Factors affecting the use of envelope interaural time differences in reverberation[a)]

Jessica J. M. Monaghan,[b)] Katrin Krumbholz, and Bernhard U. Seeber[c)]
MRC Institute of Hearing Research, University Park, Nottingham NG7 2RD, United Kingdom

(Received 31 January 2011; revised 23 January 2013; accepted 7 February 2013)

At high frequencies, interaural time differences (ITDs) are conveyed by the sound envelope. Sensitivity to envelope ITDs depends crucially on the envelope shape. Reverberation degrades the envelope shape, reducing the modulation depth of the envelope and the slope of its flanks. Reverberation also reduces the envelope interaural coherence (i.e., the similarity of the envelopes at two ears). The current study investigates the extent to which these changes affect sensitivity to envelope ITDs. The first experiment measured ITD discrimination thresholds at low and high frequencies in a simulated room. The stimulus was either a low-frequency narrowband noise or the same noise transposed to a higher frequency. The results suggest that the effect of reverberation on ITD thresholds was multiplicative. Given that the threshold without reverberation was larger for the transposed than for the low-frequency stimulus, this meant that, in absolute terms, the thresholds for the transposed stimulus showed a much greater increase due to reverberation than those for the low-frequency stimulus. Three further experiments indicated that the effect of reverberation on the envelope ITD thresholds was due to the combined effect of the reduction in the envelope modulation depth and slopes, as well as the decrease in the envelope interaural coherence.
 [http://dx.doi.org/10.1121/1.4793270]

PACS number(s): 43.66.Pn, 43.66.Ts, 43.64.Me [RYL] Pages: 2288–2300

I. INTRODUCTION

At low frequencies, interaural time differences (ITDs) provide the dominant cue for horizontal sound-source localization in humans (Wightman and Kistler, 1992; Macpherson and Middlebrooks, 2002). Low-frequency ITDs are conveyed mainly through phase-locked neural firing to the temporal fine structure (TFS) of the sound waveform. Phase locking is lost at high frequencies, so ITDs can only be conveyed through the sound's time-varying amplitude or temporal envelope. Perceptual sensitivity to envelope ITDs has been shown to depend crucially on the envelope shape. Generally, discrimination thresholds for envelope ITDs tend to be considerably higher than for fine-structure ITDs (Henning, 1974; Nuetzel and Hafter, 1981). However, envelope ITD thresholds have been shown to decrease as envelopes become sharper and more highly modulated (Bernstein and Trahiotis, 2009; Klein-Hennig *et al*., 2011; Laback *et al*., 2011). For instance, "transposed" tones (van de Par and Kohlrausch, 1997), which are high-frequency sounds with envelopes designed to mimic the TFS information in low-frequency sounds, have been shown to yield similar, or even lower, ITD thresholds as low-frequency tones (Bernstein and Trahiotis, 2002). The degree of envelope modulation has also been shown to influence the perceptual weighting of envelope ITDs in simulations of free-field listening (Macpherson and Middlebrooks, 2002; Wiggins and Seeber, 2011). The topic of envelope ITDs has particular significance for bilateral cochlear implants (CIs) because most current devices only restore envelope ITDs and interaural level differences (ILDs; Wilson and Dorman, 2008).

Ruggles *et al*. (2012) have suggested that envelope ITDs might be more useful for every-day listening than laboratory studies using anechoic conditions would imply. They present an analysis suggesting that reverberation reduces the interaural coherence (i.e., similarity across the two ears) of high-frequency envelopes to lesser degree than that of low-frequency TFS. At low frequencies, reduction in the TFS interaural coherence has been shown to reduce sensitivity to fine-structure ITDs (Jeffress *et al*., 1962; Zimmer and Macaluso, 2005; Kolarik and Culling, 2009). By analogy, reduction in the envelope interaural coherence should reduce envelope ITD sensitivity at high frequencies. While reverberation has a lesser effect on the envelope than the fine-structure interaural coherence, Rakerd and Hartmann (2010) and Devore and Delgutte (2010) found that the effect of reverberation on the perception and neural representation of envelope ITDs is more detrimental compared to fine-structure ITDs. Using narrowband noises recorded in real rooms, Rakerd and Hartmann found a greater increase in ITD discrimination threshold and decrease in the perceptual weighting of ITDs with increasing amount of reverberation at high frequencies than at low frequencies. Devore and Delgutte measured the effect of reverberation on the directional sensitivity of single neurons in the inferior colliculus (IC) and found a more detrimental effect on high-frequency neurons sensitive to envelope ITDs than on low-frequency neurons sensitive to fine-structure ITDs. The reason why

[a)]Portions of this work were presented at the 2010 spring meeting of the Acoustical Society of America and the 2011 Short Papers Meeting of the British Society of Audiology.

[b)]Author to whom correspondence should be addressed. Current address: Institute for Sound and Vibration Research, University of Southampton, UK. Electronic mail: jessica.monaghan@gmail.com

[c)]Current address: Associated Institute for Audio Information Processing, Technische Universität München, 80333 München, Germany.

 0001-4966/2013/133(4)/2288/13/$30.00

envelope ITDs are more susceptible to reverberation than fine-structure ITDs may be because reverberation not only reduces interaural coherence but also degrades the envelope shape, reducing the envelope modulation depth and the slope of the envelope flanks (Houtgast and Steeneken, 1973).

The aim of the current study was to disentangle the effects of reverberation-induced reduction in interaural coherence and degradation in envelope shape on sensitivity to envelope ITDs. The first experiment compared the effect of reverberation on ITD discrimination performance at low and high frequencies. To avoid confounding the acoustic effects of reverberation with the perceptual effects, transposition was used to present the same amount of reverberation in both frequency regions. Three subsequent experiments then assessed the effects of reduction in interaural coherence and degradation in envelope shape separately using synthetic high-frequency stimuli. The stimuli were designed in such a way as to allow each envelope property to be varied independently and a parametric relation with ITD threshold to be established.

II. EXPERIMENT 1: EFFECT OF REVERBERATION ON ITD THRESHOLDS AT HIGH AND LOW FREQUENCIES

The first experiment measured ITD discrimination thresholds for a low-frequency narrowband noise with simulated reverberation and for the same noise transposed to a higher frequency (4 kHz). In the transposed stimulus, the ITD information was only present in the stimulus envelope. The low-frequency stimulus was centered at 256 Hz because this frequency has been shown to yield relatively low ITD thresholds in both the fine-structure and envelope domains (Bernstein and Trahiotis, 2002).

A. Methods

1. Participants

Six participants (4 female) aged between 23 and 29 yr took part in Experiment 1. None of the participants had any history of hearing disorders, and all had absolute thresholds of 20 dB HL or lower at audiometric frequencies. All but one participant had previous experience in ITD discrimination and other psychoacoustic tasks. One participant (s04) was author J.J.M.M., the others were paid at an hourly rate.

2. Procedure

ITD discrimination thresholds were obtained using an adaptive three-down, one-up procedure, tracking the ITD that yields 79.4% correct responses (Levitt, 1971). The tracking algorithm started with an initial ITD of 1000 μs and used multiplicative steps. The step size was a factor of 1.584 up to the first two reversals in ITD and was decreasing to 1.122 thereafter. Threshold estimates were determined by averaging over 12 reversals once the smaller step size had been reached. Three threshold estimates were obtained for each condition.

A two-interval, two-alternative forced-choice task was used. In the first interval, the stimulus was lateralized with half the ITD to the left or right side (equal probability) and to the opposite side in the second interval. Participants were asked to report whether the stimulus moved from left to right or right to left.

3. Stimuli

The stimuli were generated digitally with a sampling rate of 44.1 kHz using MATLAB (The Mathworks, Natick, MA), converted to analogue with a 24-bit converter built into a custom headphone amplifier and presented via headphones (Sennheiser HD 600, Hannover, Germany) to the participant, who was seated in a double-walled, sound-attenuating room.

Reflections in an empty, rectangular room were simulated using a custom-made software based on the source-image method (Seeber *et al.*, 2010). The room was 4.7 m wide, 6.8 m long, and 2.5 m high. The absorption coefficients of the walls, floor, and ceiling were chosen to simulate an unfurnished, carpeted room with gypsum-cardboard walls. The reverberation radius of the room was 0.98 m, and the reverberation time was 0.27 s. The source was located toward one of the corners of the room, 1 m away from the shorter wall, 0.7 m away from the longer wall, and at a height of 1.4 m. The source was directed toward the receiver, and the line connecting the source and receiver was oriented at an angle of 68° from the shorter wall. The receiver always faced the source to minimize ITD and ILD cues in the direct sound and was set at the same elevation as the source to avoid changes in elevation-related pinna cues with increasing distance from the source. Five different source-receiver distances ranging from 0.5 to 2.5 m in 0.5-m steps (0.5, 1, 1.5, 2, and 2.5 m) were used. The binaural room impulse response (BRIR) for each distance was calculated using the head-related transfer functions (HRTFs) of participant "Bru" from the AUDIS project (Blauert *et al.*, 1998).

The BRIRs were separated into the part containing the direct sound (i.e., the sound arriving at the receiver directly from the source with no reflections) and the part containing the reverberation. A Gaussian noise with a duration of 1 s was generated and filtered in the frequency domain around a center frequency of 256 Hz and with a bandwidth of 105 Hz (corresponding to two equivalent rectangular bandwidths, or ERBs; Glasberg and Moore, 1990). The direct sound was calculated by convolving this noise with the direct-sound part of the BRIRs, which simulates the attenuation and delay of the sound between the source and receiver, as well as the effect of the HRTFs. Likewise, the reflections were calculated by convolving the noise with the reverberant part of the BRIRs. The ITDs to be discriminated were applied only to the direct sound. The experiment included a "non-reverberant" condition, where only the direct sound was used. In all other conditions, the reflections were added to the direct sound with no additional ITD. This allowed us to test sensitivity to ITDs in the direct sound only while keeping the reflections constant.

