Neural Sensitivity to Interaural Envelope Delays in the Inferior Colliculus of the Guinea Pig

Sarah J. Griffin¹², Leslie R. Bernstein³, Neil J. Ingham¹§ and David McAlpine¹²

¹Department of Physiology, University College London, Gower Street, London, WC1E 6BT, United Kingdom

²University College London Ear Institute, 330-336 Grays Inn Road, London, WC1X 8EE

³Departments of Neuroscience and Surgery, University of Connecticut Health Center, Farmington, CT, USA, 06030

Running Head: Envelope ITD Sensitivity

Corresponding Author: David McAlpine
d.mcalpine@ucl.ac.uk

Department of Physiology, University College London, Gower Street, London, WC1E 6BT, United Kingdom
(0044) (0) 2076796087
(0044) (0) 2073876368
Abstract

Interaural time differences (ITDs) are important cues for mammalian sound localization. At high-frequencies, sensitivity to ITDs, which are conveyed only by the envelope of the waveforms, has been shown to be poorer than sensitivity to ITDs at low-frequencies, which are conveyed primarily by the fine-structure of the waveforms. Recently, human psychophysical experiments have demonstrated that sensitivity to envelope-based ITDs in high-frequency transposed tones can be equivalent to low-frequency fine-structure-based ITD sensitivity. Transposed tones are designed to provide high-frequency auditory nerve fibers (ANFs) with similar temporal information to that provided by low-frequency tones. We investigated neural sensitivity to ITDs in high-frequency transposed and sinusoidally amplitude modulated (SAM) tones, in the inferior colliculus of the guinea pig. Neural sensitivity to ITDs in transposed tones was found to be greater than to ITDs in SAM tones; in response to transposed tones, neural firing rates were more modulated as a function of ITD and discrimination thresholds were found to be lower than in response to SAM tones. Similar to psychophysical findings, ITD discrimination of single neurons in response to transposed tones for rates of modulation less than 250-Hz was comparable to neural discrimination of ITDs in low-frequency tones. This suggests that the neural mechanisms that mediate sensitivity to ITDs at high- and low-frequencies are functionally equivalent provided that the stimuli result in appropriate temporal patterns of action potentials in ANFs.
Introduction

According to the duplex theory (45), human sound localization is subserved by two mechanisms. At low-frequencies (< 1500 Hz), sounds sources are localized using interaural time differences (ITDs) while at high-frequencies interaural intensity differences (IIDs) mediate localization. The dichotomy suggested by the duplex theory is, to a first approximation, reflected in the anatomy and physiology underlying sensitivity to ITDs and IIDs in mammals (28; 54). Electrophysiological recordings confirm that phase-locked action potentials converge from each cochlear nucleus onto coincidence detection neurons in the medial superior olive (MSO), generating sensitivity to ITDs (2; 22; 52; 56). Sensitivity to IIDs occurs in lateral superior olive (LSO) neurons, which receive excitatory ipsilateral and inhibitory contralateral inputs (12).

These brainstem pathways are not wholly independent, e.g. both involve glycinergic inhibition from the medial nucleus of the trapezoid body (13; 51). Physiological recordings suggest that LSO neurons do not exclusively encode IIDs and MSO neurons do not exclusively encode ITDs. Rather, some overlap occurs (2; 15; 30). The strict dichotomy suggested by the duplex theory is also contradicted by psychophysical experiments that demonstrate sensitivity to ITDs conveyed by the envelopes of high-frequency complex sounds (17; 23; 32; 39; 61).

Human sensitivity to changes in ITD conveyed by high-frequency stimuli has typically been found to be poorer than that measured with low-frequency stimuli (6; 11; 25; 61). In
addition, the lateral extent of intracranial images produced by ITDs conveyed by high-frequency, complex stimuli is typically smaller than that conveyed by low-frequency stimuli (7). The relatively poor sensitivity to ITDs at high-frequencies may be explained by a lack of specialization for the processing of envelope ITDs within central auditory centers. In the LSO, ITD sensitivity in response to high-frequency stimuli has been considered a by-product of neural circuitry specialized for IID sensitivity (30; 31; 54) whereas the MSO is viewed as being specialized for detection of ITDs in low-frequency stimuli.

Colburn & Equissaud (1976) hypothesized that the differences in ITD-sensitivity at low- and high-frequencies observed psychophysically could be accounted for by differential effects of peripheral processing on low- and high-frequency stimuli. Firing of action potentials in ANFs can be modeled by band-pass filtering, half-wave rectification, and low-pass filtering of the sound waveform (21). This produces a “phase-locked” response to low-frequency tones (Figure 1A). When a high-frequency sound is modulated in amplitude, ANF responses exhibit phase-locking to the envelope (24; 29; 43; 44). Sinusoidal amplitude modulated (SAM) tones produce firing patterns in ANFs that follow the sinusoidal envelope (Figure 1B). Although this timing information supports sensitivity to ITDs, the firing pattern lacks the distinct “off periods” (where the firing probability is zero), produced by half-wave rectification, that characterize responses to low-frequency tones (compare Figures 1 A&B).
In human psychophysical experiments, Bernstein & Trahiotis (2002) demonstrated, using “transposed” tones, that ITD-sensitivity at high frequencies could be comparable to ITD-sensitivity at low-frequencies. Transposed tones were first designed by van de Par & Kohlrausch (1997) to provide high-frequency ANFs with envelope-based information similar to the waveform-based information available from low-frequency tones. The presumption is that, despite the differences in the sound pressure waveforms, the firing probability in ANFs is similar for high-frequency transposed tones and low-frequency tones (Figure 1C).

The current study investigated neural responses to ITDs conveyed by high-frequency SAM and transposed tones; extracellular recordings were obtained from single neurons in the inferior colliculus (IC) of the guinea pig. Neural responses reflected greater sensitivity to ITDs for transposed tones than for SAM tones. Consistent with the psychophysical findings of Bernstein & Trahiotis (2002), neural discrimination of ITDs within transposed tones was comparable to neural discrimination of ITDs within low-frequency tones. Also consistent with these, and other, psychophysical (8; 39; 42) and physiological (31) findings, ITD sensitivity in neurons was limited to low modulation frequencies (< ~250 Hz) for both SAM and transposed tones. Our results suggest that, for those low rates of modulation, the central mechanisms that mediate sensitivity to envelope-based ITDs are essentially equivalent to those that mediate sensitivity to fine-structure-based ITDs.
Materials and Methods

Animals and surgery

All experiments were carried out in accordance with the Animal (Scientific Procedures) Act of 1986 of Great Britain and Northern Ireland. Single-neuron recordings were made from the central nucleus of the right IC of 27 adult guinea-pigs under urethane anaesthesia (Sigma-Aldrich, Poole, U.K.; 25% solution of 0.9% NaCl; 1g/kg, i.p.). Additional analgesia was provided using fentanyl-fluanisone (Hypnorm; Janssen-Cilag Ltd., High Wycombe, U.K.; 0.1 ml, i.m.), supplementary doses of which were administered as required. Atropine sulphate (Animalcare Ltd., York, U.K.; 0.6 mg/ml; 0.1 mls, s.c.) was administered to reduce bronchial secretions, and lidocaine hydrochloride (Martindale Pharmaceuticals, Romford, U.K; 2%; s.c.) was administered locally prior to any surgical incision. A tracheal cannula was inserted and core temperature was maintained at 37°C with a heating blanket and rectal probe (Harvard Apparatus Ltd, Kent, UK).

Animals were placed in a sound-attenuating chamber (IAC, Winchester, UK) and held in a stereotaxic frame with hollow ear speculae (modified from model 1730, David Kopf Instruments, CA). Before positioning the animal, the tragus was cut to obtain clear access to the tympanic membrane. A craniotomy was performed to expose the cortex overlying the IC, the covering dura was removed and agar (~ 2%) was applied to prevent drying and deterioration of the cortex. The bullae were vented to equalize air pressure in the middle ears, by insertion of cannulae and sealing with petroleum jelly. A parylene-coated tungsten microelectrode (1-5 MΩ; World Precision Instruments, FL; or made in house
(14)) was positioned stereotaxically 2 mm above the IC (40) and advanced ventrally from outside the recording chamber using a piezo-stepped microdrive (Burleigh Instruments, Inc., NY). At the end of each experiment, animals were administered a lethal dose of Sodium Pentobarbitone (Pentoject; Animalcare Ltd.; 60 mgs/ml; 1-2 mls, i.p. or Sagatal; 60mgs/ml).

