JARO

Contributions of Intrinsic Neural and Stimulus Variance to Binaural Sensitivity

Trevor M. Shackleton and Alan R. Palmer

MRC Institute of Hearing Research, University Park, Nottingham, NG7 2RD, UK

ABSTRACT

The discrimination of a change in a stimulus is determined both by the magnitude of that change and by the variability in the neural response to the stimulus. When the stimulus is itself noisy, then the relative contributions of the neural (intrinsic) and stimulus induced variability becomes a critical question. We measured the contribution of intrinsic neural noise and interstimulus variability to the discrimination of interaural time differences (ITDs) and interaural correlation (IAC). We measured discharge rate versus characteristic frequency (CF) tone ITD functions, and CF-centered narrowband noise ITD and IAC functions in interleaved blocks in the same units in the inferior colliculus of urethane-anesthetized guinea pigs. Ten "frozen" tokens of noise were synthesized and the responses to each token were separately analyzed to allow the relative contributions of intrinsic and stimulus variability to be assessed. ITD and IAC discrimination thresholds were determined for a simulated two-interval forced-choice experiment, based on the firing rate distributions, using receiver operating characteristic analysis. On average, between stimulus variability contributed 19% (range, 1.5–30%) of the variance in noise ITD discrimination and 27% (range, 3–50%) in IAC discrimination. Noise ITD thresholds were slightly higher than tone ITD thresholds. Taking the mean of the thresholds for individual noise tokens gave a similar result to pooling across all noise tokens. This implies that although the stimulus induced variability is measurable, it is insignificant in relation to the intrinsic noise in ITD and IAC discrimination.

Correspondence to: Trevor M. Shackleton · MRC Institute of Hearing Research · University Park, Nottingham, NG7 2RD, UK. fax: +44-115-951-8503; email: trevor@ihr.mrc.ac.uk

INTRODUCTION

The detection of faint targets in noisy backgrounds, and the discrimination of stimulus changes, are common ecological problems. Derived from similar problems in communications engineering, signal detection theory (SDT) (Green and Swets 1988) has been applied to these problems in the various sensory modalities. SDT argues that the sensitivity with which a change in stimulus can be detected is as crucially dependent on the variability in the response to a stimulus, as on the magnitude of the change in response when the stimulus is changed. If there were no environmental noise and the stimulus were perfectly deterministic then, in the absence of internal noise, this theory would predict that infinitesimally small increments in the stimulus could be discriminated. However, since under these conditions there is a minimum stimulus change that can be discriminated, there must be some source of internal, or intrinsic, variability. If the stimulus is not perfectly deterministic, or external noise is added, then the relative contributions of internal and external noise become important (Green 1964).

In auditory science the relative contributions of internal and external noise have mostly been addressed in the context of the detection of a tone in noise, either monaurally or binaurally (Swets et al. 1959; Spiegel and Green 1981; Siegel and Colburn 1983, 1989; Gilkey et al. 1985; Isabelle and Colburn 1991) and is currently the focus of study by Carney

and her colleagues (e.g., Evilsizer et al. 2002; Zheng et al. 2002; Davidson et al. 2006). Although there are many subtleties in the results, these experiments essentially show that the internal noise is of the order of between one and three times greater than the external noise. Furthermore, these experiments have shown that responses to individual, "frozen," noise tokens are dependent on factors other than the energy of the noise and is replicated across listeners. These findings are somewhat at variance with the assumptions of Colburn (1977a), who found in his modeling of the low-frequency, binaural, effects of "antiphasic detection" and "interaural correlation discrimination from a reference of unity" that the stimulus variability could be neglected.

Our motivation for making a physiological study of the relative effects of intrinsic and stimulus variability in detection was provided by our previous experiments on the neural discrimination of changes in the binaural cues of interaural time difference (ITD) and interaural correlation (IAC). We found that the best single neurons in the guinea pig inferior colliculus (IC) are capable of ITD discrimination sensitivity equivalent to humans (Skottun et al. 2001; Shackleton et al. 2003). However, in measurements of IAC sensitivity (Shackleton et al. 2005), we found that the neural thresholds did not match the behavioral thresholds, being at least three times worse. Thus, although IAC and ITD discrimination are closely linked theoretically (e.g., Colburn and Durlach 1978; Durlach et al. 1986; Koehnke et al. 1986; Shackleton et al. 2005; Louage et al. 2006), they appear to be give different results.

In our experiments, the ITD thresholds were obtained by using tones, whereas the IAC thresholds were obtained by using noise. Although these stimuli were equated in level, the neural response to noise was highly variable, because the firing patterns were often dependent on envelope shape. We used freshly computed noise for each trial, so it is possible that this variability introduced extra variance into the firing rate distributions and hence increased thresholds. This suspicion was recently partially confirmed by a psychophysical study (Goupell and Hartmann 2006), which showed that binaural noise stimuli with identical IACs of 0.99 could be discriminated from each other on the basis of their waveforms.

To determine whether this was a possible explanation for relatively poor IAC thresholds, and to determine the relative importance of intrinsic and external noise, we remeasured tone ITD and IAC discrimination thresholds, together with *noise* ITD thresholds. A limited number of tokens of frozen noise were used and spike rate distributions were obtained for each token. This allowed both within- and across-token comparisons to be made together with estimates of the stimulus induced and intrinsic neural variability in firing rate. Although the influence of stimulus variability was measurable, it was not a major factor in the elevated IAC thresholds.

METHODS AND STIMULI

Complete data sets were recorded from 25 units in the right IC of six pigmented guinea pigs weighing 469–600 g. Animals were anesthetized with urethane (1.3 g/kg i.p., in 20% solution in 0.9% saline) and Hypnorm (Janssen: 0.2 ml i.m., comprising fentanyl citrate 0.315 mg/ml and fluanisone 10 mg/ml). To reduce bronchial secretions, atropine sulfate (0.06 mg/kg, s.c.) was administered at the start of the experiment. Anesthesia was supplemented with further doses of Hypnorm (0.2 ml, i.m.), on indication by pedal withdrawal reflex. A tracheotomy was performed, and core temperature was maintained at 38°C via a heating blanket and rectal probe. Heart rate was monitored by using a pair of electrodes in the skin on either side of the animal's thorax. Animals were artificially ventilated with pure oxygen to keep the end-tidal partial pressure of CO_2 between 24 and 36 mmHg. The animals were placed inside a sound attenuating room in a stereotaxic frame in which hollow plastic speculae replaced the ear bars allowing sound presentation and direct visualization of the tympanic membrane. A craniotomy was performed over the position of the IC. The dura was reflected and the surface of the brain covered by a solution of 1.5% agar in 0.9% saline. All experiments were carried out in accordance with the United Kingdom Animal (Scientific Procedures) Act of 1986.

Recordings were made with glass-insulated tungsten electrodes (Bullock et al. 1988) advanced into the IC through the intact cortex, in a vertical penetration, by a piezoelectric motor (Burleigh Inchworm IW-700/710). Extracellular action potentials were amplified (Axoprobe 1A; Axon Instruments, Foster City, CA, USA), filtered between 300 Hz and 2 kHz, discriminated using a level-crossing detector (SD1; Tucker-Davies Technologies, Alachua, FL), and their time of occurrence recorded with a resolution of 1 μs.

Stimuli were delivered to each ear through sealed acoustic systems comprising custom-modified Radioshack 40-1377 tweeters joined via a conical section to a damped 2.5-mm-diameter, 34-mm-long tube (M. Ravicz, Eaton Peabody Laboratory, Boston, MA, USA), which fitted into the hollow speculum. The output was calibrated a few millimeters from the tympanic membrane using a Brüel and Kjær 4134 microphone fitted with a calibrated 1-mm probe tube.

All stimuli were digitally synthesized (System II; Tucker-Davies Technologies) at between 100- and

200-kHz sampling rates, and were output through a waveform reconstruction filter set at 1/4 the sampling rate (135 dB/octave elliptic: Kemo 1608/500/01 modules supported by custom electronics). If not otherwise stated, stimuli were of 50 ms duration, switched on and off simultaneously in the two ears with cosine-squared gates with 2 ms rise/fall times (10–90%).

Recording protocol

In this paper we are primarily concerned with the relative effects of stimulus and intrinsic variability. It is therefore very important for us to control stimulus variability. At the level of the IC, we would expect that units will be instantaneously responding to a narrow-band filtered, short-duration segment of the signal, so it is not appropriate to reduce sample variability by using long-duration, wideband stimuli. The importance of controlling for the sample-to-sample variability in IAC is discussed at length by Shackleton et al. (2005) along with a fuller description of this method.

The search stimulus was a 50-ms, binaural pure tone presented every 250 ms, of manually variable frequency and level and with a fixed ITD of 0.1 cycles, as this is the modal characteristic delay in the IC (McAlpine et al. 2001). When a unit was isolated the characteristic frequency (CF: frequency at which lowest discharge rate threshold was obtained) and threshold were obtained audiovisually. Frequency response areas, discharge rate versus level functions (RLFs) and peristimulus response histograms (PSTHs) were obtained by using diotic pure tones to enable units to be characterized and their location in central IC verified physiologically (see Shackleton et al. 2003 for details). RLFs were obtained by using binaurally uncorrelated, narrowband noise to determine the noise threshold. Uncorrelated noise was chosen to obtain the noise threshold because although some units are unresponsive to uncorrelated noise, for most units the response to uncorrelated noise is close to the average of the responses across the noise ITD function (Palmer et al. 1999); thus this represents a conservative estimate of noise threshold that does not require a knowledge of best delay before a representative noise threshold can be obtained.

A coarse noise ITD function (cf. Yin et al. 1986) was then obtained at a level of 20 dB above the uncorrelated noise threshold over a range of delays equivalent to ±1.5 cycles of BF in steps of 0.1 cycles (e.g., Figs. 2A and 3A). From the coarse noise ITD functions mean best phase (BP) and vector strength were calculated by using a modification of the method described by Goldberg and Brown (1969) in which the ITD function was treated like a period histogram and the strength of locking to the ITD measured. To compensate for the effective delay to

the binaural coincidence detector (in the medial superior olive; Goldberg and Brown 1969), the IAC functions were measured at an ITD corresponding to either the peak, or the trough, of the coarse noise delay curve. The decision to record at either the peak or trough was made on-line during the experiment on the basis of the coarse noise delay function. Units that appear to be "peak like" were measured at the peak, whereas units that appeared more "trough like" were measured at the trough. Peak (or trough) position was estimated visually as the position where the peak (or trough) would have occurred if the function had been continuous.

