The masking of interaural delays

Andrew J. Kolarik and John F. Culling[a)]
School of Psychology, Cardiff University, Tower Building, Park Place, Cardiff CF10 3AT, United Kingdom

(Received 5 December 2006; revised 26 January 2009; accepted 28 January 2009)

Four experiments measured discrimination of interaural time delay (ITD) thresholds for broadband noise in the presence of masking noise of the same bandwidth as the target (0.1–3 kHz for experiments 1–3 and 0–10 kHz for experiment 4). In experiments 1–3, listeners performed interaural two–interval two–alternative forced–choice (2I-2AFC) delay discrimination tasks with stimuli composed of delayed and masking noises mixed in proportions of delayed noise ranging between 1 and 0.05. Experiments 1–3 employed interaurally correlated, anticorrelated, and uncorrelated maskers, respectively. Experiment 4 measured centering accuracy for continuous noise with a range of interaural coherences (equivalent to proportion of delayed noise) obtained by mixing delayed and interaurally uncorrelated noises. Results indicate that in the presence of an interaurally correlated masker ITD thresholds doubled for every halving of the proportion of delayed noise power in the stimulus. This function became steeper as the masking noise changed from interaurally correlated, to uncorrelated, to anticorrelated. The results were compared to thresholds predicted by a model based on variations in the distribution of interaural phase differences of the stimulus components.

[DOI: 10.1121/1.3083238]

PACS number(s): 43.66.Ba, 43.66.Dc, 43.66.Mk, 43.66.Pn [RLF] Pages: 2162–2171

I. INTRODUCTION

Listeners can use three primary acoustic cues for sound localization: interaural time differences (ITDs), interaural level differences (ILDs), and spectral cues introduced by the head, torso, and pinnae. At low frequencies, listeners tend to rely on the ITD cue for sound lateralization (Wightman and Kistler, 1992). At higher frequencies, listeners tend to rely on ILDs to lateralize sounds (Kuhn, 1977). Elevation and disambiguation between the frontal and rear hemifields are provided by monaural spectral changes in the sound caused by reflections among the corrugations of the pinna (Lopez-Poveda and Meddis, 1996), head movements that change the binaural cues (Perret and Noble, 1997), and movements of the sound source that the listener controls (Wightman and Kistler, 1999). In spite of these cues, the presence of masking noise has a debilitating effect on listeners' abilities to localize sounds. In order to quantify these effects, a range of studies has investigated localization (Jeffress *et al.*, 1962; Good and Gilkey, 1996) and lateralization (Egan and Benson, 1966; Cohen, 1981; Ito *et al.*, 1982) in the context of masking noise. The effect of masking noise in the presence of delayed noise is also important when modeling binaural temporal windows using a lateralization paradigm (Wagner, 1991; Bernstein *et al.*, 2001).

Previous research has demonstrated the adverse effects of masking noise on ITD discrimination, and that different types of masking noise vary in their impact on listeners' abilities to lateralize narrow-band signals. Specifically, lateralization of a signal embedded in masking noise is least disrupted by the presence of interaurally correlated noise and most disrupted by anticorrelated noise, with disruption from uncorrelated noise falling between the two.

Egan and Benson (1966) investigated listener's abilities to lateralize a monaural tone with a duration of 250 ms at frequencies of 500 and 390 Hz embedded in binaural noise (either interaurally correlated or uncorrelated). Lateralization thresholds were obtained by varying the signal-to-noise ratio (SNR) adaptively using a 2I-2AFC task. In the presence of interaurally correlated noise, lateralization thresholds were found to be 4–5 dB lower than when an uncorrelated masker was present.

Cohen (1981) measured ITD just-noticeable-differences (jnds) around zero ITD for a 250 Hz sinusoid presented against broadband masking noise which was interaurally correlated, uncorrelated, or anticorrelated. The ITD jnd was measured as a function of SNR. For a constant SNR, jnds were highest for the anticorrelated masking condition and lowest when the masker was interaurally correlated. jnds were intermediate when the masker was uncorrelated. The slope of the thresholds versus SNR was steeper when the masker was uncorrelated than when the masker was interaurally correlated, and the two slopes converged at approximately 19 dB above N0S0 detection threshold (denoted MoSo in Fig. 2 of Cohen, 1981). Thus, at high SNRs, ITD jnds were at their lowest and equal for both interaurally correlated and uncorrelated masking noises. At lower SNRs, jnds were larger when the masker was interaurally uncorrelated. When the masker was interaurally anticorrelated, ITD jnds were larger than jnds for interaurally correlated and uncorrelated masking noises, and increased as SNR was reduced.

Finally, Ito *et al.* (1982) measured interaural time jnds for narrow-band noise (1/3 octave wide) in the presence of a broadband masker. jnds were measured at a fixed SNR for a range of interaural configurations of the masker: interaurally correlated, anticorrelated, and uncorrelated, plus quiet (no masker), and six conditions that manipulated the perceived

---

[a)] Author to whom correspondence should be addressed.

lateral position of the masker. Similar to Cohen's (1981) study using tonal stimuli, jnds were smallest for the interaurally correlated masking condition, larger for the uncorrelated masker, and largest for the anticorrelated masking condition. No consistent relation was observed between the perceived lateral position of the masker and the magnitude of the jnd. Ito *et al.* (1982) were able to successfully model the hierarchy of ITD jnds for each interaural configuration of the masker using a model based on variations in the distribution of interaural phase differences of the stimulus components.

This effect of masking noise on ITD processing has been used in studies that have investigated the temporal resolution of the binaural system. Temporal resolution can be modeled using a temporal window that is described by one or more time constants, the duration of which describes the resolution of the system. In the studies of Wagner (1991) and Bernstein *et al.* (2001), temporal windows were measured for barn owls and human participants, respectively, by presenting a burst of interaurally delayed noise (the probe) temporally contiguous with masking noise. It was assumed that the listener detects the ITD imposed on the probe by centering a temporal window at the midpoint of the probe in order for the maximum amount of interaurally delayed noise to fall within the window. The window integrates together the instantaneous interaural delays of the probe with the interaural delays of the masking noise that also falls within the window. In order to model a temporal window, it is assumed that the integration results in an internal, effective ITD that is lower than the external ITD imposed on the probe (Bernstein *et al.*, 2001). Given that the probe and surrounding noise have equal power, the integration should yield the same internal ITD value as direct mixing at the same energy ratio. As a result, the external ITD must be increased to a magnitude that will bring the internal ITD up to threshold. When the probe duration is long, ITD thresholds will be low as very little masking noise enters the window, reducing the internal effective ITD. As the probe duration decreases, more masking noise enters the window and is integrated with the probe ITD, increasing ITD threshold.