For the low-frequency conditions, a 700-ms segment was extracted from the middle of the BRIR-convolved noises to allow time for the reverberation to build up. The stimuli to the left- and the right-ear channels were gated on and off simultaneously using 50-ms Gaussian ramps, so that participants had to base their judgments on the ongoing rather than any onset or offset, ITDs.

For the transposed conditions, the BRIR-convolved noises were transposed to 4 kHz using the same procedure as used by Bernstein and Trahiotis (2002) for transposing pure tones. In particular, the stimulus was half-wave rectified and low-pass filtered below 2 kHz to restrict the bandwidth before multiplying with a 4-kHz diotic carrier tone. The transposed stimuli were gated in the same way as the low-frequency stimuli. They were presented with a continuous interaurally uncorrelated noise intended to mask audible distortion products. The continuous noise was equally exciting and low-pass filtered below 2 kHz in the frequency domain, and was presented at an overall level of 50 dB SPL. Both the low-frequency and transposed stimuli were presented at a level of 60 dB SPL.

B. Results

Figure 1 shows that the ITD thresholds increased with increasing source-receiver distance. This was true for both the low-frequency (filled symbols and solid lines) and transposed conditions (open symbols and dashed lines) and for all participants (upper panels; the lower panel shows the average). Participant s05 was excluded from further analysis because thresholds could not be obtained for all conditions for this participant. The thresholds for the transposed conditions were generally higher than those for the low-frequency conditions, and they also increased more strongly with increasing source-receiver distance: For the transposed conditions, the average thresholds increased from 100 μs without reverberation, to 1136 μs at largest source-receiver distance of 2.5 m; for the low-frequency conditions, the average thresholds increased from 61 to only 631 μs. The statistical significance of these results was tested with a two-way repeated-measures (RM) ANOVA with factors stimulus condition (low-frequency, transposed) and source-receiver distance (0.5, 1, 1.5, 2 2.5 m). The analysis revealed significant main effects of stimulus condition [$F(1,4) = 64.3$, $p = 0.001$] and distance [$F(5,20) = 40.3$, $p < 0.001$] as well as a significant interaction between these two factors [$F(5,20) = 5.49$, $p = 0.002$].

To assess the relative effect of reverberation on the ITD thresholds, we normalized the thresholds for each stimulus condition to the thresholds for the non-reverberant condition. Figure 3(a) shows that normalization virtually abolished the differences between the two stimulus conditions. This was confirmed with an RM ANOVA of the normalized thresholds, which still showed a significant main effect of distance [$F(4,20) = 21.8$, $p < 0.001$] but a non-significant main effect of stimulus condition [$F(1,4) = 1.71$, $p = 0.261$] and a non-significant interaction between these two factors [$F(5,20) = 0.39$, $p = 0.852$]. This indicates that the effect of reverberation on the ITD thresholds was multiplicative in the sense that the reverberation-induced threshold detriment was proportional to the non-reverberant threshold.

To determine the extent to which reduction in interaural coherence accounted for the detrimental effect of reverberation on the ITD thresholds, we calculated the interaural coherence, k, of each stimulus as the maximum of the normalized cross-correlation (ρ_{LR}) between the signals at the two ears, i.e.,

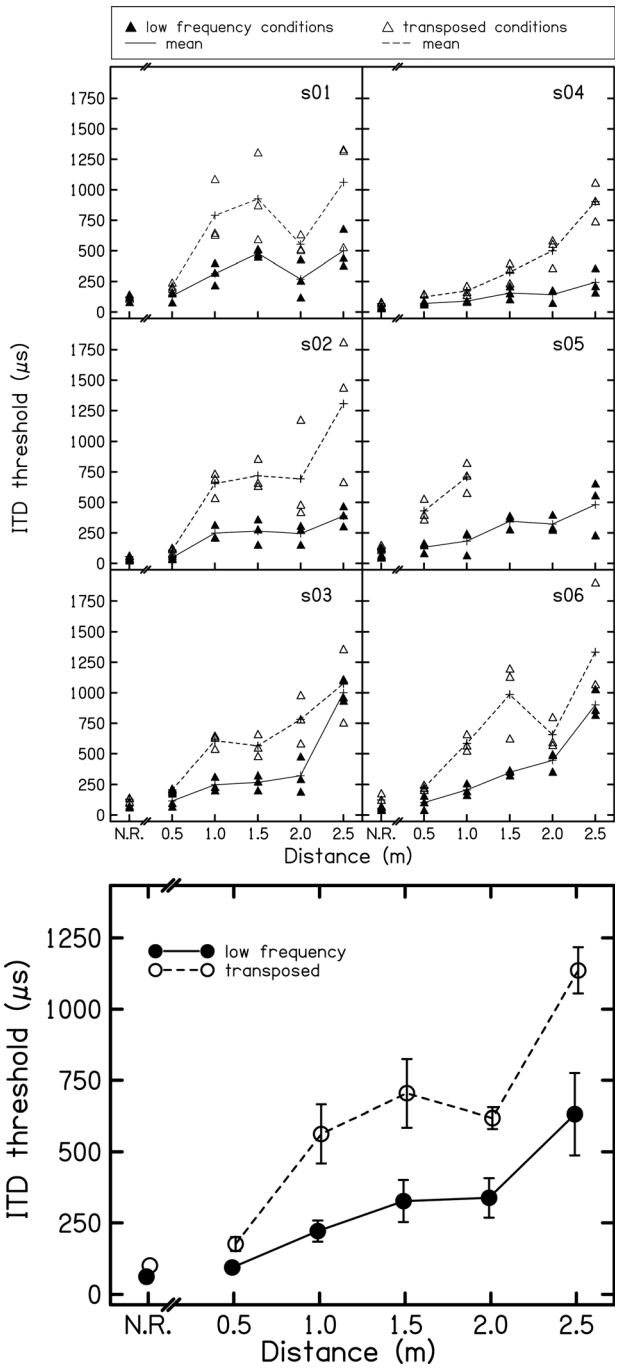

FIG. 1. Individual (top panels) and average (bottom panel) ITD discrimination thresholds from Experiment 1 as a function of the source-receiver distance; the thresholds for the non-reverberant conditions are shown on the left of the panels. The filled symbols and solid lines show the results for the low-frequency conditions and the open symbols and dashed lines show the results for the transposed conditions. In the top panels, the symbols show the three individual threshold estimates acquired for each condition and the lines connect their means; in the bottom panel, the symbols and lines both show the average thresholds. The error bars represent the standard error of the mean (SE).

$$k = \max_{\tau}(\rho_{LR}(\tau)) = \max_{\tau}\left(\frac{\int_{-\infty}^{\infty} S_L(t)S_R(t+\tau)dt}{\sqrt{\int_{-\infty}^{\infty} S_L(t)^2 dt}\sqrt{\int_{-\infty}^{\infty} S_R(t)^2 dt}}\right). \quad (1)$$

For the low-frequency conditions, S_L and S_R are the pressure waveforms of the signals arriving at the left and right ears, respectively, and for the transposed conditions, S_L and S_R are the envelopes of the left- and right-ear signals. The envelopes were extracted by band pass filtering the transposed stimuli with an ERB-wide filter centered at 4 kHz, half-wave rectifying, and low-pass filtering below 2 kHz. Figure 2 shows the data of Fig. 1 re-plotted as a function of interaural coherence. The figure indicates that as coherence decreased as a result of increasing reverberation, the ITD thresholds increased approximately linearly in both the low-frequency and transposed conditions. Therefore using a maximum likelihood procedure we fitted the data with a linear model of the form

$$\begin{aligned}\text{threshold}_i &= \alpha + \beta_1 \times \text{coherence} \\ &\quad + (\beta_2 + \gamma \times \text{coherence}) \times \text{trans} + a_i \\ &\quad + N(0, \sigma_e) \end{aligned} \tag{2}$$

where trans = 0 for the low-frequency and trans = 1 for the transposed conditions. The fixed effects of coherence, β_1, and of stimulus condition, β_2, were included with an interaction term, γ, and a fixed intercept, α. Inter-individual differences were modeled by including a random intercept a_i where i indexes participants with a mean of zero and a standard deviation that was estimated from the data. The last term in Eq. (2), $N(0,\sigma_e)$, is the measurement error, which was modeled as a normal distribution with zero mean and standard deviation σ_e. The solid and dashed lines in Fig. 2 show the model fit for the low-frequency and transposed conditions, respectively.

Despite the fact that the envelope coherence for the transposed stimuli varied over a much smaller range (0.89-1) than the waveform coherence for the low-frequency stimuli (0.66-1), the threshold detriment was much greater for the transposed than for the low-frequency conditions. The maximum likelihood fit confirmed this result; it showed that the effect of coherence was about six times greater for the transposed conditions than for the low-frequency conditions (−7240 versus −1160 μs per unit increase in interaural coherence). Comparison of the model's negative log likelihood with that of the appropriate "null" model (i.e., no dependence on coherence) indicated that the explanatory power of the model was significantly increased by including the effect of coherence only when the effect of stimulus condition, β_2, was also included [$\chi^2(1) = 17.4$, $p < 0.001$]. On its own, coherence was not a good predictor of performance [$\chi^2(1) = 1.12$, $p = 0.291$] because the effect of coherence differed significantly between stimulus conditions [$\chi^2(1) = 27.8$, $p < 0.001$]. The full model accounted for 46% of the variance in the data for the low-frequency conditions, and for 64% in the transposed conditions.

The difference in the effect of coherence between the low-frequency and transposed conditions remained significant [$\chi^2(1) = 13.4$, $p < 0.001$] even when the thresholds were normalized to the thresholds for the non-reverberant condition [coherence = 1.0; Fig. 3(b)]. This is in contrast to the interaction with distance that disappeared after normalization. The effect was about four times greater for the transposed conditions than for the low-frequency conditions (−82.3 versus −21.7 change in normalized threshold per unit increase in coherence). The explanatory power of the full model was similar for the normalized as for the unnormalized thresholds, accounting for 50% of the variance in the data for the low-frequency conditions and 59% for the transposed conditions.