Noise ITD and discharge rate versus interaural correlation (rIAC) functions were obtained by using the same 10 tokens of noise presented to the left ear (Fig. 1), the results from each token were stored separately so that functions could be obtained for each token individually or with all tokens pooled together (e.g., Fig. 2B, G). Higher-resolution ITD functions (cf. Kuwada and Yin 1983) were obtained for both BF tones at 20 dB above tone threshold (e.g., Figs. 2F and 3F) and noise at 20 dB above uncorrelated noise threshold (e.g., Figs. 2B and 3B) over a range passing through zero ITD from trough to peak in steps of 0.01 cycles of BF (between 15 and 31 steps, depending on the peak to trough separation). All ITDs, were presented, for all 10 tokens, in random order, before the same ITD and token combinations were repeated in a different order. A block consisted of 10 repeats of noise ITD functions and 20 repeats of tone ITD functions. rIAC functions (e.g., Figs. 2G and 3G) were obtained from an interaural correlation of −1 to +1 in 0.1 steps (21 steps) with an ITD value set to either the peak or the trough of the coarse noise ITD function. All IAC and token combinations were presented, in random order, before the same IAC and token combinations were repeated in a different order. A block consisted of 10 repeats.

A block of 10 repeats of the noise ITD function was collected first, followed by a block of 10 repeats of a rIAC function and, finally, a block of 20 repeats of a tone ITD function. This sequence was then followed from two to five times (typically five) for each unit, resulting in between 20 and 50 repeats of the noise ITD and rIAC functions, and 40–100 repeats of the tone ITD function. Recording times of up to 120 min were typical, so this interleaving was essential to ensure that changes in neural responsiveness did not affect comparisons between functions. Results were monitored on-line and checked off-line to ensure that the neural responsiveness did not change during recording. The interleaving of functions meant that if neural responsiveness changed, recording could be stopped without loosing the ability to compare data. It was felt preferable to

A: Left waves for ITD & IAC

B: Right waves for ITD

C: Right waves for IAC
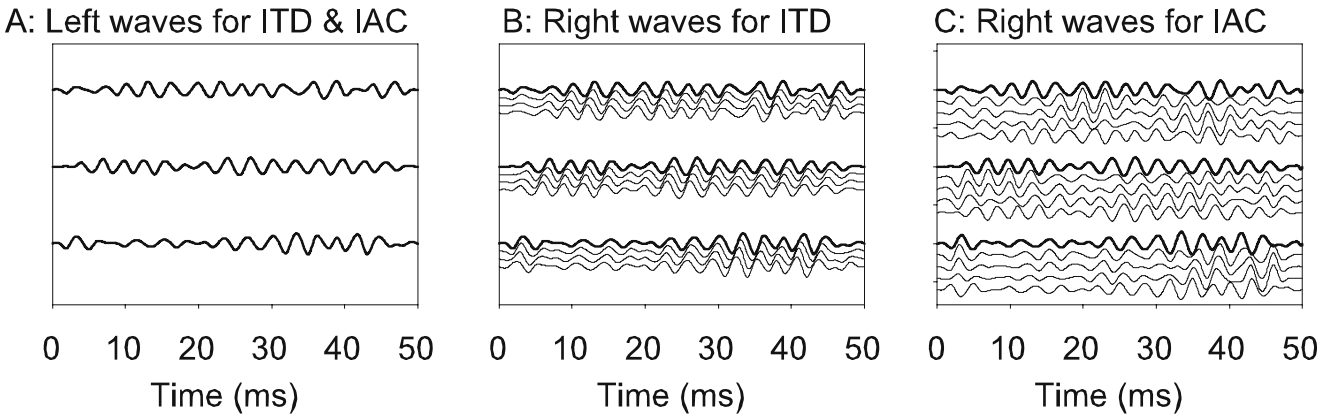
FIG. 1. (A) Waveforms presented to the left ear in both the ITD and IAC paradigms. Three of the 10 tokens are shown. These are the actual waveforms presented to the unit illustrated in Figure 2. (B) Waveform presented to the right ear in the noise ITD discrimination paradigm. Three of the ten different tokens are shown along with delayed tokens. The range of ITDs shown is −1.0 to 0 ms (−0.3 cycles of BF to 0), where the dark lines correspond to zero ITD. (C) Waveform presented to the right ear in the IAC discrimination paradigm. Three of the ten different tokens are shown along with decorrelated tokens. The correlations shown range from −1 to +1 in steps of 0.5, the dark lines correspond to a correlation of +1.

reduce the number of repeats rather than collect nonhomogeneous data. Spikes were included in the spike count if they occurred between 5 ms and 70 ms after the stimulus onset.

Noise generation

All the noise stimuli used in this paper were generated by using the fast Fourier transform and were narrowband, with infinitely steep filter cutoffs, centered on CF, with a bandwidth equal to the guinea pig ERB:

$$ERB(f_c) = 6.477 f_c^{0.56}$$

where ERB and f_c are measured in Hz (Evans et al. 1992; Evans 2001). For the noise RLFs and coarse noise ITD functions, independent noise tokens were generated freshly each trial ("rolling" noise). For the rIAC functions and fine noise ITD functions "frozen" noise was used as described below.

For each *unit*, 10 new tokens of "frozen" noise were generated; a different set of tokens was used for each unit. The same set of tokens were presented to the left (contralateral) ear while collecting *both* noise ITD functions and rIAC functions (e.g., Fig. 1A). The signal to the right ear was simply an advanced or delayed version of the left ear signal while collecting noise ITD functions (Fig. 1B). For rIAC functions, the signal presented to the right ear was a combination of that presented to the left ear, and an independent noise in the proportion $r : \sqrt{(1 - r^2)}$ (Jeffress and Robinson 1962), where r is the expected IAC (Fig. 1C).

In generating noise tokens with a given *expected* IAC using the usual methods there is a problem insofar as the *actual* IACs of the tokens will be distributed around the *expected* IAC with a standard deviation of $(1 - r^2)/\sqrt{2\mathrm{WT}}$, where W is bandwidth and T is signal duration (see Shackleton et al. 2005). This problem arises because there is usually a small degree of correlation between the original two noise samples that are added together to produce the tokens used in the experiment. We used the Gram–Schmidt procedure (Culling et al. 2001) to remove this correlated component from the independent

FIG. 2. Full analysis for a peak unit with CF of 309 Hz, threshold of 46 dB SPL and best phase of 0.06 cycles of CF (best delay = 194 μs). (A) Coarse noise delay function recorded with "rolling" noise. Peak position estimated at 0 μs. (B) Gray lines with dots: high-resolution noise ITD ▶ functions for individual tokens. Thick black line with open symbols and error bars: noise ITD function averaged over all tokens. Error bar shows standard deviation of token means. (C) High-resolution noise ITD function pooled across all tokens (black and white circles and error bars). The error bars show the standard deviation of the spike count distribution. The vertical bar graphs show the distributions of spike counts corresponding to the adjacent black points. (D) Gray lines with dots: variance of distributions of individual token spike counts with ITD. Thin black line: average of individual token variances. Thick black line: variance of pooled data. Dashed thick black line: variance of token means. (E) Proportion of total pooled variance that is attributable to difference between token means for ITD discrimination. Horizontal line is summary mean across all values. (F) As in (C), but for single tones. (G–J) Responses as a function of IAC in the same format as (B–E). Functions recorded with an ITD of 0 μs. (K) Neurometric function showing percent correct in simulated two interval forced-choice ITD discrimination away from a reference of zero. Gray lines with dots: individual noise tokens. Thick black line with dots: pooled noise tokens. Thin black line with small open circles: tone. Large circle emphasizes reference point. Downward triangles emphasize points closest to 25% threshold, solid is for noise, open for tone. (L–N) IAC discrimination with references of −1, 0, and +1 in the same format as (K). Upright triangle emphasizes point closest to 75% discrimination threshold, whereas inverted triangle emphasizes point closest to 25% discrimination threshold.
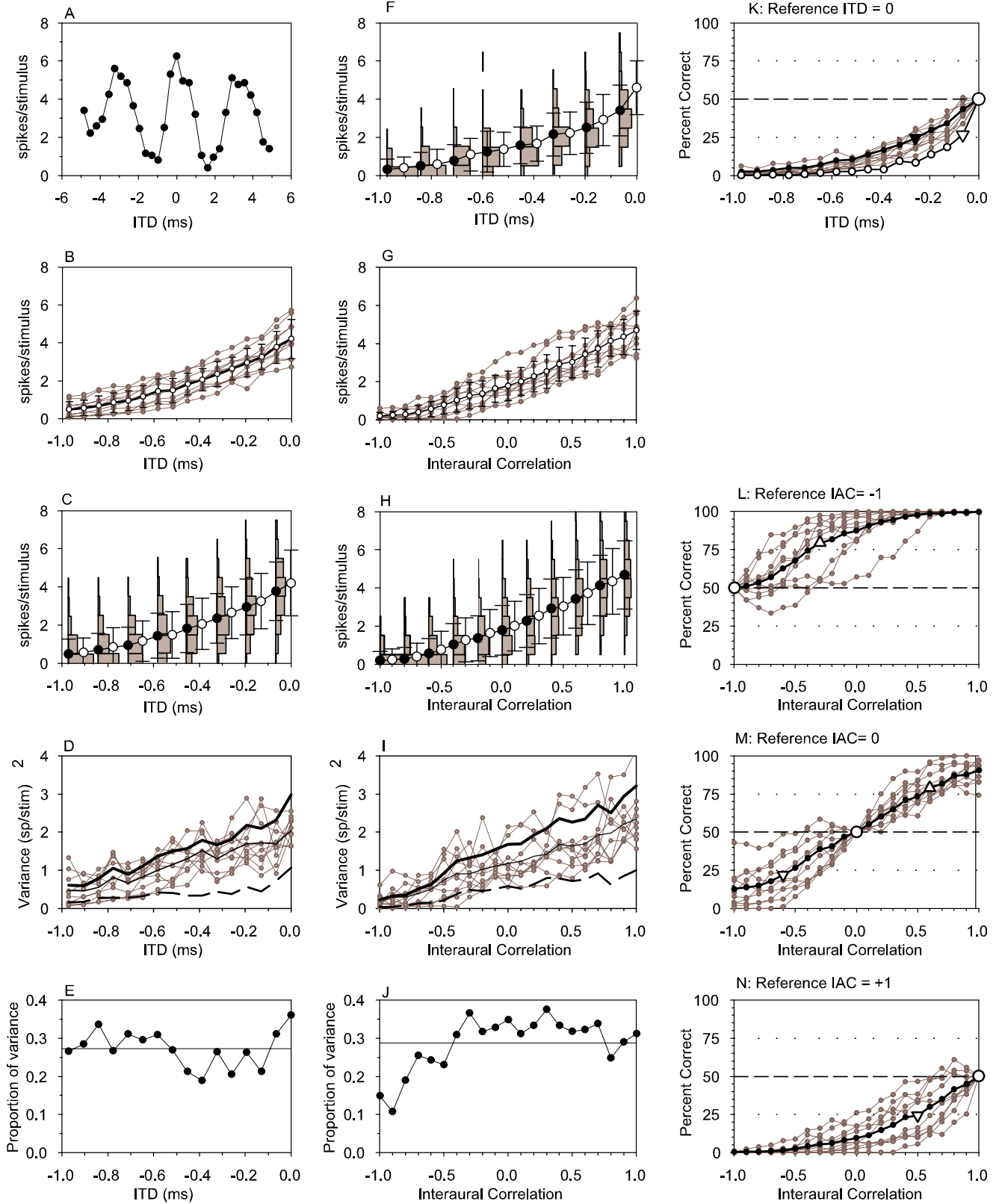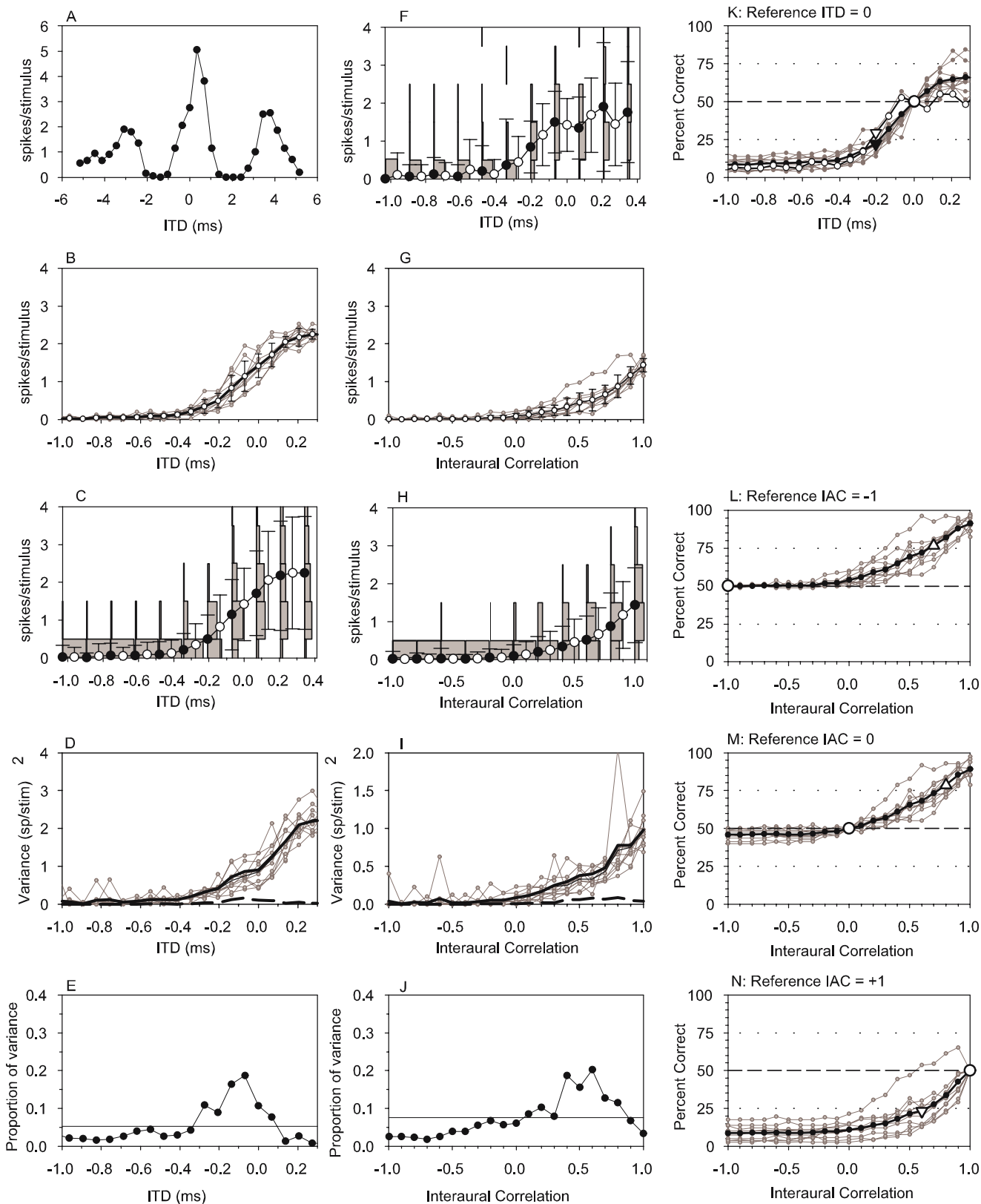
FIG. 3. Full analysis for a peak unit with CF of 291 Hz, threshold of 49 dB SPL and best phase of 0.09 cycles of CF (best delay = 309 μs). Details as in Figure 2; ITD for IAC functions was 344 μs.

