



Published in final edited form as:

J Neurophysiol. 2005 August ; 94(2): 1180–1198.

Neural Correlates and Mechanisms of Spatial Release From Masking: Single-Unit and Population Responses in the Inferior Colliculus

Courtney C. Lane^{1,2} and Bertrand Delgutte^{1,3}

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

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

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

Abstract

Spatial release from masking (SRM), a factor in listening in noisy environments, is the improvement in auditory signal detection obtained when a signal is separated in space from a masker. To study the neural mechanisms of SRM, we recorded from single units in the inferior colliculus (IC) of barbiturate-anesthetized cats, focusing on low-frequency neurons sensitive to interaural time differences. The stimulus was a broadband chirp train with a 40-Hz repetition rate in continuous broadband noise, and the unit responses were measured for several signal and masker (virtual) locations. Masked thresholds (the lowest signal-to-noise ratio, SNR, for which the signal could be detected for 75% of the stimulus presentations) changed systematically with signal and masker location. Single-unit thresholds did not necessarily improve with signal and masker separation; instead, they tended to reflect the units' azimuth preference. Both how the signal was detected (through a rate increase or decrease) and how the noise masked the signal response (suppressive or excitatory masking) changed with signal and masker azimuth, consistent with a cross-correlator model of binaural processing. However, additional processing, perhaps related to the signal's amplitude modulation rate, appeared to influence the units' responses. The population masked thresholds (the most sensitive unit's threshold at each signal and masker location) did improve with signal and masker separation as a result of the variety of azimuth preferences in our unit sample. The population thresholds were similar to human behavioral thresholds in both SNR value and shape, indicating that these units may provide a neural substrate for low-frequency SRM.

INTRODUCTION

Normal-hearing listeners have a remarkable ability to attend to a single sound source in an environment containing multiple competing sources. In contrast, hearing-impaired listeners and artificial speech-recognition systems often have difficulty in such environments, making it important to understand the normal mechanisms used by the auditory system in the hope of improving assistive devices. Here, we study the neural mechanisms underlying one aspect of noisy-environment listening, termed “spatial release from masking” (SRM), which refers to the observation that a signal is more easily detected when separated in space from a masking noise compared with when the signal and noise are co-located (Gilkey and Good 1995; Good et al. 1997; Lane et al. 2004; Saberi et al. 1991).

Address for reprint requests and other correspondence: C. C. Lane, ECE Department, Rice University, MS 380, P.O. Box 1892, Houston, TX 77251-1892 (E-mail: court@rice.edu).
Present address of C. C. Lane: Rice University, Electrical and Computer Engineering Dept., MS 380, PO Box 1892, Houston, TX 77251-1892. E-mail: court@rice.edu.

Psychophysical studies have shown that SRM occurs at all stimulus frequencies, although the mechanisms appear to differ at low and high frequencies. At high frequencies (>1.5 kHz), SRM appears to be achieved primarily through the monaural changes in signal-to-noise ratio (SNR) resulting from directionally dependent filtering by the head and pinna. In contrast, signal detection at low frequencies appears to be based on binaural hearing because masked thresholds in binaural listening are substantially better than those obtained by listening monaurally through either ear (Gilkey and Good 1995; Good et al. 1997). Here, we focus on low frequencies, which are important for speech recognition and are often spared in hearing-impaired listeners.

The ability of the binaural system to improve signal detection at low frequencies has been studied extensively by psychophysicists using the binaural masking level difference (BMLD) paradigm (for a thorough review see Durlach and Colburn 1978). In the most common BMLD experiments, identical noise is presented to both ears (N0), and the masked threshold for a signal (usually a pure tone) presented in phase at the two ears (S0) is compared with the threshold for a signal presented out of phase at the two ears ($S\pi$). These are the NOS0 and NOS π conditions, respectively, and the NOS π condition yields better thresholds because of the additional interaural phase cue. The difference in masked thresholds (the BMLD) between these two conditions is about 12–15 dB for a low-frequency ($\lesssim 500$ Hz) pure tone in broadband noise. For the case when the noise is antiphasic instead of the signal (i.e., N π S0 compared with NOS0), there is also a substantial, although slightly smaller BMLD of around 9–10 dB. Because there are no differences in signal-to-noise ratios at either ear between the BMLD conditions, the improvements in detectability must be attributed to the listener's ability to use the binaural system to exploit interaural timing cues.

A number of mathematical models are able to predict most psychophysical results on BMLDs (Colburn 1996; Colburn and Durlach 1978). Here, we focus on the cross-correlator model originally envisioned by Jeffress (1948) and developed by Colburn (1973, 1977a,b) because processing similar to the processing described by the model appears to occur in the medial superior olive (MSO) and because predictions of this model for the BMLD paradigm have been extensively studied. In this model, the stimulus waveform to each ear is processed by an auditory nerve fiber (ANF) model. Pairs of ANFs with the same characteristic frequencies (CFs) from each ear provide inputs to an array of delay lines and coincidence detectors, which fire when the neural inputs from the two sides coincide. The delay lines allow each coincidence detector to respond maximally to a different interaural time delay (ITD), its characteristic delay. Effectively, this model performs an instantaneous cross-correlation of the ANF responses, with each coincidence detector evaluating the correlation at its characteristic delay.

Colburn (1973, 1977a,b) studied how a cross-correlator network responds to BMLD stimuli. Briefly, coincidence detector units with a characteristic delay of 0 (0-ITD units) maximally respond to the N0 stimuli, whereas the remaining units produce a weaker response. When a signal is added in phase (NOS0), the firing rate for 0-ITD units tuned to the tone frequency increases, allowing the signal to be detected. If the signal is instead added out of phase (NOS π), then the response of the 0-ITD units tuned to the tone frequency *decreases* because the signal decorrelates the inputs to the 0-ITD units at that frequency. This decrease in firing rate is more easily detected than the rate increase in the NOS0 case, thereby giving rise to the BMLD. In contrast, for the N π stimulus, units with characteristic delays equal to one half the period of their CF (" π -units") respond maximally. When the signal is added in phase (N π S0), it causes the response of the π -units tuned to the signal frequency to decrease as the signal reduces the correlation seen by the π -units. By looking at the response changes across the neural population, Colburn's model successfully predicts most psychophysical results for BMLDs, including the threshold hierarchy where the NOS π threshold is best, followed by N π S0 and then NOS0.

Although controversy exists over the precise neural mechanisms, the neurons in the MSO appear to perform a cross-correlation operation similar to the one hypothesized by Jeffress and Colburn (Batra and Yin 2004; Batra et al. 1997; Yin and Chan 1990). As expected for a coincidence detector, units in the MSO receive binaural excitatory inputs, are sensitive to the ITD of tones and noise, and their best interaural phase can be predicted from the phases of the monaural responses (Batra et al. 1997; Goldberg and Brown 1969; Yin and Chan 1990). ITD-sensitive units in the MSO form a major projection to the inferior colliculus (IC; Loftus et al. 2004), where units with similar ITD sensitivity are also found (see Spitzer and Semple 1998; Yin and Chan 1990 for a comparison between MSO and IC unit responses). Because essentially all of the ascending auditory pathways synapse in the tonotopically organized central nucleus of the IC, the IC represents the first nucleus after the cochlear nucleus that contains nearly all of the information available to the auditory system and is the first to receive inputs from binaural nuclei. This convergence of information makes the IC an interesting and convenient nucleus for studying the neural mechanisms of SRM. Furthermore, the responses of the units in the IC almost certainly reflect additional processing beyond that of the MSO, and this processing, whose nature is still poorly understood, is likely to be behaviorally relevant.

An extensive series of neurophysiological studies of BMLD (e.g., Jiang et al. 1997a,b; McAlpine et al. 1996; Palmer et al. 1999, 2000) tested some of the predictions of the Colburn model for low-frequency units in the IC of anesthetized guinea pigs. Using 500-Hz pure tones in broadband noise, the authors showed that the masked thresholds for individual units in the anesthetized guinea pig vary with the interaural phases and interaural delays of the signal and masker (McAlpine et al. 1996). Furthermore, they measured neural thresholds for the NOS0 and NOS π conditions for both individual units and populations of units (Jiang et al. 1997a,b). They showed that individual units could show positive or negative BMLDs, but that when averaged across their sample, the thresholds were better for the NOS π condition than for the NOS0 condition, similar to the psychophysical results. As expected from the Colburn model, the unit population with the best average thresholds had an excitatory response in the N0 condition and showed a *decrease* in firing rate when the antiphase signal S π was added. Additionally, for a majority of neurons, the decreases in rate caused by the addition of S π were similar to the changes in response seen for noise when the interaural correlation is reduced (Palmer et al. 1999), validating the general concept of the cross-correlator model. Finally, for a different unit sample, Palmer et al. (2000) showed that the N π S0 thresholds averaged across all units were better than those seen for NOS0; however, the BMLDs were smaller than those seen with NOS π . Overall, the average unit NOS0 neural thresholds are worse than the N π S0 thresholds, which are worse than the NOS π thresholds, consistent with the psychophysical threshold hierarchy. Additionally, the responses of low-frequency IC units are generally consistent with the cross-correlator model. However, these authors reported that some units seemed to reflect the effects of additional inhibition; the responses of these units could not be entirely predicted by a simple cross-correlator model. Also, in some cases, the *individual* unit thresholds did not match the trends seen in the average unit thresholds.

The physiological studies of BMLD provide a nice link between the cross-correlator model and the behavioral results in BMLD experiments, but the BMLD paradigm is somewhat unnatural. First, the signal in BMLD experiments is a pure tone, which contains only one frequency and has a flat temporal envelope, whereas most natural sounds are broadband and have large amplitude modulations. Moreover, for the antiphase condition used in the BMLD experiments, the stimulus is out of phase for all frequencies, essentially giving a different ITD for each frequency. This stimulus condition differs from the case when a broadband sound source is placed at a location off the median vertical plane, which gives a constant ITD across all frequencies. In this case, the constant ITD gives rise to a different interaural *phase* difference (IPD) for each frequency component. Figure 1F shows the sloping phase responses in the left and right ears for a stimulus placed at -90° for a spherical model of the cat head. Clearly, the

different slopes resulting from the fixed ITD yield different IPDs at each frequency. Consequently, it is not entirely clear how the physiological BMLD results generalize to the more natural situation where broadband signals and maskers are placed at different spatial locations. Caird et al. (1991) previously addressed some of these issues by looking at the effect of masker ITD on masked thresholds using either a pure tone or a vowel. The signal was placed at the best delay for each neuron, and the masked threshold was measured as a function of noise ITD. The masked thresholds were shown to reflect sensitivity of the neuron to the noise ITD, but the effects of placing the signal at different delays were not explored. Consequently, the effect of signal and masker *separation* was not directly addressed, nor was the effect of placing the signal at the *worst* delay, the condition predicted by the cross-correlator model to give the best thresholds.

Here, we explore IC unit responses using broadband signals and maskers placed at different (virtual) spatial locations. Consistent with the results of Caird et al. (1991) and McAlpine et al. (1996), we find that the single-unit thresholds depend on the locations of the signal and the masker relative to the unit's preferred azimuth. To find a neural correlate of SRM, we turned to the response of the neural population as a whole, which includes units with a variety of preferred azimuths. We show that the best thresholds across the population of IC units are similar to human behavioral thresholds for similar stimuli, in both SNR values and dependency on signal and masker separation. We also show that a substantial number of the units have responses to the signal and masker that are not predictable from the cross-correlator model. Moreover, in contrast to Jiang et al. (1997a,b), we find that the best masked thresholds for broadband chirp trains in noise occur for units where adding the signal *increases* the overall rate response to the noise. We argue that our results do not contradict those of Jiang et al. (1997a,b), but instead reflect differences between the temporal and/or spectral properties of the signals used in the two studies. Preliminary reports of this work have been presented (Kopco et al. 2003; Lane 2003; Lane et al. 2003, 2004).

METHODS

Recording techniques

Responses of single units in the anesthetized cat inferior colliculus were recorded using methods similar to those of Litovsky and Delgutte (2002). Healthy, adult cats were initially anesthetized with an intraperitoneal injection of Dial in urethane (75 mg/kg), and additional doses were provided throughout the experiment to maintain deep anesthesia. Dexamethasone was injected intramuscularly to prevent swelling of neural tissue. A rectal thermometer was used to monitor the animal's temperature, which was maintained at 37–38°C. A tracheal cannula was inserted, both pinnae were partially dissected away, and the ear canals were cut to allow insertion of acoustic assemblies. A small hole was drilled in each bulla, and a 30-cm plastic tube was inserted and glued in place to prevent static pressure from building up in the middle ear. The animal was placed in a double-walled, electrically shielded, sound-proof chamber. The posterior surface of the IC was exposed through a posterior fossa craniotomy and aspiration of the overlying cerebellum. Parylene-insulated tungsten stereo microelectrodes (Micro Probe, Potomac, MD) were mounted on a remote-controlled hydraulic microdrive and inserted into the IC. The electrodes were oriented nearly horizontally in a parasagittal plane, approximately parallel to the isofrequency planes (Merzenich and Reid 1974). To improve single-unit isolation, the difference between the signals recorded from the two electrodes, which were separated by 125 μm , was often used as the input to the amplifier and spike timer. Spikes from single units were amplified and spike times measured with 1- μs resolution were stored in a computer file for analysis and display.

Histological processing for reconstruction of an electrode track was performed for one cat with a particularly large data yield. Only one track was made in this experiment. Every third 40-

μm parasagittal section of the IC was immunostained for calretinin to visualize putative projections from the MSO (Adams 1995), and the remaining sections were Nissl-stained. Staining for calretinin is thought to reveal terminals of MSO axons because the MSO is the only auditory structure projecting to the IC in which calretinin labeling is extensive. The electrode track was evident in the Nissl slice and one of the calretinin slices, and the track traversed the calretinin region. The microelectrode depths at which we found units indicate that we were recording from the calretinin region, suggesting that these units received inputs from the MSO. The other experiments had similar electrode placements and single-unit responses including ITD sensitivity. Therefore most of the units in our sample are likely to receive MSO inputs, as expected from the anatomical results of Loftus et al. (2004) showing large projections of MSO to the low-frequency IC.

Stimuli

The signal used was a 200-ms-long train of broadband chirps with a 40-Hz repetition rate presented in continuous broadband noise (see Fig. 1, A and C). Each chirp's frequency was swept from 300 Hz to 30 kHz logarithmically and had an exponentially increasing envelope designed to produce a flat power spectrum. Consequently, both signal and noise had a relatively flat spectrum (Fig. 1, B and D), from 300 Hz to 30 kHz, before they were shaped by the frequency response of the head-related transfer functions (see following text). In some cases, we also used 100-Hz click trains as signals similar to the stimuli used in the psychophysical literature on SRM (e.g., Gilkey and Good 1995; Saberi et al. 1991); however, units in the IC often responded with higher, more sustained rates to the chirp trains, presumably because of the lower repetition rate of the chirp train. Only results obtained with the 40-Hz chirp trains are presented here.