It is important to quantify the effect of the masking noise on ITD threshold, as this is explicitly modeled during the window fitting procedure. Best-fit ITD thresholds are calculated by using a coefficient that relates threshold to proportion of delayed noise present. Bernstein *et al.* (2001) effectively assumed that when a broadband interaurally delayed noise target (N_τ) is heard in the presence of masking interaurally correlated or uncorrelated noise (N_0 or N_u), the effective internal ITD is doubled for each halving of the proportion of delayed noise present within the window; that is, that the masking noise dilutes the interaurally delayed noise. Having fit windows using the dilution assumption in the case of interaurally correlated masking noise, Bernstein *et al.* (2001) assessed how well the same windows fitted data in the case when the masking noise was interaurally uncorrelated. They were modeled using the same assumption, but with an adjustment to account for a change in sensitivity, as thresholds in the presence of interaurally uncorrelated masking noise were higher than corresponding thresholds when the masker was interaurally correlated. Experiments 1 and 3 investigate whether the dilution assumption is valid for interaurally correlated and uncorrelated masking noises, respectively.

In the latter case, the dilution assumption appears to run contrary to the findings of Jeffress *et al.* (1962), who measured participants' centering judgments at a range of interaural correlations by mixing interaurally correlated and uncorrelated noises. The data of Jeffress *et al.* (1962) show that for each halving of the coherence (which is equivalent to the proportion of delayed noise present), thresholds are substantially *less than* doubled (see Figs. 4 and 5, bottom right panel).

The following experiments investigate the effects of masking interaurally correlated (experiment 1), anticorrelated (experiment 2), and uncorrelated (experiments 3 and 4) noises on participants' ability to lateralize broadband noise using the localization cue provided by ITD, by directly mixing the target and masking noise in various proportions.[1] The aims of these experiments were fourfold.

The first aim was to test the dilution assumption in the case of interaurally correlated, anticorrelated, and uncorrelated maskers; when a broadband interaurally delayed noise target (N_τ) is heard in the presence of masking noise, does the detectable ITD double for each halving of the proportion of delayed noise present?

The second aim was to compare listeners' abilities to lateralize a signal embedded in noise to previous studies. The results of Cohen (1981) and Ito *et al.* (1982) found the smallest ITD jnds for interaurally correlated masking conditions, larger jnds for uncorrelated maskers, and the largest jnds for anticorrelated masking conditions. If the results of the current experiments follow this ordinal relationship among masking noises, then ITD threshold slope as a function of proportion of delayed noise present for interaurally uncorrelated masking noise will fall between the anticorrelated and the correlated masking noise slopes. Results with an interaurally uncorrelated masker can be directly compared to those of Jeffress *et al.* (1962), but comparisons with Egan and Benson (1966), Cohen (1981), and Ito *et al.* (1982) are complicated by the difference in bandwidth of the delayed signal.

The third aim was to quantify the effects of masking noise in order to apply the findings to the modeling techniques employed when measuring the binaural temporal window using a lateralization paradigm. This was achieved by measuring masking coefficients (the slope describing ITD threshold versus proportion of delayed masking noise) for interaurally correlated, anticorrelated, and uncorrelated masking noises.

The fourth aim was to compare the results of the experiments to thresholds predicted by the model of Ito *et al.* (1982), which is based on interaural phase difference.

II. METHODS COMMON TO EXPERIMENTS 1–3

The stimuli were generated digitally with a sampling rate of 44.1 kHz and 16-bit sample depth using MATLAB, band-limited between 100 and 3000 Hz prior to presentation, and gated with 10-ms raised-cosine onset and offset ramps. Listeners were presented with the stimuli over Sennheiser

HD590 headphones at an overall sound level of 75 dB (A), played through a 24-bit Edirol UA-20 sound card and passed through an MTR HPA-2 headphone amplifier in a single-walled Industrial Acoustics Company (IAC) sound-attenuating booth within a sound-deadened room. Trial-by-trial feedback was provided. Interaural time delays were imposed on the stimuli by adding a ramp function to the phase spectrum at one ear, resulting in an ongoing ITD (i.e., no onset-time difference).

III. EXPERIMENT 1: INTERAURALLY CORRELATED MASKING NOISE

In the first experiment, participants were given a 2I-2AFC discrimination task, where noise with an interaural correlation of 1 was mixed with interaurally delayed noise so that each interval was composed of delayed noise and correlated noise mixed at a predetermined proportion. Half the delay was imposed on interval 1, and half the delay, in the opposite direction, was imposed on interval 2. Consistent with the model of Bernstein *et al.* (2001), it was hypothesized that for each halving of the proportion of delayed noise present, the threshold ITD would double. Thus, on a log-log plot, according to the dilution assumption, as the proportion of delayed noise decreased and more interaurally correlated noise was mixed into the stimuli, thresholds would increase in a linear manner on a 1:1 slope. The slope describing ITD threshold versus proportion of delayed interaurally correlated noise was defined here as the correlated masking coefficient (CMC).

A. Stimuli and procedure

Six participants (including one of the authors) took part in the experiment. One was male and five were female, aged between 18 and 25. All participants except for ET and RH had previous experience with psychophysical experiments. Untrained participants received at least 5 h of training before data collection. Excluding the author, they were paid upon completion.

Listeners' performed a two-interval discrimination task. Three independent Gaussian noises were generated (N_1 and N_3), and a delayed copy of $N_1(N_2)$. The magnitude of the delay imposed on N_1 was half that of the delay difference between the two intervals. To keep the power constant, the noises were scaled in amplitude in the following manner, such that p is the proportion of the total power of the stimulus that is made up of delayed noise.

$$\text{Left channel:} \sqrt{p}N_1 + \sqrt{1-p}N_3. \tag{1}$$

$$\text{Right channel:} \sqrt{p}N_2 + \sqrt{1-p}N_3. \tag{2}$$

The second interval was created using the same procedure, except that the delayed noise was presented to the opposite ear. The adapted variable was the difference in ITD between the two intervals. The inter-stimulus interval was 500 ms, and each interval had a duration of 100 ms. ITD thresholds were obtained for $p=1$, 0.75, 0.56, 0.42, 0.32, 0.24, and 0.18.

