for pure tones continues to vary considerably with frequency <1,000 Hz, whereas the SI is stable in that frequency range (see also Louage et al. 2004). The $PH_{\max}$ curve for pure tones also shows an unexpected dip at 1,000 Hz. In addition, the $PH_{\max}$ curves for SAM and transposed tones are less parallel to each other than the corresponding $SI_{\max}$ curves. In general, the $PH_{\max}$ measure tends to amplify the differences between stimulus conditions relative to SI when phase locking is strong. These differences are likely to reflect the compressive nature of the SI , which cannot exceed unity, whereas peak height can in principle reach arbitrarily large values. Despite these differences, the general trends of the effect of frequency on maximum phase locking across stimulus types hold for both metrics.

Although the SAC peak height measures the extent to which spikes occur in a certain temporal relationship with respect to each other across stimulus presentations, the half-width measures the temporal precision with which spikes occur at these preferred intervals. For periodic stimuli, the two metrics are expected to be negatively correlated because precise phase locking implies both a large peak height and a narrow half-width. However, the two measures might not be perfectly correlated if the relation between the two depended on the stimulus wave shape. Figure 12 shows a scatterplot of peak height against normalized half-width for the same data

as in Fig. 10. The figure shows a clear, monotonically decreasing relationship between the two metrics. The data were fit by a hyperbolic function

$$HW = \frac{1}{b(PH - 1) + c} \quad (2)$$

where HW is the normalized peak half-width and b and c are free parameters.² As in Fig. 10, different versions of the model were tested by grouping the responses to the three stimuli in various ways. We chose the model in which the responses to all three stimuli are grouped together (i.e., the values of b and c are constrained to be the same for all three stimuli). The best-fitting parameter values for this model were $b = 1.369$ and $c = 0.857$. Although the model with three separate curves fit the data significantly better than the single-curve model [$F(4,1085) = 490.5$, $P < 0.001$], we prefer the simpler single-curve model because the fraction of the variance accounted was only slightly greater for the three-curve model than for the single-curve model (93.62% vs. 93.30%) and the parameter estimates were essentially identical for all three curves in the more complex model. Thus the three phase-locking measures (synchronization index, peak height, and peak width) seem to be essentially equivalent for SAM and transposed tones, and the relationship between peak height and half-width is the same regardless of the stimulus. On the other hand, the relationship between peak height and synchronization index differs somewhat for pure tones compared with the other two stimuli. However, the differences in this relationship are fairly small and do not seem to substantially alter conclusions based on the traditional SI measure (Figs. 10 and 11).

DISCUSSION

We have investigated the neural mechanisms underlying ITD-based lateralization at high frequencies by characterizing the precision of phase locking for high-frequency SAM and transposed tones and for low-frequency pure tones at the level of the auditory nerve using both the traditional synchronization index and autocorrelation-based measures. Our results show similar precision of phase locking to pure tones and transposed tones only for a restricted set of conditions: stimulus levels close to thresholds, low modulation frequencies, and in low/medium-SR fibers. Phase locking to transposed and SAM tones fell faster with increasing stimulus level than phase locking to pure tones, although phase locking to transposed tones always exceeded phase locking to SAM tones. Maximum synchrony to transposed tones was similar to pure-tone maximum synchrony for modulation frequencies <250 Hz in low/medium-SR fibers, but fell faster than maximum synchrony to pure tones at higher modulation frequencies. The low-pass dependency of maximum synchrony on modulation frequency was similar for transposed and SAM tones. Responses to SAM and transposed tones with carrier frequencies below and above the CF generally showed a similar synchrony-level dependency as on-CF responses when the level was expressed relative to threshold at each frequency.

Comparison of phase locking to pure and transposed tones

Transposed tones are intended to produce similar temporal patterns of discharge in high-CF AN fibers as occur in low-CF fibers in response to low-frequency pure tones. Our results show that phase locking to transposed tones resembles phase locking to pure tones in only a restricted set of conditions. The rationale for stimulus transposition is based on a simple model of mechano-electrical transduction in hair cells consisting of ideal half-wave rectification followed by low-pass filtering (van de Par and Kohlrausch 1997). Although this model is a useful first approximation, there are several differences in the way pure tones and transposed tones are

²Because the half width is normalized to the stimulus period, the parameter c in Eq. 2 should in principle always be 1 so that the normalized half-width is 1 when the peak height is zero (no phase locking). In practice, the estimates of the half-width are inaccurate when the peak height is near 1 and a better fit is obtained by letting c freely vary than by holding it at 1.

processed in the cochlea, which lead to different auditory nerve discharge patterns. These differences involve 1) mechano-electrical transduction, 2) cochlear compressions, and 3) cochlear filtering.

For pure tones, phase locking in the auditory nerve ultimately derives from the morphological polarization of the IHC ciliary bundle and resultant asymmetrical modulation of the hair-cell receptor potential (Hudspeth and Corey 1977). Because some of the transducer channels are open at rest (Ashmore 1991; Fettiplace and Fuchs 1999), a sinusoidal displacement of the cilia causes a modulation of the IHC receptor potential around rest, which in turn results in modulation in the release of neurotransmitter by the synapses at the base of the hair cells, and ultimately modulation of the instantaneous firing rate of the afferent neurons around SR in spontaneously active fibers (Johnson 1980). Because the fraction of transducer channels open at rest is $\ll 50\%$, the receptor potentials are asymmetrical (depolarizations are larger than hyperpolarizations) for all but the lowest stimulus levels, and modulations of the instantaneous firing rate are correspondingly asymmetric. The situation is different for transposed tones because there is no hyperpolarization of the IHC membrane potential and no resulting decrease in the instantaneous firing rate below SR during the half-cycle when the stimulus has zero amplitude (Fig. 1) if possible adaptation effects are neglected (see following text). This lack of negative modulation of the instantaneous rate should result in a lower synchronization index for transposed tones compared with pure tones in spontaneously active fibers, but not in fibers with low spontaneous activity. Consistent with this reasoning, we found that maximum synchrony to pure tones is significantly greater than maximum synchrony to transposed tones in high-SR fibers, but not in low/medium-SR fibers. However, these differences between SR groups are pronounced only at low stimulus levels.