signal, so that adding the left ear and independent samples in the proportion r : $\sqrt{(1 - r^2)}$ produced *exactly* the expected IAC. A consequence of using the Gram–Schmidt procedure is that in addition to the tokens having exactly the required correlation, they also are absolutely equal in spectral level.

RESULTS

Full data were obtained from 25 single, well-discriminated, units in the right ICs of six animals. IAC functions were obtained at the ITD corresponding to the peak of the coarse noise-ITD function for 21 units and at the trough ITD for four units. The thresholds obtained from peak or trough units were not different, so the data are plotted together (peak: upright triangle, trough: inverted triangle) in Figures 4, 5, 6, 7, 8, 9 and 11.

Figure 2 shows data from a unit with a large amount of variability in response to different noise tokens. It had a 309-Hz CF, and a threshold of 46 dB SPL. The coarse noise ITD function (Fig. 2A) was cyclic, at the period of the CF, as expected (Yin et al. 1986). The central peak of the function was at zero, so this was the ITD used for collection of rIAC functions. The central peak of the function was slightly skewed, so the calculated best delay was 194 μs. Another unit, with less variability in response to different tokens, is shown in Figure 3. It has a

similar CF and threshold (291 Hz, 49 dB SPL). The peak of the noise ITD function was 344 μs, so this was the ITD used in obtaining rIAC function curves. Calculated best delay was 309 μs, somewhat outside the physiological range predicted from the head-width.

High-resolution noise ITD functions between the trough and peak of the functions are shown for 10 tokens of noise as gray lines with dots in Figures 2B and 3B. The means of these functions as well as their standard *deviation* are shown as the thick black line with error bars. The functions for individual tokens are distributed over quite a wide range of firing rates in Figure 2B, whereas in Figure 3B the distribution is smaller. In Figure 2B, in particular, the individual functions are approximately parallel, indicating that the units have the same ITD sensitivity to different noise tokens, but differ in overall mean response. This difference in mean firing rate must be attributable to some factor other than spectral level, because the Gram–Schmidt procedure ensures all tokens had the same level. The same data are represented differently in Figures 2C and 3C. Here, the data are pooled together, ignoring the different tokens. The line with black and white circles shows the pooled noise ITD function. This is identical to the mean of the individual token functions (Figs. 2B and 3B), as it should be. The error bars show the standard *deviation* of the pooled firing rates. Close examination will reveal that these (Figs. 2C and 3C) are larger than the
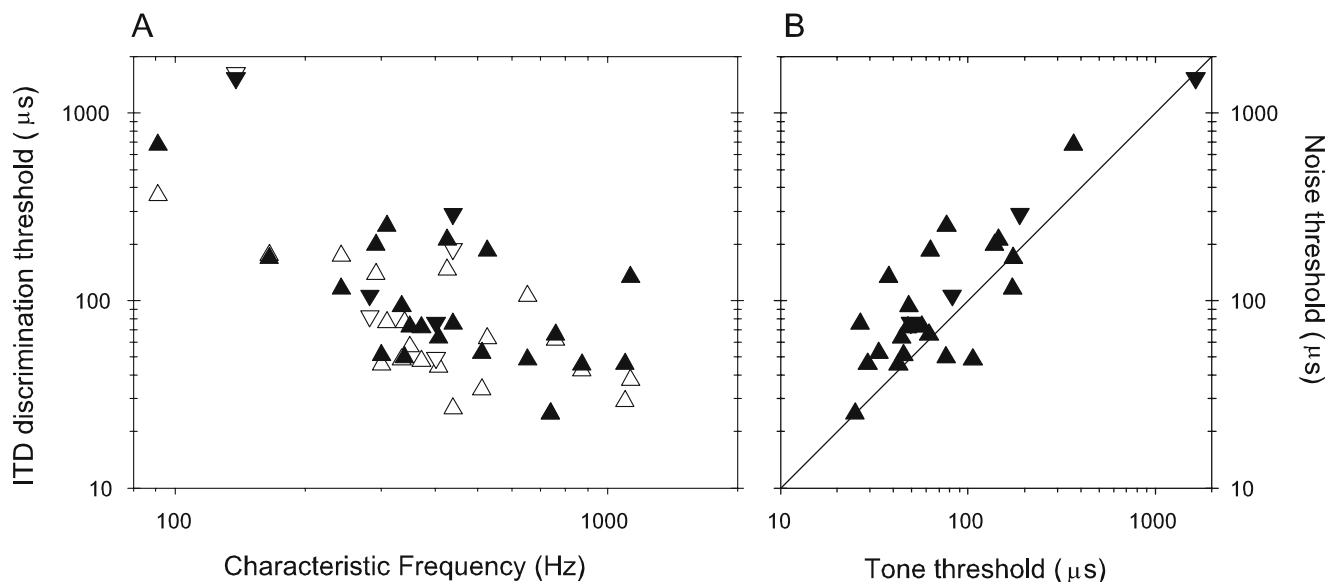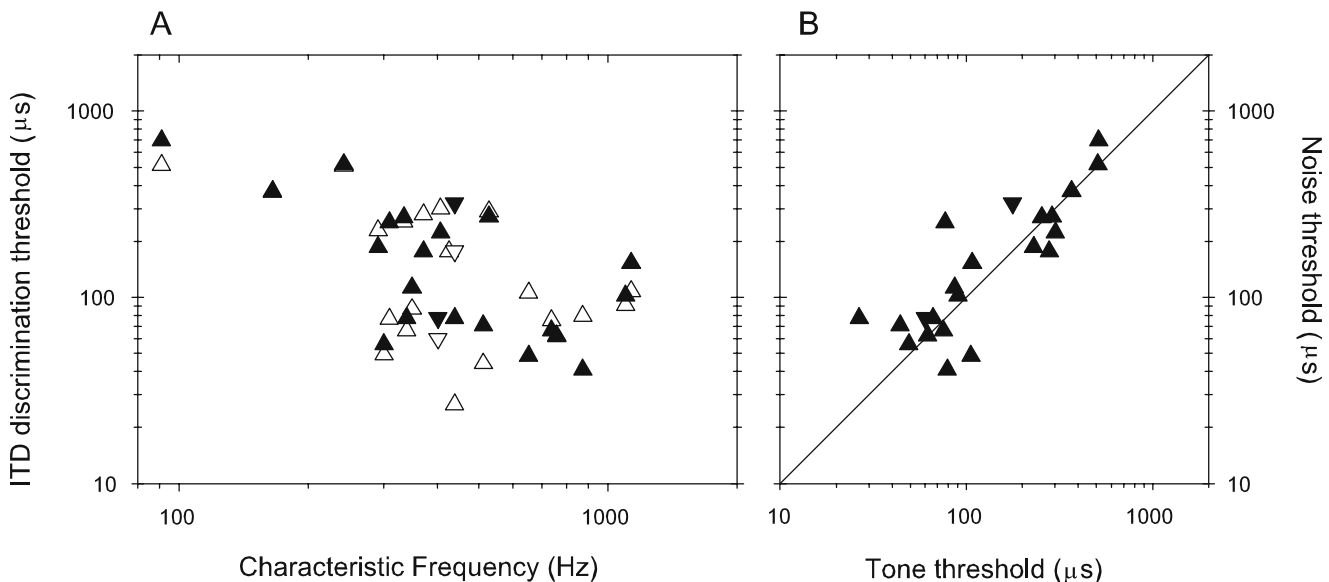