Because SRM occurs for stimuli in all frequency ranges (Gilkey and Good 1995), we use head-related transfer functions (HRTFs) to simulate sounds at different azimuths. Using HRTFs allows us to simulate sounds in the free-field while still allowing complete control over the inputs to the two ears, thereby enabling us to easily present more traditional stimuli, such as monaural stimuli or binaural beats. The HRTFs represent the directionally dependent transformations of sound pressure from a specific location in free field to the ear canal (see Fig. 1, E and F). Virtual-space stimuli were synthesized by filtering the stimuli with the same HRTFs used by Litovsky and Delgutte (2002). The nonindividualized cat HRTFs were measured by Musicant et al. (1990) for frequencies >2 kHz and were simulated by a spherical-head model for frequencies <2 kHz. (The HRTF measurements were valid only for frequencies >2 kHz because of the limitations of the sound system and anechoic room.) The low-frequency HRTFs were the product of two components: 1) a directional component representing acoustic scattering by the cat head was provided by a rigid-sphere model with a diameter of 6.8 cm (Morse and Ingard 1968); and 2) a nondirectional, frequency-dependent gain representing the sound pressure amplification by the external ear was derived from measurements of acoustic impedance in the cat ear canal (Rosowski et al. 1988). Using a frequency-dependent weighting function, the model HRTF for frequencies <2 kHz was joined with the measured HRTF >2 kHz to obtain an HRTF covering the 0 to 40-kHz range.

This paper focuses on low-frequency neurons that are sensitive to ITD, the primary sound localization cue at low frequencies. Consequently, the spherical-head model provides most of the information in the HRTF for this work. Here the phase response of the HRTF was nearly a straight line for all azimuths, as expected for a pure delay, and the magnitude of the HRTF was relatively constant for different azimuths at these low frequencies (see example in Fig. 1, E and F). We expect our results to be similar to those obtained if only ITD was varied, provided the stimuli were appropriately shaped by the nondirectional HRTF magnitude (see RESULTS).

However, the use of HRTFs would allow the present study of SRM to be easily extended to high frequencies in the future.

Experimental procedure

Search stimuli were either 200-ms chirp trains or broadband noise bursts. Both the azimuth and the mode of stimulation (binaural or monaural) of the search stimulus were varied in an effort to find a larger number of units and a more varied sample. Once a single unit was isolated, a frequency-tuning curve was measured by an automatic tracking procedure (Kiang and Moxon 1974) to determine the characteristic frequency (CF_{TC}).

A noise-delay function was also measured: the unit response was measured as a function of the ITD of 200-ms bursts of “frozen” noise (Fig. 2A, solid line with error bars). The ITD was usually varied from $-2,000$ to $2,000 \mu s$ with a step size of $400 \mu s$, although ITDs inside the physiological range (-290 to $290 \mu s$ as determined using our HRTFs) were often sampled more finely.

The primary measurements in these experiments were the responses to the signal and noise together and to the noise alone as a function of noise level (e.g., Fig. 3). From these measurements, we determined the single-unit masked thresholds. The signal level was fixed near 40 dB SPL (measured over the entire 30-kHz bandwidth), and the noise level was varied in 6-dB steps in randomized order. We used a fairly low signal level to ensure that the signal response could be fully masked without reaching maximum output of our sound system (threshold SNRs were often as low as -20 dB). We chose to vary the noise level so that we could study the effects of the noise masker on a fixed signal response, eliminating some of the confounds caused by nonmonotonic rate-level functions in the IC units. The 200-ms signal was presented at a repetition rate of 2.5/s, whereas the noise was presented continuously so that responses were obtained for both a 200-ms signal-plus-noise interval and an immediately following 200-ms noise-alone interval. The stimulus was repeated 16 or more times at each noise level. A different noise waveform was used on each trial (signal-plus-noise and noise-alone pair), but the same set of waveforms was used for all noise levels. The response was measured as a function of noise level for several noise azimuths, as time permitted. The signal azimuth was initially fixed at a location giving a strong excitatory response, usually on the side contralateral to the recording site (positive azimuths). The responses for other signal azimuths, including some at unfavorable azimuths, were then measured if time permitted. The signal azimuths used were -90 , 0 , 45 , and 90° , and the noise azimuths were -90 , -54 , -45 , 0 , 18 , 36 , 45 , 54 , 72 , and 90° .

Data analysis

A unit was included in this study if it had a low CF_{TC} (≤ 2.5 kHz), gave a sustained response to chirp trains at some signal azimuth, and was sensitive to ITD. We considered a unit ITD sensitive if the noise-delay function was modulated by $\geq 50\%$ (i.e., if the minimum discharge rate was less than half of the maximum rate). We measured the rate in a window that began 5 ms after the onset of the 200-ms noise burst and lasted 190 ms.

To determine the best ITD and best frequency (BF_{ITD}) for each unit, we fit the noise-delay function with a Gabor function (McAlpine and Palmer 2002b), which is a sinusoid with a Gaussian envelope

$$G = Ae^{-\frac{(ITD - BITD)^2}{2s^2}} \cos[2\pi BF_{ITD}(ITD - BITD)] + B$$

The least-squares fit was obtained using the Levenberg–Marquardt algorithm (Matlab’s *leastsq* function). The Gabor fit gives estimates of the best frequency (BF_{ITD} in the equation

above) and the best ITD ($BITD$) of the unit. The additional parameters, A , B , and s , give the amplitude of the curve, its DC offset, and the rate of decay of the Gaussian envelope away from the best ITD, respectively. We constrained the best ITD for the fine structure (inside the cosine) to be equal to the best ITD for the envelope (in the exponential) to facilitate modeling studies described elsewhere (Lane 2003); this constraint amounts to assuming that the characteristic phase (Yin and Kuwada 1983) is zero. Additionally, the Gabor function was half-wave rectified to ensure that the rate is never negative. A noise-delay function and its Gabor fit are shown in Fig. 2A. We show the Gabor function without half-wave rectification to emphasize the location of the worst ITD, which is not obvious from the rate function itself.¹

To facilitate comparisons between the responses of units having different CFs and best ITDs, we define a *relative IPD* (ϕ) in cycles by the equation

$$\phi = (ITD - BITD) BF_{ITD}$$

Normalizing the axes by both the best ITD and the BF_{ITD} gives a metric where the unit's preferred phase is 0 cycles, corresponding to the unit's best ITD, and the unit's worst phase is -0.5 cycles, corresponding to the unit's worst ITD given by

$$WITD = BITD - 1 / (2 BF_{ITD})$$

Use of a minus sign yields the $WITD$ closest to the midline because the best ITD is usually positive (contralateral). In Fig. 2A, the top axis shows the noise ITD in microseconds, whereas the bottom axis shows the relative phase in cycles. We use cycles for relative IPD to avoid confusion with azimuth, which is expressed in degrees.

In a few cases (four out of 31) for which the noise-delay functions were not sampled finely enough, the Gabor fit predicted best and worst azimuths at obviously incorrect locations. In these cases, the best ITD and BF_{ITD} were adjusted manually to give appropriate best and worst ITDs to match the best and worst azimuths.

Masked thresholds

To obtain neural thresholds that can be directly compared with psychophysical thresholds, which are based on a percentage correct criterion near 75%, masked threshold was defined as the lowest signal-to-noise ratio (SNR) at which the signal can be detected for 75% of the stimulus repetitions. Two different response metrics, mean rate and synchronized rate, were used to define detection thresholds. Mean rate is simply the number of spikes in the measurement window (from 5 to 195 ms post-stimulus onset), and the synchronized rate (Kim and Molnar 1979) is the Fourier component of the peristimulus time histogram at the signal repetition rate, 40 Hz. The synchronized rate, which is also the mean rate multiplied by the synchronization index or vector strength (Goldberg and Brown 1969), contains information about the spike timing as well as the number of spikes. Figure 3 shows both the mean rate (*third row*) and the synchronized rate (*fourth row*) as a function of noise level for one unit. The 200-ms chirp-train signal was held at 43 dB SPL, and the locations of the signal and the masker differ for each column of panels. To determine the masked threshold, we calculate the percentage of stimulus presentations for which the detection metric (mean rate or synchronized rate) is greater in the signal-plus-noise window compared with the noise-alone window (Fig. 3, *bottom row*). To improve the reliability of threshold estimates, the percentage-correct values were converted to z-scores by a Gaussian transform (Green and Swets 1974), smoothed with

¹In some experiments, for the noise-delay function, the noise burst was shorter in the left ear (20 ms instead of 200 ms) than in the right ear because of a programming error. In this case, the rate window for the ITD-sensitivity analysis was adjusted to match the length of the shorter noise burst. A Gabor function could still be fit for all but one unit, which was discarded from the database.

a three-point triangular filter, and then converted back to a percentage value. Because a signal can be detected through either an increase or decrease in rate (Jiang et al. 1997b), thresholds (circles in Fig. 3, *bottom row*) can occur when the percentage curve crosses either 75% (an increase in rate, see dashed lines) or 25% (a decrease in rate). This criterion gives the highest noise level or, equivalently, the lowest SNR, where the signal can still be detected 75% of the time.

We determined confidence intervals for the masked thresholds using bootstrapping methods (Efron and Tibshirani 1993). For each noise level, we sample the responses to each stimulus presentation with replacement, obtaining a new, “bootstrapped” set of spike trains. We then recompute the percentage curves for the new set of spike trains and recalculate the thresholds. The threshold is recomputed in this way 100 times. The error bars for the masked thresholds are then the range between the 10th and the 90th percentiles. Reliably estimating the thresholds was difficult because the percentage curves could be nonmonotonic, especially when the signal suppressed the noise response. Consequently, we required that, for a 25% threshold to be accepted, the signal had to decrease the overall rate below 25% for 80 out of 100 of the bootstrapped percentage curves. If this criterion was not met, then the 75% threshold was used. This requirement eliminated very low threshold SNRs that occurred as a result of spurious estimates of percentage-correct points. For the actual threshold estimate, we took the median of all the bootstrapped percentage curves and determined the threshold for this median curve.²

RESULTS

We measured masked thresholds for 60 units in the inferior colliculus of 15 animals. After eliminating the units that did not show a sustained response to the chirp train at some signal azimuth, were not held for long enough to measure thresholds for at least four signal/masker configurations, were not ITD sensitive, and/or did not have CF_{TC} values <2.5 kHz, there remained 31 units with characteristic frequencies (CF_{TC}) between 100 and 2,300 Hz. As found in both guinea pigs (McAlpine et al. 2001) and cats (Hancock and Delgutte 2004; Joris et al. 2004), the best ITDs for our units tend to decrease with increasing BF_{ITD} , and a majority of the best ITDs and worst ITDs are outside the physiological range. For our purposes, the physiological range here is defined as $\pm 290 \mu s$, which corresponds to the ITDs measured at $\pm 90^\circ$ in our HRTFs.

When azimuth is varied, other localization cues present in the HRTFs (interaural level differences and spectral cues) vary as well as ITD. For the majority of our units ($n = 19$), we compared the response for changes in noise *azimuth* and changes in noise *ITD*. Figure 2B shows one unit’s rate response as a function of both noise azimuth (solid line) and noise-ITD (dash-dot line). The two responses were similar (Fig. 2B), although the rate was higher for the ITD-only condition. For the 19 units in our sample for which this measurement was taken, Fig. 2C compares the rate when only ITD was varied to the rate when the noise azimuth was varied. Except for very low discharge rates, the two responses are similar for all of these units, indicating that ITD largely determines these units’ azimuth sensitivities.³

²This somewhat unusual technique was selected because the bootstrap threshold distributions were sometimes bimodal, making their median an unreliable measure of central tendency; in contrast, the percentage-correct curves had well-behaved, unimodal distributions.

³To obtain the noise-alone response to changing ITD only, the noise waveform was filtered through the first principle component of the HRTFs (essentially a version of the original HRTFs that has been smoothed in the frequency domain) to remove ILD and spectral cues, and the ITD was varied. At each azimuth, the ITD was adjusted to match the delay yielding the maximum of the interaural cross-correlation of the smoothed HRTFs for each azimuth (see Litovsky and Delgutte 2002 for details).

Dependency of single-unit masked thresholds on signal and masker azimuths

Based on previous physiological results (Caird et al. 1991;McAlpine et al. 1996), we expect that masked thresholds would change with signal and masker azimuth. Fig. 3 shows a typical unit's responses to the signal in noise and the noise alone as a function of noise level for three signal and masker configurations. The unit had a BF_{ITD} of 740 Hz and a best ITD of 290 μ s, which corresponds to +90°; because the unit's worst ITD (about -380 μ s) was outside the physiological range, its worst azimuth was -90°.

The first row sketches the three signal and masker configurations: signal and masker co-located at +90° (*column A*, S90, N90); signal at +90° and noise at -90° (*column B*, S90, N-90); and signal at -90° and noise at +90° (*column C*, S-90, N90). The *second row* in Fig. 3 shows the temporal discharge patterns for the signal-plus-noise interval (S+N) and the noise-alone interval (N) as a function of noise level. In these dot rasters, every dot represents a spike, and the solid lines separate the blocks of stimulus presentations for each noise level. As the noise level is raised, the signal response can be either overwhelmed by the noise response (excitatory or "line-busy" masking, *columns A* and *C*) or suppressed by the noise (suppressive masking, *column B*).

These rasters show a wide variety of potential cues for detecting the signal in noise. The types of cues available depend on the signal and masker configuration. For the signal at +90° in low-level noise (Fig. 3, *columns A* and *B*), the unit shows a highly synchronized response to the 40-Hz repetition rate of the chirp train. For this signal azimuth, the response to the signal plus noise is always greater than the noise-alone response. In contrast, the response to the signal at -90° (*column C*) is much weaker, consisting of only an onset response at the lowest noise level. At moderate noise levels for S-90, the signal suppresses the noise response, and a weak response at the signal repetition rate can be discerned, possibly reflecting a recovery from suppression during the silent periods between individual chirps in the train (see Fig. 1A). The signal can also alter the distribution of spike arrival times without causing a change in mean firing rate (*column A*). It is possible that any or all of these cues could be used to detect the signal, and an optimal central processor would use the best combination of cues, perhaps through the use of a signal template. Because we had only a few stimulus presentations for each stimulus condition, developing a reliable signal template was not feasible; instead, we chose to detect the signal through more traditional methods involving changes in mean rate and spike synchrony.