A second, more important, factor underlying differences in responses to pure and transposed tones is the strong compression manifested primarily in the saturation of rate-level functions. Despite rate saturation, phase locking to pure tones remains fairly stable over a wide range of stimulus levels (Fig. 3; Johnson 1980). In contrast, phase locking to the envelope of SAM tones reaches a maximum only a few decibels above threshold and then falls steeply with increasing level (Cooper et al. 1993; Javel 1980; Joris and Yin 1992; Smith and Brachman 1980; Yates 1987). Similar degradations in phase locking have been observed for nonsinusoidal envelope modulations (Delgutte 1980; Wang and Sachs 1993). Our results show that responses to transposed tones behave similarly to responses to SAM tones in this respect, although the fall in synchrony with increasing level is not as steep for transposed tones as for SAM tones, particularly for low modulation frequencies (Figs. 3 and 5). At higher modulation frequencies, the level dependency of synchrony becomes more similar for SAM and transposed tones (Figs. 5 and 4B), consistent with the idea that the cochlear filters increasingly attenuate the additional side bands of the transposed tones, effectively making the two stimuli more alike.

The degradation in phase locking to the envelope of SAM tones with increasing stimulus level is qualitatively consistent with the saturating character of rate-level functions because a given modulation in stimulus amplitude creates a smaller modulation of the firing rate when the mean intensity is in the saturated part of the rate-level function. At first sight, this explanation does not hold for transposed tones because the output of any instantaneous compression would remain zero regardless of stimulus level during the entire envelope half-cycle for which the stimulus amplitude is zero (the "off period"). However, the slow decay time constant of the cochlear filter (and other filters arising in the hair cell membrane and synaptic transmission) prolongs the response to transposed tones into the stimulus off period. Figure 2 clearly shows this encroachment of responses to transposed tones into the stimulus off period at high stimulus levels for a modulation frequency of 250 Hz. The duration of this encroachment at the highest level (>1 ms) is consistent with the duration of click responses for fibers with similar CFs (Kiang et al. 1965). This prolongation of the responses to each cycle of a transposed tone should

be most significant at higher modulation frequencies, where they occupy an increasingly large fraction of the modulation cycle. Consistent with this prediction, the slopes of synchrony-level functions decrease with increasing f_m and, for very low f_m (60 Hz), the degradation in phase locking with level for transposed tones is comparable to that for low-frequency pure tones (Fig. 4B).

Our discussion so far has neglected possible effects of adaptation. It has been argued that rapid adaptation generally enhances the coding of dynamic stimuli and, in particular, accounts for the observation that the maximum synchrony to the envelope of SAM tones is greater, and occurs at higher stimulus levels, than expected based on the static rate-level function (Cooper et al. 1993; Smith and Brachman 1980; Yates 1987). Although we did not quantitatively compare phase locking to transposed tones with predictions based on static rate-level functions, transposed tones are likely to behave similarly to SAM tones in this respect. Adaptation could also decrease the instantaneous firing rate below the SR during the stimulus off period and thereby enhance the coding of SAM and transposed tones in spontaneously active fibers. Although these effects of adaptation are likely to occur, they are apparently too weak to counteract the larger effects of cochlear tuning and compression on responses to transposed tones, with the net result that phase locking to transposed tones generally falls below synchrony to pure tones.

Cochlear frequency selectivity is likely to be responsible for the result that synchrony to SAM and transposed tones falls faster with increasing modulation frequency than synchrony to pure tones falls with frequency (Fig. 4). As modulation frequency increases, the side bands of SAM and transposed tones are increasingly attenuated by the cochlear filters, so the effective modulation in the mechanical drive to the hair cells also decreases. Joris and Yin (1992) argued that the upper frequency limit of phase locking to SAM tones is determined by cochlear filtering for CFs ≤ 10 –15 kHz, and by a temporal limitation at higher CFs. Because most (37 of 39 fibers, 66 of 69 responses) of our transposed tone data were from fibers with CFs < 10 kHz, and because the upper cutoff of phase locking is similar for transposed tones and SAM tones (Fig. 4A), phase locking to transposed tones is likely to be largely limited by cochlear filtering in our data. In contrast, phase locking to pure tones is limited by an inability of synapses to follow fast changes in the hair-cell receptor potential (Weiss and Rose 1988). The different mechanisms for pure and transposed tones partly explain why the upper limit of phase locking is lower for transposed tones than for pure tones (Fig. 4). This difference might be even more pronounced in humans if, as some data suggest (Shera et al. 2002; but see Ruggero and Temchin 2005), cochlear frequency resolution is significantly sharper in humans than in cats.

In short, various cochlear mechanisms including frequency selectivity, hair cell transduction, and synaptic transmission all contribute to the temporal discharge patterns in the auditory nerve for transposed tones differing from those for pure tones in all but a restricted set of conditions. Consequently, psychophysical and central physiological studies should be cautious about assuming a strict similarity in the peripheral representations of the two stimuli.

Relation to studies of ITD coding in central auditory neurons

The only single-unit study of auditory neurons that used transposed tones is that of Griffin et al. (2005), which compared ITD sensitivity to SAM and transposed tones for high-frequency neurons in the inferior colliculus (IC) of anesthetized guinea pigs. Griffin et al. found a greater proportion of neurons sensitive to ITD for transposed tones than for SAM tones and, among sensitive neurons, a greater modulation of firing rate with ITD for transposed tones. In addition, neural ITD just noticeable differences (JNDs) estimated using signal detection theory (ROC analysis) were smaller for transposed tones than for SAM tones at the same modulation frequency. At low modulation frequencies (< 280 Hz), the best neural ITD JNDs for transposed tones were comparable to the JNDs measured by Shackleton et al. (2003) for their pure-tone

counterparts. However, ITD sensitivity was too poor to allow estimation of neural JNDs for either SAM or transposed tones with $f_m > 280$ Hz, whereas pure-tone neural ITD JNDs improve with increasing frequency ≤ 500 Hz (Shackleton et al. 2003). The ITD sensitivity of the IC neurons studied by Griffin et al. is thought to be largely inherited from that of neurons in the medial and lateral superior olives (MSO and LSO) that, in turn, ultimately depends on phase locking in the afferent inputs from both ears. Because Griffin et al. used low stimulus levels (about 15 dB above pure-tone thresholds at the carrier frequency), their finding of comparable ITD sensitivity for pure tones and transposed tones at low frequencies is consistent with our observation that maximum synchrony in AN fibers is similar for the two stimuli at low frequencies (Fig. 4). However, our finding that synchrony to transposed tones falls much faster with increasing level than synchrony to pure tones raises the question of how the ITD sensitivity of IC neurons to the two stimuli would compare at higher stimulus levels.