FIG. 4. ITD discrimination thresholds from reference points chosen to minimize threshold (best reference). The average best reference for tone and noise was used where these differed. (A) Tone (open symbols) and noise (filled) thresholds as a function of characteristic frequency. Upright triangles: peak units; inverted triangles: trough units. (B) Direct comparison of tone and noise best thresholds for the same units. Upright triangles: peak units; inverted triangles: trough units.
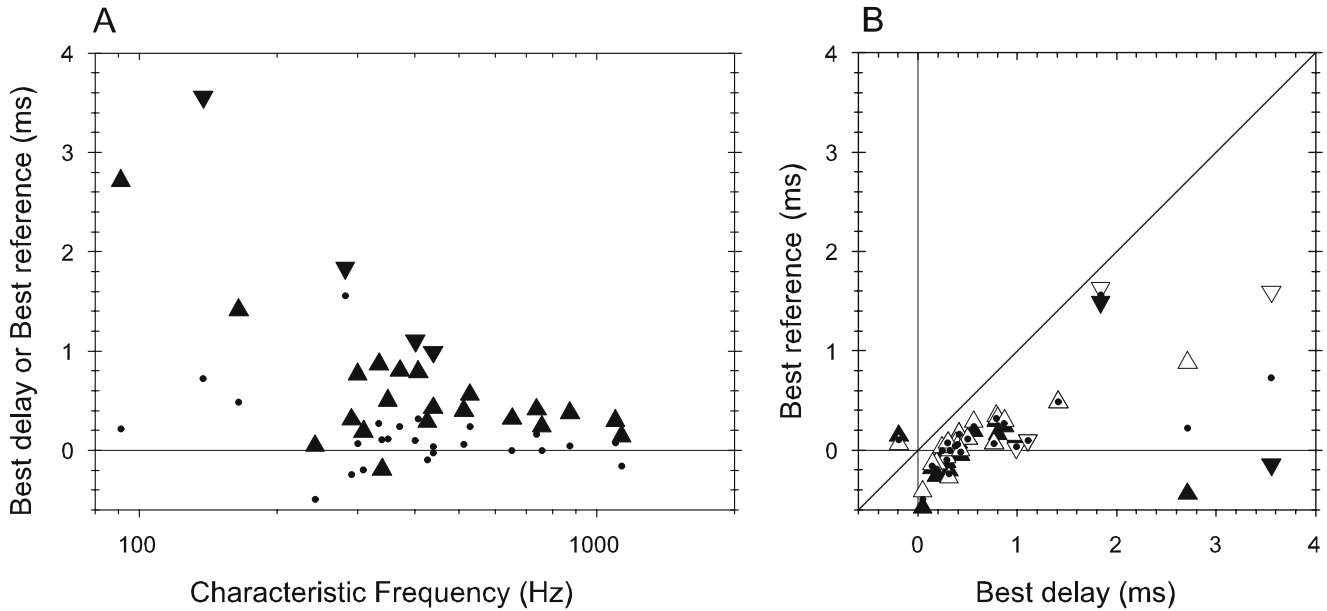
FIG. 5. Comparison of best delay measured from coarse noise ITD functions and the best-reference position for noise ITD and tone ITD discrimination. (A) Filled triangles show best delay measured from coarse noise-ITD functions. Dots show the average of best-reference ITD for noise and tone ITD discrimination. (B) Comparison of best-reference position relative to best delay. Diagonal line represents equality. Open symbols show best-reference for tone ITD discrimination; filled symbols show best-reference for noise ITD discrimination. Dots show average of noise and tone best-reference ITDs.

standard deviations of the individual token functions (Figs. 2B and 3B), illustrating the additional effect of the intrinsic neural variability. The relative magnitudes of the between-token and intrinsic variabilities can be seen more clearly in Figures 2D and 3D, where the variances of the individual tokens are shown as gray lines. The mean of these is shown as the thin black line (barely visible underneath the thick black line in Figure 3D). This is the best estimate of the

intrinsic variability of the unit. The variance between the individual token ITD functions, a measure of the between-token variability, is shown as the thick dashed curve (see Appendix for further details). In Figure 2D, this is clearly nonzero, reflecting the spread of functions in Figure 2B. However, the between-token variability is close to zero in Figure 3D, only being noticeable between about −0.2 and 0.1 ms, again reflecting the spread of functions in Figure 3B. If the

FIG. 6. As in Figure 4, except that ITD discrimination thresholds away from a reference of zero ITD are plotted.
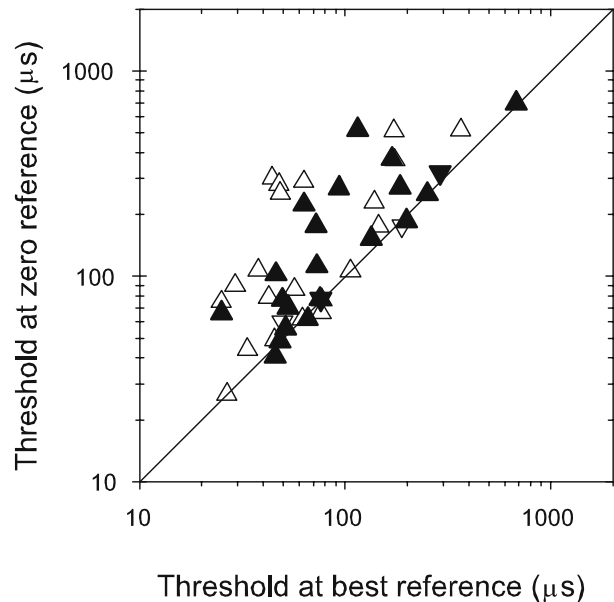
FIG. 7. Comparison of thresholds at zero and best reference points for tones (open symbols) and noise (filled symbols). Upright triangles: peak units; inverted triangles: trough units.

distinction between different tokens is ignored, and the total variance measured (like in Figs. 2C and 3C), then the thick solid line in Figures 2D and 3D is obtained. This is bigger than the intrinsic variability, because it reflects both the intrinsic and the between-tokens variabilities. In fact, it is the sum of these two variances (cf. analysis of variance, e.g., Howell 1987 and Appendix). To produce a single measure of the relative influence of between-token variability, we calculated the ratio of the variance of the mean individual-token rates to the total, pooled, variance (Fig. 2E). It can be seen that for ITD discrimination, in this particular unit, the average proportion of variance attributable to different stimuli is about 27%. Although this is quite high, it is obvious that the larger proportion of variance (73%) is attributable to the intrinsic variability of the unit itself. The unit illustrated in Figure 3 has a lower susceptibility to the differences between tokens, as demonstrated by the lower overall mean proportion of variance (Fig. 3E).

The same analyses for IAC are also shown in Figures 2 and 3. There is very similar between-tokens and total variability for IAC compared with ITD (cf. Figs. 2 and 3B–D with Figs. 2 and 3G–I), and the proportion of variance attributable to the stimulus is only slightly higher (Figs. 2 and 3J).

Discrimination thresholds for single units

Receiver operating characteristic (ROC) analysis (Cohn et al. 1975; Bradley et al. 1987; Green and Swets 1988; Shackleton et al. 2003) was used to determine the smallest change in IAC and ITD,

which the neuron could correctly indicate by a change in its firing rate. The details of, and justification for, this procedure are discussed at length by Shackleton et al. (2003); however, we will give a short summary here based on Figures 2 and 3.

ROC analysis allows threshold estimates to be made without assuming any particular distribution of spike counts. This is especially important for very low spike rates, where the distribution is significantly different from the Gaussian distribution assumed in d' analysis (Green and Swets 1988). The analysis simulates a two-interval, forced-choice, psychophysical task. It is assumed that the firing rates in two intervals are compared and the target is chosen as that occurring in the interval with the higher firing rate.

The variability in firing rate for a constant stimulus makes it difficult to reliably determine whether a stimulus has changed based on the firing rate. In Figures 2 and 3C are shown the distribution of firing rates pooled across tokens for noise ITDs corresponding to the filled circles. Similarly, Figures 2F and 3F show the distributions for tones. For example, by looking at Figure 2F, it can be seen that a rate of two spikes/stimulus can be elicited by any ITD, but it is far more likely to occur for an ITD of −0.3 ms than for an ITD of either −1 ms or zero. This means that if it is desired to discriminate whether a sound has moved from, say, straight ahead (zero ITD) to one side, then firing rate will only give the correct answer on a proportion of trials, and this proportion will increase as the separation increases. This qualitative argument is quantified in the neurometric function (Shackleton et al. 2003), the proportion of trials in which a correct discrimination from a reference can be made as a function of ITD. In Figures 2K and 3K are shown the pooled neurometric functions for noise and tone ITDs with a reference of zero. Results for individual noise tokens are shown as gray lines, for all tokens pooled together as the black line with solid dots, and for tones as black lines with open circles. In Figure 2K, all the curves start at 50% correct at zero ITD and decline to 0% at −1 ms. This means that a decision based on an *increase* in firing rate is *consistently* wrong. However, the decision can equally be made on a *decrease* in firing rate. It can easily be shown (Shackleton et al. 2003) that decisions made on the basis of an increase in firing rate are complementary to decisions made on the basis of a decrease in firing rate. Therefore the 25% threshold can be taken as a measure of discrimination performance. For the pooled data, this is shown as a solid inverted triangle for noise and an open inverted triangle for tones. In Figure 2K, the individual token detection thresholds range from −300 to −100 μs with the pooled noise threshold toward the top of this range, at 251 μs. This pooled threshold is toward the

upper end of the individual thresholds because of the significant effect of between token variability on the pooled threshold. In contrast, the thresholds for the other unit shown in Figure 3K are more similar. Note also, that for the unit in Figure 2K the tone threshold is lower that the noise thresholds. This was not generally the case, and the similarity of thresholds illustrated for the other unit in Figure 3K is more typical.

Pooled IAC discrimination thresholds could similarly be calculated for references of −1, 0 and +1, these are comparable to those reported earlier (Shackleton et al. 2005) and, for the example in Figure 2 vary from 0.5 to 0.7. There is a greater variability in the individual token thresholds (Fig. 2L–N) for IAC compared with noise ITD (Fig. 2K). The unit illustrated in Figure 2 has a large variability between noise tokens. Near the other extreme is the unit shown in Figure 3, where the variability between tokens was much less than the intrinsic variability. Pooled IAC discrimination thresholds ranged between

0.4 and 1.7 (Fig. 3L–N), with the threshold for a reference at −1 being greatly elevated as a result of the very flat rIAC function curve between −1 and 0 (Fig. 3G).