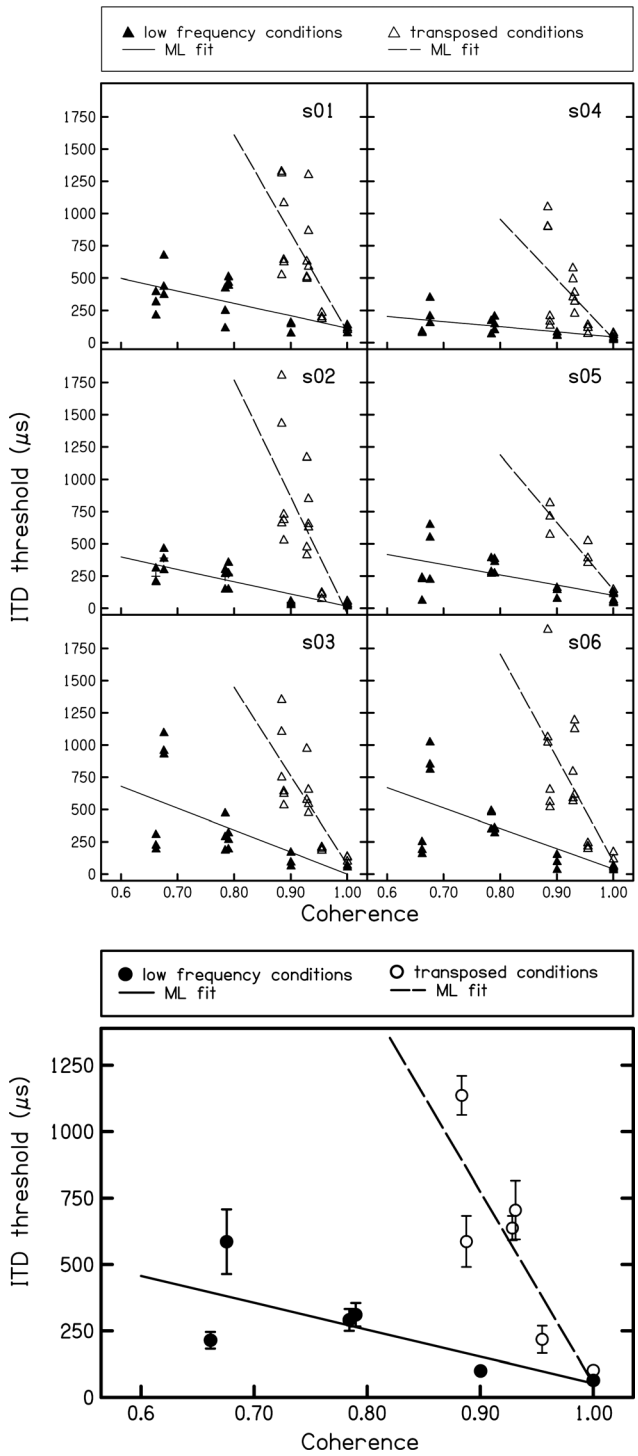

FIG. 2. Individual (top panels) and average (bottom panel) ITD discrimination thresholds from Experiment 1 as in Fig. 1 but re-plotted as a function of the interaural coherence of the stimuli at each source-receiver distance. The data are plotted in a similar way as in Fig. 1, but now the lines show the maximum likelihood linear fits (see text).

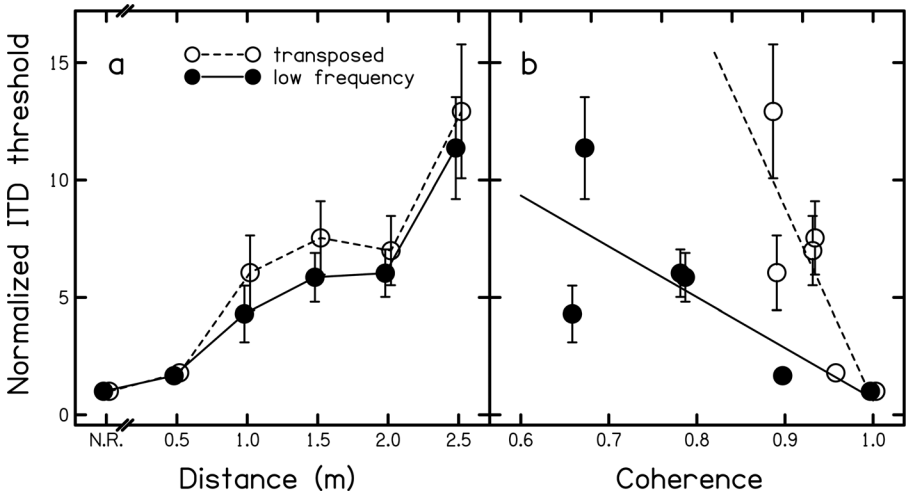

FIG. 3. (a) Average ITD discrimination thresholds from Experiment 1 normalized with respect to threshold for the non-reverberant conditions and plotted in the same way as in Fig. 1. (b) Average ITD discrimination thresholds from Experiment 1 normalized with respect to threshold for the non-reverberant conditions and plotted in the same way as in Fig. 2.

C. Discussion

The results of Experiment 1 suggest that the effect of reverberation on ITD discrimination performance is roughly multiplicative; the ITD discrimination thresholds for both the low-frequency and transposed conditions increased by factor of about 12 between the non-reverberant condition and the largest source-receiver distance used (2.5 m, corresponding to a direct-to-reverberant ratio of −11 dB). Because the non-reverberant threshold for the transposed stimuli was about twice as large as that for the low-frequency stimuli, this meant that the effect of reverberation on the transposed thresholds was much greater in absolute terms than the effect on the low-frequency thresholds. This result is consistent with the psychophysical data of Rakerd and Hartmann (2010) and the electrophysiological data of Devore and Delgutte (2010) that also show that reverberation has a greater effect on ITD sensitivity at high than at low frequencies.

The ITD thresholds were strongly correlated with interaural coherence for both the low-frequency and transposed conditions. However, the size of the effect of coherence on the ITD thresholds was much greater for the transposed than for the low-frequency conditions even when the normalized thresholds were considered. The envelope coherence for the transposed stimuli varied over a much smaller range than the waveform coherence for the low-frequency stimuli. This is consistent with the finding by Ruggles *et al.* (2012) and results from the fact that the envelope is positive, which generates a positive baseline coherence. van de Par and Kohlrausch (1998), showed that, for white noise with a waveform interaural coherence of zero, the envelope coherence never falls below 0.785. The fact that the size of the effect of coherence on the transposed ITD thresholds was nevertheless large may be because under reverberant conditions, such as those simulated here, envelope coherence covaries with other factors, namely, the slope of the envelope flanks and the envelope modulation depth, which would also be expected to influence ITD discrimination thresholds (Bernstein and Trahiotis, 2009; Klein-Hennig *et al.*, 2011). Experiments 2–4 were designed to disentangle the effects of these factors and establish the parametric relation between each factor and envelope ITD discrimination performance.

III. EXPERIMENTS 2–4: EFFECTS OF INTERAURAL COHERENCE AND ENVELOPE SHAPE ON ENVELOPE ITD THRESHOLDS

Experiments 2–4 aimed to investigate the parametric relation between the envelope ITD discrimination threshold and the three envelope properties that would be expected to be most affected by reverberation, namely, interaural coherence, slope and modulation depth. As in the studies of Bernstein and Trahiotis (2009), Klein-Hennig *et al.* (2011), and Laback *et al.* (2011), the stimuli were designed in such a way that would enable us to vary each factor as independently as possible. In particular, we used synthetic, pulsatile stimuli, for which we could control the temporal positions of the pulses to each ear as well as the pulse shape and peak-to-baseline ratio. A detailed description of the stimuli is given in the following text.

A. General methods

1. Participants

Six female participants aged between 24 and 30 yr took part in Experiments 2–4. As for Experiment 1, none of the participants had any history of hearing disorders, and all participants had audiometric thresholds of 20 dB HL or lower. All but two participants had previous experience in psychoacoustic tasks, but only one participant (s04) had also taken part in Experiment 1. All participants but s04 (who was author J.J.M.M.) were paid at an hourly rate.

2. Procedure

Experiments 2–4 measured ITD thresholds using the same task and a similar three-down, one-up adaptive procedure as used in Experiment 1. Experiments 2–4 were conducted prior to Experiment 1 and used a less efficient linear, rather than multiplicative, tracking procedure. Each track started at an initial ITD of 500 μs, and the initial step size was 200 μs. The step size was decreased to 100 and then 50 μs after the first and second reversals in ITD and was then decreased further to 30, 20 and, finally, 10 μs after the fifth, sixth, and seventh ITD reversals. As before, thresholds were

estimated by averaging over 12 further reversals with the final step size, and three threshold estimates were obtained for each condition. To avoid spending an inordinate amount of time on the most difficult conditions, a track was halted and recorded as "failed" if the ITD increased above 1000 μs in four consecutive trials. Because the chance probability of the ITD being increased was 0.5, the probability of erroneously marking a track as failed when the threshold was in fact below 1000 μs was less than 0.0625 (0.5^4). Thresholds from failed tracks were labeled "censored" as opposed to "observed" thresholds from successful tracks. Censored thresholds can be assumed to be above 1000 μs, but their exact value remains unknown.

B. Experiment 2: Effect of envelope interaural coherence

Experiment 1 showed that reverberation-induced reduction in interaural coherence had a much greater effect on the discrimination thresholds for envelope than fine-structure ITDs, but it remains unclear whether the effect was confounded by co-varying changes in envelope shape. The aim of Experiment 2 was to isolate the effect of interaural coherence by using stimuli for which coherence could be changed without concomitant changes in envelope shape.