There is scant information on the level dependency of ITD sensitivity in IC neurons to either low-frequency pure tones or high-frequency SAM tones. The few studies that even address the question (Batra et al. 1993; Fitzpatrick et al. 2005; Kuwada and Yin 1983; Yin and Kuwada 1983) are more concerned with shifts in the best ITD with level rather than with changes in sharpness of ITD tuning, which are more directly relevant to a characterization of ITD sensitivity in terms of neural JNDs. A majority of the examples of pure-tone ITD curves found in the literature (Fig. 9A in Kuwada and Yin 1983; Fig. 10B in Yin and Kuwada 1983; Fig. 10 in Kuwada et al. 1984) show no major changes in sharpness of tuning over up to a 40-dB level range, although one unit (Fig. 12 in Kuwada et al. 1984) does show substantial level-dependent changes in the shapes of ITD curves measured with binaural beats. An unpublished study (DC Fitzpatrick, personal communication) reports a mean increase of $3 \mu\text{s/dB}$ in the half-widths of composite ITD curves obtained by summing pure-tone ITD curves over a wide range of frequencies. This moderate degradation in ITD tuning appears to be related to a lowering of the frequency, which contributes most to the composite ITD curve, and does not necessarily occur in response to individual tones. Thus it is hard to confidently relate the degradation in composite ITD tuning to the small decrease in synchronization index with level we observed in AN fibers for pure tones (Fig. 3). Based on our results, the degradation in ITD tuning with level would be expected to be more substantial for SAM and transposed tones than for pure tones. Unfortunately, we are not aware of any report on the effect of level on ITD tuning widths for amplitude-modulated (AM) tones. Further studies are needed to systematically characterize the effects of stimulus level on ITD sensitivity for both low-frequency pure tones and, especially, high-frequency AM stimuli, and to determine whether the similarity in ITD tuning between pure tones and transposed tones observed at low stimulus levels by Griffin et al. (2005) also holds at higher levels.

Relation to psychophysical studies using transposed tones

The present study was motivated by the finding that ITD JNDs for transposed tones in human listeners are lower than JNDs for SAM tones and, at low modulation frequencies (≤ 128 Hz), are comparable to pure-tone JNDs (Bernstein 2001; Bernstein and Trahiotis 2002; Oxenham et al. 2004). Ideal observer (or optimal processor) models of binaural processing, based on auditory-nerve activity (Colburn 1973, 1996), predict that ITD discrimination performance improves with increasing synchronization index given a reasonable mathematical description of the shapes of period histograms of AN fibers. Therefore our characterization of the phase locking of ANFs is relevant to an understanding of the neural mechanisms underlying the psychophysical results of Bernstein and Trahiotis (2002). Their finding that ITD JNDs are lower for transposed tones than for SAM tones at all modulation frequencies is consistent with our finding that synchrony to transposed tones is better than synchrony to SAM tones (Figs. 3, 4, 5, and 11) for all frequencies and levels. However, their finding of similar ITD JNDs for pure and transposed tones is hard to interpret in the context of our physiological results when

the effect of level is taken into account. Our results suggest that, at the level (75 dB SPL) where Bernstein and Trahiotis made their measurements, virtually all AN fibers with CFs close to the carrier frequency of transposed tones would show much poorer phase locking than in response to their pure-tone counterparts, and therefore ITD JNDs predicted by ideal-observer models would be considerably poorer for transposed tones than for pure tones.

There are several possibilities for reconciling the neural data with the psychophysics. One is that the binaural processor relies on fibers with CFs far from the carrier frequency of transposed tones for ITD discrimination. We have shown that maximum phase locking to transposed tones is nearly as good off the CF as that on the CF (Fig. 8), so these remote fibers can provide precise ITD information at high stimulus levels. However, contrary to this hypothesis, an unpublished human psychophysical study (Dreyer et al. 2005) found that ITD discrimination for transposed tones remains stable at high stimulus levels even in the presence of band-reject noise that was designed to mask ITD information in all but a narrow band ($\pm 20\%$ of f_c) of CFs centered at f_c . Alternatively, ITD discrimination at high sound levels might be based on the small fraction of high-threshold fibers that would still show good phase locking to transposed tones. The binaural processor would have to somehow focus on this small fraction of informative fibers and ignore the others. In awake animals, the number of informative fibers at high levels might be increased by activation of the olivocochlear efferents, which is known to shift rate-level functions to higher levels (Guinan and Stankovic 1996; Wiederhold and Kiang 1970). However, because the effect of olivocochlear efferent on dynamic range is most obvious in the presence of background noise (May and Sachs 1992), efferents would not be expected to play a large role in the experiments of Bernstein and Trahiotis (2002) because there was no background noise near f_c . Finally, the small, but nonzero, synchrony cues available in response to transposed tones at high levels (Fig. 3) may be enhanced in the CNS before binaural interactions. Many cell types in the VCN show a general enhancement in synchrony to the envelope over synchrony seen in the auditory nerve, as well as a smaller decrease in synchrony with level (Frisina et al. 1985, 1990; Joris et al. 2004; Rhode and Greenberg 1994). This synchrony enhancement may be more significant for transposed tones, which produce small synchrony, than for pure tones, for which synchrony is already high in the auditory nerve.