Comparison of noise and tone ITD discrimination thresholds

Because a narrowband noise stimulus can be viewed as simply a modulated tone, we might expect noise discrimination to be worse than the tone discrimination by a factor equivalent to the variability introduced by the noise. When there is a significant variability in the responses to different noise tokens, we would expect that noise ITD thresholds pooled across tokens would be much larger than tone ITD thresholds. Figure 4A plots the tone (filled triangles) and noise (open triangles) ITD discrimination thresholds away from the best-reference ITD as a function of frequency. As reported earlier (Shackleton et al. 2003), the lowest thresholds follow approximately a $1/f$ pattern and are comparable to human psychophysical thresh-
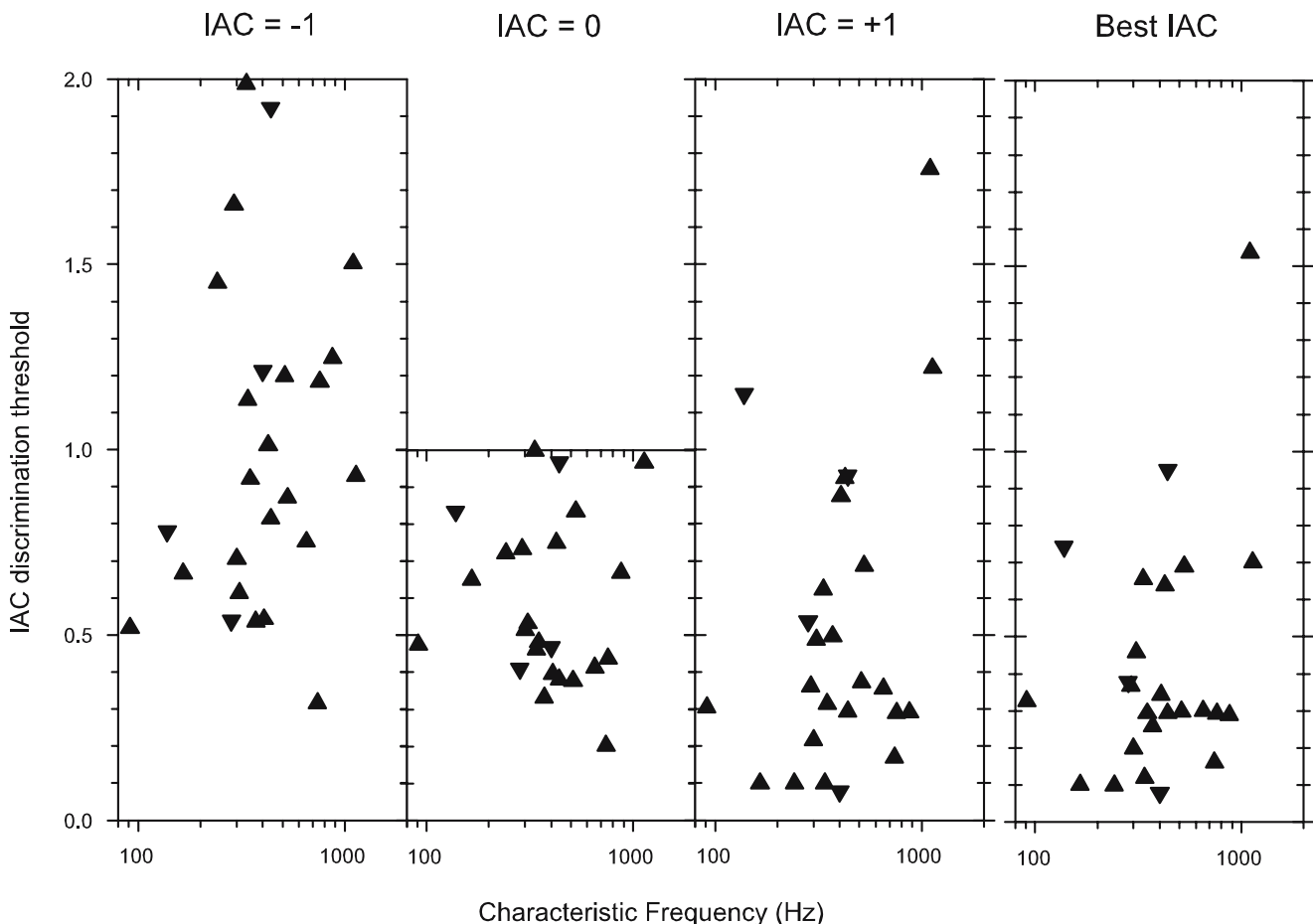
FIG. 8. IAC discrimination thresholds away from references of +1, 0, and −1; and away from the reference point chosen to give the lowest threshold ("best"). Upright triangles: peak units; inverted triangles: trough units.

olds for a comparable duration (Hafter et al. 1979). Although, on average, the noise thresholds are higher, there is a clear overlap between the noise and tone thresholds. This is emphasized when the thresholds from a single unit are plotted against each other (Fig. 4B). On average, noise thresholds are greater than tone thresholds. This is consistent with the variability in response due to different noise tokens having some effect on discrimination thresholds, but not much.

The use of the average best-reference for noise and tones requires some comment. The best-reference ITD was found to be only slightly different for tones and noise (compare filled and open triangles in Fig. 5B). Because the differences between noise and tone reference ITDs appear to be random, rather than systematic, then each can be viewed as a noisy estimate of the best-reference ITD, and a better estimate obtained by averaging them. The thresholds reported therefore correspond to the average best-reference ITD. As described previously (Shackleton et al. 2003), the position of the best reference is closer to zero ITD than the best delay (Fig. 5).

The thresholds just described represent the best discrimination performance of which the units are capable. However, because the reference point is different for different units (Fig. 5), it does not allow a comparison of how the units would behave as a population in a typical psychophysical experiment. Because psychophysical thresholds are usually determined with a reference of zero ITD, these are plotted in Figure 6. Although thresholds are generally higher, the overall trends are the same as those for the best-reference threshold, which is not surprising, given how close most of the best-references are to zero (Fig. 5). In comparing Figures 4B and 6B, there appears to be less spread, and noise thresholds are more nearly equal to tone thresholds, when a common, zero, reference is used.

Thresholds at zero- and best-reference ITDs are compared directly in Figure 7. For both tones and noise many thresholds are roughly equal, with the majority of thresholds with zero reference less than three times worse than the best. This is somewhat less than the 30-fold difference between the lowest (25 μs) and highest (680 μs) tone or noise thresholds, along each axis in Figure 7.

Finally, some comment about the stimulus levels used may be appropriate. Tone and noise stimuli were presented at 20 dB above threshold, so they ought to be equalized for equivalent energy passing through the effective filter. However, for binaurally sensitive units the firing rate, and presumably threshold, varies with the interaural phase. As discussed earlier, we measured the noise threshold by using uncorrelated noise, which we presume to equate to the threshold averaged across all ITDs. However,

tone threshold was measured at zero ITD. In this sense, then, the relative tone and noise levels at either best or zero ITD are not equated. However, the measured tone thresholds at zero ITD will be *higher* than the thresholds at best ITD, so the stimuli will always be presented *at least* 20 dB above threshold. Psychophysically, thresholds do not change greatly as a function of level once the stimulus is 10 dB above threshold (e.g., Durlach and Colburn 1978). Also, although the spike rates change, the shape of ITD tuning curves do not change greatly with level (Kuwada and Yin 1983). Therefore, although some care must be taken, when making direct comparisons between tone and noise ITD thresholds, because of the unquantified effects of level, we do not believe these to be serious.

IAC discrimination thresholds

Pooled IAC discrimination thresholds are shown in Figure 8 for reference IACs of +1, 0, and −1, and also with the reference IAC chosen to give the lowest possible threshold (best IAC). In agreement with Shackleton et al. (2005), thresholds range over the full possible extent (2 for references of ±1, and 1 for 0), but the lowest thresholds are obtained with a reference IAC of +1, followed by a reference of 0 with the thresholds for a reference of −1 being the worst. This hierarchy can be understood from the shape of the
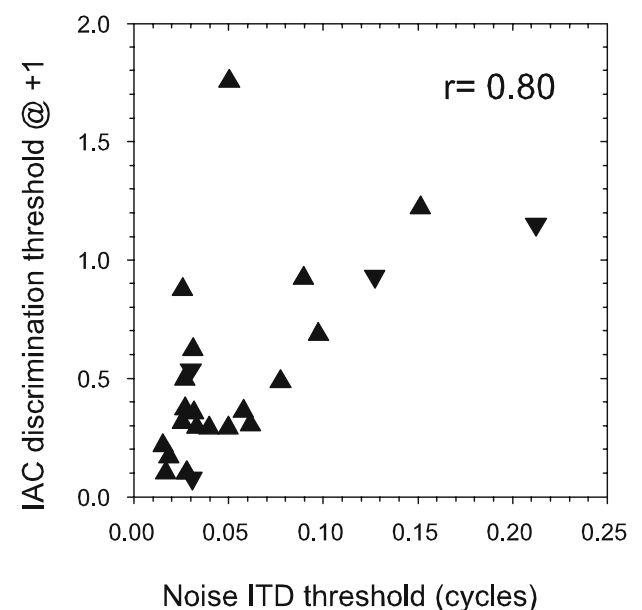

FIG. 9. Comparison of IAC discrimination threshold away from a reference of +1 and best noise ITD discrimination thresholds plotted in phase units. *r* value is Pearson correlation coefficient excluding outlying point to top left. Upright triangles: peak units; inverted triangles: trough units.

rIAC function, which is flat near IACs of −1 and becomes progressively steeper as the IAC approaches +1 (Figs. 2G and 3G). Although the thresholds in the best IAC panel ought to be the lowest, in practice they are no lower than those with a reference of +1 since, unsurprisingly, an IAC near +1 was usually also the best reference point. The very best pooled IAC thresholds are around 0.1, compared with human psychophysical thresholds for 50-ms duration, narrowband noise centered on 500 Hz, which are 0.035 (Bernstein and Trahiotis 1997). Human psychophysical thresholds for a reference of −1 are about twice those for a reference of +1 (Boehnke et al. 2002), whereas the neural thresholds are much worse.