1. Methods

The envelopes of the stimuli used in Experiment 2 were created by convolving a raised-cosine pulse with a regular train of unit impulses. The raised-cosine pulse had was 1 ms in duration (between the 0-V points). The unit impulses determined the temporal positions of the pulses within the envelope. The envelopes were then multiplied with a 4-kHz diotic carrier tone (Fig. 4). ITDs were applied by delaying the unit impulses by half the ITD in one ear and advancing them by the same amount in the other ear before convolution with the raised-cosine pulse. No ITD was applied to the first and final impulses so that participants had to base their judgments on the ITDs in the ongoing envelope rather at than the stimulus onset or offset. To reduce the envelope interaural coherence, the temporal position of each but the first and last unit impulses was jittered before convolution with the raised-cosine pulse. The amount of jitter was drawn randomly from a uniform distribution, and the jittering was performed independently for each ear. The maximal amount of jitter was limited to half the inter-impulse interval to minimize the degree to which successive pulses would overlap. This also limited the lowest coherence that could be achieved for any given pulse rate. To control for this effect, we conducted the experiment with two different rates, a higher rate of 80 pulse/s (pps) and a lower rate of 20 pps. These rates were chosen based on a preliminary experiment, which measured envelope ITD thresholds as a function of pulse rate. The stimuli in the preliminary experiment were identical to those in the main experiment, apart from the fact that no jitter was applied in this case. The participants and procedure were also the same. Figure 5 shows that the ITD thresholds were lowest for rates from about 50–100 pps, and increased steeply toward lower rates and more shallowly

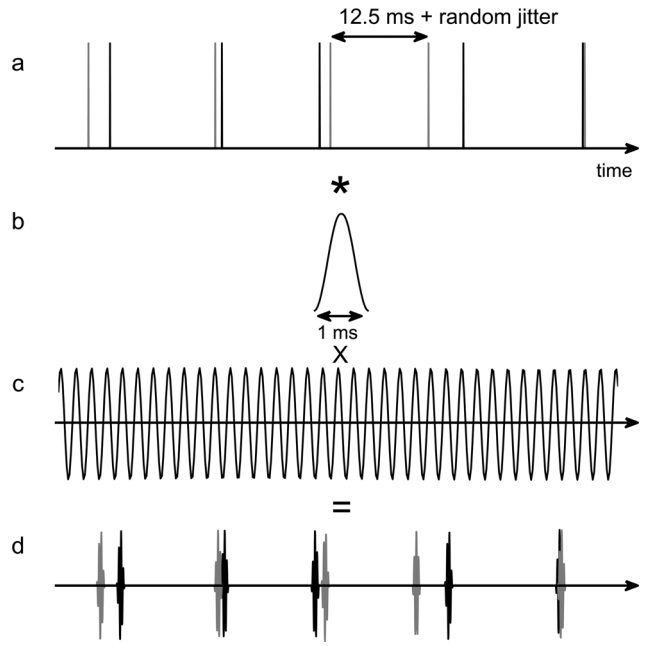

FIG. 4. Schematic explaining the stimulus generation for Experiment 2. Trains of unit impulses with independent random jitter at the left and right ears [gray and black lines in (a)] were convolved with a raised-cosine pulse (b); the resulting pulse train was then multiplied with a 4-kHz diotic carrier tone (c). (d) shows a segment of the resulting stimulus before filtering to a 1 ERB-wide passband around 4 kHz.

toward higher rates. This pattern is consistent with previous results (Hafter and Dye, 1983; Bernstein and Trahiotis, 1994, 2002). The lower pulse rate of 20 pps was chosen as the lowest rate that would still yield a reasonably low ITD threshold (i.e., $\leq\sim$300 μs), and the higher rate of 80 pps was chosen to lie within the minimum of the threshold-rate function.

The stimuli for both the preliminary and main experiments were filtered to an ERB-wide passband centered on 4 kHz using a linear-phase, 512th-order FIR filter. The filtering was performed to ensure that the vast majority of the stimulus energy would be passed by a single auditory filter

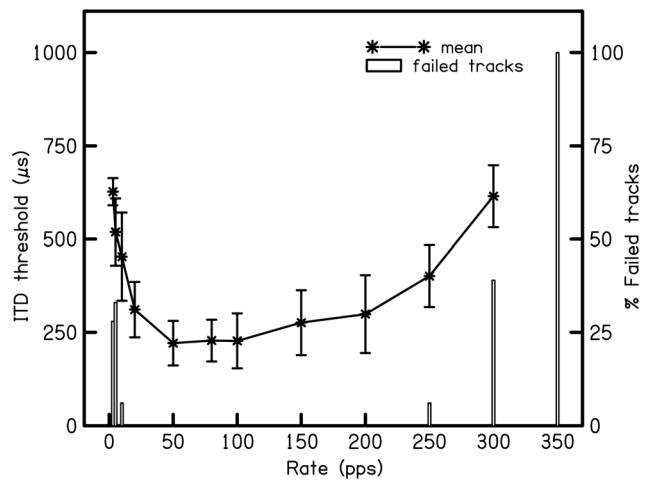

FIG. 5. Average ITD thresholds as a function of pulse rate (preliminary experiment to Experiment 2). The line and symbols show the average observed thresholds, and the vertical bars show the proportion of failed tracks. Error bars show the SE.

and thus minimize the possibility of off-frequency listening and across-frequency integration of ITDs. The band pass filtering lengthened the pulses somewhat as a result of filter ringing. For the higher pulse rate (80 pps), this introduced some degree of overlap between the nearest pulses. In the most extreme cases, the peak-to-trough ratio between successive pulses was reduced to ~25 dB (corresponding to a modulation index of approximately 0.9). However, this occurred only rarely, approximately once in each stimulus of 56 pulses for the lowest coherence value used (0.6; see following text).

The envelope interaural coherence was varied from 0.4 to 1.0 for the condition with the lower pulse rate and from 0.6 to 1.0 for the higher-rate condition. The interaural coherence was varied by changing the maximum amount of jitter. To prevent participants from learning the characteristics of individual stimuli, 100 unique envelopes were stored for each desired coherence value, selected with a tolerance of ±0.1 from a large number of previously generated envelopes. For the envelope selection procedure, coherence was calculated using envelopes extracted from the stimuli after band pass filtering with the ERB-wide filter at 4 kHz. The process for extracting the envelopes was the same as in Experiment 1 (i.e., half-wave rectification and low-pass filtering below 2 kHz).

The stimuli were presented at a sound pressure level (SPL) of 60 dB. The stimulus duration (measured from the onset of the first to the offset of the final pulse) was 700 ms. To mask potentially audible distortion products, an uncorrelated, equally exciting noise was low-pass filtered 2.5 ERBs below the 4-kHz carrier frequency (i.e., at 3003 Hz) and played continuously at an overall level of 50 dB SPL.

2. Results and discussion

Figure 6 shows the individual (upper panels) and average (lower panel) ITD thresholds as a function of envelope interaural coherence. The thresholds for the 80-pps conditions (solid lines and closed symbols) were consistently lower than those for the 20-pps conditions (dotted lines and open symbols); for a coherence of 1.0 (no jitter), the average threshold was 230 μs for 80 pps and 370 μs for 20 pps. For both rates, the observed thresholds increased, and a greater proportion of tracks failed, as the envelope coherence was reduced from 1.0. Figure 6 suggests that the effect of coherence was similar for the two rates with the observed threshold functions being parallel up to the point where they reach an asymptote. Note that the asymptote is misleading as it occurs for conditions with low coherence, where the majority of tracks were failed. The observed thresholds for these conditions represent outliers and are thus misleading.

The individual thresholds showed a considerable degree of inter-individual variability, with the more experienced participants (s04, s08) generally achieving lower thresholds than the less experienced participants. High inter-individual variability in envelope ITD discrimination at high frequencies has been documented previously, for example, by Bernstein *et al.* (1998) for narrowband (400-Hz) noise and by Zhang and Wright (2007) for sinusoidally amplitude

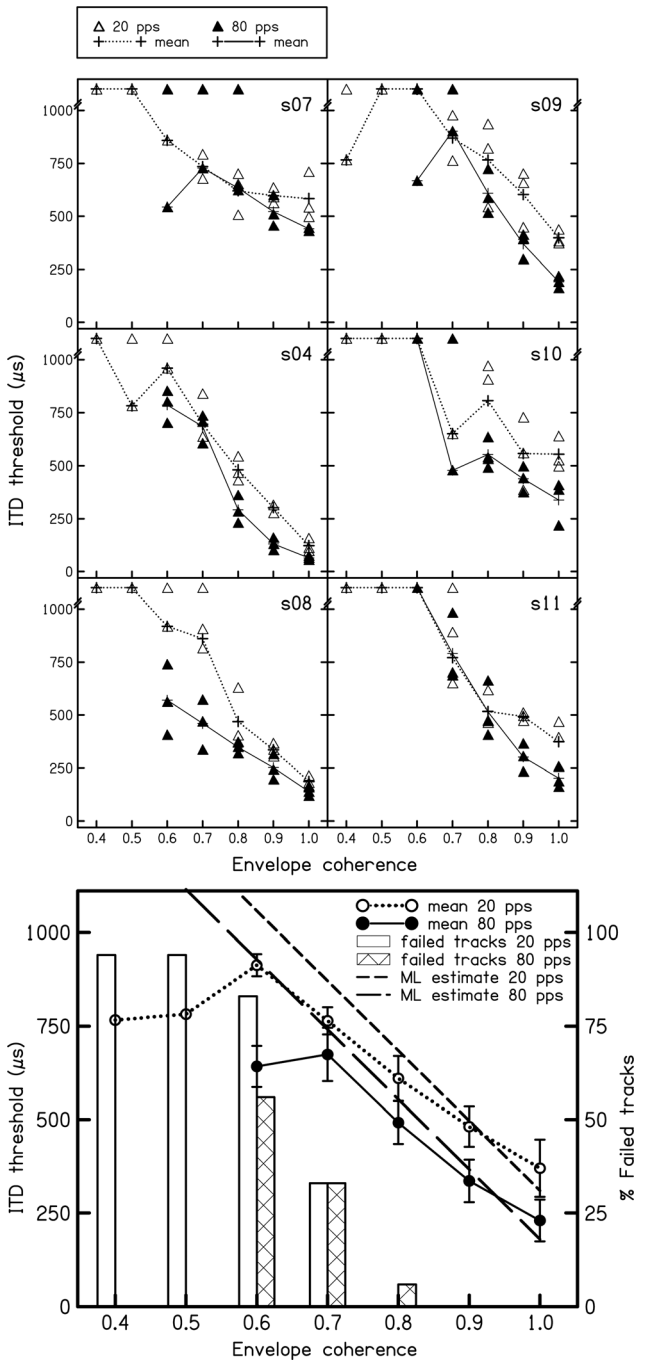

FIG. 6. Individual (top panels) and average (bottom panel) ITD discrimination thresholds from Experiment 2 as a function of the envelope interaural coherence. The filled symbols and solid lines show the observed thresholds for the 80-pps conditions, and the hashed bars show the proportion of failed tracks for these conditions. Similarly, the open symbols and dotted lines show the observed thresholds, and the open bars the proportion of failed tracks, for the 20-pps conditions. As in Fig. 1, the symbols in the top panels show the individual threshold estimates for each condition, and the lines connect their means; in the bottom panel, the symbols and lines both show the average thresholds. The long- and short-dashed lines in the bottom panel show the maximum likelihood linear fits to the 80- and 20-pps conditions, respectively. Error bars show the SE.

modulated (SAM) tones (300-Hz rate). Despite the variability, the effect of interaural coherence was highly consistent across participants.