Bernstein and Trahiotis (2002) found that psychophysical ITD JNDs for transposed and SAM tones degrade rapidly for modulation frequencies > 128 Hz, whereas pure-tone ITD JNDs continue to improve up to ≥ 512 Hz. At first sight, this frequency dependency may seem germane to our result that maximum phase locking to SAM and transposed tones begins to decrease between 125 and 250 Hz. However, the degradation in ITD discrimination performance is precipitous (some of Bernstein and Trahiotis' subjects could not do the task at all at 512 Hz), whereas the degradation in phase locking is more gradual, and there is still significant synchrony at 1,000 Hz. Moreover, we have argued that cochlear filtering is likely to determine the upper limit of phase locking to SAM and transposed tones in our data, whereas Bernstein and Trahiotis provide strong arguments against peripheral filtering being responsible for the degradation in their ITD JNDs. Thus it is likely that the frequency limitation on ITD discrimination based on the envelope is central rather than peripheral. Bernstein and Trahiotis (2002) were able to fit their JND data by introducing a 150-Hz low-pass filter operating on the envelope in their cross-correlation model of binaural processing (Bernstein and Trahiotis 1996). Although such a filter is consistent with psychophysical results on monaural envelope perception (Kohlrausch et al. 2000), the ordering of the envelope filter and the binaural processor in their model is physiologically problematic. Bernstein and Trahiotis assumed that the envelope filter *precedes* binaural processing. Yet, temporal modulation transfer functions (tMTFs) of MSO and LSO neurons, the neurons that perform initial processing of ITD, typically have considerably higher upper cutoff frequencies than the 150 Hz postulated in the Bernstein and Trahiotis model (Grothe et al. 1997, 2001; Joris and Yin 1998). The 150-Hz cutoff is more in line with the tMTFs of most IC neurons (Krishna and Semple 2000), but the ITD

sensitivity of these neurons is thought to be largely inherited from their MSO and LSO inputs (Ingham and McAlpine 2005). Nevertheless, the possibility that some form of ITD sensitivity arises anew in the IC cannot be ruled out (Malmierca et al. 2005; Oliver 1987). Alternatively, the upper frequency limit on the ability to process ITDs in the envelope of high-frequency stimuli may not originate in brain stem processing per se, but rather in the inability of higher processing stages to make use of the binaural information conveyed by brain stem neurons. Why such a limit would operate on the envelope and not the fine structure is unclear. Thus the contrast in both level and frequency dependency between the psychophysical results of Bernstein and Trahiotis (2002) and the present physiological results raises important questions about the neural mechanisms for processing envelope ITDs.

Transposed tones have also been used to study pitch perception. Oxenham et al. (2004) found that frequency JNDs are considerably poorer for transposed tones with modulation frequencies in the range 55–320 Hz than for their pure-tone counterparts, even though their listeners had good ITD discrimination with the transposed tones for modulation frequencies <150 Hz. Moreover, listeners were unable to form pitch percepts for transposed tone complexes consisting of three transposed tones with modulation frequencies corresponding to harmonics 3, 4, and 5 of a common fundamental, whereas harmonic complex tones consisting of three pure tones evoked a strong pitch percept. Oxenham et al. (2004) interpreted their results as evidence against purely temporal models of pitch processing. This interpretation depends on the assumption that transposed tones and their pure tone counterparts produce essentially equivalent temporal discharge patterns in the auditory nerve. However, at the stimulus levels used in their experiments (70 phons for simple tones; 65 dB SPL per component for complex tones), a majority of the auditory-nerve fibers would show considerably poorer phase locking to transposed tones than to pure tones, and this issue would be most severe for the transposed tone complexes because the presence of multiple components would interfere with the spread of excitation to cochlear regions that would otherwise show good phase locking.³ On the other hand, if the degradation in AN phase locking at high stimulus levels is responsible for the poor frequency discrimination performance with transposed tones compared with pure tones, it is hard to understand why ITD discrimination is comparable for the two stimuli at low frequencies. Thus the Oxenham et al. (2004) results pose problems for purely temporal accounts of pitch and binaural processing.

Phase-locking metrics

Traditionally, phase locking to pure tones and SAM tones has been characterized by the synchronization index, or vector strength (Goldberg and Brown 1969; Johnson 1980). In circular statistics, the SI is a general measure of the concentration of the probability distribution around its mean, analogous to the SD for distributions along a line (Mardia and Jupp 2000). Thus this metric is an appropriate measure of the precision of phase locking to the fundamental of periodic stimuli when the period histogram is unimodal, as was usually the case in our data. The synchronization index is also appropriate for characterizing ITD sensitivity because, given reasonable assumptions about the shapes of period histograms, performance of ideal-observer models of ITD processing increase monotonically with the SI of the input neurons (Colburn 1973). However, the synchronization index is a poor model of the computations performed by the CNS because it requires a priori knowledge of the stimulus period, information that is not explicitly available to the central processor. Louage et al. (2004) proposed alternate measures of phase locking that are more appealing from the perspective of central processing because they are based on interspike interval statistics; in addition, they apply to both periodic and

³Reported simulations by Oxenham et al. (2004) of responses to their stimuli using the Meddis and O'Mard (1997) model were run at much lower stimulus levels than those used in the psychophysical experiments, suggesting that the compressions present in the model may have degraded the temporal cues at higher levels, consistent with our physiological data.

aperiodic stimuli such as noise. Specifically, the shuffled autocorrelation (SAC) is useful to characterize phase locking because, by tallying intervals between spikes obtained in response to *different* presentations of the same stimulus,⁴ it eliminates the distortions caused by neural refractoriness in the traditional autocorrelation (Louage et al. 2004). The SAC is particularly appropriate for characterizing ITD sensitivity because the computation it performs is equivalent to that of a binaural cross-correlator receiving statistically independent, identically distributed spike train inputs from the two ears, as postulated in many models of binaural processing (Colburn 1973,1996).

We compared the traditional synchronization index measure of phase locking with two measures based on the SAC for our three stimuli. For all stimuli, the SAC had the same period as the stimulus, indicating phase locking, and each SAC period had a single peak with a symmetric, nearly triangular shape (Fig. 9). For each stimulus, the SAC peak height was monotonically related to the synchronization index and the relationship was nonlinear (expansive), reflecting the upper limit of the synchronization index at unity (Fig. 10). A similar result was previously reported for pure-tone stimuli by Louage et al. (2004).⁵ The compressive nature of SI is evident in the steeper falloff of phase locking with level seen with SAC than with SI (Fig. 2 vs. Fig. 9) and also in the frequency dependency of maximum phase locking (Fig. 4 vs. Fig. 11). Although the relationship between SI and SAC peak height was qualitatively similar for our three stimuli, the relationship for pure tones differed somewhat from that for SAM and transposed tones (Fig. 10). Model simulations [using a model similar to the Colburn (1973) description of AN activity] suggest that the modulation of the instantaneous firing rate below SR present for pure tones, but not transposed tones, may partly account for this different relationship between SI and SAC peak height for the two stimuli. In any case, the difference was small, and the overall pattern of results was similar whether the synchronization index or the SAC peak height was used to characterize phase locking. One advantage of the synchronization index is that it depends more linearly on stimulus level than the SAC peak height for our stimuli.