**Stimulus production and presentation**

Sounds were produced using Tucker Davis Technologies (TDT; Alachua, FL) digital signal processing hardware. Tonal stimuli used for the isolation and characterization of single units were generated using custom software (Trevor Shackleton and Alan Palmer, MRC Institute of Hearing Research, Nottingham, UK.; 100-kHz sampling rate) and TDT system II hardware. TDT Brainware, Real Time Processor Visual Design Studio (RPvds) and system III hardware were employed in order to generate the SAM and transposed stimuli (50-kHz sampling rate).

Stimuli were generated and scaled such that their peak voltages were at the maximum available voltage (± 10V) of the digital-to-analog converters (DACs). The outputs were attenuated to achieve the desired level for the experiments using PA4 (system II) or PA5 (system III) modules (TDT). (In initial experiments using system III, the attenuation was achieved via scaling within the signal-generation software.) Fixed amplification [ROTEL (Worthing, UK) RB971 power amplifier or a Beyerdynamic (Burgess Hill, UK) A150 Blueprint stereo-amplifier] was followed by 60 dB of “final” attenuation. Employing a high-level signal followed by attenuation maximized the signal-to-noise ratio through the
stimulus-generation pathway. Specifically, this technique yielded a greater signal-to-noise ratio than if amplification alone were used to achieve the desired intensity. Sounds were delivered via Beyerdynamic DT48A loudspeaker drivers, modified and fitted with a probe tube attachment to allow insertion and sealing into the hollow ear speculae. Knowles Acoustics (Burgess Hill, UK) FG3452 microphones [attached to steel tubes placed in the ear speculae and calibrated against a Brüel and Kjær (Stevenage, UK) Type 4136 1/8-in. microphone] were used to measure the stimulus within a few mm of the tympanic membrane in order to ensure that the sounds delivered to each ear were well matched.

*Spike collection*

Electrical signals from the electrode were conducted via a headstage to a pre-amplifier (TDT Medusa RA16PA) where they were amplified and digitized at a 25-kHz sampling rate. The signal was then conducted via a fiber-optic cable to the RA16 base station which produced fixed amplification and filtering (×1000 gain, 300-Hz high-pass filter, 10-kHz low-pass filter and 50-Hz notch filter).

Using system II hardware, a spike voltage – level discriminator (TDT ET1) was used to detect spikes from the background noise. Spikes were monitored on a Tektronix TD210 oscilloscope for continuity of spike characteristics, i.e., shape and amplitude, to ensure recordings were from a single neuron.
Using system III hardware, spike data passed from the RA16 base station to TDT Brainware and spikes that crossed a user defined trigger level were counted. Spikes were then additionally sorted, in Brainware, according to spike characteristics, to ensure data were collected from a single neuron.

**Sound stimuli - Isolation and characterisation stimuli**

Binaural “search stimuli,” consisting of 50-ms presentations of diotic single tones, repeated at a rate of five per second, were used to isolate neurons and estimate their threshold (the lowest intensity sound required to evoke firing, determined audio-visually, by the experimenter) and characteristic frequency (CF; the frequency at which the lowest threshold was obtained). The CF and threshold were confirmed by recording a frequency-versus-level response area spectrally flanking the CF estimate (2 octaves above and 4 octaves below CF) and between 10 and 90 dB of attenuation from the maximum system output of ~ 110 dB SPL. If the unit remained isolated for a sufficient length of time, neurons were further characterized by presentation of 50- or 200-ms diotic, ipsilateral and contralateral tones with frequency equal to CF. Tones were presented at 20 dB above threshold and 100 or 150 repetitions were presented.

**Sound Stimuli - SAM and transposed tones**

Two parameters of the stimuli were varied, ITD (of primary interest) and frequency of modulation ($f_m$). SAM and transposed stimuli were constructed with a carrier frequency ($f_c$; in sine phase) equal to neuronal CF and were modulated by multiplication with either a lower frequency, dc-shifted, sinusoidal waveform (in sine phase) to produce a SAM
tone (Figure 1B) or a half-rectified sinusoid (90° phase-advanced with respect to the SAM modulating tone) to produce a transposed tone (Figure 1C). Both stimuli were modulated by 100%. For each neuron, the peak voltages of SAM and transposed tones presented were equal and between 10 to 30 dB above the peak voltage of the sinusoid at CF that defined threshold. For 75/82 neurons, peak voltages were 20 dB above that of the tone at threshold. This choice resulted in the SAM and transposed tones being presented at 15.7 and 14.0 dB, respectively, above the rms level of the tone. In all cases, the rms levels of the transposed tones were within 2 dB of their SAM counterparts.

**ITD generation**

Stimuli were 500 ms in duration and were gated on and off with 2-ms cosine-squared ramps. An ITD was created by delaying the entire waveform (both fine-structure and envelope) in one channel and advancing it in the other by an equal amount. The resulting stimuli, therefore, contained onset, ongoing, and offset ITDs. A positive ITD was created by delaying the stimulus in the ipsilateral (left) ear and advancing the stimulus in the contralateral ear. A negative ITD was created by delaying the stimulus in the contralateral ear and advancing the stimulus in the ipsilateral ear. The values of ITD that were chosen depended on the $f_m$ of the stimuli such that 17 steps of an interaural phase difference (IPD) were presented over $\pm 0.5$ cycles of the $f_m$. The IPDs applied remained constant irrespective of the $f_m$ ensuring that the neural tuning was measured over a complete cycle of each $f_m$. Data are also presented from experiments in which the stimuli were either 800 ms in duration or, alternatively, were 500 ms in duration with no onset or offset ITD. In the latter case, stimuli were gated on and off at the same time in both ears.
Modulation frequency

Neural responses from one neuron, to any one $f_m$ and over all 17 IPDs, will be referred to as a single “recording”. SAM and transposed tones were both presented at each $f_m$. Initially, for each neuron, six or seven $f_m$ were presented pseudo-randomly at each IPD. Modulation frequencies varied between 10 Hz and 640 Hz (or CF/6 if this was lower) in logarithmic steps or between 10 Hz and 650 Hz (or CF/6) in linear steps. The rationale for limiting the $f_m$ to CF/6 is explained below. If the neuron remained well-isolated, further data sets were collected at more narrowly spaced $f_m$. For 11 neurons, in later experiments, a smaller range of $f_m$ was presented (between 10 Hz and 200-250 Hz) because it was over this range of $f_m$ that sensivity to ITDs was most likely. At least three (3-6, median = 3) repetitions of each stimulus condition were presented. The exact $f_m$s presented were determined by the experimenter according to the response of the neuron under investigation.

Spectral components and the $f_m$ limitation

SAM tones are characterized by three frequency components, the carrier ($f_c$) and two “sidebands” [(fc + fm) & fc − fm]. Transposed tones have additional sidebands spaced at multiples of 2×fm (Figure 2). van de Par and Kohlrausch (1997) demonstrated, in psychophysical experiments, that the five central frequency components of a high-frequency transposed stimulus were sufficient to yield improved binaural performance at high frequencies. For transmission of the temporal structure of the modulation to a single ANF, the spectral components of SAM tones and the central five spectral components of
transposed tones must fall within the ANF’s effective “filter” or spectral receptive field. As the $f_m$ increases, sidebands are attenuated by the ANF’s filter in proportion to their spectral “distance” from the CF. This alters the temporal structure of the modulation transmitted to ANFs. Accordingly, the $f_m$ in the present study was limited to a maximum of CF/6 to include $f_m$s at, and slightly above, the range over which we expected that the temporal structure would be preserved. This was judged from the shape of the frequency-versus-level response areas obtained from each neuron, and from reports of the equivalent rectangular bandwidth (ERB) of guinea pig ANFs (18). The maximum $f_m$ was also limited to 640 or 650 Hz because this included, and slightly exceeded, the range of $f_m$ at which sensitivity to ITDs has been observed at high-frequencies (e.g.(8; 31)).