In the following, we first present all the results for thresholds based on mean rate and then discuss how synchronized rate thresholds differ at the end of the RESULTS section. As described in METHODS, the rate-based masked threshold is the highest noise level (or lowest signal-to-noise ratio) where the signal can still be detected 75% of the time, based on either an increase (75% mark in Fig. 3, *row 5*) or a decrease (25% mark in Fig. 3, *row 5*) in mean rate. The rate thresholds for the unit of Fig. 3 (shown as circles in the rows) differ substantially for the three signal and masker configurations. Specifically, the threshold for the co-located condition S+90, N+90 (*column A*) is about 18 dB poorer than the threshold for S+90, N-90 (*column B*), and the threshold for S-90, N+90 (*column C*) falls between the two despite the weak signal response in this case. In *column C*, the signal causes an increase in rate at the lowest noise levels; then, once the noise level is raised a few decibels, the signal can be detected through a *decrease* in rate (see dot raster; the signal's presence is shown by the *suppression* of the noise response). The signal can still clearly be detected for noise levels >49 dB (as shown in the dot raster), making about 52 dB the masked threshold. It is apparent from this example that by only allowing the signal to be detected through *increases* in rate (the 75% mark), the threshold signal-to-noise ratios for individual neurons would be systematically overestimated (Jiang et al. 1997).

Figure 4 (*top*) shows rate-based masked thresholds as a function of noise azimuth for the unit in Fig. 3 for four different signal azimuths. When the signal is at either +45° or +90°, moving the noise away from the signal to the ipsilateral side (negative azimuths) improves thresholds by ≤ 20 dB. When the signal is at 0°, thresholds also improve as the noise is moved away from the midline to the ipsilateral side, but they become slightly worse as the noise moves to the contralateral side (positive azimuths). For these three signal locations (S90, S45, and S0), the worst thresholds occur when the noise is placed near +90°, regardless of the signal location. However, when the signal is placed at -90°, the pattern is different: the thresholds increase slightly and then decrease as the noise is moved away from the signal.

Previous experimental and theoretical findings with the BMLD paradigm (e.g., Caird et al. 1991; Colburn 1973, 1977a, Colburn b; Jiang et al. 1997a,b) suggest that finding a correlate of SRM may require looking across a population of neurons and that individual unit thresholds are determined by the unit's azimuth/ITD preference. Indeed, this unit's thresholds do not consistently improve with spatial separation, even though they can improve by up to 20 dB in some cases. Instead, with the exception of S-90, the worst thresholds for this unit tend to occur when the noise is at the best azimuth (+90°). Figure 4, *bottom*, shows a histogram of the maximum threshold change seen as the masker is separated from the signal for all the units in our sample. For a majority of units, thresholds improve with signal and masker separation, but for some units, the thresholds actually become worse. Generally, the changes in threshold seem to be a consequence of placing the signal and the masker at favorable and unfavorable azimuths, a result similar to the one shown by Caird et al. (1991).

To test the effect of signal and masker separation on the thresholds for all of our units, we determined the worst threshold for each unit and examined how this worst threshold relates to the signal and masker locations. Figure 5 shows the noise azimuth that gives rise to the worst threshold, the "worst-threshold noise azimuth," as a function of both signal azimuth (*top*) and the unit's best azimuth (*bottom*, defined as the azimuth with the relative IPD nearest to 0 within the physiological range). If separation improved thresholds, then the worst threshold should occur when the signal and noise are at the same azimuth, i.e., the worst-threshold noise azimuth and the signal azimuth should be the same. Contrary to this prediction, the correlation between a unit's worst-threshold noise azimuth and the signal azimuth is very low (0.15) and is not significant ($P = 0.2$, two-sided t -test, $n = 68$). Thus the worst thresholds do not necessarily occur when the signal and the masker are co-located. In contrast, the correlation between the worst-threshold azimuth and the best azimuth is much higher (0.57) and is highly significant ($P < 0.001$), indicating that strong excitation by the masker tends to produce poor masked thresholds. Consequently, the individual unit responses do not show a correlate of spatial release from masking, consistent with previous BMLD studies. However, as suggested by the previous results (e.g., Caird et al. 1991; Colburn 1973, 1977a, Colburn b; Jiang et al. 1997a,b), a neural correlate of spatial release from masking may still exist in the response of a population of ITD-sensitive neurons.

Neural population thresholds show SRM consistent with human psychophysics

To test the hypothesis that the population of low-frequency, ITD-sensitive units is sufficient for explaining spatial release from masking at low frequencies, we defined a population threshold based on the "lower-envelope principle" (Parker and Newsome 1998). Specifically, for each signal and noise configuration, the population threshold is the best single-unit threshold in our sample of ITD-sensitive units. The top row of Fig. 6 shows both the individual mean-rate thresholds for all the units in our sample (dot-dash lines) and the population thresholds (thick solid lines) as a function of noise azimuth for three signal azimuths (arrows). The bottom row shows the synchronized rate thresholds, which are discussed later. Unlike the single-unit thresholds, the population thresholds do show a correlate of spatial release from

masking in that they generally improve when the signal and noise are separated. Clearly, the curves do not show perfect spatial release from masking: for example, the thresholds do not improve for the signal at 45° when the noise azimuth is $>45^\circ$, and the improvement for the signal at 0° is not symmetric with respect to the midline. It is not obvious whether obtaining a larger sample of neurons would improve the correlate for the $S45^\circ$ condition, but as units in the opposite IC are expected to have mirror-imaged threshold curves, incorporating units from both ICs would almost certainly eliminate the asymmetry in population thresholds for $S0^\circ$ (see following text). Overall, it seems that the combination of all the unit responses, each with a different azimuth preference, allows for a correlate of spatial release from masking to emerge in the population response.

Figure 7 compares the neural population thresholds to human psychophysical thresholds (from Lane et al. 2004; see figure caption for details). In the psychophysical experiments, which were designed to have stimuli and methods similar to those in the physiological experiments, the signal (a 40-Hz chirp train with chirps that were swept logarithmically from 250 to 12,000 Hz) was low-pass filtered at 1,500 Hz, and the noise was low-pass filtered at 2,000 Hz to restrict listening to the low-CF region. As in the physiological experiments, the signal was fixed and the noise level was varied to find the masked threshold. To take into account differences in head size between cats and humans, we compare thresholds for similar ITDs rather than similar azimuths (see DISCUSSION). For example, a 30° azimuth in humans corresponds to an ITD of $250 \mu\text{s}$, close to the $290\text{-}\mu\text{s}$ ITD produced by a sound source at 90° in cat. Consequently, Fig. 7B shows thresholds as a function of noise ITD with $S30^\circ$ for the human thresholds and with $S90^\circ$ for the cat thresholds. Figure 7A shows thresholds for $S0^\circ$, which corresponds to the near-zero ITD in both species. The thin solid line in Fig. 7A shows the psychophysical thresholds reflected about 0° , essentially completing the curve under the assumption that the listener would have symmetric thresholds. To simulate the effects of having two ICs, we also reflected all of the individual unit thresholds about 0° and then recomputed the population thresholds using both original and reflected thresholds, making the curve symmetric about 0° . Both the shapes and the SNR values for the population threshold curves and the human behavioral data are similar. (The plotted thresholds were *not* vertically shifted to adjust the match between the two data sets, but reflect the true measured values.) There are clear differences between the functions, however. The human behavioral thresholds are better than the neural thresholds when both the signal and noise are near the midline, the neural thresholds improve more rapidly with separation than the human psychophysical thresholds, and the neural thresholds have a fine structure where the thresholds “dip” and then rise. Despite the differences (which are at least in part the result of having a limited number of neurons in the population), the overall match seems reasonably close, suggesting that the low-frequency, ITD-sensitive units in the IC probably serve as a neural substrate for spatial release from masking and that spatial release from masking at low frequencies is processed similarly across species.

Predictions of the cross-correlator model

Having identified a neural correlate of SRM in the population response of ITD-sensitive units, we now focus on whether the responses of these units to SRM stimuli can be predicted by a cross-correlator model similar to the one described by Colburn (1973,1977a,b). The example unit in Fig. 3 shows that, depending on the stimulus configuration, the signal can be detected through either an increase (*A*, *B*) or a decrease (*C*) in rate over the noise-alone response. Furthermore, masking can arise from the noise either suppressing (*B*) or overwhelming (*A*, *C*) the signal response. To test whether this diversity of responses is qualitatively consistent with the cross-correlator model, we implemented a simple cross-correlator model with parameters that matched the CF and best ITD for the unit in Fig. 3 (see caption of Fig. 8 for implementation details). The model response is shown in Fig. 8 for comparison with the unit’s rate response in row 4 of Fig. 3.

When the signal is placed at the azimuth corresponding to the model neuron's best ITD, $+90^\circ$ in this example, the cross-correlator delay compensates for the interaural delay so that the cross-correlation value is high, yielding a large response to the signal at low noise levels (Fig. 8, A and B). In contrast, when the signal is placed at the model unit's worst azimuth (-90° , Fig. 8C), the inputs to the cross-correlator arrive nearly out of phase, producing a weak correlation. When both the signal and the noise are presented simultaneously, the response gradually transitions from the signal-alone response at low noise levels to the noise-alone response at high noise levels. Therefore the overall model response depends on the relative levels and locations of the signal and the masker. Specifically, as expected from the results of Colburn (1973, 1977a,b), introducing a signal at a favorable azimuth tends to cause a higher correlation and an increase in rate over that evoked by the noise alone (Fig. 8, A and B), whereas introducing a signal at an unfavorable azimuth tends to reduce the correlation and decrease the rate (Fig. 8C). Similarly, a masker at a favorable azimuth (Fig. 8, A and C) masks by increasing the correlation so that the signal can no longer be detected, whereas a masker at an unfavorable azimuth masks by decorrelating the inputs from the two ears and bringing the rate down to zero (Fig. 8B). Overall, the model responses shown in Fig. 8 are qualitatively similar to the neural responses shown in Fig. 3. Some differences can also be seen. For example, in Fig. 8A, the signal-plus-noise rate decreases slightly as the noise level increases, an effect that reflects adaptation by the continuous noise in the auditory nerve fiber model response. A similar effect is seen for the neural response in Fig. 3A, but is more pronounced than that in the model. Also, in Fig. 8C, the model predicts that the signal, which is at the worst azimuth, should produce no response, but the unit gives an onset response to the signal. Nevertheless, the cross-correlator model predicts that both *how noise masks the signal* and *how the signal is detected* depend systematically on the positions of the signal and the masker relative to the unit's best and worst azimuths, and the results for this neuron seem qualitatively consistent with the model.

Effect of noise on signal response depends on noise azimuth

To test whether the units' behavior is *quantitatively* consistent with the cross-correlator model, we define two metrics: one that characterizes how the noise masks the signal response and one that characterizes the effect of the signal on the noise response. The first metric, the "masking type index" (MTI) quantifies whether the noise masker overwhelms or suppresses the signal response at threshold. The MTI is the difference between the signal-in-noise rate at threshold, $R(S + N_{Th})$, and the approximate signal-alone rate, $R(S)$, the signal response with the noise at the lowest level. This difference is then normalized by whichever of the two rates is larger

$$MTI = \frac{R(S + N_{Th}) - R(S)}{\max [R(S + N_{Th}), R(S)]}$$

The MTI ranges from -1 for purely suppressive masking to $+1$ for purely excitatory masking. To illustrate this metric, the third row of Fig. 3 shows the mean rates for the signal in noise and the noise alone for our example unit. In column B, the masker suppresses the signal response without exciting the unit, and the MTI is -0.86 , indicating strong suppressive masking [the MTI is not -1 because $R(S + N_{Th})$ is measured at threshold where the signal response is not completely suppressed]. We also computed the MTI for our model unit in Fig. 8. In this case (Fig. 8B), the model MTI has a similar value (-0.83), again indicating strong suppressive masking. In Fig. 3, column C, the noise response becomes so strong that it overwhelms the weak, largely suppressive signal response, and the MTI is 0.79 , indicating excitatory masking. The model response in Fig. 8 shows even stronger excitatory masking, giving an MTI of 0.90 . In Fig. 3, column A, although both the signal and masker are at the best azimuth, the masker at first decreases the signal response, then eventually overwhelms it; in this case, the MTI is slightly negative (-0.33) as a result of this initial decrease. For the model unit (Fig. 8A), the

response also shows a slight initial decrease (attributed to adaptation in the auditory nerve fiber model), but recovers more quickly, making the MTI slightly positive (0.1).

To test the predictions of the cross-correlator model for our entire sample of units, we examine how the MTI depends on noise azimuth for all of the units. The model predicts that, for the masker at a favorable azimuth, the number of coincidences increases with noise level to produce excitatory masking (MTI >0). For the noise at unfavorable azimuths, we expect the noise to decorrelate the signal response and produce suppressive masking (MTI <0), providing the signal response is sufficiently strong. We plot the MTI as a function of both the noise azimuth (Fig. 9A) and the noise relative IPD (Fig. 9B) for all the units in our sample. We show results only for favorable signal azimuths ($|\varphi_s| < 0.1$) to see the effect of the masker on a strong signal response. When the noise is in the ipsilateral hemifield (negative azimuths), the masking is usually suppressive (MTI near -1). However, noise in the contralateral hemifield (positive azimuths) can mask through either excitation or suppression. This dependency of MTI on noise azimuth may arise from the fact that most of the units have their best azimuths on the contralateral side. To test this possibility, Fig. 9B replots the MTI as a function of noise relative IPD φ_n , thereby normalizing for differences in best azimuths across units (see METHODS). By definition, favorable azimuths have relative IPDs near 0, whereas unfavorable azimuths have relative IPDs near -0.5. Figure 9B shows that the MTI across the population changes abruptly around $\varphi_n = -0.25$: when the noise is at an unfavorable azimuth ($\varphi_n < -0.25$), the masking is always suppressive, as expected; however, when the relative IPD of the noise is favorable ($\varphi_n > -0.25$), the masking can be either excitatory or suppressive, despite the fact that the signal and masker are both at favorable azimuths. It seems that the masker can reduce the overall rate even when it is placed at a “favorable” azimuth, contrary to the predictions of a simple cross-correlator model. Because the signal and the masker have similar spectra, effects such as lateral (cross-frequency) inhibition or cochlear suppression are not likely to explain this result. Instead, this result suggests that additional processing beyond cross-correlation, probably some type of temporal processing, affects the relative responses to the signal and the noise in some units. In the DISCUSSION, we propose a likely candidate for such additional processing.