We also found a monotonic relationship between the height of the SAC main peak and its half-width for our stimuli (Fig. 12). This relationship appears to be very similar for all three stimuli, although there was considerable scatter in the half-width when phase locking was poor. This result is consistent with the simple, triangular shape of the SAC for all stimuli (Fig. 9). Even though the SAC peak height and its half-width give essentially equivalent information for our stimuli, the peak height measure is preferable for two reasons. First, the half-width is more difficult to measure when data are noisy, such as when there are only a few spikes or when phase locking is poor. Second, because the peak height can in principle reach arbitrarily large values, whereas the normalized half-width is limited between 0 and 1, the peak height measure is less compressive and therefore more sensitive.

In summary, the SAC peak height appears to have a number of advantages as a measure of phase locking. Compared with the synchronization index, it performs a physiologically realistic computation, it applies to nonperiodic as well as periodic stimuli, and, for periodic stimuli, it is more sensitive (less compressive) when phase locking is precise. Its main disadvantages are

⁴A related procedure (de Cheveigné 1998) first aligns the spikes from all stimulus presentations onto the same time axis and then computes the autocorrelation of the combined spike train. The resulting histogram contains both intervals between spikes from different stimulus presentations and intervals from within the same presentation, i.e., it is the sum of the SAC and the traditional autocorrelation. For large numbers of stimulus presentations, it approaches the SAC because the contribution of the intervals from the same presentation becomes negligible.

⁵A quantitative comparison with the data reported in Louage et al. (2004) is difficult because their peak height metric was not based directly on the SAC, but on a more complex (and less physiological) form of correlation, the “difcor,” which combines responses to two stimuli with opposite polarities. Although their metric has the advantage of separating phase locking to the envelope from phase locking to the fine time structure, it was for this very reason not suitable for our purpose of comparing phase locking to the pure-tone waveform with phase locking to the envelopes of SAM and transposed tones.

that, because it is based on interspike intervals rather than absolute spike times, it discards phase information (Louage et al. 2004) and may underestimate phase locking at high frequencies (Johnson and Kiang 1976).

In conclusion, the stimulus transposition technique is intended to mimic the temporal discharge patterns produced in the auditory periphery by low-frequency tones using high-frequency, amplitude-modulated stimuli. We directly evaluated this idea by characterizing phase locking of auditory-nerve fibers to pure tones, SAM tones, and transposed tones over wide ranges of frequencies and levels. Consistent with the rationale for stimulus transposition and with psychophysical data on ITD sensitivity, phase locking to transposed tones was always better than phase locking to SAM tones. However, phase locking to transposed tones approached phase locking to pure tones only in restricted conditions (low stimulus levels, low modulation frequencies, low-SR fibers). Phase-locked responses to transposed tones differed markedly from responses to pure tones in both level and frequency dependency and, in this respect, transposed tones behaved more similarly to SAM tones. These results suggest caution in assuming a close similarity in temporal patterns of peripheral activity produced by transposed and pure tones in both psychophysical studies and neurophysiological studies of central neurons. Together with previous neurophysiological studies, they also raise fundamental questions about the central neural processing of the amplitude envelope at high sound levels, in particular with respect to the binaural processing of envelope ITDs.

Acknowledgements

The authors thank E. J. Tsai for providing the pure-tone data, C. Miller for surgical support, L. Cedolin for assistance with the experimental setup, and K. E. Hancock for software assistance. They are also grateful to K. E. Hancock, A. J. Oxenham, and two anonymous reviewers for critical readings of an earlier version of this manuscript.

GRANTS

This research was supported by National Institute of Deafness and Other Communications Disorders Grants R01 DC-002258, P30 DC-005209, and T32 DC-000038.

References

- Anderson DJ, Rose JE, Hind JE, Brugge JF. Temporal position of discharges in single auditory nerve within the cycle of a sine-wave stimulus: frequency and intensity effects. *J Acoust Soc Am* 1971;49:1131–1139. [PubMed: 4994692]
- Ashmore JF. The electrophysiology of hair cells. *Ann Rev Physiol* 1991;53:465–476. [PubMed: 2042969]
- Batra R, Kuwada S, Stanford TR. High-frequency neurons in the inferior colliculus that are sensitive to interaural delays of amplitude-modulated tones: evidence for dual binaural influences. *J Neurophysiol* 1993;70:64–80. [PubMed: 8395589]
- Bernstein LR. Auditory processing of interaural timing information: new insights. *J Neurosci Res* 2001;66:1035–1046. [PubMed: 11746435]
- Bernstein LR, Trahiotis C. Lateralization of low-frequency, complex waveforms: the use of envelope-based temporal disparities. *J Acoust Soc Am* 1985;77:1868–1880. [PubMed: 3998297]
- Bernstein LR, Trahiotis C. Detection of interaural delay in high-frequency sinusoidally amplitude-modulated tones, two-tone complexes, and bands of noise. *J Acoust Soc Am* 1994;95:3561–3567. [PubMed: 8046145]
- Bernstein LR, Trahiotis C. The normalized correlation: accounting for binaural detection across center frequency. *J Acoust Soc Am* 1996;100:3774–3784. [PubMed: 8969479]
- Bernstein LR, Trahiotis C. Enhancing sensitivity to interaural delays at high frequencies using “transposed stimuli. *J Acoust Soc Am* 2002;112:1026–1036. [PubMed: 12243151]
- Colburn HS. Theory of binaural interaction based on auditory-nerve data. I. General strategy and preliminary results on interaural discrimination. *J Acoust Soc Am* 1973;54:1458–1470. [PubMed: 4780800]