The large number of failed tracks for the lower coherence values presents a problem for standard statistical analyses

because these analyses are unable to deal with censored thresholds. Nevertheless, censored thresholds contain valid information in that a greater proportion of failed tracks indicates a reduction in task performance. To take full account of the censored thresholds, we fitted the data with a linear mixed model using a modified maximum likelihood procedure (Dempster *et al.*, 1977). This procedure allowed the censored thresholds to be taken into account by contributing to the model likelihood via the cumulative probability of observing thresholds greater than 1000 μs. In contrast, the model likelihood for the observed thresholds was the probability of those exact threshold values being observed. The model was implemented in SAS (SAS Institute, Cary, NC) using modified code from Thiébaut and Jacqmin-Gadda (2004), which involves the "nlmixed" procedure in SAS. The model was fitted to the data for coherence values of 0.6 and greater. The model was specified as

$$\begin{aligned}\text{threshold}_i = {} & \alpha_1 + \beta_1 \times \text{coherence} \\ & + (\alpha_2 + \beta_2 \times \text{coherence}) \times \text{rate} \\ & + a_i + b_i \times \text{coherence} + N(0, \sigma_e), \end{aligned} \quad (3)$$

where rate = 0 for the higher-rate (80-pps) conditions, and rate = 1 for the lower-rate (20-pps) conditions. The model included a fixed intercept (α_1) and fixed effects of rate (α_2) and coherence (β_1) as well as a fixed interaction between rate and coherence (β_2). Inter-individual differences were modeled by a random intercept, a_i, and a random effect of coherence, b_i, where i indexes the participants. The parameters a_i and b_i were described with a two-dimensional normal distribution with zero mean and covariance matrix, Σ, which was estimated from the data. The random measurement error was modeled as a normal distribution, $N(0,\sigma_e)$, with zero mean and variance $\sigma_e{}^2$. The significance of a given model parameter was determined by testing whether inclusion of that parameter in the model significantly increased the model likelihood. Including β_1 and b_i produced a significant increase in the model likelihood [$\chi^2(3) = 239.7$, $p < 0.001$], indicating a significant main effect of coherence. Including α_2 indicated a significant main effect of rate [$\chi^2(1) = 30.4$, $p < 0.001$], but including β_2 showed that the interaction between coherence and rate was non-significant [$\chi^2(1) = 0.2$, p = 0.65]. This confirms that coherence had a similar effect at both rates.

The fitted model was used to generate threshold predictions for each condition. The predicted thresholds were censored in the same way as the measured thresholds and compared with the observed thresholds. This showed that the predicted and observed thresholds were highly correlated ($r^2 = 0.85$), indicating that the assumed linear dependence of the ITD thresholds on coherence was appropriate. The size of the effect of coherence derived from the maximum likelihood fit, β_1, was 1870 μs per unit reduction in coherence. This is less than a third of the effect found in Experiment 1 (7240 μs per unit reduction in coherence). This is consistent with our hypothesis that the large effect of coherence in Experiment 1 was predominantly due to the confounding effect of envelope shape. The following two experiments were aimed to investigate the effects of the two aspects of envelope shape that would be expected to be most affected by reverberation, namely the slope of the envelope flanks and the envelope modulation depth.

C. Experiment 3: Effect of steepness of envelope flanks

Previous studies suggest that the slope of the envelope flanks is an important factor in determining envelope ITD thresholds. Many of the early studies on envelope ITD discrimination have varied the envelope slope as a result of changing other envelope properties, such as the modulation rate or the modulation depth in SAM tones (Henning, 1974). The more recent studies (Bernstein and Trahiotis, 2009; Klein-Hennig *et al.*, 2011), however, have changed the slope independently of other envelope properties. Bernstein and Trahiotis, for instance, have used "raised-sine" envelopes, which are similar to SAM envelopes but raised to a variable exponent in order to vary the slope without changing the modulation rate or depth. However, in both studies, the envelope slope is not constant across the duration of the envelope flanks. In the raised-sine envelopes of Bernstein and Trahiotis, for instance, the slope is zero in the envelope troughs and at the peaks, and is maximal in between. The same is true for the sine-squared envelope flanks used by Klein-Hennig *et al.* (2011). A change in envelope slope over the duration of the flanks makes a parametric analysis of the effects of envelope slope difficult. The current study used trapezoidal envelopes with linear flanks on a logarithmic intensity scale. Such envelopes are characterized by a constant slope in dB per ms. Thus under the assumption that the auditory system represents sound intensity in logarithmic units (see Stevens, 1955), they would be expected to yield a constant rate of change in neural activity. A large range of flank gradients (2.4–30 dB/ms) was used in order to characterize the parametric relation between slope and ITD thresholds.

1. Methods

The procedure for generating the stimuli was similar to that used in Experiment 2. The stimulus envelopes were generated by convolving a pulse with exponential flanks (i.e., linear in logarithmic units) with a regular train of unit impulses (see Fig. 7). The pulse rate was fixed at 20 pps in this case. The slope of the flanks of the exponential pulse was varied by changing the flank duration (measured from the −60-dB point) from 2 to 25 ms. Twenty-five milliseconds was the maximum value that could be used before successive pulses started to overlap. The steady-state portion of the pulse was adjusted to keep the overall duration of the pulse at half maximum (i.e., between the −6-dB points) constant at 5 ms. This meant that the pulse energy was also roughly constant across conditions.

As in Experiment 2, the ITDs were applied to the envelopes, and the envelopes were multiplied with a 4-kHz diotic carrier tone. No ITD was applied to the first and last pulses. As before, the stimuli were filtered to an ERB-wide passband centered on 4 kHz, and the stimulus duration was 700 ms. The stimulus presentation level varied somewhat with flank duration, ranging from 63 dB SPL for the 2-ms flanks to 61 dB SPL for the 25-ms flanks.

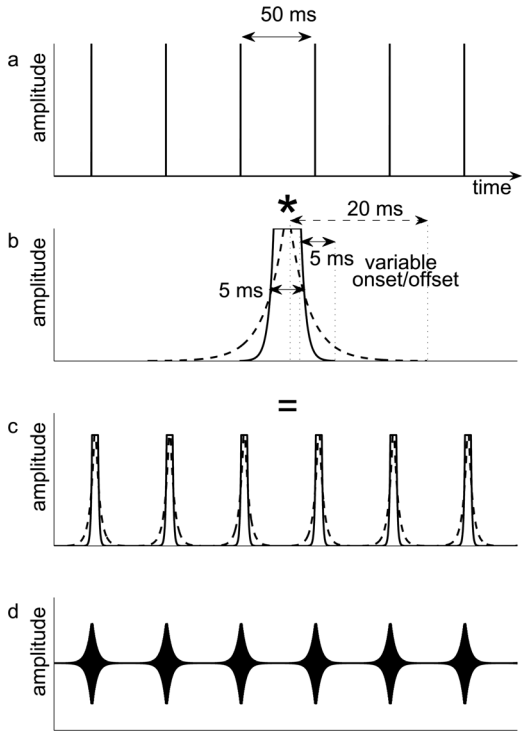

FIG. 7. Schematic explaining the stimulus generation for Experiment 3. A regular train of unit impulses (a) was convolved with an exponential pulse with variable flank duration to vary the envelope slope [see dashed and solid lines in (b)]; the resulting pulse train (c) was then multiplied with a 4-kHz diotic carrier tone to produce the stimulus [shown before filtering to a 1 ERB-wide passband around 4 kHz in (d)].

2. Results and discussion

Figure 8 shows that both the individual (upper panels) and the average (middle panel) ITD thresholds decreased with increasing envelope slope. The average threshold was 678 μs for the shallowest slope tested (2.4 dB/ms). It first decreased sharply with increasing slope, but eventually reached an asymptote for slopes between 6 and 12 dB/ms. The asymptotic ITD threshold at the steepest slope tested (30 dB/ms) was 245 μs.