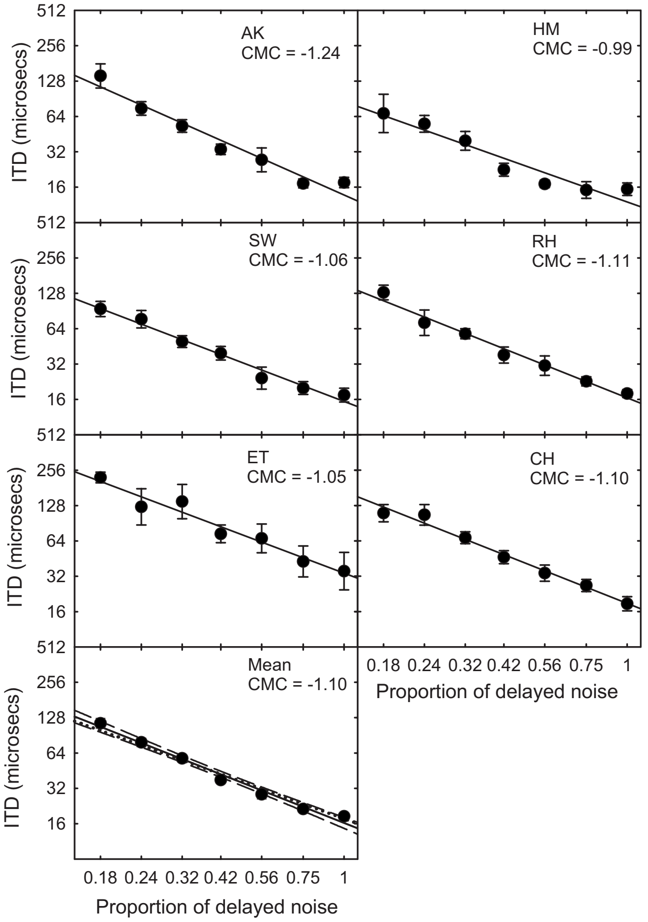

FIG. 1. Results for experiment 1 (correlated masker). Solid regression lines are plotted for individual listeners and the mean. The dotted line represents a slope of −1. Dashed lines represent 95% confidence intervals. Both axes are plotted logarithmically.

The listener's task was to identify in which direction the sound image moved from interval to interval (left to right or right to left), corresponding to the ITDs embedded within each interval. Participants pressed "1" if the noise moved from right to left, and "2" if it moved left to right. The sign of the interaural delay was randomized and trial-by-trial feedback was provided. The ITD was varied adaptively in order to obtain a 70.7% correct estimate (Levitt, 1971). Each adaptive track started with the ITD set to 500 μs. Initially, the step size of the adaptive track corresponded to a factor of 0.2 and was reduced to a factor of 0.05 following two reversals. Threshold was obtained after a further ten reversals. The last ten reversals composed a measurement phase, and the average of the reversals within the measurement phase was taken as threshold.

Each listener performed nine experimental runs. Measurement began after performance had stabilized; experimental thresholds were taken by averaging the remaining runs after the listener achieved thresholds below 500 μs at each proportion of delayed noise (nine runs for AK, CH, RH, and SW, eight runs for HM, and four runs for ET).

B. Results

Figure 1 shows threshold ITDs measured at different proportions of delayed noise mixed with interaurally correlated noise. When $p=1$, all the noise within each interval was delayed, producing the lowest thresholds for the six listeners (an average of 18.5 μs). Threshold ITD was found to approximately double for each halving of delayed noise present. The slope of the regression line of the mean observed data was -1.10 and accounted for 98% of the variance. The dotted line represents predicted values based on a CMC of -1 and falls within the 95% confidence intervals of the mean data, showing that the data do not deviate significantly from this trend.

C. Discussion

As the proportion of delayed noise decreased and more interaurally correlated masking noise was mixed into the intervals, thresholds increased in a log-linear manner on an approximately 1:1 slope so that for each halving of the delayed noise present, the threshold ITD doubled, supporting the dilution assumption of Bernstein *et al.* (2001).

The following two experiments characterized the effect of interaurally anticorrelated (experiment 2) and uncorrelated (experiment 3) masking noises on ITD threshold and examined whether thresholds change in the same manner as they do in the presence of interaurally correlated masking noise. However, we modified the method for these experiments compared to experiment 1 because we anticipated a potential problem. If one considers cross-correlation as a mathematical approximation to human ITD processing, the effect of increasing ITD will be to move a peak in the cross-correlation function further and further from zero. As ITD increases, the contrast between the cross-correlation functions for a positive and a negative ITD increases monotonically as the peaks separate. However, as there must be some limit to the range of delays over which cross-correlation is performed by the auditory system, this contrast eventually saturates and begins to decline, suggesting that discrimination performance will decline once again. Perceptually, the laterality reaches an extreme at somewhere around the ecological limit of naturally occurring ITDs, and then images start to become diffuse (Blodgett *et al.*, 1956). Adding interaurally uncorrelated noise can only be expected to make the task harder, as the noise will reduce the size of the cross-correlation peaks without altering their locations. Cohen (1981) and Ito *et al.* (1982) found that both interaurally uncorrelated and anticorrelated masking noises produced higher ITD jnds than correlated masking noise, raising the danger that, for a low proportion of delayed noise, threshold might be close to a non-monotonic section of the underlying psychometric function. With these facts in mind, we decided to obtain psychometric functions in experiments 2 and 3, rather than the adaptive track paradigm in order to avoid potential violation of assumption 1 of Levitt (1971) (p. 468): "The expected proportion of positive responses is a monotonic function of stimulus level (at least over the region in which observations are made)."

To obtain a psychometric function, stimuli were produced with a range of fixed ITDs, and the percentage of correct responses at each ITD was recorded. From these data, the 71% correct point of the psychometric function can be interpolated from a fitted function to obtain threshold values that can be directly compared to those obtained using the adaptive track method.

IV. EXPERIMENT 2: INTERAURALLY ANTICORRELATED MASKING NOISE

The log-linear slope describing ITD threshold versus the proportion of interaurally anticorrelated noise was defined here as the anticorrelated masking coefficient (AMC). This was measured in experiment 2.

A. Stimuli and procedure

Three participants (including one of the authors) took part in the experiment. One was male and two were female, aged between 18 and 28. All participants had previous experience with psychophysical experiments. Excluding the author, they were paid upon completion.

Listeners' performed a 2I-2AFC discrimination task. Stimuli were generated in the same way as in experiment 1, but with interaurally anticorrelated (correlation$=-1$) rather than correlated masking noise. That is to say, using Eqs. (1) and (2) except that N_3 was inverted in one channel. Intervals were presented so that the delayed noise in the first interval was presented to the ear favored by the interaural delay embedded in the interval, and the delayed noise in the second interval presented to the opposing ear. The inter-stimulus interval was 500 ms.