Effect of signal on masker response depends on signal azimuth

The second metric used to compare the neural responses to the predictions of the cross-correlator model is the “signal effect index” (SEI), which characterizes the effect of the signal on the noise response. The SEI is again a normalized difference, this time between the S + N rate, $R(S + N_{Max})$, and the N rate, $R(N_{Max})$, at the noise level N_{Max} where the signal causes the largest change in rate

$$N_{Max} = \arg \max [|R(S + N) - R(N)|]$$

$$SEI = \frac{R(S + N_{Max}) - R(N_{Max})}{\max [R(S + N_{Max}), R(N_{Max})]}$$

To obtain the SEI, we consider only changes in rate that have the same sign as the change in rate caused by the signal at threshold. For an SEI of +1, the signal causes the largest increase in rate when there is no response to the noise alone; for an SEI of -1, the signal completely suppresses the response to the noise. For example, in Fig. 3, columns A and B, the signal causes the largest change in rate (a positive one) at the lowest noise levels, giving SEIs of +1 in both cases, but in column C, the largest change (a negative one) occurs for noise levels around 45 dB SPL, giving an SEI of -0.63. This value does not reach -1 because the signal does not completely suppress the noise response. The model response in Fig. 8 gives similar results: the SEI values for A and B are both near +1 as in the data, whereas in C, the SEI is negative (-0.24),

indicating that the signal suppresses the noise response. Here, the suppression caused by the signal was greater for the neuron than for the model.

Figure 9 shows the SEI as a function of both signal azimuth (Fig. 9C) and the signal relative IPD (Fig. 9D). Only results for favorable noise azimuths ($|\varphi_n| < 0.1$) are shown so that the effect of adding the signal can be reliably evaluated. When the signal is near the midline or in the contralateral hemifield (positive azimuths), it is detected through an increase in rate in most cases (108 out of 122 thresholds in Fig. 9C; many of the points are plotted on top of each other, especially near 1). The median SEI (solid line) is near 1 in these cases. In contrast, when the signal is placed at -90° , it is usually detected through a decrease in rate (11 out of 14 cases). The SEIs never reach -1 , but are usually near -0.5 , indicating that the signal does not completely suppress the noise response. In Fig. 9D, the SEI is replotted against the signal relative IPD to normalize for cross-unit differences in best ITD and CF. Placing the signal at a favorable azimuth ($\varphi_s > -0.25$) almost always increases the overall rate (106 out of 122 thresholds), as expected, but occasionally decreases the rate. Signals at unfavorable azimuths ($\varphi_s < -0.25$) decrease the overall rate, as expected, in a majority of cases (nine out of 14), but can also increase the rate in some cases. When combined with the MTI results, these results suggest that, whereas the cross-correlator model gives useful predictions for many units, some additional processing is affecting the relative rates to the signal and the masker in a substantial fraction of the units.

Best thresholds occur for signal at best azimuth

For a majority of the units in our sample, $+90^\circ$ is near the best ITD, and -90° is near the worst ITD. For such units, placing the stimuli at these azimuths makes the neural inputs arrive as near to in phase or as near to out of phase as possible inside the physiological range. Therefore placing the signal and noise at $+90^\circ$ and -90° in various combinations is analogous to the well-studied N0S0, N0S π , and N π S0 conditions for units having their best ITD near 0 (0-ITD units). In the modeling studies by Colburn (1973, 1977a,b), the 0-ITD units were the most sensitive to changes in interaural correlation in the traditional BMLD conditions (N0S0 compared with N0S π). Specifically, the in-phase conditions (N0, S0) for the 0-ITD units are similar to placing the stimulus at the best azimuth in this study because the inputs from the two ears would arrive in phase at the coincidence detector. The out-of-phase conditions (N π , S π) for a 0-ITD unit are similar to placing the stimulus at the worst azimuth for our units because the inputs would arrive nearly out of phase. The psychophysical thresholds for the N0S π condition are better than the N π S0 for a wide variety of signals, and both thresholds are better than those for the N0S0 condition (Durlach and Colburn 1978). Using a 500-Hz pure-tone signal, Jiang et al. (1997) found a correlate of this threshold hierarchy in the average thresholds of IC units. Furthermore, as predicted by the Colburn model, they showed that for the majority of units, the N0S π neural thresholds were better when adding the signal *decreased* the overall response, indicating that the best thresholds occur when the signal decorrelates the noise response. If these results could be extended to our experiments, one would expect the best thresholds to occur when the signal is placed at the *worst* azimuth, usually near -90° , and the noise is placed at the *best* azimuth, usually near 90° , so that the signal is detected by decorrelating the noise response.

Figure 10 shows the thresholds plotted against CF for 11 units in three animals for which we measured responses for the signal and the noise on opposite sides of the head. Each unit's response was tested with the signal near the best azimuth and the noise near the worst (S90, N -90 , white squares) as well as the signal near the worst azimuth and the noise near the best (S -90 , N90, black circles). The thresholds for the signal and masker co-located near the best azimuth at 90° (S90, N90, \times 's) are also shown for the same units. For all the units, the thresholds for the signal placed near the *best* azimuth, the condition most like N π S0, are always at least

as good as the thresholds for the signal placed near the *worst* azimuth, the condition most analogous to $N0S\pi$ (white squares are always lower than black circles). This relationship is the reverse of the one expected from previous psychophysical and physiological studies of BMLD with pure-tone signals. The $S90$, $N-90$ thresholds are always better than the co-located thresholds as expected from the BMLD psychophysical and physiological results (white squares are always lower than \times 's); however, the $S-90$, $N90$ thresholds, which might be expected to be the best thresholds overall, are not necessarily even as good as the co-located thresholds (\times 's are sometimes lower than black circles). It should be noted, however, that we have biased our results somewhat by selecting only neurons that gave a sustained response to the chirp at some azimuth; searching for units that showed a response to the signal by suppressing the noise response would be difficult at best. Nevertheless, in contrast to previous findings and model predictions, the best thresholds for these stimuli do not seem to occur when the signal decorrelates the noise response, but rather when the signal correlates the anticorrelated noise response.⁴

Additionally, we directly compared the responses to a chirp train signal and a 500-Hz pure-tone signal for one unit (BF_{ITD} of 600 Hz, best ITD of 400 μ s). In all conditions tested, the SNR at threshold was better for the tone signal than for the chirp train (Fig. 11). Presumably, the greater sensitivity for tones is explained by the fact that only a small fraction of the chirp energy passes through the peripheral auditory filter centered at the CF, whereas almost all of the tone energy passes through the filter when the tone frequency is near the CF, as was the case here.

Focusing now on the directional dependency of thresholds, the threshold SNR for the pure-tone signal placed at -90° was *better* than the threshold for the tone placed at 90° (Fig. 11), consistent with the Jiang et al. (1997a,b) results. In contrast, at the same overall signal level, the SNR at threshold for the chirp train was better at $+90^\circ$ than at -90° (unfilled squares are lower than filled circles), consistent with the pattern seen in our units. However, because there are differences between the pure-tone and the chirp-train signals in the amount of signal energy that passes through the peripheral auditory filters, we varied the level of the chirp train to see whether the threshold hierarchy might depend on signal level. As the level of the chirp train was raised from 38 to 43 dB SPL, the threshold SNR for $S90^\circ$, $N-90^\circ$ remained about the same (Fig. 11, white squares), but the threshold SNR for $S-90^\circ$, $N90^\circ$ improved considerably (Fig. 11, black circles). At higher chirp levels, thresholds for the two configurations were similar. It seems that a low-level signal at -90° does not sufficiently suppress the noise response to be detected at any noise level, but the amount of suppression increases with signal level, so that the thresholds reach about the same values as when the signal is at 90° at higher levels. Thus although overall signal level can affect the threshold SNR, the fundamental result holds across levels: the chirp-train thresholds for the signal at favorable azimuths are always at least as good as thresholds at unfavorable ones. These differences in results for the tone and chirp train signals do not seem to be caused by differences in signal energy through the peripheral filters because we were unable to find a signal level where the $S-90^\circ$ thresholds were significantly better than the $S90^\circ$ thresholds. Instead, other differences in the spectral or temporal properties of the signals are likely to affect the threshold hierarchy for this unit.

⁴To test directly whether thresholds improve when the signal is detected by a decrease in rate, as in Jiang et al (1997a,b), we looked at the thresholds for units where the signal was detected by a decrease in rate (showed a negative SEI) at some signal azimuths and by an increase in rate (positive SEI) at others (11 units, not all the same as in Fig. 10). Because we measured thresholds for many stimulus conditions, each of these conditions showed a range of thresholds. We therefore compared the best threshold for a given unit when the signal increased the rate to the best threshold for that unit when the signal decreased the rate. In contrast to the results of Jiang et al. (1997a,b), for all the units, the best threshold when the signal increased the rate was at least as good as the best threshold when the signal decreased the rate (not shown).

Mean-rate thresholds compared with synchronized-rate thresholds

To examine the role of spike timing in these unit responses, we evaluated the effects of using synchronized rate instead of mean rate to define both the individual unit thresholds and the population thresholds. The fourth row of Fig. 3 shows the synchronized rate for the S+N response as well as the N response as a function of noise level for our example unit. Because the synchronized rate is the mean rate multiplied by the vector strength at 40 Hz, it is equal to the mean rate (row 3) when the response is perfectly synchronized to the signal (all three conditions at low levels, column B at all levels). Otherwise, the synchronized rate is *always less* than the mean rate, an effect that can make the thresholds better or worse, as discussed in the following text. Row 5 in Fig. 3 shows the percentage of stimulus presentations for which the signal-plus-noise response is greater than the noise-alone response as a function of noise level for both the mean rate (dots) and the synchronized rate (\times 's). Threshold is the noise level where the signal can be detected for 75% of the stimulus presentations through either an increase or decrease in mean rate or synchronized rate (dotted 25 and 75% lines). The circles show the thresholds for each condition: the synchronized rate thresholds can be better than (column A), the same as (column B), or worse than (column C) the mean rate thresholds.

When both the signal and the noise are excitatory (Fig. 3, column A), the use of timing information can improve thresholds because introducing the signal causes the spike times to become synchronized to the signal repetition rate without a concomitant change in mean rate compared with the noise-alone condition. In contrast, when the noise suppresses the signal response (Fig. 3, column B), the percentage curves and the masked thresholds for the mean rate and synchronized rate are similar. Here, because every spike is phase locked to the signal envelope, spike timing provides no additional information over the rate. Finally, when the noise is excitatory and the signal is suppressive (Fig. 3, column C), the synchronized rate threshold is actually *worse* than the rate threshold because the signal acts to decrease the rate but increase the vector strength, reducing the overall change in synchronized rate when the signal is added to the noise. In the case of Fig. 3, column C, the consequence of reducing the difference was considerable: the percentage curve for the synchronized rate did not reach 25%, so that the 75% threshold had to be used. This 75% synchronized rate threshold was 18 dB worse than the 25% mean rate threshold.

Figure 12, A and B, shows both the synchronized-rate thresholds and the mean-rate thresholds as a function of noise azimuth for S+90° (A) and S−90° (B) for the unit in Fig. 3. Again, the synchronized-rate thresholds can be the same as, better than, or worse than the mean-rate thresholds, depending on whether the signal and the masker are excitatory or suppressive. However, at the signal and masker locations that give the best thresholds (those with the lowest SNR; in this case, S+90° and negative noise azimuths), the mean-rate and the synchronized-rate thresholds are essentially the same. This observation is typical for our sample of units because the best thresholds generally occur when the signal is excitatory and the masker is suppressive, the very case when there is little difference between the two threshold metrics.

Figure 12, C and D, shows how the differences in synchronized-rate and mean-rate thresholds are related to the type of masking for our sample of units. Figure 12C shows the threshold difference as a function of MTI when the signal is detected through an increase in rate (SEI >0). For suppressive masking (MTI <0), there is little difference between the two thresholds; in contrast, for excitatory masking (MTI >0), the synchronized-rate thresholds are usually better because the signal can be detected through a rearrangement of the spike times even if the rate does not change. Finally, when the signal is detected through a decrease in rate (SEI <0, Fig. 12D), the synchronized-rate thresholds are usually worse than the mean-rate thresholds because, as in Fig. 3, the signal *decreases* the overall rate, but *increases* the synchronization, reducing the difference between the signal-plus-noise response compared with the noise-alone response. (In Fig. 12D, the masking cannot be strongly suppressive, i.e., MTI <0, because there

would be no measurable spike discharge to either the signal or the noise, and therefore a threshold could not be determined.)

Consistent with the above results, Fig. 6 shows that the population thresholds based on the mean rate (*top*) and the synchronized rate (*bottom*) are very similar. Except for one spurious synchrony threshold for S45°, the shapes and SNR values for both thresholds are nearly identical. Again, this similarity arises because the best thresholds tend to occur when the noise suppresses the signal response; consequently, the additional information in the spike timing does not improve the population thresholds. Thus the population thresholds appear to be robust with respect to the choice of threshold metric for our highly modulated stimuli.

DISCUSSION

In this paper, we showed that the responses of low-frequency, ITD-sensitive single units in the IC can change dramatically when signals and noise maskers are placed at different spatial locations. Although a cross-correlator model predicts such changes in the response, some unit responses appear to be influenced by additional processing beyond the cross-correlator model. Moreover, consistent with results shown for the BMLD paradigm (Caird et al. 1991; Colburn 1973, 1977a, Colburn b; Jiang et al. 1997a, b), we show that individual units' masked thresholds depend primarily on the unit's azimuth preference. Because of the variety of azimuth preferences in our population, the best thresholds in the population as a whole show neural correlates of spatial release from masking.

Unit responses compared with predictions of the cross-correlator model

Although many of the neuron responses were qualitatively consistent with a cross-correlator model of ITD processing, many others appeared to be influenced by additional processing beyond cross-correlation. In general, the deviations from the predictions of the cross-correlator model were consistent with an excess responsiveness to the chirp signal compared with the noise masker. Specifically, the signal could be detected by an increase in rate over the noise-alone response, even when it was placed at an unfavorable azimuth where the cross-correlation model predicts fewer coincidences and therefore a lower rate (Fig. 9D). Similarly, the masking could be suppressive even when the noise was placed at a favorable azimuth where it would be expected to mask by creating an overabundance of coincidences (Fig. 9B). This difference in responsiveness between the signal and the noise must reflect some physical difference between the two stimuli. Although the two stimuli have similar broadband spectral envelopes, they differ in their temporal characteristics: the chirp signal has a pronounced 40-Hz amplitude modulation (AM), whereas the noise masker has a flat envelope (although some envelope modulation would be introduced by peripheral auditory filtering). One striking property of IC units is that many have band-pass, rate-based modulation transfer functions (rMTFs), i.e., they fire maximally for a particular modulation frequency (for a thorough review see Joris et al. 2004). This property of IC neurons contrasts with those in lower auditory nuclei, which tend to have flat rMTFs. This transformation is not restricted to high-frequency units: a recent study by Sterbing et al. (2003) shows that low-frequency ITD-sensitive units can have higher or lower rates depending on the modulation frequency of the stimulus envelope. Because the 40-Hz modulation frequency of the chirp train lies within the pass band of many IC units, the band-pass sensitivity of IC unit's rMTFs may account for their greater responsiveness to the modulated chirp signal compared with the noise.