To fit the thresholds with a linear mixed model, they were re-plotted as a function of the flank duration because this made the threshold functions appear approximately linear (see bottom panel in Fig. 8). For the exponential pulses used here, the flank duration is inversely related to the slope. The linear relation between the thresholds and the flank duration thus suggests that the relation between the thresholds and the slope is hyperbolic. As before, the fit was obtained using a maximum-likelihood procedure. The model was specified as

$$\text{threshold}_i = \alpha + \beta \times \text{slope} + a_i + b_i \times \text{slope} + N(0, \sigma_\varepsilon). \quad (4)$$

This is similar to the model used to fit the data from Experiment 1 [Eq. (2)] except that β and b_i are now the fixed and random effects of the flank duration rather than the coherence. Comparison with the appropriate null model showed that the effect of flank duration was significant [$\chi^2(3) = 137.5$, $p < 0.001$]. The model fitted the observed thresholds very

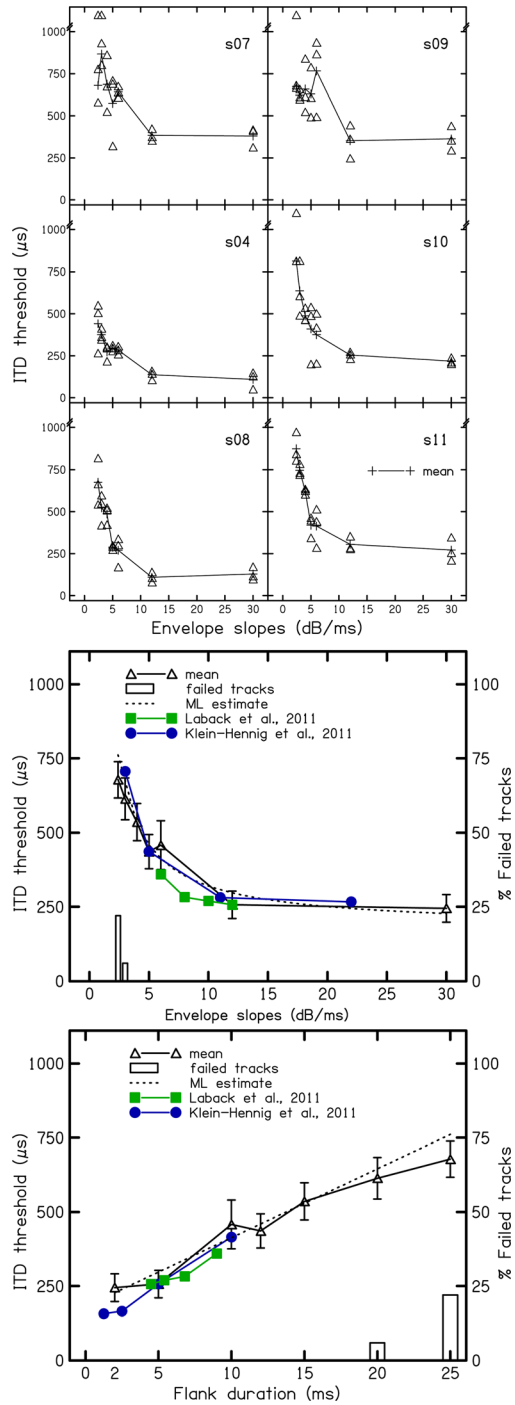

FIG. 8. (Color online) Individual (top panels) and average (middle panel) ITD discrimination thresholds from Experiment 3 as a function of the slope of the envelope flanks in dB/ms. The bottom panel shows the average thresholds re-plotted as a function of the flank duration. In the top panels, the symbols show the individual observed threshold estimates for each condition, and the lines connect their means. In the middle and bottom panels, the open triangles and solid lines show the average observed thresholds; the hashed bars show the proportion of failed tracks. The squares and circles (connected by lines) show data from Laback *et al.* (2011) and Klein-Hennig *et al.* (2011), respectively. To facilitate comparison, the data from Laback *et al.* were normalized to coincide with the current average thresholds for the 12-dB/ms slope (middle panel) or the 5-ms flank duration (bottom panel); the data from Klein-Henning *et al.* were normalized to coincide with the current average thresholds for the 5-dB/ms slope (middle panel) or the 5-ms flank duration (bottom panel). In the middle panel, Klein-Hennig *et al.* data were plotted as a function of the slope in dB/ms at the -6-dB point in their envelopes. The dotted lines in the middle and bottom panels show the maximum likelihood linear fit to the current data. Error bars show the SE.

well, explaining almost 90% of their variance ($r^2 = 0.89$). It indicated that the ITD thresholds increased by approximately 25 μs per millisecond increase in the flank duration.

The current data are in qualitative agreement with the data of Bernstein and Trahiotis (2009) in that they too found a decrease in ITD threshold with increasing envelope slope. However, the complicated nature of their stimuli (they increased the slope by increasing the exponent of their raised-sine stimuli) makes a quantitative comparison in term of dB/ms difficult. Klein-Hennig *et al.* (2011) also found a decrease in the ITD thresholds with increasing slope, and when their thresholds are plotted as a function of flank duration as for the current data, the relation also appears linear (see circles in bottom panel of Fig. 8). Under the assumption that their thresholds were determined by the slope at a particular (e.g., the steepest) point within their envelope flanks, their thresholds would also appear to be hyperbolically related to the slope steepness in dB/ms, consistent with the current data (circles in middle panel of Fig. 8). As in the current study, the stimuli of Laback *et al.* (2011) were trapezoidal pulses with flanks which were linear on a logarithmic intensity scale. Our data are consistent with theirs (squares in Fig. 8) in that they too found that the thresholds decreased with increasing slope for the steeper slopes and then reached an asymptote. The asymptote occurred for slopes steeper than 8 dB/ms which is in the same range of slope steepness as in our data (between 6 and 12 dB/ms). Taken together, Klein-Hennig *et al.*, Laback *et al.*, and the current data all suggest a hyperbolic relation between envelope ITD thresholds and envelope slope in dB/ms.

D. Experiment 4: Effect of modulation depth

In previous studies, the effect of modulation depth on envelope ITD discrimination has often been investigated by changing the modulation index, m, in SAM tones (e.g., Henning, 1974). SAM tones with different values of m differ not only in the ratio of the envelope maxima and minima but also in the envelope slope with deeper modulations leading to steeper slopes. In order to avoid this confound, we used the same stimuli as in Experiment 3 with exponential pulses, and we varied the modulation depth by filling in the gaps between the pulses with a constant DC offset (Fig. 9). This leaves the slope of the pulse flanks unchanged. Two different values of the slope were employed to investigate whether there is an interaction between the slope and the modulation depth.

1. Methods

The stimuli were generated by scaling a similar exponential pulse as used in Experiment 3 to peak amplitude of $1 + m$, convolving the pulse with a regular train of unit impulses, and then imposing a DC offset of $1 - m$ on (Fig. 9; m is analogous to the modulation index in SAM tones). The resulting envelopes had a maximum amplitude of $1 + m$ and a minimum amplitude of $1 - m$ and a constant flank steepness. As in Experiment 3, the pulse rate was 20 pps, and pulse duration at half maximum was set to 5 ms. The exponential pulse was generated with two different slopes of 2.4 and 30 dB/s (corresponding to flank durations of 25 and 2 ms). As in Experiment 3, ITDs were applied only to the envelopes, before multiplication with a 4-kHz diotic carrier tone, and zero ITD was applied to the first and last pulses. The stimulus duration was 700 ms. The stimulus level varied somewhat with modulation depth. For the 2.4-dB/ms flanks, the level ranged from 64 to 67 dB SPL and from 66 to 69 dB SPL for the 30-dB/ms flanks. As before, the stimuli were filtered to an ERB-wide passband centered on 4 kHz.

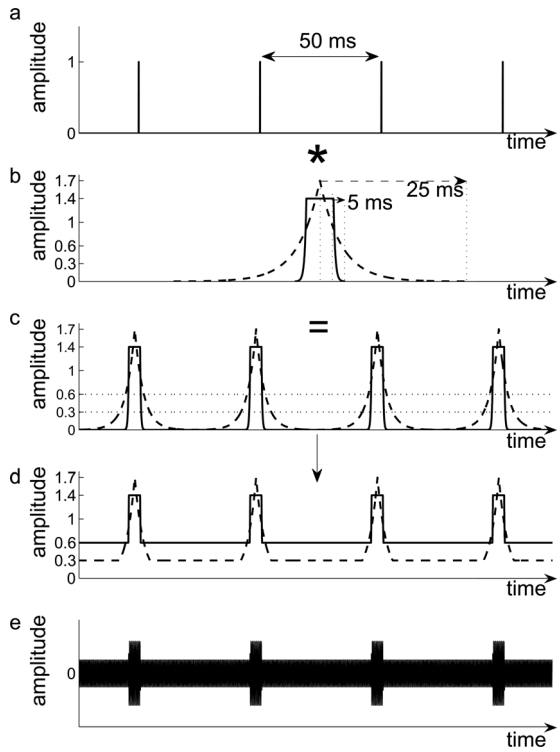

FIG. 9. Schematic explaining the stimulus generation for Experiment 4. A regular train of unit impulses (a) was convolved with an exponential pulse with two different flank durations [dashed and solid lines in (b)]. The interpulse gaps in the resulting pulse train (c) were then filled in with a DC offset (d) and multiplied with a 4-kHz diotic carrier tone to produce the stimulus [shown before filtering to a 1 ERB-wide passband around 4 kHz in (e)].

2. Results and discussion

Figure 10 shows that the ITD thresholds remained roughly constant for modulation indices down to between 0.4 and 0.5. The thresholds and the number of failed runs then increased sharply for lower modulation indices. The asymptotic threshold at high modulation depths was considerably lower for the steeper than for the shallower envelope flanks used (279 μs, on average, for the 30-dB/ms flanks compared to 540 μs for the 2.4-dB/ms flanks). Toward lower modulation depths (i.e., for modulation indices below ~0.4), the thresholds for the two slope conditions appeared to converge.