The listener's task was to identify in which direction the sound image moved between the two intervals (left to right or right to left), corresponding to the ITDs embedded within each interval. The direction of ITD change was randomized and trial-by-trial feedback was provided. The participants were tested at a stimulus duration of 100 ms with ITDs of 1024, 512, 256, 128, 64, 32, 16, and 8 μs. Initially, a block of trials was presented with proportions of 1, 0.5, 0.2, and 0.1, the order in which the eight ITDs were presented within the block was randomized, and 20 trials were repeated for each ITD. A second block of trials with proportions of 0.75, 0.25, 0.15, and 0.05 followed. Each block consisted of 8 ITDs$\times$4 proportions$\times$20 repetitions$=$640 trials in all. Following two training runs, participants performed three experimental runs at each proportion. The average of the three runs was taken as threshold. Two-parameter (slope and threshold) logistic functions were fitted to the data, and the 71% points of the functions were taken as thresholds in order to compare the thresholds with those taken in experiment 1 using the adaptive track procedure.

B. Results

Figure 2 shows threshold ITDs measured at different proportions of delayed noise mixed with interaurally anticorrelated noise. When $p=1$, all the noise within each interval was delayed, producing the lowest thresholds for the three listeners. Although listener EO's thresholds were substan-

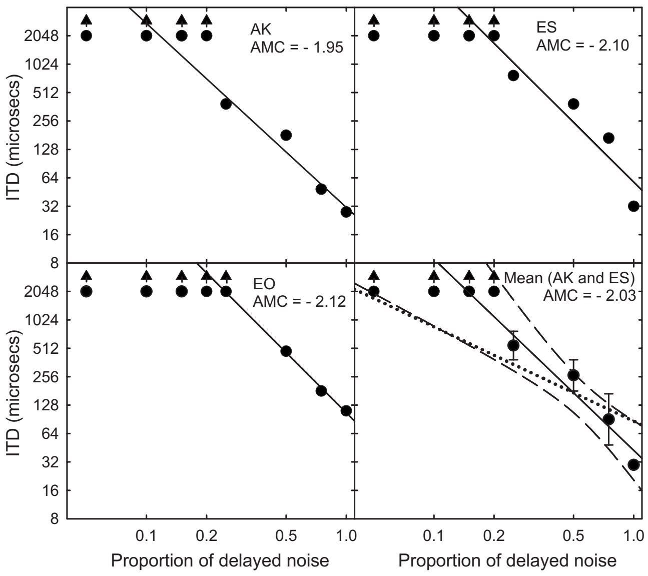

FIG. 2. Results for the experiment 2 (interaurally anticorrelated masker). Solid regression lines are plotted for individual listeners and the mean. Mean thresholds are plotted as an average across data for listeners AK and ES only, as listener EO displayed lower sensitivity than these listeners. Upward pointing arrows indicate that a measure of threshold could not be obtained for those coherences as the 71% point of the psychometric function fell outside the measured range. The dotted line represents a slope of −1. Dashed lines represent 95% confidence intervals. Both axes are plotted logarithmically.

tially higher than those of other listeners, the slope relating threshold to proportion of delayed noise present was approximately equal to that of other listeners, indicating that the sensitivity of this listener was lower than that of the others.[2] Consequently, mean thresholds in the bottom right panel of Fig. 2 are plotted as an average of listeners AK and ES only. Threshold ITD was found to more than double for each halving of delayed noise present. The slope of the regression line of the mean observed data was −2.03 and accounted for 92% of the variance. The dotted line represents predicted values based on a slope of −1 and falls outside the 95% confidence intervals of the mean data (represented by dashed lines in Fig. 2), showing that the slope of the data is significantly steeper than −1.

C. Discussion

As the proportion of delayed noise decreased and more interaurally anticorrelated noise was mixed into the intervals, thresholds increased in a log-linear manner on a slope of −2.03 so that for each halving of the delayed noise present, the threshold ITD more than doubled. As the slope of ITD thresholds was steeper than that observed when an interaurally correlated masker was employed (i.e., −1, see Fig. 1), the presence of interaurally anticorrelated noise was more disruptive to listener's abilities to lateralize a signal embedded in noise than interaurally correlated masking noise. These results follow the same hierarchy as those of Cohen (1981) and Ito *et al.* (1982), who found that ITD jnds were larger for interaurally anticorrelated than correlated masking conditions.

V. EXPERIMENT 3: INTERAURALLY UNCORRELATED MASKING NOISE

In experiment 3, the effect of interaurally uncorrelated masking noise on ITD threshold was assessed using stimuli comparable to experiments 1 and 2. The interaural coherence of each interval was manipulated by mixing interaurally delayed and uncorrelated noises. Note that when interaurally delayed and uncorrelated noises are added, the interaural coherence is equal to the proportion of delayed noise power in the stimulus, so the proportion of delayed noise is numerically equivalent to the coherence and to the interaural correlation when the noise delay is zero. The slope describing ITD threshold versus the proportion of interaurally uncorrelated noise was defined here as the uncorrelated masking coefficient (UMC).

Results obtained using an interaurally uncorrelated masker can be directly compared with those of Jeffress *et al.* (1962), who investigated the effect of interaural correlation on the ability of participants to center a noise. In that study, interaurally correlated and uncorrelated noises were mixed in each ear to obtain interaural correlations of 1, 0.75, 0.5, 0.25, 0.2, 0.15, 0.1, and 0. Participants adjusted a delay line in the presence of continuous noise in order to center the noise in their heads. The standard deviation of each participant's centering judgment was measured as a function of interaural correlation. This standard deviation should, in principle, be equivalent to an ITD threshold. The slope relating such a threshold ITD to correlation/coherence/proportion of delayed noise was −0.48 (see Fig. 4, bottom right panel). This slope is considerably shallower than the CMC values observed in experiment 1 with interaurally correlated masking noise that ranged from −0.99 to −1.24 (see Fig. 1).

To examine any effects of stimulus duration in experiment 3, participants were tested at three stimulus durations (100 ms, 500 ms, and 1 s). The long-duration stimulus conditions were included in order to be more comparable with the experiment of Jeffress *et al.* (1962) in which continuous noise was used.

A. Stimuli and procedure

Three listeners took part in experiment 3. One (the first author) was male and two were female, aged between 18 and 25. All listeners had participated in experiment 1. Excluding the author, they were paid upon completion.