- Colburn HS, Esquissaud P. An auditory-nerve model for interaural time discrimination of high-frequency complex stimuli. *J Acoust Soc Am Suppl* 1976;1:S23.
- Colburn HS. Theory of binaural interaction based on auditory-nerve data. II. Detection of tones in noise. *J Acoust Soc Am* 1977;61:525–533. [PubMed: 845314]
- Colburn, HS. Computational models of binaural processing. In: Hawkins, H.; McMullin, T.; Popper, AN.; Fay, RR., editors. *Auditory Computation*. New York: Springer Verlag; 1996. p. 332-400.
- Cooper NP, Robertson D, Yates GK. Cochlear nerve fiber responses to amplitude-modulated stimuli: variations with spontaneous rate and other response characteristics. *J Neurophysiol* 1993;70:370–386. [PubMed: 8395584]
- deCheveigne A. Cancellation model of pitch perception. *J Acoust Soc Am* 1998;103:1261–1271. [PubMed: 9514016]
- Delgutte B. Representation of speech-like sounds in the discharge patterns of auditory-nerve fibers. *J Acoust Soc Am* 1980;68:843–857. [PubMed: 7419820]
- Dreyer, A.; Delgutte, B. Comment on paper by Bernstein and Trahiotis. In: Pressnitzer, D.; de Cheveigné, A.; McAdams, S.; Collet, L., editors. *Auditory Signal Processing: Physiology, Psychoacoustics, and Models*. New York: Springer-Verlag; 2004. p. 384-386.
- Dreyer A, Oxenham A, Delgutte B. Predicting lateralization performance at high frequencies from auditory-nerve spike timing. *Abstr Assoc Res Otolaryngol* 2005;28:71.
- Fettiplace R, Fuchs PA. Mechanisms of hair cell tuning. *Ann Rev Physiol* 1999;61:809–834. [PubMed: 10099711]
- Fitzpatrick DC, Patel MR, Ebert CS, Demke JS. The effect of intensity on ITD tuning to pure tones. *Abstr Assoc Res Otolaryngol* 2005;28:659.
- Frisina RD, Smith RL, Chamberlain SC. Differential encoding of rapid changes in sound amplitude by second-order auditory neurons. *Exp Brain Res* 1985;60:417–422. [PubMed: 4054285]
- Frisina RD, Smith RL, Chamberlain SC. Encoding of amplitude modulation in the gerbil cochlear nucleus. I. A hierarchy of enhancement. *Hear Res* 1990;44:99–122. [PubMed: 2329098]
- Goldberg JM, Brown PB. Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *J Neurophysiol* 1969;32:613–636. [PubMed: 5810617]
- Greenwood DD, Joris PX. Mechanical and “temporal” filtering as codeterminants of the response by cat primary fibers to amplitude-modulated signals. *J Acoust Soc Am* 1996;99:1029–1039. [PubMed: 8609286]
- Griffin SJ, Bernstein LR, Ingham NJ, McAlpine D. Neural sensitivity to interaural envelope delays in the inferior colliculus of the guinea pig. *J Neurophysiol* 2005;93:3463–3478. [PubMed: 15703234]
- Grothe B, Covey E, Casseday JH. Medial superior olive of the big brown bat: neuronal responses to pure tones, amplitude modulations, and pulse trains. *J Neurophysiol* 2001;86:2219–2230. [PubMed: 11698513]
- Grothe B, Park TJ, Schuller G. Medial superior olive in the free-tailed bat: response to pure tones and amplitude-modulated tones. *J Neurophysiol* 1997;77:1553–1565. [PubMed: 9084619]
- Guinan JJ Jr, Stankovic KM. Medial efferent inhibition produces the largest equivalent attenuations at moderate to high sound levels in cat auditory-nerve fibers. *J Acoust Soc Am* 1996;100:1680–1690. [PubMed: 8817894]
- Henning GB. Detectability of interaural delay in high-frequency complex waveforms. *J Acoust Soc Am* 1974;55:84–90. [PubMed: 4815755]
- Hudspeth AJ, Corey DP. Sensitivity, polarity, and conductance change in the response of vertebrate hair cells to controlled mechanical stimuli. *Proc Natl Acad Sci USA* 1977;74:2407–2411. [PubMed: 329282]
- Ingham NJ, McAlpine D. GABAergic inhibition controls neural gain in inferior colliculus neurons sensitive to interaural time differences. *J Neurosci* 2005;25:6187–6198. [PubMed: 15987948]
- Javel E. Coding of AM tones in the chinchilla auditory nerve: implications for the pitch of complex tones. *J Acoust Soc Am* 1980;68:133–146. [PubMed: 7391355]
- Johnson DH. The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *J Acoust Soc Am* 1980;68:1115–1122. [PubMed: 7419827]