Elsewhere (Lane 2003; Lane et al. 2003, 2004), we show that the IC units' sensitivities to temporal envelope can quantitatively account for the differences between the data and the predictions of the cross-correlator model. Briefly, we first implemented a cross-correlator model, but adjusted the responses to account for the nonmonotonic rate level functions seen in the ICC. We found that this model accounts for the noise-alone response quite well.

However, when the (spectrally similar) signal is added, the response is not well predicted for many of the units, indicating that additional processing beyond cross-correlation is required. When processing that provides sensitivity to different modulation rates was added to the model, the model responses for both the noise-alone and the signal-plus-noise conditions matched the data well.

The departures from the traditional cross-correlation model of ITD processing are important because they may provide clues as to the functional role of the IC. Previous work shows that ITD-sensitive units in the IC do not necessarily act as the cross-correlator model predicts (e.g., Fitzpatrick et al. 2002; Joris 2003; Palmer et al. 1999); nevertheless, the literature tends to emphasize the consistency of these units' responses with cross-correlation predictions. One potential reason that evidence for additional processing is not more prevalent in the literature is the use of stimuli with flat envelopes and constant ITDs. Using stimuli with modulated envelopes is important because most natural sounds including animal vocalizations contain pronounced modulations (Singh and Theunissen 2003), but only a few studies (e.g., Sterbing et al. 2003) exist documenting interactions between ITD sensitivity and AM sensitivity. The importance of envelope modulations for ITD processing in the IC is consistent with evidence for the importance of temporal processing in the IC for human speech (Delgutte et al. 1996), musical sounds (McKinney et al. 2000), and bat echolocation sounds (Covey and Casseday 1999). One of the few well-documented departures from cross-correlation occurs in response to stimuli with time-varying ITD, and the processing involved appears to arise in the IC itself (McAlpine and Palmer 2002a; Spitzer and Semple 1998). Computational models (Borisjuk et al. 2002; Cai et al. 1998a,b) suggest that temporal processing, such as adaptation or interaction between excitation and inhibition, can explain the time-varying ITD results. As shown by Nelson and Carney (2004), similar processing may also give rise to sensitivity to AM.

Individual unit responses and thresholds

As expected from experimental and theoretical results for BMLD stimuli (Caird et al. 1991; Colburn 1973, 1977a, 1977b; Jiang et al. 1997a,b), the worst individual thresholds tend to occur when the noise produces a strong excitatory response. Furthermore, the best thresholds tend to occur when the addition of the chirp signal *increases* the overall rate, in contrast to the results of Jiang et al. (1997a,b) for pure-tone signals. However, the two sets of results are not necessarily contradictory. When tested in one unit, we found the same threshold hierarchy as that of Jiang et al. (1997a,b) and Palmer et al. (2000) for a pure-tone signal, although when the chirp-train signal was used, the thresholds were again better when the signal was placed at a favorable azimuth. By measuring the thresholds for the chirp train at different signal levels, we showed that the differences in threshold hierarchy for the two signals do not seem to be caused by differences in the amount of energy passing through the peripheral filters, but more likely result from differences in the signals' temporal or spectral characteristics. Additional data comparing responses to pure-tone and broadband signals in noise are needed to fully resolve the issue.

Population thresholds

As expected from previous modeling and physiological studies of the BMLD (Caird et al. 1991; Colburn 1973, 1977a, 1977b; Jiang et al. 1997a,b), the population of ITD-sensitive units shows a correlate of SRM even though the responses of individual units primarily reflect their azimuth preference. Because the units have a variety of preferred azimuths, the population thresholds generally improve with separation between signal and masker, and both the actual threshold values and the amounts of masking release are similar for the human behavioral data and the cat neural population data. Overall, the population of low-frequency, ITD-sensitive units in both ICs seems likely to provide a neural substrate for spatial release from masking,

and these results suggest that spatial release from masking at low frequencies may be processed similarly across species.

In this study, we recorded from units in the left IC. Because these units tend to have the best thresholds when the signal is on the right (contralateral) side, our results suggest that the chirp signal is likely to be detected by the IC contralateral to the signal location. Presumably units in the opposite (right) IC would detect signals on the left hemifield, providing a mirror image of our population threshold curves. For a signal directly in front, both ICs would be able to detect the signal. These results suggest that a patient with a lesion in one IC would have difficulty detecting broadband signals in noise when the signal is on the side *contralateral* to the lesion. In contrast, the previous results with pure-tone signals (Jiang et al. 1997) predict that the patient's deficit for a pure-tone signal would occur when the signal is *ipsilateral* to the lesion.

Anesthesia

One difficulty with our study is that neurophysiological results from anesthetized cats are compared with human behavioral data. Barbiturate anesthesia alters the responses of units in the inferior colliculus, and its effects are often described as being attributed to increased GABAergic inhibition (e.g., Kuwada et al. 1989). Our anesthetic is a mixture of Dial (a barbiturate) and urethane, which acts on different receptor channels (Hara and Harris 2002), so the overall effects are difficult to predict. For example, inhibition is known to shape rate-level functions (Sivaramakrishnan et al. 2004), which could in turn affect how the signal-plus-noise response and the noise-alone response depend on level. The best thresholds occurred for units with monotonic rate-level functions (not shown), the most common type of neuron in awake preparations. As a result, the best thresholds and therefore the population response may be robust to the effects of anesthesia. Additionally, barbiturate anesthesia has been shown to affect the ITD sensitivity of individual units in the IC (Kuwada et al. 1989). However, the rate-ITD functions in the unanesthetized rabbit show the same general types of responses as those in the anesthetized cat. Now that we have studied these units intensively in an anesthetized preparation, future work could test, in a targeted way, whether the results shown here hold in awake preparations.

Species differences

Another potential issue that arises in comparing human behavioral thresholds with cat neural population thresholds is the differences between the two species. For one, the two species have different head sizes. To some extent, we took this factor into account by comparing thresholds for similar ITDs instead of similar azimuths. Previous studies (Hancock and Delgutte 2004; Joris et al. 2004; McAlpine et al. 2001) have shown that the distribution of best ITDs is similar in cats and in guinea pigs despite the difference in head size between the two species. If humans were similar to cats and guinea pigs in this respect, differences in head size might not pose a serious problem to our results. Additionally, Shera et al. (2002) suggest that human cochlear filters may be more than twice as sharp as those of cats. Sharper frequency tuning is expected to make the random envelope modulations for the noise more pronounced, possibly reducing the difference in responsiveness we observed between the signal and the noise. Models, such as the one developed in Lane (2003), could be used to address the effects of the species differences.

Neural metrics for signal detection

In an effort to evaluate the possible contribution of spike timing to detection thresholds, we compared masked thresholds based on mean discharge rate versus synchronized rate. The synchronized-rate thresholds of individual units could be lower than, the same as, or higher than the corresponding mean-rate thresholds. It is possible that different methods of

incorporating spike timing, such as comparing the responses to a signal template, might change the results for individual unit thresholds. For example, at a higher signal level, detection methods based on spike timing may become essential because the rate responses may be saturated in a majority of units. On the other hand, in the case when the excitatory signal response is masked by suppression, the individual thresholds are unlikely to change regardless of the threshold metric used as long as every spike is phase locked to the stimulus envelope. This observation is especially important because these thresholds are usually the most sensitive and would likely remain so regardless of the threshold metric. Consequently, the population thresholds are robust to the choice of detection metric because the best thresholds across the population tend to occur when the noise suppresses the phase-locked signal response.

Population thresholds and pooling of information across neurons

We used the most sensitive unit in our population as our estimate of the population threshold, following the “lower envelope principle” reviewed by Parker and Newsome (1998). The most sensitive neurons in a neural population often predict psychophysical performance in sensory neurophysiology; examples include Mountcastle et al. (1972) in somatosensation, Bradley et al. (1987) in vision, and Delgutte (1990) in audition. An alternative to the lower envelope principle is to assume that the central processor can pool the individual neuron responses across the population. Provided the information is processed in a meaningful way, the performance based on pooled responses must be at least as good as the lower envelope performance because the central processor has access to the most sensitive neuron as well as all the others. In some cases, the performance improvements achieved by optimal pooling can be quite significant: for example, the pooled performance based on auditory nerve fiber responses for intensity discrimination predicts the “near miss” to Weber’s law for pure tones better than the lower envelope performance (for a review see Delgutte 1996). Nevertheless, pooling the unit responses was not required in this study to achieve thresholds comparable in both magnitude and shape to human behavioral thresholds, indicating that, if pooling occurs, it is not necessary to improve detection.

The population thresholds for the BMLD conditions with pure-tone stimuli (Jiang et al. 1997a; Palmer et al. 2000) were reported as averages across the population of units, rather than according to the lower envelope principle. Nevertheless, consistent with our results, only a few neurons responded to the tone signal at low signal levels, corroborating the idea that a few neurons dictate performance at threshold. Results from other studies of ITD-sensitive neurons in the IC indicate that understanding how the information is combined across neurons is not straightforward. Consistent with the lower envelope hypothesis, Shackleton et al. (2003) showed that human ITD acuity for broadband noise near the midline can be accounted for by the performance of the most sensitive neurons in the guinea pig IC. However, Shackleton and Palmer (2004) showed that changes in interaural correlation, a task closely related to ITD discrimination as well as the SRM and BMLD, do not seem to be predicted by individual unit thresholds. Additionally, Hancock and Delgutte (2004) found that pooling responses across IC neurons with different CFs seems necessary to account for ITD acuity away from the midline. To further address this issue, computational studies, such as the one in Lane (2003), allow us to quantitatively assess which neurons determine performance in the SRM task and how performance may be selectively affected by removal of some of these neurons.

In summary, we have shown that the masked thresholds of individual ITD-sensitive units in the cat inferior colliculus are sensitive to the positions of the signal and masker in space. Both the manner in which the noise masks the signal response (suppressive or excitatory masking) and the way that the signal is detected (through an increase or decrease in rate) change systematically with signal and masker azimuth, as predicted by a cross-correlator model. However, additional processing beyond cross-correlation appears to influence the units’

response, perhaps because of sensitivity to the stimulus temporal envelope. Individual unit threshold curves depend on the locations of the signal and masker relative to the neurons' preferred azimuths. Because of the variety of azimuth preferences in the population, the neural population threshold, which is based on the most sensitive unit at each signal and masker configuration, shows a correlate of spatial release from masking. The population threshold curves were similar in both shape and magnitude to human behavioral thresholds, indicating that ITD-sensitive IC units may provide a neural substrate for low-frequency human spatial release from masking. These neurophysiological results, combined with computational studies described elsewhere (Lane 2003; Lane et al. 2003, 2004), take us several steps closer to a quantitative understanding of the neural mechanisms of spatial release from masking and of listening in complex environments in general.

Acknowledgements

The authors thank C. Miller and L. Liberman for surgical work and the assistance of K. Brinsko and K. Hancock for histological preparation. We thank N. Kopco and B. Shinn-Cunningham for the related psychophysical work, and S. Colburn and J. Guinan for many thoughtful discussions related to this paper. We also thank J. Guinan, K. Hancock, Z. Smith, and the three anonymous reviewers whose suggestions have improved this manuscript.

GRANTS

This work was supported by a National Science Foundation Graduate Student Fellowship and by National Institute on Deafness and Other Communication Disorders Grants DC-00361, DC-02258, DC-00038, and DC-05209.

References

- Adams J. Cytochemical mapping of the inferior colliculus. *Assoc Res Otolaryngol Abstr* 1995;160
- Batra R, Kuwada S, Fitzpatrick DC. Sensitivity to interaural temporal disparities of low- and high-frequency neurons in the superior olivary complex. II. Coincidence detection. *J Neurophysiol* 1997;78:1237–1247. [PubMed: 9310415]
- Batra R, Yin TC. Cross correlation by neurons of the medial superior olive: a reexamination. *J Assoc Res Otolaryngol* 2004;5:238–252. [PubMed: 15492883]
- Borisuyuk A, Semple MN, Rinzel J. Adaptation and inhibition underlie responses to time-varying interaural phase cues in a model of inferior colliculus neurons. *J Neurophysiol* 2002;88:2134–2146. [PubMed: 12364535]
- Bradley A, Skottun BC, Ohzawa I, Sclar G, Freeman RD. Visual orientation and spatial frequency discrimination: a comparison of single neurons and behavior. *J Neurophysiol* 1987;57:755–772. [PubMed: 3559700]
- Brown, TJ. MS thesis. Cambridge, MA: Massachusetts Institute of Technology; 2000. Characterization of Acoustic Head-Related Transfer Functions for Nearby Sources.
- Cai H, Carney LH, Colburn HS. A model for binaural response properties of inferior colliculus neurons. I. A model with interaural time difference-sensitive excitatory and inhibitory inputs. *J Acoust Soc Am* 1998a;103:475–493. [PubMed: 9440334]
- Cai H, Carney LH, Colburn HS. A model for binaural response properties of inferior colliculus neurons. II. A model with interaural time difference-sensitive excitatory and inhibitory inputs and an adaptation mechanism. *J Acoust Soc Am* 1998b;103:494–506. [PubMed: 9440335]
- Caird DM, Palmer AR, Rees A. Binaural masking level difference effects in single units of the guinea pig inferior colliculus. *Hear Res* 1991;57:91–106. [PubMed: 1774216]
- Chan JC, Yin TC, Musicant AD. Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. II. Responses to band-pass filtered noises. *J Neurophysiol* 1987;58:543–561. [PubMed: 3655882]
- 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. Theory of binaural interaction based on auditory-nerve data. II. Detection of tones in noise. *J Acoust Soc Am* 1977a;61:525–533. [PubMed: 845314]