Listeners' performed a 2I-2AFC discrimination task. To manipulate coherence, two independent Gaussian noises (N_1 and N_2) were generated and mixed: the left channel was presented containing noise N_1, and the right channel contained a mixture of two noises N_1 (delayed) and N_2 mixed in the proportion (p).

$$\text{Left channel:} N_1. \tag{3}$$

$$\text{Right channel:} pN_1 + \sqrt{1-p^2}N_2. \tag{4}$$

An interaural delay was added to one channel after noise mixing. Participants were tested at stimulus durations of 100 ms, 500 ms, and 1 s. Following two training runs, participants performed three experimental runs for each coherence

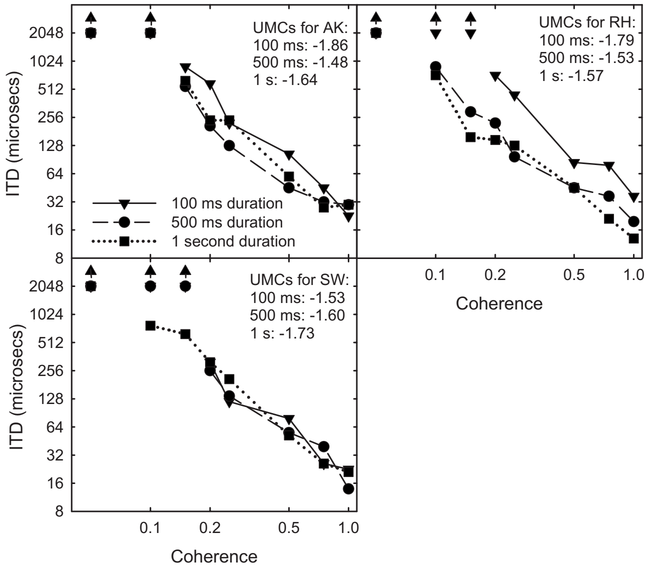

FIG. 3. Individual data for experiment 3 (interaurally uncorrelated masker) for stimulus durations of 100 ms (downward pointing filled triangles joined by solid lines), 500 ms (filled circles joined by dashed lines), and 1 s (filled squares joined by dotted lines). Thresholds represent 71% discriminability. Upward pointing arrows indicate that a measure of threshold could not be obtained for those coherences as the 71% point of the psychometric function fell outside the measured range. Both axes are plotted logarithmically.

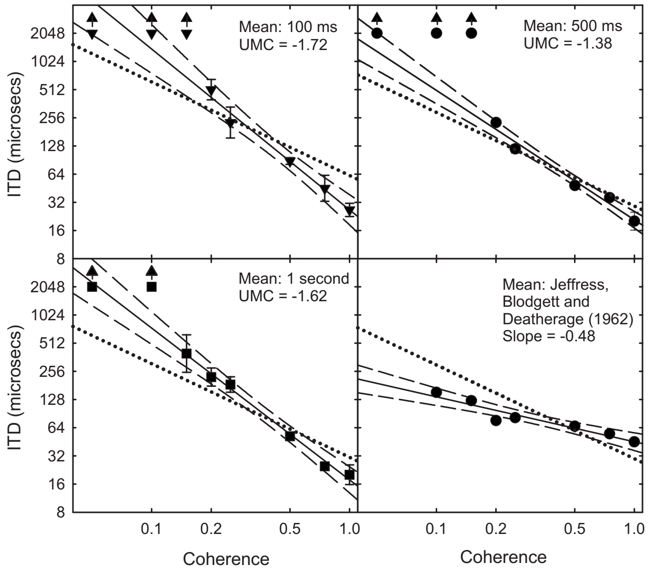

FIG. 4. Mean results for experiment 3 (interaurally uncorrelated masker). Each panel shows results at each of the three stimuli durations, and the bottom right panel shows mean data replotted from Jeffress *et al.* (1962), who used continuous stimuli. Solid regression lines are plotted for each duration. The plotted thresholds are 71% points averaged from the last three experimental runs. The dotted line represents thresholds with a slope of −1. Dashed lines represent 95% confidence intervals. Both axes are plotted logarithmically.

at each stimulus duration. Participants were tested in blocks of coherences of 1, 0.5, 0.2, 0.1, 0.75, 0.25, 0.15, and 0.05 by measuring psychometric functions using the procedure described in experiment 2.

B. Results

Figure 3 shows individual data for the three stimulus durations. UMCs ranged from −1.48 to −1.86 (for clarity, regression lines representing the UMC for each participant and duration are not plotted). The mean of the UMCs measured for individual participants across all listeners and stimulus durations was −1.64. When the coherence was unity, all the noise within each interval was delayed, producing the lowest thresholds for the three listeners. Threshold ITD was found to more than double for each halving of coherence. Thresholds at the lowest coherences could not be measured as the 71% point of the psychometric function fell outside the measurable range.

Figure 4 shows the mean of the data across listeners for the three stimulus durations. The dotted lines represent predicted values based on a UMC of −1. At all three stimulus durations the dotted lines fall outside of the 95% confidence intervals of the regression (represented by dashed lines in Fig. 4), in each case indicating that the slope of the data was significantly steeper than −1. The mean UMCs for each duration ranged from −1.38 to −1.72, which were substantially higher than the slope of −0.48 obtained by Jeffress *et al.* (1962). No consistent effect of stimulus duration was observed, and the range of UMCs obtained across stimulus duration and listener was wide (−1.48 to −1.86).

C. Discussion

As coherence decreased, thresholds rose so that each time the coherence was halved, thresholds more than doubled (see Figs. 3 and 4). Once again, this finding contrasts with the effect on ITD of interaurally correlated masking noise observed in experiment 1, where thresholds were observed to double, and indicates that the dilution hypothesis does not hold for the case where an interaurally uncorrelated masker is present.

An explanation of the differing influences of interaurally correlated, anticorrelated, and uncorrelated masking noises was put forward by Ito *et al.* (1982) based on their model. They measured interaural time jnds for narrow-band noise in the presence of a broadband masker. They were able to model the hierarchy of ITD jnds at a fixed SNR of approximately +10 dB for various interaural configurations of the masker (smallest for the interaurally correlated masker, larger for the uncorrelated masker, and largest for the anticorrelated masker). Their model was based on variations in the distribution of interaural phase differences of the stimulus components and similar to the models proposed by Webster (1951) and Jeffress *et al.* (1956).