- Johnson DH, Kiang NY. Analysis of discharges recorded simultaneously from pairs of auditory nerve fibers. *Biophys J* 1976;16:719–734. [PubMed: 938715]
- Joris PX, Carney LH, Smith PH, Yin TCT. Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. *J Neurophysiol* 1994;71:1022–1036. [PubMed: 8201399]
- Joris PX, Schreiner CE, Rees A. Neural processing of amplitude-modulated sounds. *Physiol Rev* 2004;84:541–577. [PubMed: 15044682]
- Joris PX, Yin TC. Envelope coding in the lateral superior olive. III. Comparison with afferent pathways. *J Neurophysiol* 1998;79:253–269. [PubMed: 9425196]
- Joris PX, Yin TCT. Responses to amplitude-modulated tones in the auditory nerve of the cat. *J Acoust Soc Am* 1992;91:215–232. [PubMed: 1737873]
- Kiang NY, Eddington DK, Delgutte B. Fundamental considerations in designing auditory implants. *Acta Otolaryngol (Stockh)* 1979;87:204–218. [PubMed: 443001]
- Kiang NYS, Moxon EC. Tails of tuning curves of auditory-nerve fibers. *J Acoust Soc Am* 1974;55:620–628. [PubMed: 4819862]
- Kiang, NYS.; Watanabe, T.; Thomas, EC.; Clark, LF. *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve*. Cambridge, MA: MIT Press; 1965.
- Kohlrausch A, Fassel R, Dau T. The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers. *J Acoust Soc Am* 2000;108:723–734. [PubMed: 10955639]
- Krishna BS, Semple MN. Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus. *J Neurophysiol* 2000;84:255–273. [PubMed: 10899201]
- Kuwada S, Yin TCT. Binaural interaction in low-frequency neurons in inferior of the cat. I. Effects of long interaural delays, intensity, and repetition rate on interaural delay functions. *J Neurophysiol* 1983;50:981–999. [PubMed: 6631473]
- Kuwada S, Yin TCT, Syka J, Buunen TJ, Wickesberg RE. Binaural interaction in low-frequency neurons in inferior of the cat. IV. Comparison of monaural and binaural response properties. *J Neurophysiol* 1984;51:1306–1325. [PubMed: 6737032]
- Leakey DM, Sayers M, Cherry C. Binaural fusion of low- and high-frequency sounds. *J Acoust Soc Am* 1958;30:222–222.
- Lieberman MC. Auditory-nerve responses from cats raised in a low-noise chamber. *J Acoust Soc Am* 1978;63:442–455. [PubMed: 670542]
- Louage DH, van der Heijden M, Joris PX. Temporal properties of responses to broadband noise in the auditory nerve. *J Neurophysiol* 2004;91:2051–2065. [PubMed: 15069097]
- Macpherson EA, Middlebrooks JC. Listener weighting of cues for lateral angle: the duplex theory of sound localization revisited. *J Acoust Soc Am* 2002;111:2219–2236. [PubMed: 12051442]
- Malmierca MS, Saint Marie RL, Merchan M, Oliver DL. Laminar inputs from dorsal cochlear nucleus and ventral cochlear nucleus to the central nucleus of the inferior colliculus: two patterns of convergence. *Neuroscience* 2005;136:883–394. [PubMed: 16344158]
- Mardia, KV.; Jupp, PE. *Directional Statistics*. New York: Wiley; 2000.
- May BJ, Sachs MB. Dynamic range of neural rate responses in the ventral cochlear nucleus of awake cats. *J Neurophysiol* 1992;68:1589–1602. [PubMed: 1479432]
- McFadden D, Pasanen EG. Lateralization at high frequencies based on interaural time differences. *J Acoust Soc Am* 1976;59:634–639. [PubMed: 1254790]
- Meddis R, O'Mard L. A unitary model of pitch perception. *J Acoust Soc Am* 1997;102:1811–1820. [PubMed: 9301058]
- Oliver DL. Projections to the inferior colliculus from the anteroventral cochlear nucleus in the cat: possible substrates for binaural interaction. *J Comp Neurol* 1987;264:24–46. [PubMed: 2445792]
- Oxenham AJ, Bernstein JG, Penagos H. Correct tonotopic representation is necessary for complex pitch perception. *Proc Natl Acad Sci USA* 2004;101:1421–1425. [PubMed: 14718671]
- Palmer AR, Evans EF. Intensity coding in the auditory periphery of the cat: responses of cochlear nerve and cochlear nucleus neurons to signals in the presence of bandstop masking noise. *Hear Res* 1982;7:305–323. [PubMed: 7118733]

- Palmer AR, Russell IJ. Phase-locking in the cochlear nerve of the guinea pig and its relation to the receptor potential of inner hair cells. *Hear Res* 1986;24:1–15. [PubMed: 3759671]
- Rhode WS, Greenberg S. Encoding of amplitude modulation in the cochlear nucleus of the cat. *J Neurophysiol* 1994;71:1797–1825. [PubMed: 8064349]
- Rhode WS, Smith PH. Characteristics of tone-pip response patterns in relationship to spontaneous rate in cat auditory nerve fibers. *Hear Res* 1985;18:159–168. [PubMed: 2995298]
- Ruggero MA, Temchin AN. Unexceptional sharpness of frequency tuning in the human cochlea. *Proc Natl Acad Sci USA* 2005;102:18614–18619. [PubMed: 16344475]
- Shackleton TM, Skottun BC, Arnott RH, Palmer AR. Interaural time difference discrimination thresholds for single neurons in the inferior colliculus of guinea pigs. *J Neurosci* 2003;23:716–724. [PubMed: 12533632]
- Shera CA, Guinan JJ Jr, Oxenham AJ. Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. *Proc Natl Acad Sci USA* 2002;99:3318–3323. [PubMed: 11867706]
- Smith RL, Brachman ML. Response modulation of auditory-nerve fibers by AM stimuli: effects of average intensity. *Hear Res* 1980;2:123–133. [PubMed: 7364668]
- Sumner CJ, O'Mard L, Lopez-Poveda EA, Meddis R. A nonlinear filter-bank model of the guinea-pig cochlear nerve: rate responses. *J Acoust Soc Am* 2003;113:3264–3274. [PubMed: 12822799]
- Trahiotis, C.; Stern, RM. Models of binaural interaction. In: Moore, BC., editor. *Handbook of Perception and Cognition*. Orlando, FL: Academic Press; 1995. p. 387-424.
- Tsai EJ, Delgutte B. Neural mechanisms underlying intensity discrimination: responses of auditory-nerve fibers to pure tones in band-reject noise. *Assoc Res Otolaryngol Abstr* 1997;20:154.
- van de Par S, Kohlrausch A. A new approach to comparing binaural masking level differences at low and high frequencies. *J Acoust Soc Am* 1997;101:1671–1680. [PubMed: 9069634]
- Wang X, Sachs MB. Neural encoding of single-formant stimuli in the cat. I. Responses of auditory nerve fibers. *J Neurophysiol* 1993;70:1054–1075. [PubMed: 8229159]
- Weiss TF, Rose C. Stages of degradation of timing information in the cochlea: a comparison of hair-cell and nerve-fiber responses in the alligator lizard. *Hear Res* 1988;33:167–174. [PubMed: 3397326]
- Westerman L, Smith R. A diffusion model of the transient response of the cochlear inner hair cell synapse. *J Acoust Soc Am* 1988;83:2266–2276. [PubMed: 3411018]
- Wiederhold ML, Kiang NYS. Effects of electric stimulation of the crossed olivocochlear bundle on single auditory-nerve fibers in the cat. *J Acoust Soc Am* 1970;48:950–965. [PubMed: 5480390]
- Wightman FL, Kistler DJ. The dominant role of low-frequency interaural time differences in sound localization. *J Acoust Soc Am* 1992;91:1648–1661. [PubMed: 1564201]
- Yates GK. Dynamic effects in the input/output relationship of auditory nerve. *Hear Res* 1987;27:221–230. [PubMed: 3610850]
- Yin TCT, Kuwada S. Binaural interaction in low-frequency neurons in inferior of the cat. II. Effects of changing rate and direction of interaural phase. *J Neurophysiol* 1983;50:1000–1019. [PubMed: 6631458]
- Zhang X, Carney LH. Response properties on an integrate-and-fire model that receives subthreshold inputs. *Neural Comput* 2005;17:2571–2601. [PubMed: 16212763]

Fig 1.