- Colburn, HS. Detection of tones in noise (Supplementary Material. 1977b. Theory of binaural interaction based on auditory-nerve data. II. AIP document no. PAPS and JASMA-610525–98
- Colburn, HS. Computational models of binaural processing. In: Hawkins, HL.; McMullen, TA.; Popper, AN.; Fay, RR., editors. *Auditory Computation*. New York: Springer-Verlag; 1996.
- Colburn, HS.; Durlach, NI. Models of binaural interaction. In: Carterette, EC.; Friedman, MP., editors. *Handbook of Perception*. New York: Academic Press; 1978.
- Covey E, Casseday JH. Timing in the auditory system of the bat. *Annu Rev Physiol* 1999;61:457–476. [PubMed: 10099697]
- Delgutte B. Physiological mechanisms of psychophysical masking: observations from auditory-nerve fibers. *J Acoust Soc Am* 1990;87:791–809. [PubMed: 2307776]
- Delgutte, B. Physiological models for basic auditory percepts. In: Hawkins, HL.; McMullen, TA.; Popper, AN.; Fay, RR., editors. *Auditory Computation*. New York: Springer-Verlag; 1996.
- Delgutte, B.; Hammond, BM.; Cariani, PA. Neural coding of the temporal envelope of speech: relation to modulation transfer functions. In: Palmer, AR.; Reese, A.; Summerfield, AQ.; Meddis, R., editors. *Psychophysical and Physiological Advances in Hearing*. London: Whurr; 1998.
- Durlach, NI.; Colburn, HS. Binaural phenomena. In: Carterette, EC.; Friedman, MP., editors. *Handbook of Perception*. New York: Academic Press; 1978.
- Efron, B.; Tibshirani, RJ. *An Introduction to the Bootstrap*. New York: Chapman & Hall; 1993.
- Fitzpatrick DC, Kuwada S, Batra R. Transformations in processing interaural time differences between the superior olivary complex and inferior colliculus: beyond the Jeffress model. *Hear Res* 2002;168:79–89. [PubMed: 12117511]
- Gilkey RH, Good MD. Effects of frequency on free-field masking. *Hum Factors* 1995;37:835–843. [PubMed: 8851779]
- 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]
- Good, MD.; Gilkey, RH.; Ball, JM. The relation between detection in noise and localization in noise in the free field. In: Gilkey, RH.; Anderson, TR., editors. *Binaural and Spatial Hearing in Real and Virtual Environments*. Mahwah, NJ: Erlbaum; 1997.
- Green, DM.; Swets, JA. *Signal Detection Theory and Psychophysics*. New York: Krieger; 1974.
- Hancock KE, Delgutte B. A physiologically based model of interaural time difference discrimination. *J Neurosci* 2004;24:7110–7117. [PubMed: 15306644]
- Hara K, Harris RA. The anesthetic mechanism of urethane: the effects on neurotransmitter-gated ion channels. *Anesth Analg* 2002;94:313–318. [PubMed: 11812690]
- Irvine, DRF. Physiology of the auditory brainstem. In: Popper, AN.; Fay, RR., editors. *The Mammalian Auditory Pathway: Neurophysiology*. New York: Springer-Verlag; 1992.
- Jeffress L. A place theory of sound localization. *J Comp Physiol Psych* 1948;41:35–39.
- Jiang D, McAlpine D, Palmer AR. Detectability index measures of binaural masking level difference across populations of inferior colliculus neurons. *J Neurosci* 1997a;17:9331–9339. [PubMed: 9364078]
- Jiang D, McAlpine D, Palmer AR. Responses of neurons in the inferior colliculus to binaural masking level difference stimuli measured by rate-versus-level functions. *J Neurophysiol* 1997b;77:3085–3106. [PubMed: 9212259]
- Joris PX. Interaural time sensitivity dominated by cochlea-induced envelope patterns. *J Neurosci* 2003;23:6345–6350. [PubMed: 12867519]
- Joris, PX.; Heijden, M.; Louage, D.; van de Sande, B.; van Kerckhoven, C. Dependence of binaural and cochlear “best delays” on characteristic frequency. In: Pressnitzer, D.; de Cheveigne, A.; McAdams, S.; Collet, L., editors. *Auditory Signal Processing: Physiology, Psychoacoustics, and Models*. New York: Springer-Verlag; 2004.
- Kiang NY, Moxon EC. Tails of tuning curves of auditory-nerve fibers. *J Acoust Soc Am* 1974;55:620–630. [PubMed: 4819862]

- Kim DO, Molnar CE. A population study of cochlear nerve fibers: comparison of spatial distributions of average-rate and phase-locking measures of responses to single tones. *J Neurophysiol* 1979;42:16–30. [PubMed: 430109]
- Kopco N, Lane CC, Shinn-Cunningham BG. Spatial unmasking of chirp trains in a simulated acoustic environment: behavioral results and model predictions. *Assoc Res Otolaryngol Abstr* 2003:541.
- 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, Batra R, Stanford TR. Monaural and binaural response properties of neurons in the inferior colliculus of the rabbit: effects of sodium pentobarbital. *J Neurophysiol* 1989;61:269–282. [PubMed: 2918355]
- Kuwada S, Yin TC. Binaural interaction in low-frequency neurons in inferior colliculus of the cat. I. Effects of long interaural delays, intensity, and repetition rate on interaural delay function. *J Neurophysiol* 1983;50:981–999. [PubMed: 6631473]
- Lane, CC. Signal Detection in the Auditory Midbrain: Neural Correlates and Mechanisms of Spatial Release from Masking (PhD dissertation). Cambridge, MA: Harvard–MIT Division of Health Sciences and Technology; 2003.
- Lane CC, Delgutte B, Colburn HS. A population of ITD sensitive units in the cat inferior colliculus shows correlates of spatial release from masking. *Assoc Res Otolaryngol Abstr* 2003:960.
- Lane, CC.; Kopco, N.; Delgutte, B.; Shinn-Cunningham, BG.; Colburn, HS. A cat's cocktail party: psychophysical, neurophysiological and computational studies of spatial release from masking. In: Pressnitzer, D.; de Cheveigne, A.; McAdams, S.; Collet, L., editors. *Auditory Signal Processing: Physiology, Psychoacoustics, and Models*. New York: Springer-Verlag; 2004.
- Langner G, Schreiner CE. Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *J Neurophysiol* 1988;60:1799–1822. [PubMed: 3236052]
- Litovsky RY, Delgutte B. Neural correlates of the precedence effect in the inferior colliculus: effect of localization cues. *J Neurophysiol* 2002;87:976–994. [PubMed: 11826062]
- Litovsky RY, Fligor BJ, Tramo MJ. Functional role of the human inferior colliculus in binaural hearing. *Hear Res* 2002;165:177–188. [PubMed: 12031527]
- Litovsky, RY.; Lane, CC.; Atencio, CA.; Delgutte, B. Physiological measures of the precedence effect and spatial release from masking in the cat inferior colliculus. In: Breebart, DJ.; Houtsma, AJM.; Kohlrausch, A.; Prijs, VF.; Schoonhoven, R., editors. *Physiological and Psychophysical Bases of Auditory Function*. Maastricht, The Netherlands: Shaker Publishing BV; 2000.
- Loftus WC, Bishop DC, Saint Marie RL, Oliver DL. Organization of binaural excitatory and inhibitory inputs to the inferior colliculus from the superior olive. *J Comp Neurol* 2004;472:330–344. [PubMed: 15065128]
- McAlpine D, Jiang D, Palmer AR. Binaural masking level differences in the inferior colliculus of the guinea pig. *J Acoust Soc Am* 1996;100:490–503. [PubMed: 8675842]
- McAlpine D, Jiang D, Palmer AR. A neural code for low-frequency sound localization in mammals. *Nat Neurosci* 2001;4:396–401. [PubMed: 11276230]
- McAlpine D, Palmer AR. Blocking GABAergic inhibition increases sensitivity to sound motion cues in the inferior colliculus. *J Neurosci* 2002a;22:1443–1453. [PubMed: 11850471]
- McAlpine D, Palmer AR. Binaural bandwidths of inferior colliculus neurones measured using interaurally-delayed noise. *Assoc Res Otolaryngol Abstr* 2002b:155.
- McKinney, MF.; Tramo, MJ.; Delgutte, B. Neural correlates of the dissonance of musical intervals in the inferior colliculus. In: Breebart, DJ.; Houtsma, AJM.; Kohlrausch, A.; Prijs, VF.; Schoonhoven, R., editors. *Physiological and Psychophysical Bases of Auditory Function*. Maastricht, The Netherlands: Shaker Publishing BV; 2000.
- Merzenich MM, Reid MD. Representation of the cochlea within the inferior colliculus of the cat. *Brain Res* 1974;77:397–415. [PubMed: 4854119]
- Morse, PM.; Ingard, KU. *Theoretical Acoustics*. New York: McGraw-Hill; 1968.
- Mountcastle VB, LaMotte RH, Carli G. Detection thresholds for stimuli in humans and monkeys: comparison with threshold events in mechanoreceptive afferent nerve fibers innervating the monkey hand. *J Neurophysiol* 1972;35:122–136. [PubMed: 4621505]

- Musicant AD, Chan JC, Hind JE. Direction-dependent spectral properties of cat external ear: new data and cross-species comparisons. *J Acoust Soc Am* 1990;87:757–781. [PubMed: 2307774]
- Nelson PC, Carney LH. A phenomenological model of peripheral and central neural responses to amplitude-modulated tones. *J Acoust Soc Am* 2004;116:2173–2186. [PubMed: 15532650]
- Oliver DL, Beckius GE, Bishop DC, Kuwada S. Simultaneous antero-grade labeling of axonal layers from lateral superior olive and dorsal cochlear nucleus in the inferior colliculus of cat. *J Comp Neurol* 1997;382:215–229. [PubMed: 9183690]
- Oliver DL, Morest DK. The central nucleus of the inferior colliculus in the cat. *J Comp Neurol* 1984;222:237–264. [PubMed: 6699209]
- Palmer AR, Jiang D, McAlpine D. Desynchronizing responses to correlated noise: a mechanism for binaural masking level differences at the inferior colliculus. *J Neurophysiol* 1999;81:722–734. [PubMed: 10036273]
- Palmer AR, Jiang D, McAlpine D. Neural responses in the inferior colliculus to binaural masking level differences created by inverting the noise in one ear. *J Neurophysiol* 2000;84:844–852. [PubMed: 10938311]
- Parker AJ, Newsome WT. Sense and the single neuron: probing the physiology of perception. *Annu Rev Neurosci* 1998;21:227–277. [PubMed: 9530497]
- Rosowski JJ, Carney LH, Peake WT. The radiation impedance of the external ear of cat: measurements and applications. *J Acoust Soc Am* 1988;84:1695–1708. [PubMed: 3209774]
- Saberi K, Dostal L, Sadralodabai T, Bull V, Perrott DR. Free-field release from masking. *J Acoust Soc Am* 1991;90:1355–1370. [PubMed: 1939900]
- 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]
- Shackleton, TR.; Palmer, DM. Sensitivity to changes in interaural time difference and interaural correlation in the inferior colliculus. In: Pressnitzer, D.; de Cheveigne, A.; McAdams, S.; Collet, L., editors. *Auditory Signal Processing: Physiology, Psychoacoustics, and Models*. New York: Springer-Verlag; 2004.
- 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]
- Singh NC, Theunissen FE. Modulation spectra of natural sounds and ethological theories of auditory processing. *J Acoust Soc Am* 2003;114:3394–3411. [PubMed: 14714819]
- Sivaramakrishnan S, Sterbing-D'Angelo SJ, Filipovic B, D'Angelo WR, Oliver DL, Kuwada S. GABA (A) synapses shape neuronal responses to sound intensity in the inferior colliculus. *J Neurosci* 2004;24:5031–5043. [PubMed: 15163696]
- Spitzer MW, Semple MN. Responses of inferior colliculus neurons to time-varying interaural phase disparity: effects of shifting the locus of virtual motion. *J Neurophysiol* 1993;69:1245–1263. [PubMed: 8492161]
- Spitzer MW, Semple MN. Transformation of binaural response properties in the ascending auditory pathway: influence of time-varying interaural phase disparity. *J Neurophysiol* 1998;80:3062–3076. [PubMed: 9862906]
- Sterbing SJ, D'Angelo WR, Ostapoff EM, Kuwada S. Effects of amplitude modulation on the coding of interaural time differences of low-frequency sounds in the inferior colliculus. I. Response properties. *J Neurophysiol* 2003;64:2827–2836. [PubMed: 12890793]
- Yang L, Pollak GD, Resler C. GABAergic circuits sharpen tuning curves and modify response properties in the mustache bat inferior colliculus. *J Neurophysiol* 1992;68:1760–1774. [PubMed: 1479443]
- Yin TC, Chan JC. Interaural time sensitivity in medial superior olive of cat. *J Neurophysiol* 1990;64:465–488. [PubMed: 2213127]
- Yin TC, Kuwada S. Binaural interaction in low-frequency neurons in inferior colliculus of the cat. III. Effects of changing frequency. *J Neurophysiol* 1983;50:1020–1042. [PubMed: 6631459]
- Zhang X, Heinz MG, Bruce IC, Carney LH. A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression. *J Acoust Soc Am* 2001;109:648–670. [PubMed: 11248971]