The explanation of Ito *et al.* (1982) essentially encompassed signal detection notions of how changing the ratio between the mean (μ) and the standard deviation (σ) is predicted to affect sensitivity, that discriminability depends on the differences in the means relative to the standard deviation of the two ITD distributions. When no masking noise was present, thresholds were approximately equal for all types of masking noise (see Figs. 1–4), but when masking noise was added the distribution of ITDs became wider and the standard deviation increased, requiring an increase in the magnitude of ITD to support discrimination. According to this model, the slopes of the CMC, UMC, and AMC differ because, as Ito *et al.* (1982) showed, the variability of short-term estimates of ITD about their mean of zero was comparably small in the presence of an interaurally correlated masker, larger when an uncorrelated masker was present, and largest when an anticorrelated masker was present. The predictions of the model of Ito *et al.* (1982) are evaluated more fully in Sec. VI.

Steeper UMCs at all three stimulus durations were observed in the current experiment compared to the slope observed by Jeffress *et al.* (1962) (see Fig. 4). These results suggest that the shallow slope obtained in their experiment was probably not due to their use of continuous stimuli, as the UMC obtained with each stimulus duration was considerably steeper than the slope obtained by Jeffress *et al.* (1962). The following experiment was carried out in order to replicate the experiment of Jeffress *et al.* (1962), and explored the possibility that the shallow slope of their results stemmed from the narrow range of ITDs that were employed in their study.

VI. EXPERIMENT 4: REPLICATION OF Jeffress *et al.* (1962)

As described in the introduction to experiment 3, Jeffress *et al.* (1962) asked listeners to center a noise in their heads by means of adjusting a delay line with an ITD range of ±450 μs. It is not clear whether an adjustment beyond the ITD limit resulted in the ITD reaching a hard limit where the delay line reached its maximum and the knob consequently stopped moving, or whether the sign of the interaural delay was reversed at this point, flipping the lateralization of the stimulus from one side of the head to the other. It seems reasonable to assume that the latter was the case[3] as the former would give the listener a rather obvious cue. Even so, it is possible that an unintended cue was still present. The distance from the control knob's "flipping point" when the ITD limit was reached could have been used by the listener as a method of producing relatively accurate centering. Increasing the range of the ITDs available would reduce the usefulness of this cue because the laterality produced by ITD decreases above about 1 ms (Blodgett *et al.*, 1956) which would prevent the stimulus "flipping" in perceived laterality. This possibility was addressed in the following replication of the experiment of Jeffress *et al.* (1962), where participants performed the task with ITDs ranging from ±450 μs, 1 ms, and 2 ms. An increase in threshold slope with increasing ITD range would support the hypothesis that a restricted range provided a cue based on the stimulus flipping in laterality when the ITD limit was reached.

A. Stimuli and procedure

Three listeners took part in experiment 4. One (the first author) was male and two were female, aged between 18 and 25. The author had participated in experiments 1–3, and the others were naïve participants. Excluding the author, they were paid upon completion.

The stimuli consisted of continuous broadband (0–10 kHz) noise generated by concatenating 102.4-ms segments of noise drawn at random from a 5-s buffer. The coherence of the stimuli was manipulated using the same mixing method implemented in experiments 1 and 2, but with N_3 replaced by an independent noise at one ear. All noises were obtained from the 5-s buffer. The participants' task was to center the noise in their heads by using keyboard controls. Pressing '1' resulted in a decrement in stimulus ITD of 36 μs, '2' a decrement of 9 μs, "3" an increment of 9 μs, and "4" an increment of 36 μs. Pressing "x" terminated the program. If the ITD limit was reached, the sign of the interaural delay was reversed. Each stimulus trial was initially presented with a random ITD within the specified range. Initially, participants were presented with a block of trials with coherences of 1, 0.75, 0.5, 0.25, 0.2, 0.15, 0.1, and 0 with an ITD range of ±450 μs in order to familiarize them with the experimental setup. Following this training, participants were given another block of trials with the ITD ranging from ±450 μs, and blocks of ±1 and ±2 ms followed. This procedure was repeated three times, and the standard deviation of the three runs at each ITD range was taken as threshold.

All stimuli were generated on-line at a sampling frequency of 20 kHz using a TDT AP2 array-processor card and presented via a TDT System-2 psychoacoustic rig (DD1, PA4, FT5-9, and HB6) and presented over Sennheiser HD650 headphones at an overall sound level of 70 dB (A) in a single-walled IAC sound-attenuating booth within a sound-deadened room. Filtering was performed using an FT6 reconstruction filter with a low pass cut-off frequency of 10 kHz. No feedback was given.

B. Results

Figure 5 shows results from the replication of Jeffress *et al.* (1962) at ITD ranges of ±0.45, ±1, and ±2 ms. The dotted lines represent predicted values based on a UMC of −1. The slope of the data at an ITD range of ±0.45 ms (−0.54) was comparable to the slope obtained by Jeffress *et al.* (1962), which was −0.48. As the range increased, the slope of the data became steeper. At a range of ±2 ms, the slope of the data (−1.67) was comparable to the UMCs observed in experiment 3.

C. Discussion

The data from experiment 4 show that the shallow slope of ITD thresholds obtained by Jeffress *et al.* (1962) was probably due to the limited range of ITDs (±0.45 ms) that were employed in that study. The increase in threshold slope with increasing ITD range supports the hypothesis that a restricted ITD range provides a cue based on the stimulus flipping in laterality when the ITD limit was reached. Increasing the range progressively reduced the usefulness of this cue because the laterality produced by ITD decreases above about 1 ms (Blodgett *et al.*, 1956).

VII. PREDICTIONS OF THE MODEL OF Ito *et al.* (1982)

The model developed by Ito *et al.* (1982) was implemented computationally in order to compare the experimen-