In their psychophysical experiments, Bernstein & Trahiotis (2002) limited the spectra of transposed stimuli to prevent the use of energy outside the psychophysically-determined “auditory filter” [see (41)] surrounding the centre frequency of the signal under investigation. They low-pass filtered their half-wave rectified tones before multiplication with high-frequency carriers. It was not necessary to impose such a spectral limitation in the present study because we investigated the response of single IC neurons. The spectral components influencing single neurons’ responses are necessarily limited by their spectral receptive fields. Spectral components distant from $f_c$ (outside the central 5 components) are unlikely to contribute any appreciable energy to a neuron with a CF equal to $f_c$ and at the intensities at which transposed tones were presented (10 to 30 dB above pure tone threshold).
High-resolution functions

“High-resolution” IPD functions were collected from 14 neurons. A \( f_m \) was selected to which a given neuron exhibited a substantial modulation in its firing rate as a function of IPD (i.e., was found to be IPD-sensitive). SAM and transposed tones with 101 (or 102) IPDs between +0.5 and -0.5 cycles, including an onset and offset ITD, were presented at this \( f_m \). The duration of the stimulus was 500 ms and at least 6 repetitions (6 – 13, median = 10) were presented, pseudo-randomly, of each stimulus condition.

Data analysis

Spike times were exported into Matlab 6.5 (The Mathworks Ltd., Natick, MA) for off-line analysis.

Data analysis – Sensitivity to ITDs

Spike rates were determined by counting spikes over two time windows. First, spike rates were calculated by counting the number of spikes occurring during a 600-ms window from the onset of the stimulus at one ear to the offset of the stimulus at the other ear. This took account of the fact that for the lowest modulation rate used (10 Hz) with an ITD equivalent to 0.5 cycles of interaural phase difference, the 500 ms stimulus was gated on at one ear 50 ms before the stimulus was gated on at the other ear, and gated off 50 ms before the stimulus. Thus, spikes counted over this time period included the onset and offset response to the stimuli, which could include responses to monaural stimulation at large values of ITD. Second, spikes were also counted over the middle 350 ms of the stimulus presentation, and over a time period that ensured a whole number of modulation
periods within the 350 ms. This middle time period was employed in order to exclude neural responses to the onset and offset of the stimuli and to eliminate periods of monaural stimulation at large values of ITD.

In a small number of early experiments the stimulus duration was either 800 ms or was 500 ms but IPDs were generated with no onset ITD. Recordings from these experiments (from 13 neurons) were analyzed over the middle 350 ms of the stimulus, and are included in the current study.

Tuning to ITDs was calculated from the mean spike rates at each \( f_m \) and over ± 0.5 cycles of IPD. The following calculations were performed for each \( f_m \) (i.e. for each recording) in response to SAM and transposed tones. The mean phase vector was calculated using the method of Goldberg and Brown (1969). At every \( f_m \), there were two measurements for an IPD of 0.5 cycles (at +0.5 and -0.5 cycles) and the mean of these was used in the calculation. The mean phase vector was termed the ‘best’ phase (BP); this was considered the IPD to which the recording was tuned. The normalized length of the vector (the vector strength, which varied between 0 and 1) was used to calculate the Rayleigh coefficient (35). The Rayleigh coefficient was the test static for a chi-squared distribution with two degrees of freedom (36). An individual recording was classified as “ITD-sensitive” if the Rayleigh coefficient was greater than 13.815, \( p < 0.01 \) (36; 58), indicating that firing rates were not evenly distributed around a cycle of IPD. Individual neurons were classified as ITD-sensitive if one or more recordings were classified as ITD-sensitive.
Data analysis – High-resolution functions

High-resolution, IPD functions were obtained in order to calculate neural discrimination thresholds. ROC (Receiver Operating Characteristic) analysis was used to estimate the smallest ITD that could be discriminated from 0 ITD, using the firing rates of a single neuron. IPDs were converted into ITDs with reference to the modulation frequency. ROC analysis has been used previously (49; 50) to compare neural ITD-sensitivity to human psychophysical performance (See Shackleton et al. 2003 for a detailed description of the method.) For this analysis, spike rates were calculated over the middle 350 ms analysis period. To perform ROC analysis, a neurometric function was calculated for each high-resolution IPD function. The distribution of firing rates obtained at 0 ITD was compared with the distribution of firing rates at every IPD presented. The spike rates recorded from all repetitions (6-10) of each ITD form a sample distribution of the spike rates at that ITD. If 101 ITDs were presented, 101 pairs of spike rates are compared, including the comparison of the spike rates at 0 ITD with itself. The neurometric functions describe the probability of randomly selecting a spike rate from each pair of distributions and finding the spike rate at 0 ITD to be lowest.

Neurometric functions were smoothed by averaging over three consecutive ITDs. ITDs were considered discriminable from zero ITD if the probability (from the smoothed function) was >= 0.75 or <= 0.25 (giving two discrimination thresholds). We defined the just noticeable difference (JND) in ITD as the smallest ITD that is discriminable from zero ITD. No interpolation was carried out to estimate the point where the neurometric
functions crossed 0.75 or 0.25. JNDs were, therefore, if anything, slightly overestimated. For 5 high-resolution functions, 102 (instead of 101) values of IPD were presented. In these cases no response to zero ITD was collected, and discrimination was then calculated from the closest negative ITD to zero.

Data analysis - Period histograms

Entrainment of spike times to the modulation period for SAM and transposed stimuli was examined by binning spike times over one cycle of the modulation rate. Spike times over the middle 350 ms of the stimulus presentation, and over a time period that ensured a whole number of modulation periods within the 350 ms, were analyzed. The cycles of the modulation period at which spikes occurred were adjusted such that 0 cycles was always with reference to the onset of sound in the contralateral ear. To quantify the degree of phase-locking to the stimulus period, at each IPD and $f_m$, the vector strength was calculated. Phase-locking was considered significant if the Rayleigh coefficient was > 13.815 (p < 0.01) (36).
Results

Sensitivity to envelope ITDs in SAM and transposed tones

Responses to SAM and transposed tones were obtained from 82 IC neurons with CFs > 2 kHz (2.0 – 13 kHz, median = 4.5 kHz). For 55/82 neurons, peri-stimulus time histograms (PSTHs) were obtained to monaural and diotic CF tones. All neurons exhibited greater firing rates in response to tones at the contralateral ear than at the ipsilateral ear. For 70% of the sample (39/55), ipsilateral stimulation resulted in inhibition, exhibited as either a reduction in the spontaneous firing rate in response to ipsilateral stimulation or as a reduction in the rate of firing in response to a contralateral sound when stimulation was binaural.

Sixty-nine neurons were investigated using SAM and transposed tones of 500-ms duration with ITDs generated by delaying the entire waveform. Over both analysis windows (350 ms and 600 ms), more recordings (the response of a single neuron at one $f_m$) showed sensitivity to ITDs within transposed tones than in SAM tones. Correspondingly, more neurons were sensitive to ITDs within transposed than SAM tones. Responses from a further 13 neurons were recorded using stimuli which were 800 ms in duration or had no onset ITDs. The responses of these 13 neurons were analyzed over the middle 350 ms of the stimulus duration and appeared qualitatively similar to responses recorded from neurons using stimuli of 500 ms (with onset ITDs). They are therefore included in all subsequent analyses examining responses to the middle 350 ms of the stimulus. The first column in Table 1 shows the number of neurons from which
recordings were made, as well as the total number of recordings, using analysis windows of 600 ms and 350 ms. The second column in Table 1 shows the number of neurons and recordings in which ITD sensitivity was observed for transposed, but not SAM, stimuli. The third column shows the number of neurons and recordings in which ITD sensitivity was observed for both transposed and SAM stimuli, and the fourth column shows the number of neurons and recordings in which ITD sensitivity was observed for SAM stimuli only. The final column shows the number of neurons and recordings in which ITD sensitivity was not observed for either transposed or SAM stimuli.

Although the entire SAM and transposed waveforms were delayed, ITD sensitivity was conveyed by the envelope structure of the stimuli; neurons are not sensitive to carrier-based ITDs at frequencies above 2 kHz (27). In addition, the modulation in firing rates occurred over a cycle of interaural phase with respect to the modulation frequency and not the carrier frequency.

For the 69 neurons from which recordings were made using stimuli of 500-ms duration, more recordings were classified as ITD sensitive when spikes were analysed over the full (600-ms) analysis window than the middle 350-ms analysis window. Recordings were considered ITD sensitive if the Rayleigh coefficient was >13.815 (see Methods). However, the vector strengths of the recordings were, on average, lower over the 600-ms analysis window as compared to the 350-ms analysis window (Wilcoxon rank-sum test; p < 0.001 in response to both SAM and transposed tones). Because the Rayleigh coefficient increases as the vector strength and/or the spike count increases, the likely reason for the
increased number of neurons classified as ITD sensitive when the full 600-ms window was analysed is a result of an increase in the number of spikes counted. This appears to compensate for the reduced vector strength over the 600-ms analysis window as compared to the 350-ms analysis window. All subsequent analyses pertain to responses from all 82 neurons, analysed over the middle 350-ms window of the stimulus.