As for Experiment 2, the data were fitted with a mixed model using a maximum-likelihood procedure to take account of the failed tracks. To capture the different patterns of the data at high and at low modulation depths, a piecewise linear model was used. The model consisted of two linear sections, one for the asymptotic threshold range at high modulation depths and one for the sharp increase in the thresholds towards lower modulation depths. The break point

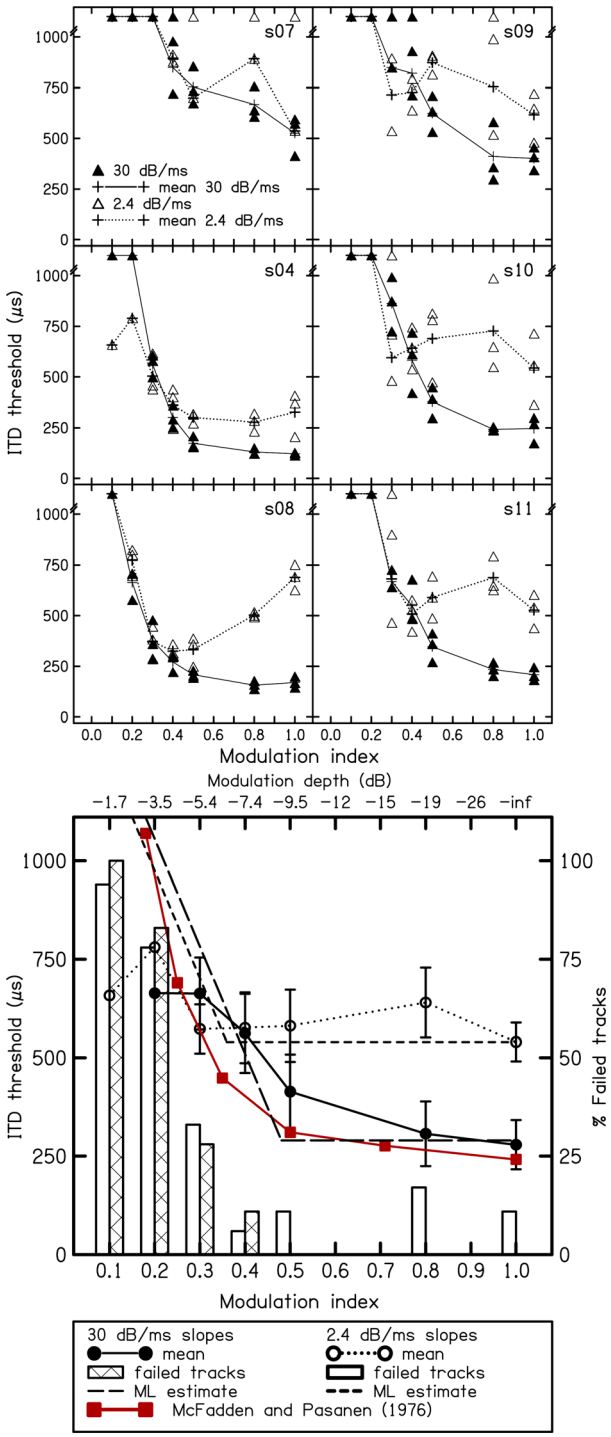

FIG. 10. (Color online) Individual (top panels) and average (bottom panel) ITD discrimination thresholds from Experiment 4 as a function of the modulation index. The filled symbols and solid lines show the observed thresholds for the steeper (30-dB/ms) slope conditions, and the hashed bars show the proportion of failed tracks for these conditions. Similarly, the open symbols and dotted lines show the observed thresholds, and the open bars the proportion of failed tracks, for the shallower (2.4-dB/ms) slope conditions. The symbols in the top panels show the individual threshold estimates for each condition, and the lines connect their means; in the bottom panel, the symbols and lines both show the average thresholds. The long- and short-dashed lines in the bottom panel show the maximum-likelihood piecewise-linear fits to the steep and shallow slope conditions, respectively. The squares and lines in the bottom panel show the thresholds from participant LR from McFadden and Pasanen (1976). The data were normalized to coincide with the average maximum likelihood threshold for $m = 0.18$ to facilitate comparison. Error bars show the SE.

between the two sections was a free parameter and was allowed to vary independently for each slope condition. The model for each linear section was of a similar form as that used for Experiment 2 with fixed and random effects of modulation depth and slope and a fixed interaction between these two factors [see Eq. (3)]. Additional constraints were imposed by the necessity for the two sections to coincide at the break point.

As for Experiment 2, the significance of a given model parameter was tested by assessing the influence of that parameter on the model likelihood. This confirmed that the main effects of modulation depth and slope were both significant [modulation depth: $\chi^2(3) = 132$, $p < 0.001$; flank steepness: $\chi^2(3) = 103$, $p < 0.001$]. The model likelihood was not significantly improved by including an interaction between slope condition and modulation depth, either below [$\chi^2(1) = 0.2$, $p = 0.6547$] or above the break points [$\chi^2(1) = 1.8$, $p = 0.1797$] and so these terms were not included in the final model.

The final model produced an excellent fit to the observed thresholds, explaining 85% of their variance. The fit generated a break point at $m = 0.48$ (9.1 dB) for the steeper (30-dB/ms) flanks, and at $m = 0.36$ (6.5 dB) for the shallower (2.4-dB/ms) flanks. For modulation indices above the break points, the effect of modulation depth was not significantly different from zero [$t(4) = 0.02$, $p = 0.9826$]. This is consistent with the impression that the thresholds reached an asymptote at high modulation depths. The difference between the asymptotic threshold levels for the two slopes was significant [$t(4) = 12.23$, $p < 0.001$]. For modulation indices below the break points, the effect of slope was nonsignificant [$F(4,1) = 1.77$, $p = 0.25$]. This is consistent with the impression that the threshold functions for the two slopes converge at low modulation depths.

The relation between the ITD thresholds and the modulation depth observed in the current study is similar to that found in two previous studies (McFadden and Pasanen, 1976; Nuetzel and Hafter, 1981). This is despite the fact that these studies did not vary modulation depth independently of slope as is the case in the current study. In both previous studies, the ITD thresholds also reached an asymptote at high modulation depths and increased sharply toward lower modulation depths. Re-plotting McFadden and Pasanen's data as a function of modulation index (squares in Fig. 10) suggests that our piecewise linear model does not fully capture the threshold behavior for the middle range of modulation indices. Both McFadden and Pasanen's and Nuetzel and Hafter's data suggest that a power-law model may have been more appropriate to fit the current data. Unfortunately, this alternative model failed to converge, likely because the data had insufficient power.

Taken together, the current and the previous data suggest that the effect of modulation depth decreases towards high modulation depths but is strong at low modulation depths; so strong, in fact, that the effect overrides other factors, such as the effect of slope as tested in the current study. This dichotomous characteristic might be explained by the change in the synchrony of neuronal firing as a function of modulation depth that has been observed in physiological studies (reviewed in Joris *et al.*, 2004). These studies have shown that neural firing in the inferior colliculus (IC) in

response to amplitude modulated stimuli follows the shape of the modulating envelope at low modulation depths but becomes more tightly time-locked to a particular point within the modulation cycle at higher modulation depths.

Contrary to our results, Bernstein and Trahiotis (2009) did find an effect of flank steepness at low modulation depths. They increased flank steepness by increasing the exponent in their raised-sine stimuli. Unlike the current study, Bernstein and Trahiotis did not restrict the spectral width of their stimuli by band pass filtering. The spectral width of raised-sine stimuli increases with increasing exponent. This would be expected to make them more conducive to off-frequency listening. Bernstein and Trahiotis (2010) showed that raised-sine stimuli can produce a higher modulation depth in off- than on-frequency channels and suggested that off-frequency listening might explain the apparent sensitivity to flank steepness at low modulation depths in their 2009 study.

IV. GENERAL DISCUSSION

The aim of this study was to investigate the effect of reverberation on ITD discrimination. The results from Experiment 1 suggest that the effect is multiplicative rather than additive; the discrimination thresholds for both fine-structure ITDs at low frequencies and envelope ITDs at high frequencies increased by a factor of about 12 as the amount of reverberation was increased from no reverberation to a direct-to-reverberant ratio of −11 dB. Given that the envelope ITD threshold for no reverberation was almost twice as large as that of the fine-structure ITD threshold, this meant that, in absolute terms, the increase in the envelope ITD thresholds owing to reverberation was much greater than the increase in the fine-structure ITD thresholds. This is consistent with the results of Rakerd and Hartmann (2010), who also found a greater effect of reverberation on ITD discrimination at high than at low frequencies. Rakerd and Hartmann used free-field stimuli, so the amount of reverberation may have been different at high and low frequencies. The current study used transposition to ensure that the amount of reverberation was the same in the two frequency regions tested. A model fit to the data from Experiment 1 showed that the increase in the ITD thresholds with increasing reverberation was strongly correlated with the concomitant reduction in interaural coherence. Interestingly, the effect of coherence was much greater for the envelope than for the fine-structure ITD thresholds; per unit reduction in interaural coherence, the envelope thresholds increased by ∼7000 μs, compared to only ∼1000 μs for the fine-structure thresholds. The results of Experiment 2 indicated that this was due to the envelope ITD thresholds being affected by other, co-varying, factors, specifically, the reverberation-induced degradation in envelope shape. In Experiment 2, interaural coherence was varied independently of the envelope shape, and this yielded a much smaller effect of coherence (less than 2000 μs per unit reduction in coherence). Reverberation degrades the envelope shape by making the flanks in the envelope shallower and by filling in the dips, thereby reducing the envelope modulation depth. Experiments 3 and 4 showed that both of these effects have a strong influence on the envelope ITD thresholds. Experiment 3 showed that when the flanks are linear on a logarithmic scale, the ITD thresholds increase approximately linearly with the flank duration (i.e., the inverse of the slope), with a 25-μs increase in ITD threshold for every millisecond increase in flank duration. Experiment 4 showed that the effect of slope is overridden by a strong effect of modulation depth when the modulation depth is low. The effect of modulation depth decreases toward higher modulation depths, and the thresholds reach an asymptote.

It is well-known that envelope ITDs are a weaker sound localization cue than fine-structure ITDs (e.g., Wightman and Kistler, 1992; Macpherson and Middlebrooks, 2002). The current results indicate that reverberation makes envelope ITDs even less useful as a localization cue; even for a moderately reverberant room, and a moderate distance between the source and the receiver, the average discrimination threshold for envelope ITDs rose above 1000 μs, which is well beyond the physiological range of ITDs (i.e., the time taken for sound to traverse the head). This suggests that in reverberant environments, envelope ITDs would be unlikely to mediate much more than a mere impression of left or right. This is consistent with the findings of Rakerd and Hartmann (2010) and runs counter to the hypothesis proposed by Ruggles *et al.* (2012) that envelope ITDs might become relatively *more* useful in reverberant environments. It remains possible, however, that envelope ITDs might be more reliable when the stimulus contains information across many auditory channels. Many naturally occurring stimuli, such as speech, cover a wide range of frequencies, and the envelopes within different frequency channels tend to be comodulated. Such stimuli provide redundant ITD information across many channels and promote across-frequency integration of ITD information (Wenzel, 1976; Hartmann, 1983).