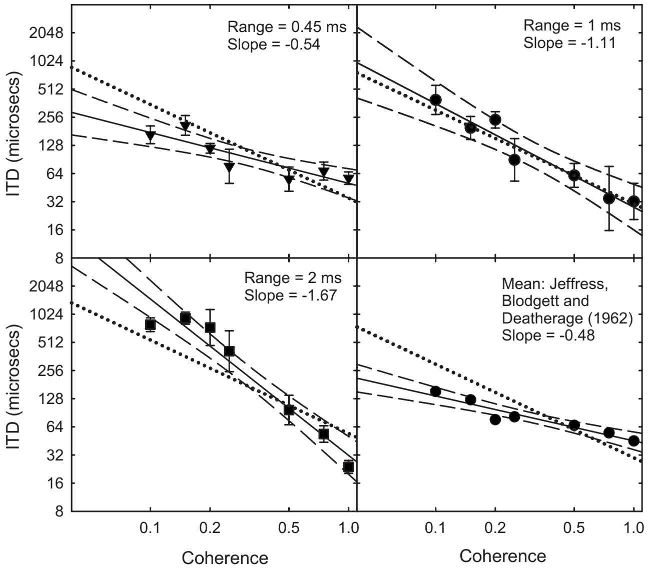

FIG. 5. Mean results for experiment 4. Each panel shows results with continuous stimuli at ITD ranges of ±0.45 μs [the range employed by Jeffress *et al.* (1962)], ±1 ms, and ±2 ms. Solid regression lines are plotted for each range of ITD. The plotted thresholds are averaged standard deviations for the three participants. The dotted line represents thresholds with a slope of −1. Dashed lines represent 95% confidence intervals. Both axes are plotted logarithmically.

tal results with thresholds produced by the model (see Appendix). Figure 6 (right panel) shows the effect of p on the predicted thresholds produced by our implementation of the model (each data point is the mean of four simulations) and compares it to the mean coefficients observed for experiments 1–3 (left panel). Interestingly, the ordinal relationship among the three types of masking noise changes as a function of p. For p below about 0.5 (corresponding to 0-dB SNR), interaurally uncorrelated noise was predicted to be the most effective masker. The fitted slope to the thresholds simulated for an interaurally correlated masker (−1.07) was very similar to the CMC obtained in experiment 1 (−1.10). The fitted slope for an interaurally uncorrelated masker (−1.52) was very similar to the mean UMC observed in the current study (−1.57), but the fitted slope for an anticorrelated masker (−1.35) was considerably shallower than the mean AMC observed in experiment 2 (−2.03). Thresholds simulated for masking interaurally uncorrelated and anticorrelated noises were observed to be curvilinear in the region of high correlation (see right panel of Fig. 6). Filled symbols indicate thresholds taken from Fig. 1 of Ito *et al.* (1982) for the first set of four listeners in interaurally correlated, uncorrelated, and anticorrelated conditions at a single SNR of +10.3 dB, equivalent to a proportion of delayed noise (p) of 0.915. Despite differences in the stimuli employed [the stimuli of Ito *et al.* (1982) were band-limited to 1/3 octave around 500 Hz (445–561 Hz) with a duration of 350 ms, whereas stimuli in the current experiment were broadband (0–3 kHz) with durations ranging between 100 ms and 1 s], the ordinal relation between thresholds obtained with interaurally correlated, uncorrelated, and anticorrelated masking noises and the magnitude of the ITD thresholds is similar across the two studies.

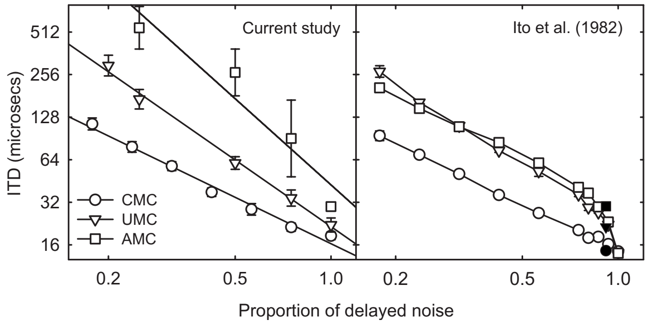

FIG. 6. Mean masking thresholds from the current study (left panel) and predicted thresholds produced by implementation of the model developed by Ito *et al.* (1982) (right panel). Circles indicate an interaurally correlated masker, triangles indicate an uncorrelated masker, and squares an anticorrelated masker. Solid regression lines are plotted for coefficients produced in the current study. The plotted CMC thresholds are averaged across all listeners, and the AMC thresholds are averaged over listeners AK and ES. Plotted UMC thresholds are averaged across all listeners and all stimuli durations. Thresholds plotted from the model of Ito *et al.* (1982) are averaged across four simulations. Filled symbols are the plotted means of thresholds taken from Fig. 1 of Ito *et al.* (1982). Both axes are plotted logarithmically.

In summary, implementation of the model of Ito *et al.* (1982) produced a slope of thresholds for interaurally correlated masking noise that closely matched the CMC observed in experiment 1, and produced the ordinal relationship among the three types of masking noises for p values above approximately 0.5. However, the functions relating proportion of delayed noise to ITD threshold for interaurally uncorrelated and anticorrelated noises were incorrectly predicted to be curvilinear, and below $p=0.5$ the ordinal relationship changed, with uncorrelated noise predicted to produce more effective masking than anticorrelated noise.

VIII. GENERAL DISCUSSION

The results of the experiments described in this study indicated that the presence of masking noise resulted in an increase in the ITD threshold of delayed noise. In the case of an interaurally correlated masker, for each halving of the proportion of delayed noise present, threshold ITD doubled (i.e., ITD threshold slope=−1). ITD threshold slope was steeper than −1 in the case of an interaurally uncorrelated masker (−1.64), and steeper still in the case of an interaurally anticorrelated masker (−2.03). Thus, the experiments described in this paper indicate that masking noise with an interaural correlation of 1, anticorrelated noise, and uncorrelated noise would be integrated within a binaural temporal window in different ways. Modeling the temporal window should thus depend on the type of masker incorporated in the design of the experiment.

For example, Bernstein *et al.* (2001) measured temporal windows by presenting the listener with a burst of interaurally delayed noise (the probe) temporally fringed with masking noise. Best-fit thresholds were calculated in the following manner. First, the integral of the area occupied by the probe was computed and divided by the total integral of the window. This value represented the proportion of delayed noise within the window and is affected by the duration and the assumed shape of the window. Best-fit ITD thresholds were then calculated by using a coefficient that related threshold to proportion of delayed noise present. By using the assumption that masking noise diluted the delayed noise, Bernstein *et al.* (2001) effectively determined predicted threshold ITDs using a coefficient of -1. The experiments described in this paper have allowed temporal windows to be measured using stimuli of a similar design to Bernstein *et al.* (2001) employing uncorrelated masking noise and determining predicted threshold ITDs using the UMC (Kolarik and Culling, 2009). When interaurally correlated masking noise is present, then the CMC (-1) should be modeled to predict thresholds, for uncorrelated masking noise the UMC (-1.64) should be modeled, and for anticorrelated masking noise the AMC (-2.03) should be modeled.

IX. CONCLUSIONS

The data described within this paper demonstrate the following conclusions.