Responses of a typical IC neuron sensitive to ITDs in the envelope of SAM and transposed tones are shown in Figure 3. The neuronal CF was 3.4 kHz, which is confirmed by the frequency-vs-level response area (Figure 3A). Peri-stimulus-time histograms (PSTHs) for monaural and diotic CF tones indicate the neuron to be excited by contralateral stimulation but unresponsive to ipsilateral stimulation (Figure 3B). The response to diotic stimulation was characterized by a peak spike rate lower than that to contralateral stimulation alone, but longer in duration.

Raster plots (Figures 3C-H) show responses to SAM or transposed tones which were interaurally delayed over a range of ITDs equivalent to interaural phase differences ± 0.5 cycles of the $f_m$ for $f_m$ between 10 Hz and 510 Hz. Because the range of ITDs encompassed by ± 0.5 cycles of IPD differs for different modulation rates, all functions are plotted with respect to IPD. Phase-locking to the envelope of the stimulus is evident for the response to both stimuli. The neuron was broadly tuned for IPD, favoring IPDs around zero. At 110 Hz $f_m$, IPD tuning was enhanced compared with 10 Hz, with discharge rates increasing at favorable IPDs and decreasing at unfavorable IPDs. The response at favorable IPDs was greater for transposed than for SAM tones. As $f_m$
increased, discharge rates fell, particularly during the latter portion of the stimulus, and responses became less modulated with envelope delay. The criteria for ITD-sensitivity were met in response to transposed tones at $f_m$ between 10-310 Hz and for SAM tones at 110 and 210 Hz.

Responses to SAM (left) and transposed (right) tones across all $f_m$ and IPDs are displayed as 3D mesh plots in Figure 3I. The neuron had a band-pass rate modulation transfer function (rMTF), with maximum firing rates at 110 Hz, a preference for IPD close to zero, and enhanced responses to transposed tones compared with SAM tones; peak discharge rates at the preferred $f_m$ were almost double those evoked by SAM tones.

Responses to IPDs imposed on SAM and transposed stimuli for a second IC neuron, with a CF of 4.3 kHz, are shown in Figure 4. PSTHs in Figure 4B indicate that contralateral stimulation evoked higher discharge rates than binaural stimulation; ipsilateral stimulation caused discharge rates to fall below the spontaneous rate of the neuron. Raster plots (Figure 4C-H) show spike times for six recordings in response to both SAM and transposed tones from 10 to 650 Hz $f_m$. Spike times were clearly phase-locked to the envelope at 10 Hz modulation. The neuron was sensitive to IPDs in both stimuli at 10 Hz; the pattern of the phase-locking to the modulation period changed with IPD and the spike rate fell around 0 IPD. At 138 Hz $f_m$ (Figure 4D), the neuron was IPD-sensitive to transposed, but not to SAM, tones. The neuron was insensitive to IPDs in either stimulus at the higher modulation rates (Figure 4E-H).
Figure 4I indicates that, at 10 Hz $f_m$, discharge rates were greater at favorable IPDs and lower at unfavorable IPDs, for transposed tones than for SAM tones. IPD sensitivity was low-pass with respect to $f_m$, unlike the band-pass response in Figure 3I. The responses also differ from those in Figure 3 in another important way. In Figure 4, discharge rates were minimal for values of IPD near zero whereas they were maximal near 0 IPD in Figure 3. Specifically, the mean best IPD in response to transposed tones was 0.39 cycles at $138 \, \text{Hz} \, f_m$ (equivalent to an ITD of $+2826 \, \mu\text{s}$) whereas, in Figure 3, the mean best IPD was 0.03 cycles at $110 \, \text{Hz} \, f_m$ (equivalent to an ITD of $+272 \, \mu\text{s}$).

Examples of recordings from two ITD-insensitive neurons are shown in Figures 5 and 6. The neuron referred to in Figure 5 (CF = 3.1 kHz) exhibited a strong onset response to diotic tones followed by a pause and then a more sustained response (Figure 5B). The PSTH to contralateral stimulation was similar but with a less well-defined pause, whereas ipsilateral stimulation resulted in a small onset response followed by inhibition of spontaneous firing (Figure 5B). Raster diagrams (Figures 5C-H) and the 3D mesh plot (Figure 5I) indicate that the neuron’s discharge rate increased with increasing $f_m$. That is, discharge rate was high-pass with respect to $f_m$. At higher values of $f_m$, the discharge rate was greater for transposed than for SAM tones while at the lower values of $f_m$, the opposite was the case. Thus, the discharge rate exhibited greater modulation as a function of $f_m$ in response to transposed tones. Visual inspection suggests some form of binaural interaction occurred at $f_m$s of 10 and 20 Hz; phase-locking appeared weaker near zero IPD (Figures 5C & D). Further, at 320 Hz and below, the discharge rate was lower for values of IPD near zero but the criteria for ITD-sensitivity were not met (Figure 5I).
Neural responses in Figure 6 (CF = 2.8 kHz) were also insensitive to ITDs in SAM or transposed tones. PSTHs to pure tones (Figure 6B) indicate the neuron to be excited by contralateral stimulation and the number of spikes to be facilitated by binaural stimulation. The PSTH to ipsilateral stimulation shows inhibition of spontaneous firing, following a small onset response. In contrast to the onset response of the neuron in Figure 3 and the onset-pause-sustained responses in Figures 4B and 5B, this neuron had an adapting PSTH to binaural stimulation. Raster plots in Figure 6C-H indicate strong phase-locking to the $f_m$. At 10 Hz (Figure 6C) it is clear that the spike times were more tightly phase-locked to the $f_m$ in response to transposed tones than SAM tones. Although discharges were entrained to the envelopes of both SAM and transposed tones, the response to the transposed tones occurred over a shorter duration of the modulation period, reflecting the shorter “on-period” (where the value of the sound pressure waveform is $> 0$) per cycle of the transposed tone. Mean firing rates in Figure 6I were not modulated with either $f_m$ or IPD.

Mean firing rates of a further nine neurons, chosen to illustrate the variety of responses obtained, are displayed as 3D mesh plots in Figure 7. Neurons classified as sensitive to envelope-based ITDs and with firing rates that peak close to zero IPD are arranged in the top row and neurons with firing rates that peak close to IPDs of 0.5 cycles in the middle row. The bottom row shows neurons insensitive to envelope-based ITDs. For each ITD-sensitive neuron, the maximum Rayleigh coefficient (which was used as a measure of sensitivity to ITDs) was determined across all recordings in response to SAM and
transposed tones and both the top and middle rows (Figures 7A-C and D-F) are arranged from left to right in order of decreasing maximum Rayleigh coefficient.

All three neurons in the top row of Figure 7 showed greater modulation of their discharge rate in response to transposed tones than SAM tones, as a function of both IPD and $f_m$. The neuron in Figure 7A was very strongly sensitive to ITDs conveyed by transposed tones over a specific range of modulation rates (138 to 266 Hz), with the greatest Rayleigh coefficient at 266 Hz, whereas it was largely unresponsive to ITDs conveyed by the SAM stimulus. The other two neurons (Figures 7B & C) responded strongly to both transposed and SAM tones and were sensitive to ITDs over a specific range of $f_m$ (around 20 & 40 Hz in Figure 7B and around 115 & 220 Hz in 7C in response to transposed tones and around 40 Hz in Figure 7B in response to SAM tones). Generally, when neurons were sensitive to ITDs in both stimuli, sensitivity occurred at the same $f_m$ but extended to a wider range of $f_m$ in response to transposed tones (e.g. in 7B). Overall, discharge rates were more highly modulated to ITDs conveyed by transposed stimuli than they were to ITDs conveyed by SAM stimuli. In addition, rates were consistently higher at favorable ITDs and lower at unfavorable ITDs for ITDs conveyed by transposed as compared to SAM tones. At any $f_m$, preferred IPD tuning (where such sensitivity occurred) was similar in response to both stimuli i.e. peaks and troughs in firing rates occurred at similar IPDs. As a function of $f_m$, the neural response in Figure 7C had a trough in its firing rate at 115 Hz $f_m$ which was more pronounced (i.e. reaching a lower firing rate) in response to transposed tones than SAM tones.
Discharge rates as a function of IPD were also more modulated in response to transposed tones for the responses in Figures 7D & E. Neurons were insensitive to ITDs in the envelope of SAM tones whereas in Figures 7D & E, sensitivity occurred for more than one $f_m$ in response to transposed tones. Atypically for our sample of neurons, the firing rate of the neuron in Figure 7F was more modulated as a function of IPD in SAM tones as compared to transposed tones at 250 Hz $f_m$ (although the reverse was true at 330 Hz). There was a response minimum at 250 Hz $f_m$ which was lower for transposed tones and may explain the reduced modulation in firing rate as a function of IPD to transposed tones. As in Figures 7B & C, the IPD tuning in Figures 7F was similar for SAM and transposed tones, i.e. peaks and troughs in firing rate were similarly positioned and, for each neuron, there was a modulation rate at which sensitivity was greatest (210 Hz in 7D, 10 Hz in 7E and 250 Hz in 7F).