The current results have implications for bilateral CIs because most current devices only convey envelope ITDs and ILDs. If anything, envelope ITD discrimination performance would be expected to be worse in electric than in acoustic hearing owing to other factors, such as mismatched electrode positions between the ears and reduced neural survival. This conjecture is consistent with the results of Laback *et al.* (2011), who found considerably larger envelope ITD thresholds in CI than in normal-hearing listeners under comparable conditions. This may be why CI users rely mainly on ILDs for sound localization (van Hoesel and Tyler, 2003; Seeber and Fastl, 2008). Laback *et al.* found that despite the large absolute difference in ITD discrimination performance, the relative effects of some aspects of envelope shape were remarkably similar between CI and normal-hearing listeners. This suggests that enhancing the envelope shape by using appropriate signal processing strategies might help to improve envelope ITD perception in CI users. The results of Experiment 3 indicate that algorithms designed to alter the envelope shape, so as to enhance the transmission of ITD cues, should ensure that the envelope flanks are as steep as possible. The results of Experiment 4 indicate that such algorithms should also ensure the modulation depth lies in the range over which ITD thresholds are fairly constant. For our narrowband stimuli, this required a modulation depth of 9 dB or larger, although one might

expect this to depend on bandwidth and rate. Our results suggest that any improvements in envelope ITD discrimination in non-reverberant conditions would be amplified in reverberant conditions. Envelope enhancement might thus help to make envelope ITDs a more useful localization cue to CI users. Various monaural envelope-enhancing strategies have been proposed to improve the perception of speech and music in CI users. Green *et al*. (2004) encoded information about the voice pitch information using very steep envelope modulations at the fundamental frequency (F0). The benefit of enhancing F0 information could be increased by synchronizing steep temporal modulation across channels such as in the algorithms of Milczynski *et al*. (2009) and Vandali and van Hoesel (2011). Strategies to improve envelope ITD perception (e.g., Seeber and Monaghan, 2012) might be similar to these monaural envelope-enhancing strategies but would first have to identify those portions of the stimulus that are dominated by the direct sound.

ACKNOWLEDGMENTS

We thank Oliver Zobay for help with the statistical analysis and David Moore and Michael Akeroyd for their valuable comments on an earlier version of the manuscript. We would also like to thank three anonymous reviewers for their help in improving the manuscript. This work was funded by the Intramural Programme of the UK Medical Research Council (MRC U135097132).

Bernstein, L. R., and Trahiotis, C. (**1994**). "Detection of interaural delay in high-frequency sinusoidally amplitude-modulated tones, two-tone complexes, and bands of noise," J. Acoust. Soc. Am. **95**, 3561–3567.

Bernstein, L. R., and Trahiotis, C. (**2002**). "Enhancing sensitivity to interaural delays at high frequencies by using 'transposed stimuli,' " J. Acoust. Soc. Am. **112**, 1026–1036.

Bernstein, L. R., and Trahiotis, C. (**2009**). "How sensitivity to ongoing interaural temporal disparities is affected by manipulations of temporal features of the envelopes of high-frequency stimuli," J. Acoust. Soc. Am. **125**, 3234–3242.

Bernstein, L. R., and Trahiotis, C. (**2010**). "Accounting quantitatively for sensitivity to envelope-based interaural temporal disparities at high frequencies," J. Acoust. Soc. Am. **128**, 1224–1234.

Bernstein, L. R., Trahiotis, C., and Hyde, E. L. (**1998**). "Inter-individual differences in binaural detection of low-frequency or high-frequency tonal signals masked by narrow-band or broadband noise," J. Acoust. Soc. Am. **103**, 2069–2078.

Blauert, J., Bruggen, M., Hartung, K., Bronkhorst, A., Drullmann, R., Reynaud, G., Pellieux, L., Krebber, W., and Sotteck, R. (**1998**). "The Audis catalog of human HRTFs," in *16th International Congress on Acoustics* (New York).

Dempster, A. P., Laird, N. M., and Rubin, D. B. (**1977**). "Maximum likelihood from incomplete data via the EM algorithm," J. R. Stat. Soc. Ser. B (Methodol.) **39**, 1–38.

Devore, S., and Delgutte, B. (**2010**). "Effects of reverberation on the directional sensitivity of auditory neurons across the tonotopic axis: Influences of interaural time and level differences," J. Neurosci. **30**, 7826–7837.

Glasberg, B. R., and Moore, B. C. J. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hear. Res. **47**, 103–138.

Green, T., Faulkner, A., and Rosen, S. (**2004**). "Enhancing temporal cues to voice pitch in continuous interleaved sampling cochlear implants," J. Acoust. Soc. Am. **116**, 2298–2310.

Hafter, E. R., and Dye, R. H. (**1983**). "Detection of interaural differences of intensity in trains of high-frequency clicks as a function of interclick interval and number," J. Acoust. Soc. Am. **73**, 644–651.

Hartmann, W. M. (**1983**). "Localization of sound in rooms," J. Acoust. Soc. Am. **74**, 1380–1391.

Henning, G. B. (**1974**). "Detectability of interaural delay in high-frequency complex waveforms," J. Acoust. Soc. Am. **55**, 84–90.

Houtgast, T., and Steeneken, H. J. M. (**1973**). "Modulation transfer-function in room acoustics as a predictor of speech intelligibility," Acustica **28**, 66–73.

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.

Joris, P. X., Schreiner, C. E., and Rees, A. (**2004**). "Neural processing of amplitude-modulated sounds," Physiol. Rev. **84**, 541–577.

Klein-Hennig, M., Dietz, M., Hohmann, V., and Ewert, S. D. (**2011**). "The influence of different segments of the ongoing envelope on sensitivity to interaural time delays," J. Acoust. Soc. Am. **129**, 3856–3872.

Kolarik, A. J., and Culling, J. F. (**2009**). "The masking of interaural delays," J. Acoust. Soc. Am. **125**, 2162–2171.

Laback, B., Zimmermann, I., Majdak, P., Baumgartner, W. D., and Pok, S. M. (**2011**). "Effects of envelope shape on interaural envelope delay sensitivity in acoustic and electric hearing," J. Acoust. Soc. Am. **130**, 1515–1529.

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

Macpherson, E. A., and Middlebrooks, J. C. (**2002**). "Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited," J. Acoust. Soc. Am. **111**, 2219–2236.

McFadden, D., and Pasanen, E. G. (**1976**). "Lateralization of high frequencies based on interaural time differences," J. Acoust. Soc. Am. **59**, 634–639.

Milczynski, M., Wouters, J., and van Wieringen, A. (**2009**). "Improved fundamental frequency coding in cochlear implant signal processing," J. Acoust. Soc. Am. **125**, 2260–2271.

Nuetzel, J. M., and Hafter, E. R. (**1981**). "Discrimination of interaural delays in complex waveforms: Spectral effects," J. Acoust. Soc. Am. **69**, 1112–1118.

Rakerd, B., and Hartmann, W. M. (**2010**). "Localization of sound in rooms. V. Binaural coherence and human sensitivity to interaural time differences in noise," J. Acoust. Soc. Am. **128**, 3052–3063.

Ruggles, D., Bharadwaj, H., and Shinn-Cunningham, B. G. (**2012**). "Why middle-aged listeners have trouble hearing in everyday settings," Curr. Biol. **22**, 1417–1422.

Seeber, B. U., and Fastl, H. (**2008**). "Localization cues with bilateral cochlear implants," J. Acoust. Soc. Am. **123**, 1030–1042.

Seeber, B. U., Kerber, S., and Hafter, E. R. (**2010**). "A system to simulate and reproduce audio-visual environments for spatial hearing research," Hear. Res. **260**, 1–10.

Seeber, B. U., and Monaghan, J. J. M. (**2012**). "Improving hearing with cochlear implants in reverberant spaces," in *Frontiers in Computational Neuroscience. Conference Abstract: Bernstein Conference 2012*. doi: 10.3389/conf.fncom.2012.55.00034.

Stevens, S. S. (**1955**). "The measurement of loudness," J. Acoust. Soc. Am. **27**, 815–829.

Thiébaut, R., and Jacqmin-Gadda, H. (**2004**). "Mixed models for longitudinal left-censored repeated measures," Comput. Methods Programs Biomed. **74**, 255–260.

Vandali, A. E., and van Hoesel, R. J. (**2011**). "Development of a temporal fundamental frequency coding strategy for cochlear implants," J. Acoust. Soc. Am. **129**, 4023–4036.

van de Par, S., and Kohlrausch, A. (**1997**). "A new approach to comparing binaural masking level differences at low and high frequencies," J. Acoust. Soc. Am. **101**, 1671–1680.

van de Par, S., and Kohlrausch, A. (**1998**). "Analytical expressions for the envelope correlation of narrow-band stimuli used in CMR and BMLD research," J. Acoust. Soc. Am. **103**, 3605–3620.

van Hoesel, R. J., and Tyler, R. S. (**2003**). "Speech perception, localization, and lateralization with bilateral cochlear implants," J. Acoust. Soc. Am. **113**, 1617–1630.

Wenzel, E. M. (**1976**). "Lateralization of high-frequency clicks based on interaural time: Additivity of information across frequency" (University of California, Berkeley), pp. 1–154.

Wiggins, I. M., and Seeber, B. U. (**2011**). "Dynamic-range compression affects the lateral position of sounds," J. Acoust. Soc. Am. **130**, 3939–3953.

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.

Wilson, B. S., and Dorman, M. F. (**2008**). "Cochlear implants: Current designs and future possibilities," J. Rehabil. Res. Dev. **45**, 695–730.

Zhang, Y., and Wright, B. A. (**2007**). "Similar patterns of learning and performance variability for human discrimination of interaural time differences at high and low frequencies," J. Acoust. Soc. Am. **121**, 2207–2216.

Zimmer, U., and Macaluso, E. (**2005**). "High binaural coherence determines successful sound localization and increased activity in posterior auditory areas," Neuron **47**, 893–905.