(1) When interaurally correlated masking noise is mixed with delayed noise, ITD threshold doubles for each halving of delayed noise present.
(2) When interaurally anticorrelated or uncorrelated masking noise is mixed with delayed noise, ITD threshold more than doubles for each halving of delayed noise present. The presence of interaurally anticorrelated noise is more debilitating to the listener than the presence of uncorrelated noise.

ACKNOWLEDGMENTS

The work was supported by the EPSRC. Thanks are due to Steve Colburn and an anonymous reviewer for their helpful comments relating to earlier drafts of the manuscript. We would also like to thank Mike Akeroyd for his discerning advice and to Torsten Marquardt for his thoughtful comments. Summaries of these findings were presented at the September 2005 BSA and May 2006 ASA meetings.

APPENDIX

The model of Ito *et al.* (1982) of sensitivity to interaural delay under masking was implemented computationally using MATLAB. Stimuli similar to those used by Ito *et al.* (1982) in their experiments were generated and band-limited to 1/3 octave around 500 Hz (445–561 Hz) with a duration of 350 ms. To create the stimulus, delayed and masking noises were added together in the appropriate ratio. Ito *et al.* (1982) employed a single SNR of +10.3 dB equivalent to a proportion of delayed noise (p) of 0.915. The mean interaural phase difference was derived by calculating the analytic signal and thus the instantaneous phase for each channel, subtracting the instantaneous phase of the left channel from the right, wrapping values outside the range $\pm\pi$ back into that range, and then taking the average across the entire 350-ms stimulus. We found that implementing internal noise as additive interaural phase noise with a standard deviation of 1.5 rad as independent samples yielded good predictions of the data of Ito *et al.* (1982). If the resulting mean instantaneous interaural phase was positive, the model reported the stimulus on the right, and otherwise on the left. Predictions of percent correct were produced by repeating these operations for 200 such stimuli. Predicted thresholds were derived by generating psychometric functions for seven noise delays logarithmically spaced between 10 and 316 μs and taking the 75% point from a fitted logistic function.

---

[1] It is assumed that the binaural system does not distinguish between the temporal differences between sequential presentation of noises (as per the stimuli presented by Bernstein *et al.* (2001) and simultaneously presented delayed target and masking noise (as in the current experiments). For example, the temporal sequence of ITDs in the experiment of Bernstein *et al.* (2001) with interaurally correlated masking noise is described by a series of zero values during presentation of the interaurally correlated noise preceding the probe, followed by a series of values at the test ITD, followed by a series of zero values during the interaurally correlated noise lagging the probe. In the current experiments where delayed and masking noises are directly mixed, the sequence of ITD values is always random due to the combination of the delayed and masking noises. When delayed noise is mixed with interaurally correlated masking noise, the interaction of the two statistically independent noises results in a random distribution of noises with a mean determined by the relative power of the noise components, where the mean is between zero and the ITD of the delayed noise. When the delayed target noise and the masking noise are mixed within a hypothetical temporal window, it is assumed that it is still the proportion of delayed noise within a window that dictates threshold ITD, regardless of how it is temporally distributed.

[2] Participant EO received an additional 4 h of data collection in order to establish whether further training would result in lower thresholds. As no further improvement was observed, and as this listener was already previously experienced in psychophysical experiments, it was concluded that she was less sensitive than the others.

[3] Jeffress *et al.* (1962) were able to accurately predict listeners' chance performance at a correlation of zero by assuming a rectangular chance distribution among the points of the switch that listeners used to respond in a condition that was included as check on the experimental method. Had the end-points of the delay line been marked by a fixed limit to the knob movement, listeners would have been able to at least avoid setting the knob close to this limit.

Bernstein, L. R., Trahiotis, C., Akeroyd, M. A., and Hartung, K. (**2001**). "Sensitivity to brief changes of interaural time and interaural intensity," J. Acoust. Soc. Am. **109**, 1604–1615.

Blodgett, H. C., Wilbanks, W. A., and Jeffress, L. A. (**1956**). "Effect of large interaural time differences upon the judgment of sidedness," J. Acoust. Soc. Am. **28**, 639–643.

Cohen, M. F. (**1981**). "Interaural time discrimination in noise," J. Acoust. Soc. Am. **70**, 1289–1293.

Egan, J. P., and Benson, W. (**1966**). "Lateralization of a weak signal presented with correlated and with uncorrelated noise," J. Acoust. Soc. Am. **40**, 20–26.

Good, M. D., and Gilkey, R. H. (**1996**). "Sound localization in noise: I. The effect of signal-to-noise ratio," J. Acoust. Soc. Am. **99**, 1108–1117.

Ito, Y., Colburn, H. S., and Thompson, C. L. (**1982**). "Masked discrimination of interaural time delays with narrow-band signal," J. Acoust. Soc. Am. **72**, 1821–1826.

Jeffress, L. A., Blodgett, H. C., and Deatherage, B. H. (**1962**). "Effect of interaural correlation on the precision of centering a noise," J. Acoust. Soc. Am. **34**, 1122–1123.

Jeffress, L. A., Blodgett, H. C., Sandel, T. T., and Wood, C. L. (**1956**). "Masking of tonal signals," J. Acoust. Soc. Am. **28**, 416–426.

Kolarik, A. J., and Culling, J. F. (**2009**). "Measurement of the binaural temporal window using a lateralization task," Hear. Res. **248**, 60–68.

Kuhn, G. F. (**1977**). "Model for the interaural time differences in the azimuthal plane," J. Acoust. Soc. Am. **62**, 157–167.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

Lopez-Poveda, E. A., and Meddis, R. (**1996**). "A physical model of sound diffraction and reflections in the human concha," J. Acoust. Soc. Am. **100**, 3248–3259.

Perret, S., and Noble, W. (**1997**). “The effect of head rotations on vertical plane sound localization,” Percept. Psychophys. **59**, 1018–1026.

Wagner, H. (**1991**). “A temporal window for lateralization of interaural time differences in barn owls,” J. Comp. Physiol. [A] **169**, 281–289.

Webster, F. A. (**1951**). “The influence of interaural phase on masked thresholds I. The role of interaural time-deviation,” J. Acoust. Soc. Am. **23**, 452–462.

Wightman, F. L., and Kistler, D. J. (**1992**). “The dominant role of low-frequency interaural time differences in sound localization,” J. Acoust. Soc. Am. **91**, 1648–1661.

Wightman, F. L., and Kistler, D. J. (**1999**). “Resolution of front-back ambiguity in spatial hearing by listener and source movement,” J. Acoust. Soc. Am. **105**, 2841.