The vector strengths of all ITD-sensitive recordings in response to transposed tones were, on average, higher than the vector strengths of the ITD-sensitive recordings in response to SAM tones (Wilcoxon rank-sum test, $p = 0.0011$), as were the Rayleigh coefficients (calculated using the Rayleigh coefficients from all recordings; Wilcoxon rank-sum test, $p < 0.001$). Both observations are consistent with transposed tones evoking discharge rates that are better modulated with envelope ITDs than SAM tones.

Neurons for which responses are shown in the bottom row of Figure 7 were not sensitive to envelope-ITDs in SAM or transposed tones. However, their discharge rates were modulated as a function of $f_m$. For each of the three neurons, changes in discharge rate as
a function of $f_m$ were more pronounced in response to transposed tones than SAM tones. Although not of central interest to the current study, neural sensitivity to the $f_m$ was examined, in the form of rate modulation transfer functions (rMTFs). Responses of 71 neurons were examined where SAM and transposed tones were presented with zero ITD and at $f_m$ between 10 and at least 300 Hz. Responses were considered modulated as a function of $f_m$ if discharge rates varied by $\geq 70\%$ of the maximum discharge rate over the range of $f_m$s tested. In response to SAM tones, 12 neurons did not meet this criterion while, in response to transposed tones, only one neuron did not meet the criterion ($p<0.01$, test for independence of paired proportions). Further characterization of the rMTFs is not shown because a comprehensive range of $f_m$s was not examined for all neurons. We conclude, however, that neural responses to transposed tones were more modulated as a function of $f_m$.

Sensitivity to ITD as a function of $f_m$

The number of ITD-sensitive recordings as a proportion of the total number of recordings is displayed in Figure 8, calculated as a function of $f_m$, in 50-Hz bins. Sensitivity to ITDs could be described as band-pass as a function of $f_m$, with the greatest proportion of ITD sensitive recordings between 60 and 310 Hz $f_m$. The sharpest reduction in the proportion of ITD-sensitive recordings occurred as the $f_m$ increased above 310 Hz and no recordings were considered sensitive to ITDs at $f_m$ above 550 Hz. The total number of recordings in each $f_m$ bin is shown in the inset of Figure 8.
The distribution of best phases (BPs) for all ITD-sensitive recordings is shown in Figure 9. Positive BPs indicate a preference for sounds leading at the contralateral ear; negative BPs indicate a preference for sounds leading at the ipsilateral ear. BPs to both transposed tones (right) and SAM tones (left) cluster around 0 and 0.5 cycles, with more positive than negative values. This is consistent with previous physiological investigations of neural sensitivity to ITDs which indicate that IC neurons have, predominantly, BPs corresponding to sounds leading at the contralateral ear (16; 38; 57; 59; 60).

For some ITD-sensitive neurons in this study it was possible to make recordings that were closely spaced in \( f_m \), and the ITD tuning of these neurons as a function of \( f_m \) was investigated. Responses from the eight neurons with the highest number of ITD-sensitive recordings in response to transposed tones are shown in Figure 10A. Discharge rates were normalized to maximum at each \( f_m \) to emphasize the ITD tuning. Responses to SAM tones are shown on the left and responses to transposed tones are shown on the right. Note that abscissae are scaled to take account of the range of modulation rates at which sensitivity to ITDs was observed in each neuron. Neurons are organized by eye, according to whether maximum discharge rates were aligned around a common ITD for different \( f_m \)s (top two rows) or whether minimum discharge rates were aligned around a common value of ITD for different \( f_m \)s (bottom row). For most neurons, the positions of response maxima and response minima as a function of \( f_m \) were similar for SAM and transposed tones.
Neurons in which response maxima are aligned are referred to as “peak” neurons and neurons in which response minima are aligned are referred to as “trough” neurons (3; 59). Phase plots (BP plotted as a function of $f_m$) are shown in Figure 10B for the same neurons for which responses are shown in Figure 10A. Theoretically, for neurons described as “peak” or “trough” neurons, BP changes systematically with pure-tone frequency or modulation frequency (3; 26; 46). The slope of a linear regression describing the relation between BP and frequency is referred to as the characteristic delay (CD) and the intercept with the ordinate axis is referred to as the characteristic phase (CP) (46; 59). For “peak” neurons, the CD is equal to the ITD at which the neuron responds maximally. A CP of 0 would be predicted from perfect coincidence of excitatory (EE) inputs. A CP of 0.5 would be predicted from coincidence of excitatory and inhibitory inputs (EI). Although it is possible to fit a linear regression to some of the phase plots in Figure 10B (e.g., Figures 10Bd & h in response to SAM tones, and in Figure 10Ba in response to transposed tones have $R^2 > 0.6$), for other neurons the linear regression is not appropriate (e.g., Figure 10Bd). Previous investigations of sensitivity to ITDs of pure tones and high-frequency SAM tones in the IC reported a similar range of neural responses; phase plots are often not well-fitted by linear regression and, where phase plots are well fitted by linear regression, CPs may not be close to 0 or 0.5 (3; 35; 37; 57; 59).

Detection thresholds for ITDs in the envelope of transposed tones: comparison with SAM tones and low-frequency pure tones.

Of primary interest for this investigation was the neural sensitivity to ITDs in SAM as compared to transposed tones. In order to quantify the sensitivity to ITDs in response to
transposed tones and SAM tones, high-resolution IPD functions were obtained and neural discrimination thresholds were calculated using ROC analysis (see Methods).

High-resolution IPD functions were obtained in response to SAM and transposed tones from 14 neurons. For 12 neurons, functions were obtained only at a single $f_m$. For two neurons, functions were obtained at two $f_m$s (for one of these two neurons, a second high-resolution function was obtained in response to transposed, but not SAM, tones). This gave a total of 16 functions obtained in response to transposed tones, and 15 functions obtained in response to SAM tones. Six examples of such functions are shown in Figures 11A-F. On the left, the mean discharge rates and standard deviations (error bars) are shown in response to SAM (blue) and transposed (red) tones. The abscissae differ across plots because they reflect the presentation of different values of $f_m$ which, given the constant range of IPDs presented, resulted in different ranges of ITDs being presented.

Neurons in Figures 11A, C & E responded best to IPDs around zero cycles whereas those in Figures 11B, D & F responded best to IPDs around 0.5 cycles. Consistent with the responses shown in Figures 3, 4 & 7, discharge rates were more highly modulated to transposed tones as compared to SAM tones.

The probability of correctly discriminating between zero ITD and all other ITDs presented is shown in the neurometric functions in each of Figures 11A-F (right panels). Thresholds for discrimination (see Methods) are marked by black horizontal lines. From these ITDs, the single value closest to zero was taken to be a measure of the Just
Noticeable Difference (JND). Neural JNDs are shown by dashed lines for SAM (blue) and transposed (red) tones.