Comparison of noise ITD and IAC thresholds

As pointed out in the Introduction and by Shackleton et al. (2005), there is a tight theoretical link between ITD processing and IAC. Most contemporary models of ITD processing actually measure the cross-correlation of the input signals (e.g., Stern and Trahiotis 1995; Colburn 1996). It can easily be shown that as an ITD is applied to a sinusoidal signal, then the IAC is proportional to the cosine of the interaural *phase* difference (IPD). Because the narrow-band passed signals we used are quasi-sinusoidal, we would also expect the IAC to be proportional to the cosine of the IPD. We would therefore expect that IAC thresholds ought to be closely linked to the ITD thresholds ex-

pressed as a function of *phase* (Fig. 9). If the clear outlier to the top left is excluded, there is a correlation of 0.8 between these variables. This demonstrates that, to a reasonable degree, IAC thresholds are predictable from ITD thresholds, and *vice versa*. The prediction is not perfect, because in addition to experimental error the nature of the IAC is qualitatively different in the two cases. For an ITD shift, the IAC is static, whereas for an IAC shift the IAC dynamically varies around the expected value.

Comparison of pooled and individual token thresholds

As shown in Figures 2 and 3K–N, it is possible to obtain thresholds from the individual noise tokens. These thresholds will be somewhat less accurate than those obtained from the pooled data because they are only based on 20–50 repeats; however, their spread provides information about the variability between tokens. If there were a significant variability attributable to differing tokens, then each of the thresholds from individual tokens ought to be lower than the threshold for the pooled data. For example, in Figure 2K the mean of the individual token thresholds is slightly lower than that of the pooled results (most of the individual neurometric functions are to the right of the pooled neurometric). Box plots of the distribution of individual noise token thresholds are shown as a function of the pooled
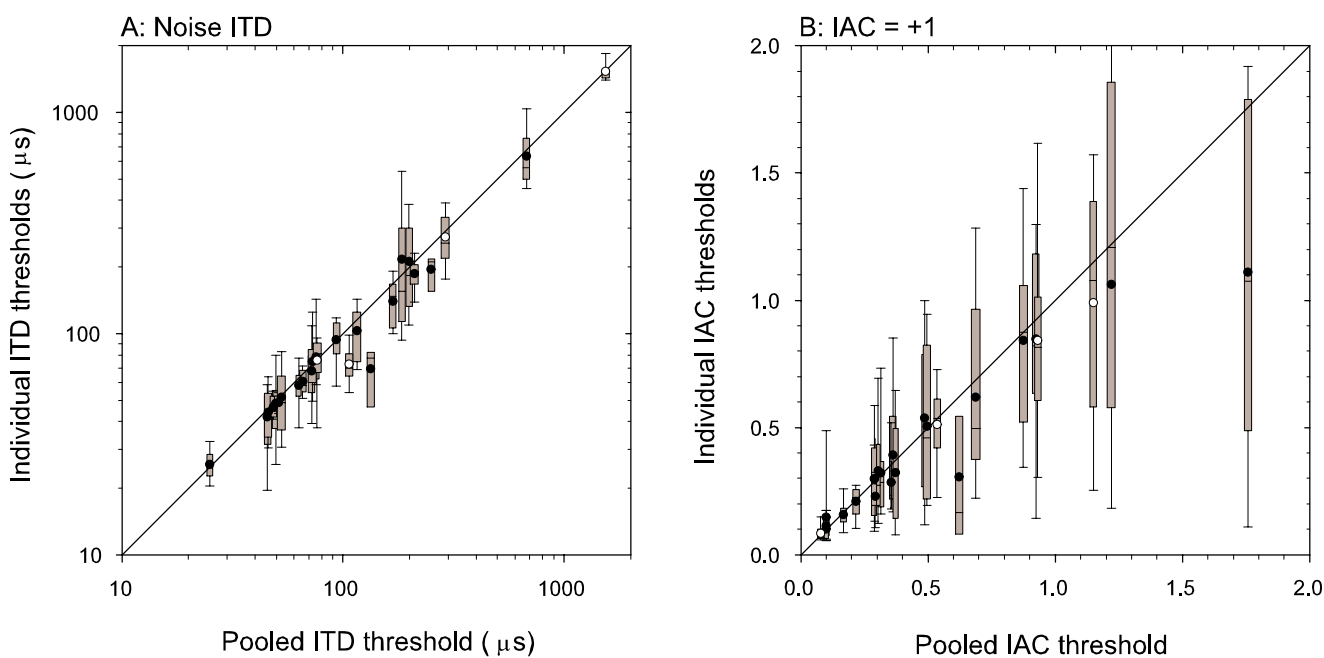
FIG. 10. Comparison of pooled discrimination thresholds with thresholds derived from individual tokens. Box plots are used to represent the distribution of token thresholds. The gray box shows the interquartile range (25% to 75% percentiles), the whiskers show the 10% and 90% percentiles. The circles show the mean of the individual token thresholds. Open: peak units; filled: trough units. (A) Noise ITD discrimination thresholds away from best reference. (B) IAC discrimination thresholds away from a reference of +1.
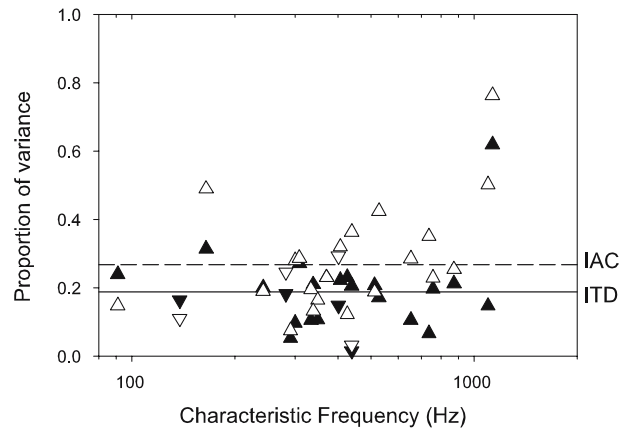
FIG. 11. Proportion of total variance that can be attributed to variance between different tokens for both ITD and IAC discrimination. Upright triangles: peak units; inverted triangles: trough. Filled: ITD; unfilled: IAC. Solid line: mean for ITD. Dashed line: mean for IAC.

thresholds in Figure 10. In each box, the central point shows the mean threshold, the bar the range of thresholds between the 25% and 75% percentiles and the whiskers show the 10% and 90% percentiles. The diagonal line is the line of equality between pooled and mean thresholds. In Figure 10A, the distribution of individual ITD thresholds are tightly grouped around the mean, which are consistently slightly below the line of equality. However, this deviation from equality is not great, indicating that the variability due to different noise tokens does not greatly affect the ITD threshold. The IAC thresholds from a reference of +1 are shown in panel B. These show a great deal of variability around the mean, but the means are consistently below the line of equality. This suggests that the variability between the responses to different tokens is having an effect in increasing the pooled threshold, and that this effect is greater than for the ITD results.

Proportion of variance due to different stimulus tokens

The analysis shown in Figures 2 and 3 allows a direct measure of the proportion of variance in spike counts attributable to the different stimulus tokens. Panels D and I show the proportion of variance as a function of ITD or IAC. Also shown is the mean proportion of variance across all ITDs or IACs. These means are plotted for all units in Figure 11. Apart from one unit, the proportion of variance accountable to stimulus tokens is between 1.5% and 30% for ITDs (solid symbols) and between 3% and 50% for IAC (open symbols). The mean proportion is about 19% for ITD and 27% for IAC. These values show that the variability attributable to different stimulus

tokens is measurable, but it is less than the intrinsic neural variability. Also, the proportion is slightly greater for IAC discrimination. These findings are consistent with the effects on individual versus pooled thresholds reported in Figure 10, although the conclusion from the discrimination thresholds is that the increased variance only has a minor effect on discrimination.

Becasue the variability of envelope peak height and number of peaks of a narrowband noise is related to its bandwidth, we might expect there to be a frequency dependence of the proportion of variance. However, there is no noticeable trend in Figure 11, possibly because the intrinsic neural variance is greater than the between tokens variability.

DISCUSSION

How individual neural responses are combined to determine a behavioral ITD discrimination threshold has recently been a subject for debate. Fitzpatrick et al. (1997) have argued that if discrimination is based on the movement of the position of the peak in an array of coincidence detectors like those of the classical "Jeffress delay line" (Jeffress 1948), then responses of around 30–40 neurons would need to be pooled together. However, we (Skottun 1998; Skottun et al. 2001; Shackleton et al. 2003) have argued that if discrimination occurs in the region of a neuron's response which is most sensitive to change, then the ITD discrimination of single guinea pig neurons and human psychophysics closely match. This argument is supported by the finding that the distribution of neural characteristic delays appears to be such that the ITD function slopes are at their steepest through zero ITD (McAlpine et al. 1996; McAlpine et al. 2001; Brand et al. 2002; McAlpine and Grothe 2003; Palmer 2004), corresponding to the straight-ahead position that has been shown psychophysically to be the position of greatest acuity (e.g., Domnitz 1973). The ITD at which each neuron was maximally sensitive is plotted in Figure 5; for the most part, these best-reference positions are very close to zero ITD. It is useful shorthand, although not entirely accurate, to describe the best discrimination as happening at the position of maximum slope. In fact, the best discrimination generally occurs closer to the trough of the function than the ITD value at the position of greatest slope, since the variance is normally proportional to the firing rate and the threshold is determined by the ratio of the slope to the variance. This comparison is shown in Figure 4 of Shackleton et al. (2003).

Previously (Skottun 1998; Skottun et al. 2001; Shackleton et al. 2003), we argued that it was the

most sensitive guinea pig neurons that matched human sensitivity; a version of the lower envelope hypothesis (Parker and Newsome 1998). However, in the barn owl it is suggested that the behavioral discrimination threshold is the average of the neural thresholds (Bala and Takahashi 2003; Bala et al. 2003); whereas in the rabbit it has been found that the behavioral threshold is much higher than most neural thresholds (Patel et al. 2005). Furthermore, measurements and modeling by Hancock and Delgutte (2004) have shown that certain features of human psychophysical ITD thresholds can only be explained by pooling units across frequency. Clearly, more work is required to determine which of these results represent differing ecological requirements and which different neural mechanisms.

Although it can be argued that the binaural system has evolved to facilitate discrimination at the midline by the most sensitive units, this is at the expense of the ability to discriminate correlation in single units. We measured the correlation sensitivity at the peak (or trough) of the noise ITD function which maximizes the amplitude and slope (modulation) of the IAC function. However, as just discussed, the peaks of the ITD functions are generally outside the physiological range of the guinea pig and it is the maximum slope that is close to zero. We previously showed (Shackleton et al. 2005) that the IAC functions measured at zero ITD are somewhat less modulated, and discrimination thresholds (at least at a reference of +1) are much higher. Even when measured under optimum conditions the neural IAC thresholds are somewhat higher than those obtained from humans, so measuring thresholds in ecologically valid conditions (i.e., zero ITD) will make the mismatch even worse.