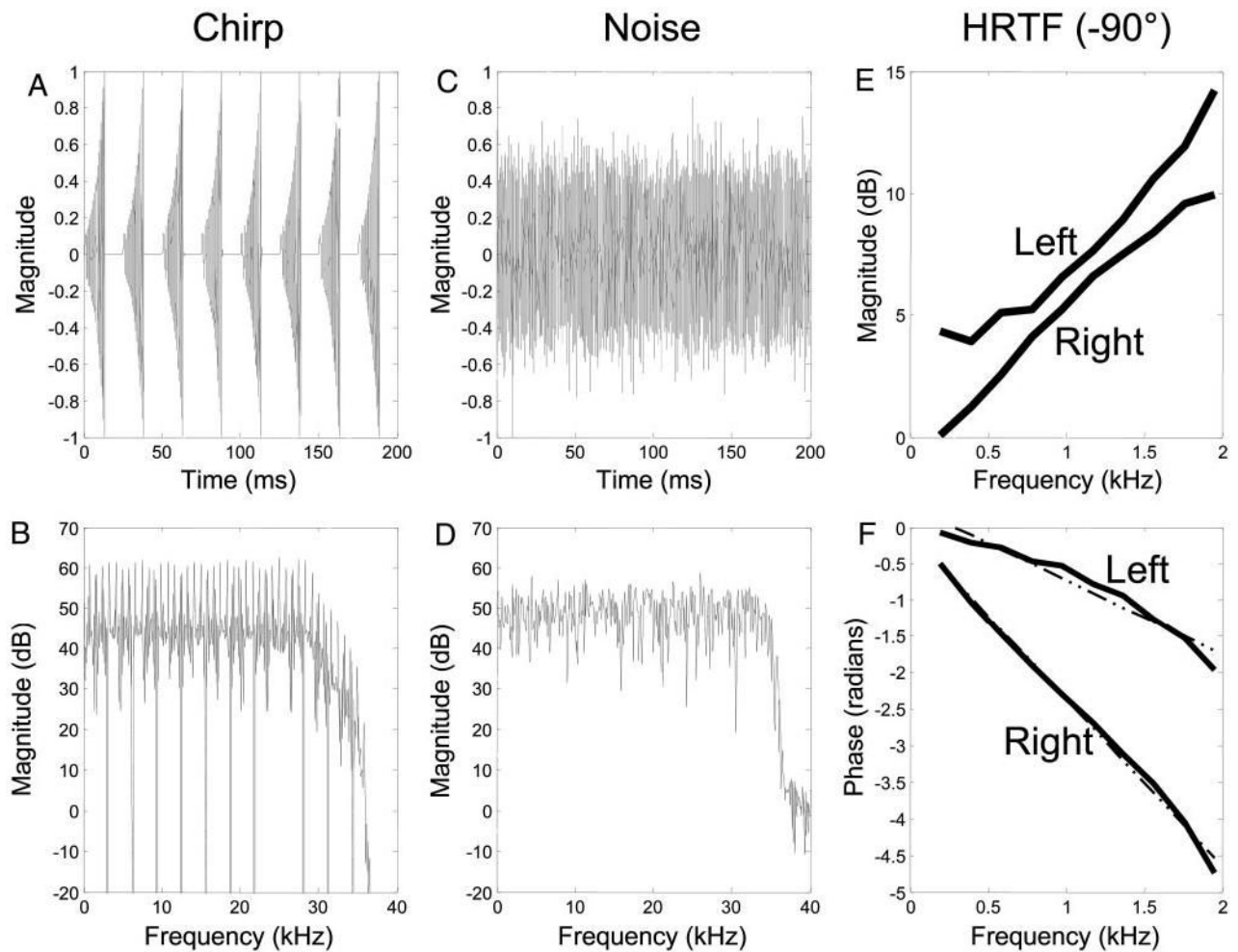
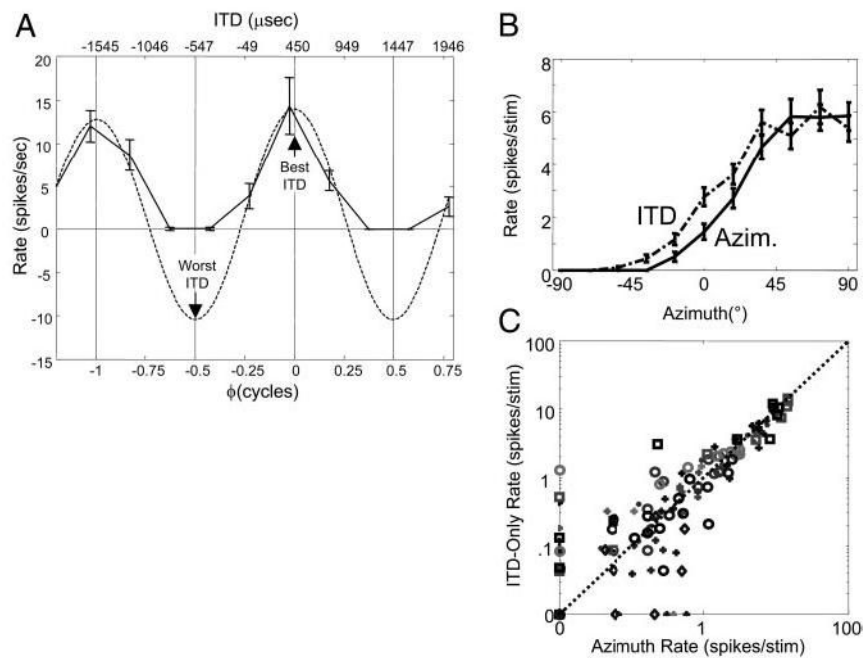


Fig 1. Stimuli and head-related transfer functions (HRTFs). Waveforms and power spectra of chirp-train signal (*A* and *B*) and noise masker (*C* and *D*). Frequency of each chirp is swept from 300 Hz to 30 kHz logarithmically; the envelope increases exponentially to flatten the spectrum. *E* and *F*: example HRTF for a stimulus at -90° , near the left ear. The gain for the left ear is larger, and the delay for the left ear (slope of the phase) is smaller. Slope of the magnitude with frequency was similar for all azimuths, although the gains changed by a few dB. This example shows how the differences in slopes (dash-dot line is a linear fit, as expected from a simple delay) create different interaural phase differences at different frequencies.

**Fig 2.**

A: Gabor fit to noise-delay function for an inferior colliculus (IC) unit. Solid curve shows rate as a function of noise interaural time delay (ITD; *top axis*) and relative interaural phase difference (IPD; *bottom axis*). Error bars: ± 1 SD. Relative-IPD (ϕ) axis normalizes the response so that the best ITD (450 μs) is 0 cycles, and the worst ITD (-547 μs) is -0.5 cycles. Best frequency (BF_{ITD}) of this unit is 500 Hz. Dashed line is the Gabor fit without half-wave rectification to show the worst ITD. B: rate vs. noise azimuth (solid line) and rate vs. noise ITD (dash-dot line) for one unit. Error bars: ± 1 SE. ITDs matched those in the HRTFs, and a detailed description of the method for the ITD manipulation, which was based on principle-component analysis, can be found in Litovsky and Delgutte (2002). C: rate from azimuth curve vs. rate from ITD-only curve for 19 units. Units on the axes have zero rate. Different combinations of marker shade and shape indicate different units.

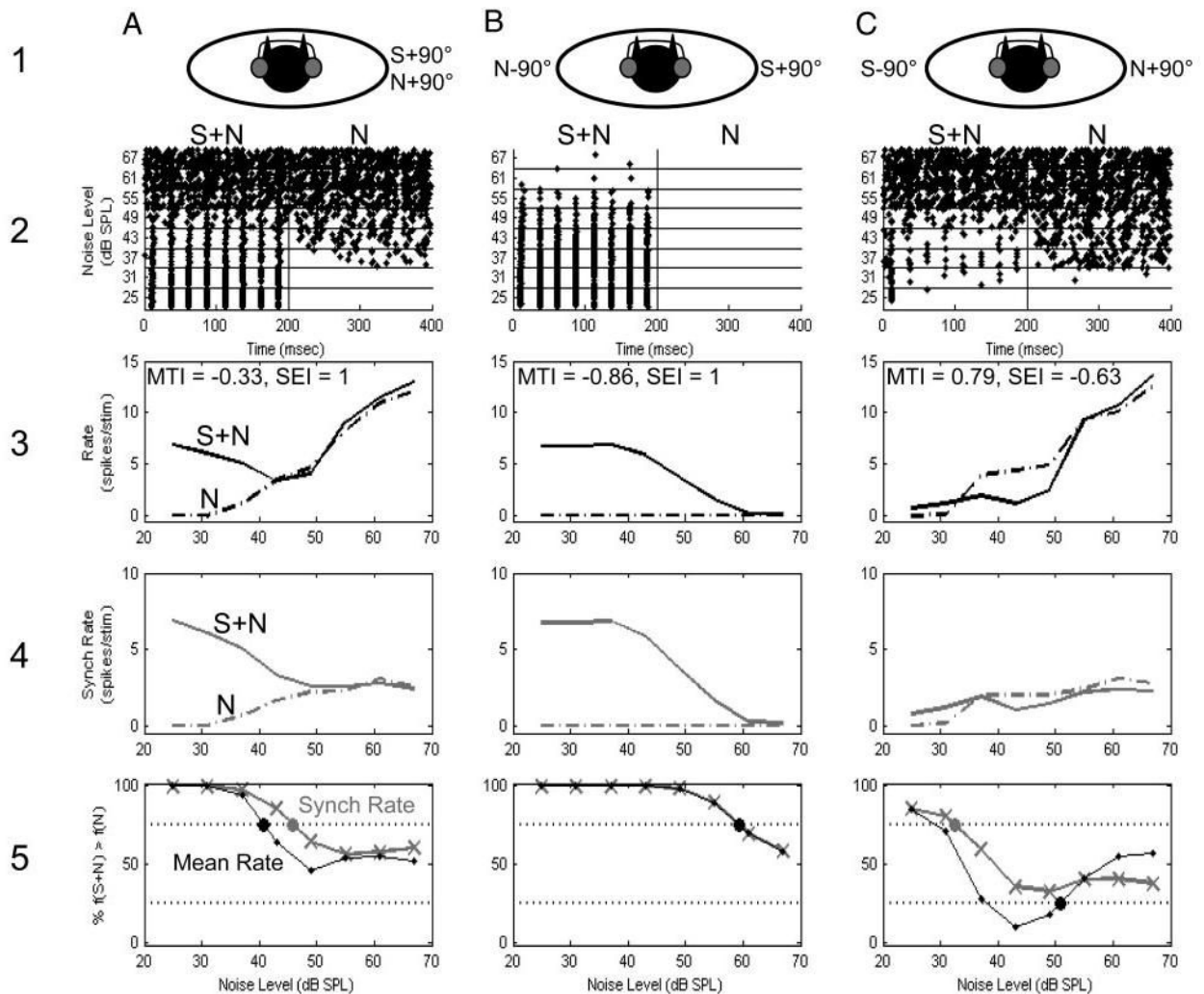


Fig 3.

Response for a single unit ($BF_{ITD} = 740$ Hz, best $ITD = 290 \mu s$) for 3 signal and noise configurations. Best azimuth for this unit is $+90^\circ$ ($\phi_{90^\circ} = 0.0$ cycles) and the worst azimuth is -90° ($\phi_{-90^\circ} = -0.43$ cycles). Signal level is 43 dB SPL. *First row*: configurations of signal and noise. *Second row*: dot rasters as a function of noise level. Every dot represents a spike. Signal is present for the first 200 ms of each trial, whereas the noise is continuous. Unit entrains to the chirp train at low levels (*bottom of plots*) and is masked as noise level is raised, either by swamping or suppressing the signal response. *Third row*: mean rate vs. noise level for signal-plus-noise window (S+N, solid lines) and noise-alone window (N, dash-dot lines). *Fourth row*: synchronization rate vs. noise level for signal-plus-noise (S+N, solid lines) and noise-alone (N, dash-dot lines). *Fifth row*: percentage of signal presentations where the mean rate (thin lines with dots) or synchronization rate (thick lines with \times 's) is larger for the signal-plus-noise window than the noise-alone window. Threshold is defined to be when the signal can be detected 75% of the time through either an increase or a decrease in rate, corresponding to either 75 or 25% greater (dotted lines). Filled circles indicate thresholds.

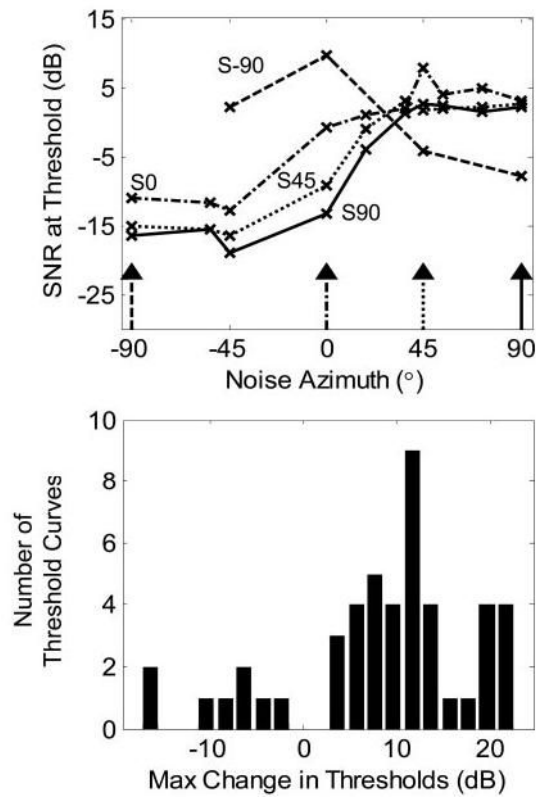


Fig 4.

Top: signal-to-noise ratio at mean rate threshold as a function of noise azimuth for 4 signal azimuths for same unit as in Fig. 3. Arrows indicate the signal azimuths and the tails of the arrows indicate the corresponding threshold curve. *Bottom:* for each unit and signal azimuth, we show the largest change in threshold at one masker azimuth compared with the co-located condition. We included the threshold curves for all the signal azimuths measured for every unit so there are more points than units in our sample (47 threshold curves for 31 units).

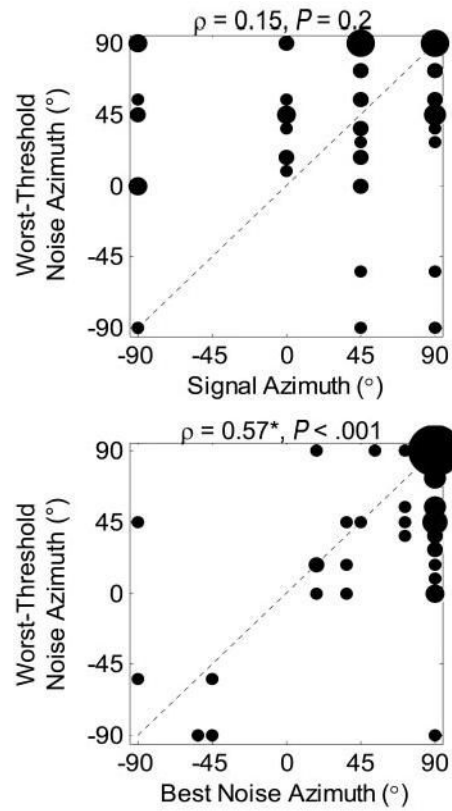


Fig 5. Bubble plot showing worst-threshold azimuth as a function of the signal azimuth (*top*) and the best azimuth (*bottom*). Radius of each bubble indicates the number of points at that graph location. Here $n = 68$, one for each signal azimuth for each unit.