Figure 11A depicts data obtained with a $f_m$, of 100 Hz. At that rate of modulation, the neural JNDs were 644 µs for SAM, and 258 µs for transposed tones. Figure 11C shows data obtained from the same neuron in Figure 11A, but for a $f_m$ of 140 Hz. For that rate of modulation, JNDs were 672 µs and 177 µs for SAM and transposed tones, respectively. Consistent with the greater modulation in firing rates in response to transposed tones as a function of ITD, JNDs derived from these functions were correspondingly lower for transposed than for SAM tones. This is also the case in Figures 11B, D and E. For the responses shown in Figure 11E, the neural JNDs were large, being 10 ms and 3.3 ms in response to SAM and transposed tones, respectively. These large values were, at least in part, a result of the large variability observed in the measures of firing rate compared to the mean firing rate change over a cycle of IPD. For the responses depicted in Figure 11F, the JND in response to transposed tones was > 2 ms and, in response to SAM tones there were no ITDs that were determined to be discriminable from an ITD of zero.

JNDs $\leq 4$ ms for SAM and transposed tones are plotted as a function of $f_m$ in Figure 11G. All 16 JNDs calculated in response to transposed tones were $\leq 4$ ms and are marked by red diamonds in Figure 11G. For two neurons, JNDs were calculated from high-resolution functions at two $f_m$s. These are indicated by a square or a circular black outline.
around the red diamonds. For one of these neurons (the JND marked by a circular outline) a high-resolution IPD function was obtained in response to SAM tones at only one \( f_m \). Of the 15 JNDs obtained in response to SAM tones, only 10 were \( \leq 4 \) ms, and these are plotted in Figure 11G as blue circles. The remaining 5 JNDs were either greater than 4 ms or were immeasurable. The horizontal dotted lines in Figure 10G indicate estimates of the maximum ITD experienced by the guinea pig. The lower value is from a theoretical consideration of the maximum difference in time for sound traveling around a spherical head (38) and the higher value is from head-related transfer functions measured in the guinea pig (53) (see Discussion).

Also shown in Figure 11 are neural JNDs for ITD measured with low-frequency tones below 400 Hz (48; 49) for low-CF neurons in the guinea-pig IC (squares). Median JNDs in response to SAM (blue), transposed (red) and low-frequency tones (black), are marked by asterisks. Overall, JNDs in ITD are lower for transposed tones than for SAM tones (Wilcoxon rank-sum test for equal medians, \( z = 2.6, p < 0.05 \)). In the case of just one neuron (also shown in Figure 7F), and only at a \( f_m \) of 200 Hz, was the JND lower in response to SAM tones than transposed tones (0.5 ms vs. 1.4 ms). When \( f_m \) was increased to 280 Hz, the JND for that same neuron was 320 \( \mu \)s in response to transposed tones, lower than either JND at 200 Hz.

The lowest JNDs in response to transposed tones were comparable to JNDs obtained in response to low-frequency tones, at corresponding frequencies, whereas the lowest JNDs in response to SAM tones were substantially higher. For \( f_m \) over 280 Hz, no neural JNDs
could be calculated. As previously discussed (see Figure 8), responses indicating sensitivity to changes in ITD were rarely recorded at these rates and when such responses were observed they were only weakly modulated with variations in ITD. In contrast, neural JNDs in response to low-frequency tones decrease as the tone frequency increases up to ~ 500 Hz (48). JNDs in response to low-frequency tones between 280 and 400 Hz contribute to the low median JND for low-frequency tones.

*Phase-locking to the envelope modulation*

Transposed tones are designed to evoke a similar temporal pattern of action potentials from high-frequency ANFs to that evoked in low-frequency fibers by low-frequency tones. As previously discussed, high-frequency ANFs would not be expected to show this pattern in response to SAM tones. It is predicted, therefore, that ANFs should show “tighter” phase-locking to the period of a transposed than a SAM waveform (compare Figure 1B and 1C, right). We examined phase-locking to the modulation rate in our sample of IC neurons.

Period histograms of the number of spikes occurring throughout the period of SAM and transposed waveforms were calculated for the four example neurons in Figures 3 – 6 and are displayed in Figure 12A-D, respectively. Paired responses to SAM (top) and transposed (bottom) tones are shown in response to each modulation rate, with the response to 10 Hz modulation on the left. The response is clearly phase-locked to 10 Hz in Figure 12A. In response to transposed tones, spikes occur over a narrower range of the
stimulus cycle than they do in response to SAM tones. In the second pair of histograms, measured at 110-Hz modulation, phase-locking occurred in response to both stimuli but was tighter in response to transposed tones. As the modulation rate increased, phase-locking became weaker and was no longer apparent at 310 Hz. A similar pattern is observed in Figures 12B, C & D; the phase-locking is limited to the lower modulation rates in response to both stimuli and in response to transposed tones is “tighter” than to SAM tones. This likely reflects the shorter “on-period” per cycle of the transposed tones.

To summarize phase-locking to the modulation period for all neurons in our sample, the vector strength was calculated at each value of $f_m$ and IPD that yielded significantly phase-locked responses (Rayleigh coefficient $> 13.815$, $p < 0.01$). A three factor ANOVA for stimulus type, $f_m$ and IPD revealed each factor had a significant effect on the phase-locking to the modulation rate (all $p < 0.001$). Changes in phase-locking, as a function of IPD, can be seen in the raster plot in Figure 4C.

In order to examine the effect of stimulus type and $f_m$ on the vector strength, the vector strength values (for all recordings that showed significant phase-locking) were grouped into logarithmically-spaced $f_m$ bins. The means for each $f_m$ bin are shown in Figure 13. Overall, phase-locking was greater in response to transposed than SAM tones ($t$-test, $p<0.001$) and was low-pass as a function of $f_m$ for both stimuli. The vector strengths in response to both stimuli decreased from similar $f_m$s (around 100 Hz) and the characteristics of the reduction were qualitatively similar. Phase-locking is essential for coding temporal information that mediates sensitivity to ITDs in binaural neurons. The
reduction in phase-locking with increasing $f_m$ might underlie the limit on sensitivity to envelope-based ITDs observed at around 300 Hz (see Figure 8 & 10) (see Discussion).
Discussion

The principal finding of this study is that neural sensitivity to ITDs conveyed by the envelopes of high-frequency sounds can be as great as that observed for ITDs conveyed by the fine-structure of low-frequency sounds. The neural data presented here, recorded from the IC of the guinea pig, are consistent with the enhanced ITD-sensitivity yielded by transposed tones observed in psychophysical data with human listeners (9). More neurons were sensitive to ITDs within the envelope of transposed as compared to SAM tones. Firing rates were more modulated as a function of ITD, and JNDs for ITD were consistently lower, in response to transposed tones as compared to SAM tones. We conclude that neurons are more sensitive to ITDs within transposed tones than SAM tones.

Comparison with human psychophysics using transposed tones

The current study of neural responses to ITDs was motivated by the finding that, for human listeners, transposed tones confer an improvement in sensitivity to ITDs at high carrier frequencies compared to SAM tones (9). The neural JNDs we observed follow a similar pattern to those obtained from human listeners; threshold-ITDs obtained with transposed tones were smaller than those obtained with SAM tones. It is noteworthy that Bernstein and Trahiotis (2002) found threshold-ITDs to be immeasurable with the high-frequency stimuli for rates of modulation above 256 Hz, $f_m$. A similar limitation was observed in the neural data presented in the current study.
Bernstein and Trahiotis (2002) used a 300-ms presentation and a two-cue, two-alternative forced choice adaptive task to determine ITD- thresholds. Clearly, one should be cautious when comparing human psychophysical data with neural recordings from guinea pigs, given the species differences and questions concerning the contribution of the response of single neurons to behavior. (These are well discussed in Skottun et al. 2001). Additionally, guinea pigs were subject to urethane anesthesia. Despite this, a similar dependence of ITD-discrimination thresholds on $f_m$ and the type of stimulus was observed between neural responses and human listeners (9).

**Comparison with previous in vivo electrophysiology**

Previous authors have reported mammalian neural sensitivity to ITDs in the envelopes of high-frequency stimuli including data using SAM tones recorded from the MSO (2; 3), LSO (26; 30) and IC (4; 5), noise recorded from the IC (27), clicks recorded from the IC (16) and high-frequency tones with trapezoid envelopes recorded from the IC (60). The neural responses to SAM and transposed tones, in the current study, were qualitatively similar to these previous studies; BPs were predominantly leading at the contralateral ear and phase plots were not necessarily well-described by linear regression. Clustering of BPs close to 0 and 0.5 cycles was reminiscent of EE and EI interactions described previously in response to high- (2) and low-frequency sounds (59).