Although ITD discrimination is normally described by subjects in psychophysical experiments as being guided by a change in position, in IAC discrimination subjects are generally not quite so sure, suggesting changes in image width or "quality" as the cues they use (e.g., Gabriel and Colburn 1981). Alternatively, IAC can be viewed as time-varying ITD and interaural level difference (ILD) cues. From this viewpoint, it is clear that changing IAC changes the distribution of instantaneous image positions. Because the variation is rapid, this is perceived as a widening of the image, rather than a rapidly moving image. As noted earlier, the change in neural firing rate caused by a change in ITD is time-invariant, and would vary coherently across an array of coincidence detectors like those of the "Jeffress delay line" (Jeffress 1948). This is clearly different to the case of IAC discrimination, where the internal correlation as seen by a single unit varies over time, and the variation across different units is incoherent. Because a change in ITD results in a coherent change across many neurons, it is possible

to detect this change by using a single neuron, and the only effect of using more than a single neuron would be to reduce intrinsic noise. However, the time-varying change in a single unit firing caused by an IAC would be harder to distinguish from random fluctuation in the stimulus, so we might not expect single neuron thresholds to match behavioral performance. This may be achieved by comparing the changes in firing rate across many neurons, or by viewing the change in the pattern of activity of a group of neurons. Additionally, IAC is by its very nature a statistical and variable cue. Unless the stimulus is highly controlled, as in our experiment, then the token IAC will vary greatly between trials. As discussed previously (Shackleton et al. 2005), IAC also varies considerably throughout the duration of a long stimulus and, if it is wideband, across frequency. Therefore, comparison across many neurons and across long integration times would be advantageous. This probably explains why although we find neural ITD thresholds that are comparable with human performance, neural IAC performance is much worse.

In previous papers (Shackleton et al. 2003, 2005), we reported results from any binaurally sensitive unit that we could isolate. However, in this study we used a more stringent criterion of only recording from units with excellent signal-to-noise ratio that we felt would be stable for up to 2 h recording. This may have biased recording to only units of a certain type. The only hint from the physiology that this may have happened was a smaller proportion of trough units in the sample than previously. However, the lower limit and spread of discrimination thresholds, and the response areas and PSTHs recorded suggest that there was a similar heterogeneity of responses in the current sample as in the previous experiments. Additionally, the stereotaxic coordinates from which we recorded were no different.

The technique of interleaving the three different discrimination measurements in short blocks proved very useful. Although there were shifts in the responsiveness of units over long periods, these applied to all measurements thus allowing for uncomplicated comparison of different measurements. Where a unit was lost, or signal-to-noise became too poor, before the complete set of measurements was complete, valid comparisons could still be made, even if the measures were a little noisier.

Throughout this article, we have only considered the possibility that IAC and ITD might be coded by the firing rate of neurons. This is in line with most of the modeling of psychophysical binaural effects. However, we should stress that valuable information might be contained within the firing patterns of neurons. The data we have may allow this possibility to be addressed, but it is beyond the scope of the current paper.

We were able to measure the variability in the neural responses to different noise tokens. The variance of the data pooled across all noise tokens was decomposed into a component due to intrinsic neural variance and a component due to the difference between tokens. For ITD discrimination, the between tokens variability was found to be between 1.5% and 30% of the total variance, with a mean of 19%. The proportion of variability due to tokens was higher for IAC discrimination, being between 3% and 50% with a mean of 27%. Although these proportions of variability are measurable, they are much smaller than the intrinsic neural variability. In some units, the variability is completely dominated by the intrinsic neural variability.

It could be argued that one reason for the small influence of between tokens variability is because in using the Gram–Schmidt procedure, we have removed a major component of between token variability, i.e., the random fluctuations of IAC for individual tokens around their expected value. As pointed out in several papers (e.g., Pollack and Trittipoe 1959a, b; Gabriel and Colburn 1981; Shackleton et al. 2005), the variance in the IAC between tokens is a function of the expected IAC, being minimal at IACs of ±1 and maximal near zero. One of our initial motivations for controlling for IAC variability (Shackleton et al. 2005) was that the shape of the IAC discrimination function as a function of IAC (Pollack and Trittipoe 1959a) so closely paralleled the token variance that removing it might reveal the underlying neural sensitivity, which may have been a great deal more constant across IAC. We cannot rule out the possibility that stimulus variability would be much larger if we had not controlled for random IAC and level variations. However, what is of interest is that there still is a measurable component of stimulus variability even when these factors are controlled for. This is consistent with the recent finding that humans are able to discriminate between tokens having identical IAC (Goupell and Hartmann 2006).

We might argue that variability in the IAC is not important, because in an earlier experiment (Shackleton et al. 2005) we found similar IAC discrimination results for both narrowband noises (which had well-controlled IAC) and wideband noises (which did not). This could argue that either IAC variability is negligible compared with other forms of waveform variability, or that we have not succeeded in controlling for IAC variability. In the first case, IAC variability is negligible compared to the other components of stimulus variability. In the second case, our measurements already include the influence of IAC variability.

An increase in total variance due to variability between noise tokens would be expected to increase discrimination thresholds. We found that pooled noise thresholds were greater than pure tone thresholds; however, many units had very similar thresholds and few were elevated by more than three times. The waveforms of individual noise tokens are quasi-sinusoidal (see Fig. 1), being well modeled as slowly modulated sine waves. We might expect discrimination thresholds for this type of frozen waveform to either be higher or lower than for pure sinusoids depending on the characteristics of the unit. For units that are more sensitive to the envelope than the fine structure, we might expect the threshold to be lower for the frozen noise token, because its envelope is richer than that of a pure tone. However, a unit which is more sensitive to the fine structure than the envelope would be expected to have a lower threshold for the pure tone, because the noise token has many cycles with lower peak amplitudes. Unfortunately, the number of repeats of each noise token was too small to allow this question to be addressed directly.

Discrimination thresholds were measurable for individual noise tokens, but were very variable presumably, in part, because of the low number of repeats in addition to the differences between the token waveforms. The large spread in both ITD and IAC individual token discrimination thresholds is therefore not very meaningful; however, the mean of the individual token thresholds is based on the same number of measurements as the pooled thresholds, and so should be just as accurate. The individual token thresholds are only affected by the intrinsic neural variability, whereas the pooled thresholds are affected by both the intrinsic neural variability and the variability due to different stimuli. We found for ITD discrimination and IAC discrimination from a reference of +1 that the mean of the individual token thresholds was generally slightly less than the pooled thresholds, indicating that the between token variability was indeed increasing the threshold for stimuli with some token to token variability. However, this increase was not great, presumably because, as discussed above, the intrinsic neural variability was always greater (sometimes much greater) than the between tokens variability. To conclude, although between tokens response variability is measurable, it does not appear to greatly increase discrimination thresholds.

ACKNOWLEDGMENTS

APPENDIX: PARTITIONING OF RESPONSE VARIANCE INTO INTRINSIC AND STIMULUS EVOKED VARIANCES.

The derivation shown below is an abstract of the theory behind the analysis of variance (ANOVA). We make no claims of originality in including this here; the same concepts with slightly different notation are discussed elsewhere (e.g., Colburn 1977b; Rieke et al. 1999). It is included to demonstrate that although ANOVA makes several assumptions, which ideally need to be verified before the statistical test can be used, the partitioning used in this paper is assumption-free.

We measured the firing rates in 25 units, for several IACs or ITDs, for 10 noise tokens with up to 50 repeats each. Let the number of repeats be n_r and an individual repeat be indexed by the subscript r; similarly the number of tokens will be n_t, and the index t. Units will be indexed by u, and ITDs or IACs by i. The firing rate of the u'th unit, in response to the r'th repeat of the t'th token of the i'th ITD or IAC will be x_{rtiu}.

The variance of the firing rate for a given unit and ITD or IAC across all tokens and repeats (i.e., the total variance) is:

$$s_{iu}^2 = \frac{1}{n_t n_r} \sum_t^{n_t} \sum_r^{n_r} \left(x_{rtiu} - \overline{\overline{x_{iu}}} \right)^2$$

where

$$\overline{\overline{x_{iu}}} = \frac{1}{n_t n_r} \sum_t^{n_t} \sum_r^{n_r} x_{rtiu},$$

i.e., the mean firing rate across all tokens and repeats, so

$$s_{iu}^2 = \frac{1}{n_t n_r} \sum_t^{n_t} \sum_r^{n_r} x_{rtiu}^2 - \frac{1}{n_t^2 n_r^2} \left(\sum_t^{n_t} \sum_r^{n_r} x_{rtiu} \right)^2$$

However, we can also compute the variance for each token individually: where

$$s_{tiu}^2 = \frac{1}{n_r} \sum_r^{n_r} \left(x_{rtiu} - \overline{x_{tiu}} \right)^2$$

where

$$\overline{x_{tiu}} = \frac{1}{n_r} \sum_r^{n_r} x_{rtiu},$$

i.e., the mean firing rate for each token across all repeats, so

$$s_{tiu}^2 = \frac{1}{n_r} \sum_r^{n_r} x_{rtiu}^2 - \frac{1}{n_r^2} \left(\sum_r^{n_r} x_{rtiu} \right)^2$$

so the average within-token variance (intrinsic variance) is given by:

$$\overline{s_{iu}^2} = \frac{1}{n_t} \sum_t^{n_t} s_{tiu}^2$$

$$= \frac{1}{n_t} \sum_t^{n_t} \left(\frac{1}{n_r} \sum_r^{n_r} x_{rtiu}^2 - \frac{1}{n_r^2} \left(\sum_r^{n_r} x_{rtiu} \right)^2 \right)$$

$$= \frac{1}{n_t n_r} \sum_t^{n_t} \sum_r^{n_r} x_{rtiu}^2 - \frac{1}{n_t n_r^2} \sum_t^{n_t} \left(\sum_r^{n_r} x_{rtiu} \right)^2$$

The variance between different token means (stimulus variance):

$$s_{tokens,iu}^2 = \frac{1}{n_t} \sum_t^{n_t} \left(\overline{x_{tiu}} - \overline{\overline{x_{iu}}} \right)^2$$

$$= \frac{1}{n_t} \sum_t^{n_t} \left(\overline{x_{tiu}} \right)^2 - \frac{1}{n_t^2} \left(\sum_t^{n_t} \left(\overline{x_{tiu}} \right) \right)^2$$

$$= \frac{1}{n_t} \sum_t^{n_t} \left(\frac{1}{n_r} \sum_r^{n_r} x_{rtiu} \right)^2 - \frac{1}{n_t^2} \left(\sum_t^{n_t} \left(\frac{1}{n_r} \sum_r^{n_r} x_{rtiu} \right) \right)^2$$

$$= \frac{1}{n_t n_r^2} \sum_t^{n_t} \left(\sum_r^{n_r} x_{rtiu} \right)^2 - \frac{1}{n_t^2 n_r^2} \left(\sum_t^{n_t} \sum_r^{n_r} x_{rtiu} \right)^2$$

If we sum these latter variances, we can see that they are mathematically the same as the original, total variance:

$$\overline{s_{iu}^2} + s_{tokens,iu}^2 = \frac{1}{n_t n_r} \sum_t^{n_t} \sum_r^{n_r} x_{rtiu}^2 - \frac{1}{n_t n_r^2} \sum_t^{n_t} \left(\sum_r^{n_r} x_{rtiu} \right)^2$$

$$+ \frac{1}{n_t n_r^2} \sum_t^{n_t} \left(\sum_r^{n_r} x_{rtiu} \right)^2 - \frac{1}{n_t^2 n_r^2} \left(\sum_t^{n_t} \sum_r^{n_r} x_{rtiu} \right)^2$$

$$= \frac{1}{n_t n_r} \sum_t^{n_t} \sum_r^{n_r} x_{rtiu}^2 - \frac{1}{n_t^2 n_r^2} \left(\sum_t^{n_t} \sum_r^{n_r} x_{rtiu} \right)^2,$$

since the middle two terms cancel

$$= s_{iu}^2$$

So, we have shown that the variance pooled across tokens (s_{iu}^2) can be partitioned into two factors, which can be identified as the average within-token variance (intrinsic variance: $\overline{s_{iu}^2}$), and the variance between different token means (stimulus variance: $s_{tokens,iu}^2$).