Acoustic waveforms and hypothetical outputs of the peripheral auditory processor for a pure tone, a sinusoidally amplitude-modulated (SAM) tone, and a transposed tone. (Adapted from Bernstein 2001.)

Fig 2.

A: period histograms as a function of stimulus level for a pure tone, a SAM tone, and a transposed tone. Pure-tone responses (black squares) are from a low characteristic frequency (CF, 344 Hz), medium spontaneous rate (SR) fiber (SR = 2 spikes/s and threshold of 28 dB SPL) using a frequency near the CF (354 Hz). Responses to SAM (red triangles) and transposed (blue circles) tones are from a high-CF (7,000 Hz), medium-SR fiber (5.2 spikes/s and threshold 20 dB SPL) using a carrier at the CF and a modulation frequency of 250 Hz. Stimulus waveforms are shown at the *bottom* on a normalized timescale. B: synchronization index calculated from the period histograms as a function of level with respect to rate threshold for the 3 stimuli. C: average discharge rate as a function of level with respect to rate threshold.

Fig 3.

Synchronization index (SI) as a function of level with respect to rate threshold in response to pure tones (black dotted), transposed tones (blue solid), and SAM tones (red dashed) for entire data set presented at the CF. Only pure-tone responses <1,000 Hz are included. For each stimulus type, a bold line was drawn using the mean values of the slope, maximum SI, and SPL of maximum SI from the best linear fit to each SI-level function.

Fig 4.

A: maximum values of synchronization index ($SI_{\max}$) over all levels as a function of frequency or modulation frequency for pure tones (dotted black), transposed tones (solid blue), and SAM tones (dashed red). Error bars represent ± 1 SE. B: mean slopes of SI-level functions (expressed in dB^{-1}) as a function of modulation frequency. In both panels, pure-tone responses are binned to the closest frequency shown.

Fig 5. SI vs. level functions for 4 modulation frequencies. Each panel shows responses and best-fitting line (bold) to SAM (dashed red) and transposed (solid blue) tones for one modulation frequency. Best-fit lines are computed using same method as in Fig. 3.

Fig 6.

SI_{max} values over all levels as a function of frequency or modulation frequency for pure tones, transposed tones, and SAM tones differentiated by high-SR (dashed lines) and low/medium-SR (solid lines) groups of fibers. Error bars represent ± 1 SE.

Fig 7.

A: period histograms to transposed tones ($f_m = 250$ Hz) with carrier frequencies (f_c) below, on, and above the CF as a function of stimulus levels. Carrier frequencies were 5,603 Hz (below-CF, light blue), 6,920 Hz (on-CF, blue), and 7,490 Hz (above-CF, green). Fiber CF = 6,920 Hz, SR = 1.0 spikes/s, threshold at CF = 31 dB SPL. B: synchrony level functions derived from period histograms for each carrier frequency. Level is expressed relative to rate threshold at f_c . SI value for lowest level on-CF response is not shown because it did not reach statistical significance. C: rate level functions.

Fig 8.

Comparison of slope of synchrony level functions off CF and on CF. *A, top and bottom:* SI vs. level expressed with respect to level maximum of each SI-level function. Three colors represent responses to f_c on CF (dashed black), above CF (solid magenta), and below CF (dotted green). *B, top and bottom:* change in SI (ΔSI) level functions relative to on-CF condition for f_c above (solid magenta) and below CF (dotted green). ΔSI -level curves are obtained by binning level axis for the synchrony level functions in *A* in 5-dB steps and subtracting the on-CF response from the off-CF response for each level.

Fig 9. A: shuffled autocorrelograms (SACs) as a function of stimulus level in response to a pure tone (black squares), SAM tone (red triangles), and a transposed tone (blue circles). Figure is based on the same data as in Fig. 2. B: normalized autocorrelation peak height vs. sound level with respect to rate threshold for the 3 stimuli. C: normalized autocorrelation half-width against sound level with respect to rate threshold for the 3 stimuli.

Fig 10.

Normalized SAC peak height (see METHODS) against SI for all recordings in response to low-frequency pure tones (black), transposed tones (blue), and SAM tones (red). Only pure-tone frequencies <1,000 Hz are included. Best-fitting hyperbolic curves (solid lines) were fit separately to pure-tone data and to combined SAM and transposed tone data.

Fig 11. Maximum SAC peak heights across all levels as a function of frequency or modulation frequency for pure tones (black dotted), transposed tones (blue solid), and SAM tones (red dashed). Error bars represent ± 1 SE. Pure-tone responses are binned to the closest frequency shown.

Fig 12. SAC peak height against normalized half-width for all low-frequency pure tone (black circles), transposed tone (blue triangles), and SAM tone (red asterisks) data. Solid line shows the best-fitting hyperbolic function (see RESULTS) for the entire data set.

Table 1Slopes and intercepts of Δ SI-level functions for SAM and transposed tones

Stimulus	Carrier Range	Intercept	Slope, dB ⁻¹
SAM	$f_c < \text{CF}$	0.029	0.0040
SAM	$f_c > \text{CF}$	-0.108*	0.0059*
Transposed	$f_c < \text{CF}$	-0.135*	0.0034
Transposed	$f_c > \text{CF}$	-0.021	0.0005

Slopes and intercepts of the best-fitting lines to the Δ SI-level functions shown in Fig. 8B for each of the two stimuli (SAM and transposed tones) and two carrier frequency ranges (below and above CF). Asterisks indicate values significantly different from 0 at the 0.05 level.