We compared neural JNDs obtained in the current study with neural JNDs obtained in response to ITDs in low-frequency tones in the IC of the guinea pig by Shackleton and colleagues (48-50). Neural JNDs in response to transposed tones were comparable to
JNDs in response to low-frequency tones at $f_m < \sim 300$ Hz, consistent with human psychophysics (9). The current study used a longer stimulus duration (500 ms vs. 50 ms) and a smaller number of repetitions (6-13 vs. 100) than were used by Shackleton and colleagues. Onset and offset responses were excluded in the calculation of spike rates for ROC analysis in the current study, in contrast to the inclusion of the whole stimulus-driven response in the calculation of low-frequency JNDs by Shackleton and colleagues. The stimulus used by Shackleton and colleagues had no onset ITD. Despite the differences in stimulus characteristics, neural JNDs from transposed and low-frequency tones were similar. Finally, no qualitative differences between JNDs calculated over the 600-ms analysis window and the 350-ms analysis window were observed in response to SAM and transposed tones.

**Phase locking and a $f_m$ limitation on sensitivity to ITDs**

We suggest that the enhanced neural ITD sensitivity observed with transposed stimuli was a result of the temporal pattern of action potential firing in high-frequency ANFs, stemming from the temporal signature of the envelopes of those stimuli. Transposed tones were originally designed to overcome the nature of peripheral auditory processing hypothesized to be responsible for the relatively poor binaural performance observed with high-frequency stimuli (55). The transposed stimulus was designed to produce an output at the level of high-frequency ANFs which, in terms of its temporal pattern, mimics the output normally observed at low-frequencies (Figure 1). Data describing ANF responses to transposed tones are not yet published but the analysis of phase-locking in the IC presented here indicates that phase-locking is enhanced in response to transposed as
compared to SAM tones (Figure 13). This enhancement in phase-locking does not require any increased ability to entrain to the transposed waveform but can be explained simply by the different envelopes of the stimuli. If a neuron’s pattern of firing followed exactly the time course of SAM and transposed tones, the phase-locking to the transposed waveform would be “tighter” because the transposed waveform itself occurs over a narrow range of the modulation period.

The tighter phase-locking in response to transposed tones may explain the enhancement in sensitivity to ITDs observed in behavioral experiments with human listeners and in the physiological recordings presented here. Specifically, binaural neurons in the superior olivary complex (SOC) may receive temporally more precise patterns of action potentials in response to transposed tones than in response to SAM tones. Inputs arriving in phase at the SOC are more likely to produce coincident action potentials, with a high output from EE (peak) neurons, whereas inputs arriving out of phase produce fewer coincident action potentials, and reduced output from EE neurons. Responses to SAM tones were still highly phase-locked to the SAM waveforms and an alternative explanation for the improvement in ITD-sensitivity might involve, in some direct fashion, the “off period” in response to transposed and low-frequency tones. Binaural inputs might not be any more likely to arrive coincidently in response to transposed tones but rather the “off period” might be crucial for the efficient binaural processing of ITDs. In any case, the notion that the enhancement observed with transposed tones is directly tied to the properties of the stimulus is consistent with Bernstein and Trahiotis’ (2002) finding that the enhancements
observed psychophysically were well accounted for by changes in the normalized interaural correlation of the waveforms as processed by the auditory periphery.

Although transposed tones are designed to provide a pattern of responses in high-frequency ANFs similar to that seen in low-frequency fibers in response to low-frequency tones, the response patterns are, in fact, unlikely to be identical. In response to low-frequency tonal stimulation auditory hair cells are hyperpolarized during the “off” or non-preferred half-cycle of the motion of the basilar membrane and stereocilia are deflected towards their shortest row and the mechanically gated ion channels are closed (20). During these “off” periods, the response of the auditory nerve can fall below its spontaneous rate. Although such a reduction would not be expected in response to high-frequency transposed tones, neural JNDs for modulation rates below 300 Hz are remarkably similar for transposed and low-frequency tones.

While phase-locking is greater in response to transposed tones, phase-locking decreases with $f_m$ in the same manner in response to either SAM or transposed tones (Figure 12). This is consistent with the rate-limitation of phase-locking to amplitude modulated sounds measured previously (31; 34), and with behavioral accounts of monaural temporal modulation transfer functions (19; 33). Not surprisingly, given that ITD-processing is dependent on phase-locking, few neurons were ITD-sensitive to rates of modulation above 300 Hz (Fig 8A & 11G). In contrast, ITD sensitivity in response to low-frequency tones occurs at frequencies above 300 Hz; human JNDs are minimal around 1 kHz (32) and guinea pig IC neural JNDs are minimal around 500 Hz (48). Previous neural
investigations (2; 31; 60) and psychophysical investigations (8; 9; 39; 42) at high-
frequencies have observed a $f_m$ limitation on sensitivity to ITDs.

As the modulation frequency of an amplitude-modulated stimulus increases such that the
spectral components (see Figure 2) fall outside the spectral receptive field of an ANF,
phase-locking to the amplitude modulation is lost. This forms an absolute limit on the
modulation rates at which sensitivity to envelope ITDs is possible. For transposed tones,
the middle 5 spectral components are required to gain a binaural advantage (55). As $f_m$ is
increased, therefore, the components that are most spectrally distant from CF are
attenuated most. Thus, phase-locking and ITD sensitivity would be predicted to decrease.
We might predict that the $f_m$ at which ITD-sensitivity can occur would increase with
increasing CF because ANF spectral receptive fields widen with increasing CF (18).
Bernstein & Trahiotis (1994) tested this hypothesis and found that this was not the case.
They suggested instead, the existence of some other limit on the $f_m$s at which ITD-
sensitivity occurs. To account quantitatively for their psychophysical data, Bernstein and
Trahiotis (2002) included a 150-Hz low-pass filter within an interaural correlation-based
model of ITD sensitivity to simulate the limitation on ITD sensitivity imposed by the
modulation rate. The filter they employed was similar to that employed in accounts of
monaural temporal modulation transfer functions (19; 33). Consistent with the findings
of these psychophysical studies, the data from the present study revealed no clear relation
between CF and the maximum $f_m$ at which ITD-sensitive recordings could be obtained.
Joris and Yin (1998) proposed that a neural limit on sensitivity to amplitude modulation and to ITDs in high-frequency sounds arises in the LSO. It is likely that any such limitation results from the membrane properties of LSO neurons e.g. the presence or absence of ion channels capable of altering the ability of neurons to follow the temporal structure of their inputs (1).

**Neural JNDs and physiological detection of ITDs**

Natural sounds, such as human speech, have an envelope structure replete with onsets, offsets and “off-periods” (47). ITDs conveyed by the envelope of high-frequency stimuli could be relevant to binaural tasks such as sound localization and signal detection in noise (10). Would the neural JNDs measured in the current study be useful to the guinea pig? In Figure 10G, two measures of the “maximum” ITD that a guinea pig might experience are marked by dotted lines. The value of 330 µs is the maximum ITD, at low-frequencies, in HTRFs measured using noise stimuli by Sterbing et al. 2003 whereas the value of 150 µs is the result of a theoretical consideration of the time sound takes to travel around a spherical head (38). Below 300 Hz, there is only one JND for low-frequency tones that is <= 150 µs, and below 250 Hz there are few JNDs <= 330 µs. Similarly, in response to transposed tones, there are few JNDs below 330 µs. If single neurons are sufficient to encode ITDs, this suggests that the usefulness of these tuning curves is limited because the JNDs are close to the maximum ITD that can be experienced. Shackleton et al. (2003) found that the lowest JNDs in response to low-frequency tones were not generated using 0 ITD as a reference. In general, they found that JNDs were minimal with reference ITDs ipsilateral to the IC from which the neuron
was recorded. Similarly, there may be regions of the IPD-tuning curves in response to transposed tones that have greater accuracy for ITD discrimination. We do not propose that single neurons are sufficient to encode ITDs; it is more likely that the responses of many neurons are pooled.

Despite limitations in the temporal coding of neurons to high-frequency sounds, the results presented here demonstrate that under conditions in which binaural neurons receive appropriate spike patterns, sensitivity to ITDs conveyed by high-frequency stimuli can be equivalent to that observed in response to low-frequency stimuli. This suggests, as first conjectured by Colburn and Esquissaud (1976) that mechanisms underlying ITD sensitivity in low- and high-frequency channels of the auditory system are, to a first approximation, equivalent.
Acknowledgements

§Present Address for NJI:- Centre for the Neural Basis of Hearing, Physiological Laboratory, Downing Street, Cambridge, CB2 3EG, United Kingdom.