To obtain the summary measure of proportion variance accounted for by the stimulus, we calculate the ratio of the stimulus variance to the total variance averaged across all ITDs or IACs:

$$\overline{\eta^2} = \frac{1}{n_i} \sum_i^{n_i} \eta^2 = \frac{1}{n_i} \sum_i^{n_i} S_{tokens,iu}^2 \Big/ S_{iu}^2$$

REFERENCES

BALA A, TAKAHASHI T. Behavioral discrimination of auditory space is predicted by average neuronal discrimination. Assoc. Res. Otolaryngol. Abs. 26:177, 2003.

BALA ADS, SPITZER MW, TAKAHASHI TT. Prediction of auditory spatial acuity from neural images on the owl's auditory space map. Nature 424:771–774, 2003.

BERNSTEIN LR, TRAHIOTIS C. The effects of randomizing values of interaural disparities on binaural detection and on discrimination of interaural correlation. J. Acoust. Soc. Am. 102:1113–1120, 1997.

BOEHNKE SE, HALL SE, MARQUARDT T. Detection of static and dynamic changes in interaural correlation. J. Acoust. Soc. Am. 112:1617–1626, 2002.

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. 57:755–772, 1987.

BRAND A, BEHREND O, MARQUARDT T, MCALPINE D, GROTHE B. Precise inhibition is essential for microsecond interaural time difference coding. Nature 417:543–547, 2002.

BULLOCK DC, PALMER AR, REES A. Compact and easy-to-use tungsten-in-glass microelectrode manufacturing workstation. Med. Biol. Eng. Comput. 26:669–672, 1988.

COHN TE, GREEN DG, TANNER WP. Receiver operating characteristic analysis. Application to the study of quantum fluctuation effects in optic nerve of *Rana pipiens*. J. Gen. Physiol. 66:583–616, 1975.

COLBURN HS. Theory of binaural interaction based on auditory-nerve data. II. Detection of tones in noise. J. Acoust. Soc. Am. 61:525–533, 1977a.

COLBURN HS. (1977B) Theory of binaural interaction based on auditory-nerve data. II. Detection of tones in noise. Supplementary material. In: AIP document PAPS jasma-61-525-98. American Institute of Physics, Physics Auxiliary Publication Service, New York.

COLBURN HS. (1996) Computational models of binaural processing. In: Hawkins HL, McMullen TA, Popper AN, and Fay RR (eds) Auditory Computation. Springer-Verlag, New York, pp. 332–400.

COLBURN HS, DURLACH NI. (1978) Models of binaural interaction. In: Carterette EC and Friedman MP (eds) Handbook of Perception, Vol IV, Hearing. Academic Press, New York, pp. 467–518.

CULLING JF, COLBURN HS, Spurchise M. Interaural correlation sensitivity. J. Acoust. Soc. Am. 110:1020–1029, 2001.

DAVIDSON SA, GILKEY RH, Colburn HS, Carney LH. Binaural detection with narrowband and wideband reproducible noise maskers. III. Monaural and diotic detection and model results. J. Acoust. Soc. Am. 119:2258–2275, 2006.

DOMNITZ R. The interaural time jnd as a simultaneous function of interaural time and interaural amplitude. J. Acoust. Soc. Am. 53:1549–1552, 1973.

DURLACH NI, COLBURN HS. (1978) Binaural phenomena. In: Carterette EC and Friedman MP (eds) Handbook of Perception, Vol IV. Hearing. Academic Press, New York, pp. 365–466.

DURLACH NI, GABRIEL KJ, COLBURN HS, TRAHIOTIS C. Interaural correlation discrimination. II. Relation to binaural unmasking. J. Acoust. Soc. Am. 79:1548–1557, 1986.

EVANS EF. (2001) Latest comparisons between physiological and behavioural frequency selectivity. In: Houtsma AJM, Kohlraush A, Prijs VF, Schoonhoven R (eds) Physiological and Psychophysical Bases of Auditory Function. Shaker Publishing, Maastricht, pp. 382–387.

EVANS EF, PRATT SR, SPENNER H, COOPER NP. (1992) Comparisons of physiological and behavioural properties: Auditory frequency selectivity. In: Cazals Y, Demany L, Horner K (eds) Auditory Physiology and Perception. Pergamon, Oxford, pp. 159–170.

EVILSIZER ME, GILKEY RH, MASON CR, COLBURN HS, CARNEY LH. Binaural detection with narrowband and wideband reproduc-

ible noise maskers: I. Results for human. J. Acoust. Soc. Am. 111:336–345, 2002.

FITZPATRICK DC, BATRA R, STANFORD TR, KUWADA S. A Neuronal population code for sound localization. Nature 388:871–874, 1997.

GABRIEL KJ, COLBURN HS. Interaural correlation discrimination. I. Bandwidth and level dependence. J. Acoust. Soc. Am. 69:1394–1401, 1981.

GILKEY RH, ROBINSON DE, HANNA TE. Effects of masker waveform and signal-to-masker phase relation on diotic and dichotic masking by reproducible noise. J. Acoust. Soc. Am. 78:1207–1219, 1985.

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. 32:613–636, 1969.

GOUPELL MJ, HARTMANN WM. Interaural fluctuations and the detection of interaural incoherence: bandwidth effects. J. Acoust. Soc. Am. 119:3971–3986, 2006.

GREEN DM. Consistency of auditory detection judgments. Psychol. Rev. 71:392–407, 1964.

GREEN DM, SWETS JA. (1988) Signal Detection Theory and Psychophysics. Los Altos, CA, Peninsula Publishing.

HAFTER ER, DYE RH, GILKEY RH. Lateralization of tonal signals which have neither onsets nor offsets. J. Acoust. Soc. Am. 65:471–477, 1979.

HANCOCK KE, DELGUTTE B. A physiologically based model of interaural time difference discrimination. J. Neurosci. 24:7110–7117, 2004.

HOWELL DC. (1987) Statistical Methods for Psychology, 2nd edn. PWS-Kent Publishing, Boston.

ISABELLE SK, COLBURN HS. Detection of tones in reproducible narrow-band noise. J. Acoust. Soc. Am. 89:352–359, 1991.

JEFFRESS LA. A place theory of sound localization. J. Comp. Psychol. 44:35–39, 1948.

JEFFRESS LA, ROBINSON DE. Formulas for the coefficient of correlation for noise. J. Acoust. Soc. Am. 34:1658–1659, 1962.

KOEHNKE J, COLBURN HS, DURLACH NI. Performance in several binaural-interaction experiments. J. Acoust. Soc. Am. 79:1558–1562, 1986.

KUWADA S, YIN TCT. 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. 50:981–999, 1983.

LOUAGE DHG, JORIS PX, VAN DER HEIJDEN M. Decorrelation sensitivity of auditory nerve and anteroventral cochlear nucleus fibers to broadband and narrowband noise. J. Neurosci. 26:96–108, 2006.

MCALPINE D, GROTHE B. Sound localization and delay lines—do mammals fit the model? Trends Neurosci. 26:347–350, 2003.

MCALPINE D, JIANG D, PALMER AR. Interaural delay sensitivity and the classification of low best-frequency binaural responses in the inferior colliculus of the guinea pig. Hear. Res. 97:136–152, 1996.

MCALPINE D, JIANG D, PALMER AR. A neural code for low-frequency sound localisation in mammals. Nat Neurosci. 4:396–401, 2001.

PALMER AR. Reassessing mechanisms of low-frequency sound localisation. Curr. Opin. Neurobiol. 14:457–460, 2004.

PALMER AR, JIANG D, MCALPINE D. Desynchronizing responses to correlated noise: a mechanism for binaural masking levels differences at the inferior colliculus. J. Neurophysiol. 81:722–734, 1999.

PARKER AJ, NEWSOME WT. Sense and the single neuron: probing the physiology of perception. Annu. Rev. Neurosci. 21:227–277, 1998.

PATEL M, COFFEY C, MARSHALL A, SKAGGS J, FITZPATRICK D. Neural correlates of behavioural thresholds to interaural time differences. Assoc. Res. Otolaryngol. Abs. 28:39, 2005.

Pollack I, Trittipoe WJ. Binaural listening and interaural noise cross correlation. J. Acoust. Soc. Am. 31:1250–1252, 1959a.

Pollack I, Trittipoe WJ. Interaural noise correlations: Examination of variables. J. Acoust. Soc. Am. 31:1616–1618, 1959b.

Rieke F, Warland D, de Ruyter van Steveninck R, Bialek W. (1999) Spikes. Cambridge, MA, MIT Press.

Shackleton TM, Arnott RH, Palmer AR. Sensitivity to changes in interaural correlation in the inferior colliculus of the guinea pig. JARO—Journal of the Association for Research in Otolaryngology 6:244–259, 2005.

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. 23:716–724, 2003.

Siegel RA, Colburn HS. Internal and external noise in binaural detection. Hear. Res. 11:117–123, 1983.

Siegel RA, Colburn HS. Binaural processing of noisy stimuli: Internal/external noise ratios for diotic and dichotic stimuli. J. Acoust. Soc. Am. 86:2122–2128, 1989.

Skottun BC. Sound localization and neurons. Nature 393:531, 1998.

Skottun BC, Shackleton TM, Arnott RH, Palmer AR. The ability of inferior colliculus neurons to signal differences in interaural delay. Proc. Natl. Acad. Sci. USA 98:14050–14054, 2001.

Spiegel MF, Green DM. Two procedures for estimating internal noise. J. Acoust. Soc. Am. 70:69–73, 1981.

Stern RM, Trahiotis C. (1995) Models of binaural interaction. In: Moore BCJ (ed) Hearing. Academic Press, New York, pp. 347–386.

Swets JA, Shipley EF, McKey MJ, Green DM. Multiple observations of signals in noise. J. Acoust. Soc. Am. 31:514–521, 1959.

Yin TCT, Chan JCK, Irvine DRF. Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. I. Responses to wideband noise. J. Neurophysiol. 55:280–300, 1986.

Zheng L, Early SJ, Mason CR, Idrobo F, Harrison JM, Carney LH. Binaural detection with narrowband and wideband reproducible noise maskers. II. Results for rabbit. J. Acoust. Soc. Am. 111:346–356, 2002.