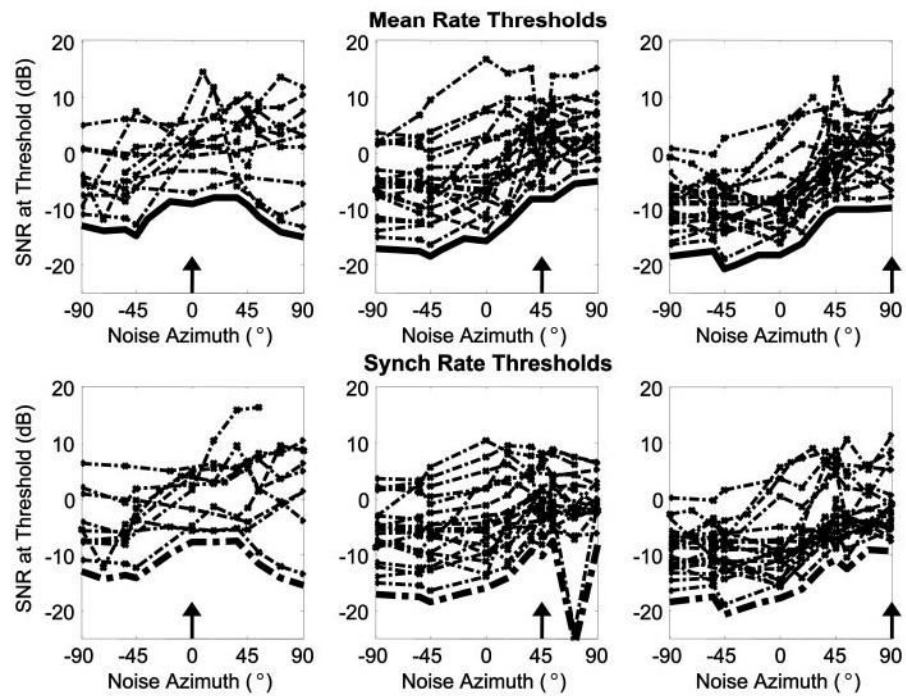
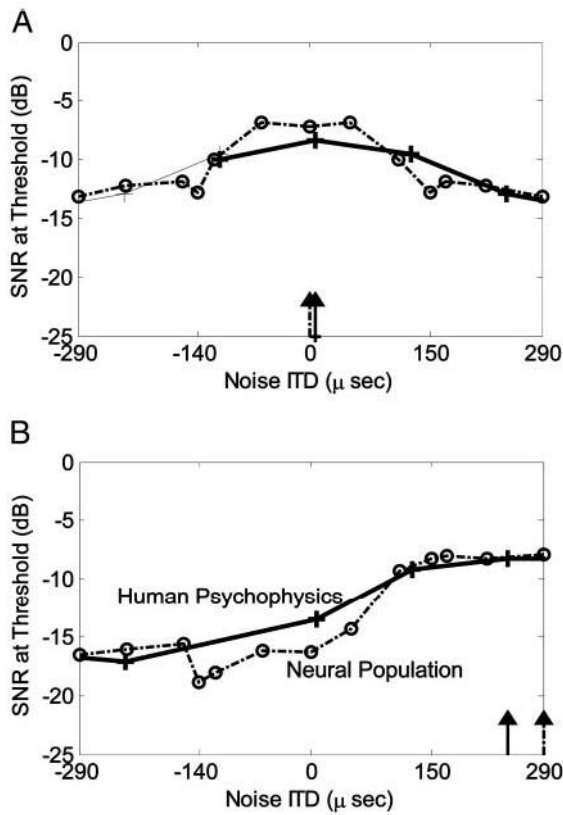
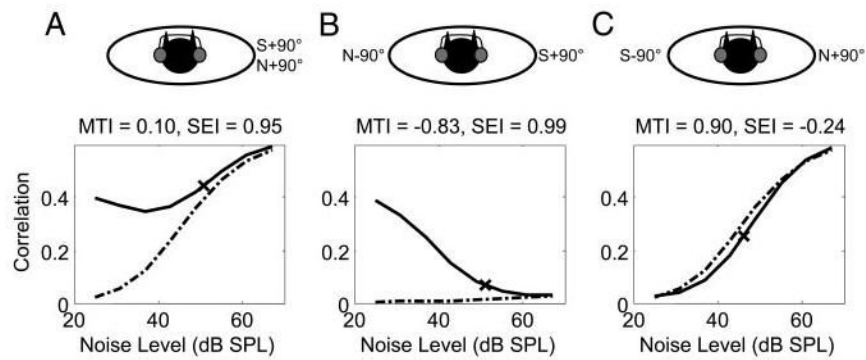


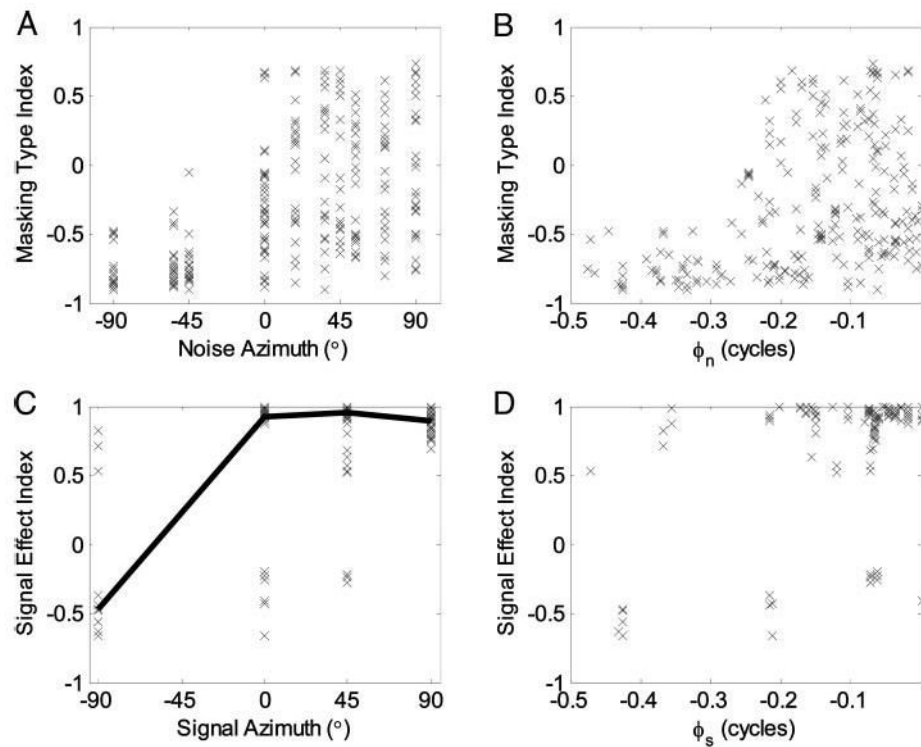
Fig 6. Single-unit thresholds (dash-dot lines) and population thresholds (solid lines) based on mean rate (*top*) and synchronized rate (*bottom*) plotted against noise azimuth for 3 signal azimuths (arrows). Individual unit thresholds were interpolated across azimuth so that the thresholds could be compared for all units. Population thresholds are offset vertically by 2 dB to reveal the single-unit thresholds.

**Fig 7.**

Human psychophysical thresholds (solid lines with crosses) compared with cat neural population thresholds (dash-dot lines with circles) as a function of noise ITD for 2 different signal locations. Arrows indicate signal azimuth; arrow tail indicates corresponding threshold curve. Psychophysical thresholds were measured for 3 normal-hearing subjects using low-pass stimuli so that spatial masking release would be based primarily on differences in ITD (Lane et al. 2003). Signal was a 200-ms, 40-Hz chirp train, band-pass filtered between 200 and 1,500 Hz, and the Gaussian noise masker was band-pass filtered between 200 and 2,000 Hz. Stimulus azimuth was simulated using human non-individualized HRTFs (Brown 2000), and stimuli were delivered by insert earphones. Signal was fixed in both azimuth and spectrum level (14 dB re $20 \mu\text{Pa}/\sqrt{\text{Hz}}$, about 45 dB SPL over entire bandwidth). For each masker azimuth, the masker level was adjusted using a 3-down, 1-up procedure to estimate the signal-to-noise ratio (SNR) yielding 79.4% correct detection performance. *A*: signal at 0° . Thin solid line shows the thresholds reflected about 0° , essentially completing the curve assuming that the listener has symmetric thresholds. Individual unit thresholds were also reflected about 0° to simulate the effect of having units from both ICs, and the population response was recomputed from the original and reflected data, making the response symmetric and different from Fig. 6. *B*: signal at 30° for humans, 90° for cat neural population, which give similar ITD values (250 and 290 μs , respectively).

**Fig 8.**

Cross-correlator predictions for unit in Fig. 3 ($BF_{ITD} = 740$, best ITD = $290 \mu s$) with a best azimuth of $+90^\circ$ and a worst azimuth of -90° . Stimulus waveforms (identical to the ones used experimentally) for both the left and right ears were passed through a realistic auditory nerve fiber (ANF) model (Zhang et al. 2001; 50 spikes/s spontaneous rate), whose CF matched the unit's BF_{ITD} . Output of one ear was delayed by the unit's best ITD and then the 2 ANF outputs were multiplied and the product was integrated over time to get the cross-correlation evaluated at the best ITD. Because the output of the ANF model is proportional to the probability of discharge, the unnormalized correlation is always positive and depends on stimulus level. Solid curve is the signal-plus-noise correlation, and dash-dot curve is the noise-alone correlation. The x's show the simulated masked thresholds for a criterion change in correlation of 0.05, which was selected to give thresholds similar to those found physiologically. Masking type index (MTI) and signal effect index (SEI) for each condition are given in the titles.

**Fig 9.**

MTI and SEI as a function of noise azimuth and noise relative IPD. *A*: MTI as a function of noise azimuth for favorable signal azimuths ($|\varphi_s| < 0.1$). *B*: MTI as a function of noise relative IPD for favorable signal azimuths ($|\varphi_s| < 0.1$). *C*: SEI as a function of signal azimuth for favorable noise azimuths ($|\varphi_n| < 0.1$). Many of the points overlap because the signal was in the same location for multiple noise azimuths. Line shows the median value. *D*: SEI as a function of signal relative IPD for favorable noise azimuths ($|\varphi_n| < 0.1$).

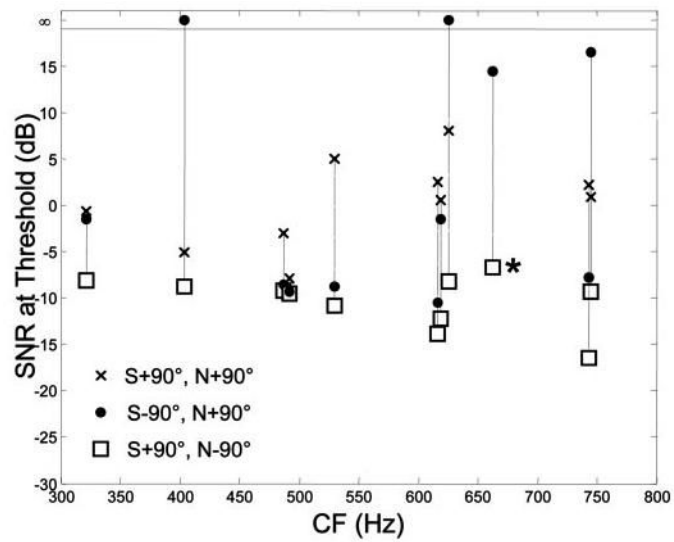


Fig 10.

Masked thresholds of 11 units for S+90, N-90 (open squares), for S-90, N+90 (filled circles), and for S+90, N+90 (x's). Horizontal axis is the difference in absolute value of the relative IPD for -90° and $+90^\circ$. Thresholds marked infinite could not be measured because the signal could not be detected at any masker level. For the unit marked (*), we did not measure responses in the S+90, N+90 condition because of a short holding time.

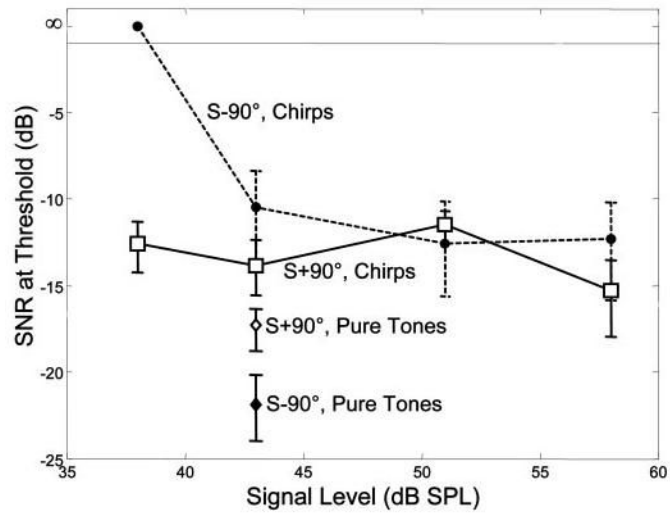


Fig 11.

Masked thresholds for one unit (BF_{ITD} of 600 Hz, best ITD of 400 μ s) as a function of signal level for a chirp train and a 500-Hz pure tone. Thresholds for chirp trains at $+90^\circ$ (squares) are fairly constant with signal level, but chirp trains at -90° (circles) could not be detected when presented at 38 dB. -90° chirp thresholds improve with increasing signal level to give thresholds similar to the $+90^\circ$ condition. Threshold hierarchy for the pure tone is reversed: a 500-Hz tone placed at $+90^\circ$ (open diamond) gives a worse threshold than the same tone placed at -90° (filled diamond).

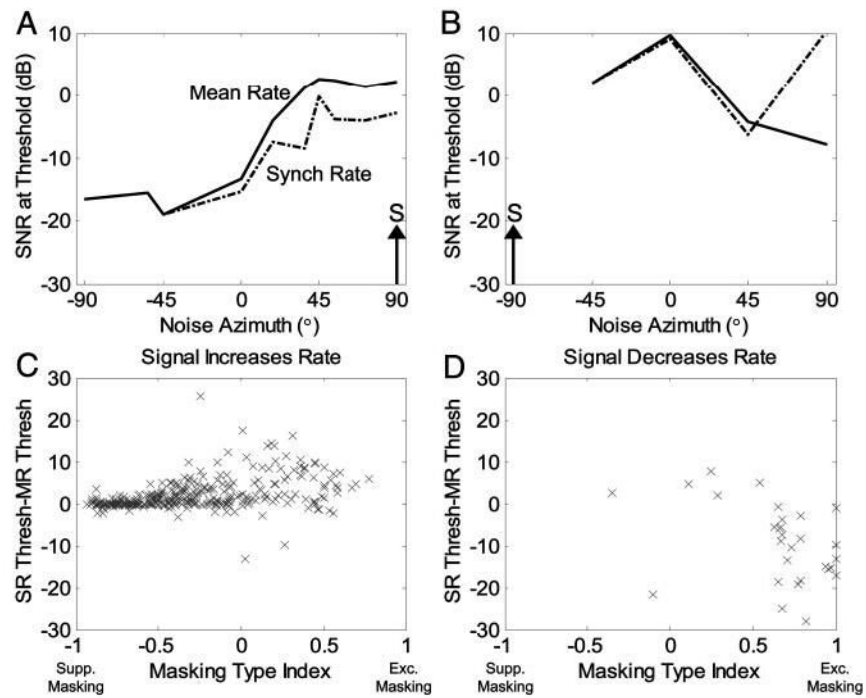


Fig 12.

Mean-rate thresholds compared with synchronized-rate thresholds. *A* and *B*: mean rate (solid lines) and synchronized rate (dash-dot lines) thresholds (same unit as Fig. 3) for $S+90^\circ$ (*A*) and $S-90^\circ$ (*B*). *C* and *D*: difference between synchronized rate threshold and mean rate threshold as a function of MTI when the signal is detected by an increase in rate ($SEI > 0$, *C*) and a decrease in rate ($SEI < 0$, *D*).