We acknowledge Trevor Shackleton and Alan Palmer for use of their software.

Grants

SJG is sponsored by the Wellcome Trust four year programme in Neuroscience.

This work was funded by the MRC (DM).

Previous Reports

Parts of this work have previously published in abstract form.


Reference List


25. **Jones PJ and Williams RP.** An experiment to determine whether the interaural time differences used in lateralizing middle and high-frequency complex tones is dependent in any way on fine structure information. *Acustica* 47: 164-169, 1981.


Legends

Figure 1. The time-waveforms of, and modeled auditory nerve fiber responses to, three sounds. Left: The sound pressure waveform of a low frequency tone (A), SAM tone (B) and a transposed tone (C). Right: The results of peripheral auditory processing of each of the sounds (modeled by half-wave rectification and low-pass filtering). Adapted from Bernstein & Trahiotis (2002).

Figure 2. The power spectra of exemplars of the SAM (blue) and transposed (red) tones used in the experiment. Spectra were recorded by presenting SAM (blue) and transposed (red) tones \( f_c = 4 \) kHz and \( f_m = 160 \) Hz to a spectrum analyser.

Figure 3. An IC neuron (28104; CF = 3.4 kHz; threshold = -67 dB re maximum system output) that was classified as ITD-sensitive to both SAM and transposed tones. (A) The neuron’s frequency-versus-level response area. Spike counts are indicated by gray level. (B) PSTHs to 50 ms diotic, contralateral (Contra) and ipsilateral (Ipsi) tones at CF. (C) – (H) Raster plots for recordings at 10, 110, 210, 310, 410 and 510 Hz \( f_m \) respectively. All axes are scaled as in G. Each band on the ordinate shows spike times in response to repeat presentations of different IPDs from -0.5 to +0.5 cycles. Black bars below the abscissae indicate the middle 350 ms of the stimulus. (I) 3D mesh plots of mean spike rate, calculated over the period indicated by black bars in (C)-(H), as a function of IPD and \( f_m \). Mean spike rates were averaged over 3 consecutive IPDs.
Figure 4. Another IC neuron (25916; CF = 4.3 kHz; threshold = -68 dB re maximum system output) that was classified as ITD-sensitive to both SAM and transposed tones. The figure follows the same format as Figure 4. (C) - (H) Raster plots for recordings at 10, 138, 266, 394, 522 and 650 Hz \( f_m \). Black bars indicating the middle 350 ms of the stimulus are not shown.

Figure 5. An IC neuron (20108; CF = 3.4 kHz; threshold = -65 dB re maximum system output) that was not classified as ITD-sensitive to either SAM or transposed tones. The Figure follows the same format as Figure 4. (B) PSTHs to 200 ms stimulation (C) – (H) Raster plots for recordings at 10, 20, 40, 80, 160 and 320 Hz \( f_m \) respectively. Raster plots are not displayed for the responses at 640 Hz \( f_m \) which are included in (I). Black bars indicating the middle 350 ms of the stimulus are not shown.

Figure 6. Another IC neuron (25805; CF = 2.8 kHz; threshold = -79 dB re maximum system output) that was not classified as ITD-sensitive to either SAM or transposed tones. The figure follows the same format as Figure 4. (C) – (H) Raster plots for recordings at 10, 30, 50, 70, 110 and 150 Hz \( f_m \) respectively. Raster plots are not displayed for the responses at 90 or 130 Hz \( f_m \) which are included in (I). Black bars indicating the middle 350 ms of the stimulus are not shown.

Figure 7. 3D mesh plots (same format as in Figure 4I) for 9 IC neurons. Firing rates in response to ITDs conveyed by SAM (left) and transposed (right) tones, at different \( f_m \) are
shown. Top row: ITD-sensitive neurons with peak responses close to 0 ITD. (A) Neuron 25806, CF = 7.8 kHz, threshold = -44 dB re maximum system output. (B) Neuron 12505, CF = 3.6 kHz, threshold = -81 dB re maximum system output. (C) Neuron 19202, CF = 2 kHz, threshold = -52 dB re maximum system output. Middle row: ITD-sensitive neurons with a trough in response close to 0 ITD. (D) Neuron 27204, CF = 2.2 kHz, threshold = -55 dB re maximum system output. (E) Neuron 12821, CF = 8.8 kHz, threshold = -62 dB re maximum system output. (F) Neuron 20818, CF = 4.8 kHz, threshold = -70 dB re maximum system output. Bottom row: Neurons classified as insensitive to ITDs. (G) Neuron 22507, CF = 5.1 kHz, -85 dB re maximum system output. (H) Neuron 20804, 5.4 kHz, -81 dB re maximum system output. (I) Neuron 25605, CF = 3.8, threshold = -73 dB re maximum system output.

Figure 8. The proportion of recordings that were considered sensitive to ITDs plotted as a function of modulation rate. Inset: The total number of recordings at each modulation rate.

Figure 9. Best Phases (BPs) of all ITD-sensitive recordings in response to SAM (left) and transposed (right) tones.

Figure 10. ITD tuning as a function of modulation frequency. (A) Normalized firing rates, at different $f_m$, plotted as a function of ITD for 8 IC neurons (a)-(h) in response to transposed (left of each pair) and SAM (right of each pair) tones. (a) 19202, (b) 22207, (c) 18808, (d) 20709, (e) 22205, (f) 20818, (g) 20810 & (h) 27204. (B) Best Phases (BPs) plotted as a function of modulation rate for the same 8 neurons in (A) in response to
transposed (closed circles) and SAM (open circles) tones. Note that the y-axes in a-d and in e-h are different.

Figure 11. The derivation of neural ITD discrimination thresholds (just noticeable differences, JNDs). (A)-(F) Left: The mean firing rates and standard deviations in response to SAM (blue) and transposed (red) tones plotted as a function of ITD from 6 “high-resolution” functions (see methods). Right: The (three-point averaged) neurometric functions associated with the firing rates shown to the left. The y-axis shows the probability of correctly discriminating the ITD of a transposed or SAM tone from zero ITD. Black lines indicate the criterion level for discrimination from zero. All discriminable ITDs are marked by a filled circle for SAM tones (blue) and transposed tones (red). The smallest discriminable ITD is marked by the dotted blue line (SAM) and dotted red line (transposed). (A) Neuron 22508, $f_m = 100$ Hz. (B) Neuron 20810, $f_m = 180$ Hz. (C) Neuron 22508, $f_m = 140$ Hz. (D) Neuron 27204, $f_m = 180$ Hz. (E) Neuron 23008, $f_m = 40$ Hz. (F) Neuron 25916, $f_m = 120$ Hz. (G) JNDs in response to SAM (blue), transposed (red) and low-frequency tones (black (from Shackleton & Palmer (2004))). Median JNDs are marked by asterisks at $x = 0$. The dotted black lines indicate, approximately, the maximum ITD a guinea pig could experience, calculated from two sources; lower line (150 us), maximum ITD estimated from a spherical head model (38), upper line (330us), maximum ITD from HTRF measurements (53). See discussion. Data points from the same neuron are indicated by either an unfilled black circle, for one neuron, or an unfilled black square, for another neuron.
Figure 12. Phase locking as a function of modulation rate. (A) Period histograms of spikes grouped over one cycle of modulation rate in response to SAM (top) and transposed (bottom) tones for the neuron in Figure 4. The corresponding modulation rates ($f_m$), in Hz, are marked in italics atop the histograms of responses to SAM tones. (B) Period histograms for the neuron in Figure 5, (C) Period histograms for the neuron in Figure 6 and (D) Period histograms for the neuron in Figure 7.

Figure 13. Phase locking as a function of modulation rate. The mean and standard deviation vector strengths of phase locking, from all neurons, to SAM (open circles) and transposed (closed circles) tones are plotted with error bars showing ± 1 standard deviation of the mean. The vector strengths obtained with the transposed tones have been slightly offset horizontally.

Table 1. The number of neurons and recordings sensitive to ITDs, analysed over a 600 ms and 350 ms period.
Figure 1

Figure 2
Figure 4
Neural Sensitivity to Interaural Envelope Delays

Figure 5
Figure 7
Figure 8

Figure 9
Figure 10
Figure 11
Figure 12
Figure 13