Lateralization of large interaural delays

Jennifer E. Mossop and John F. Culling
University Laboratory of Physiology, Parks Road, Oxford OX1 3PT, United Kingdom

(Received 11 April 1997; revised 7 May 1998; accepted 28 May 1998)

Two experiments explored the limits of listeners’ abilities to interpret large interaural time delays (ITDs) in terms of laterality. In experiment 1, just-noticeable differences (jnd’s) were measured, using an adaptive procedure, for various reference ITDs of Gaussian noise between 0 and 3000 μs. The jnd’s increased gradually with reference ITD for reference ITDs between 0 μs and 700 μs, then rose sharply to plateau at much higher jnd’s for the remainder of the standard ITDs tested (1000–3000 μs). The second experiment tested left/right discrimination of Gaussian noise that was interaurally delayed up to 10 000 μs, and high-pass filtered to cutoff frequencies between 0 Hz (broadband) and 3000 Hz. There was good discrimination (62%; significantly above chance, p < 0.05) for broadband and 500-Hz high-pass cutoff stimuli for all ITDs up to 10 000 μs, and for ITDs up to at least 3000 μs for higher high-pass cutoff frequencies. These results indicate that laterality cues are discriminable at much larger ITDs than are experienced in free-field listening, even in the absence of energy below 3 kHz. © 1998 Acoustical Society of America.

PACS numbers: 43.66.Ba, 43.66.Qp, 43.66.Pn [RHD]

INTRODUCTION

The size of the human head creates maximum interaural time delays (ITDs) of approximately 600 μs, depending on the frequency components of the sound (Kuhn, 1977). ITDs of broadband noise stimuli within this range and up to about 1 ms, presented through headphones, produce images that are described as fused and lateralized to some intracranial position (for a review, see Durlach and Colburn, 1978). As ITDs are increased beyond the physiological range, the sound image of continuous broadband noise remains at the leading ear but gradually becomes more diffuse until the image either appears at both ears or fills the head entirely and the perception of laterality is abolished (Blodgett et al., 1956). It is well known that the perception of laterality persists for ITDs much greater than those experienced ecologically, but there is relatively little data on the limits of this ability. It is important to establish such limits as they may give insights into binaural mechanisms. They may, for instance, have particular relevance to estimates of the range of lengths of the delay lines which are widely believed to underlie binaural processing (Jeffress, 1948; Colburn, 1996).

From a variety of studies, Durlach and Colburn (1978) estimate that this loss of laterality occurs, for continuous wide-band signals, between 10 and 30 ms of ITD. In a more recent review Colburn (1996, p. 340) suggests a figure of about 15 ms. Blodgett et al. (1956) used an adjustable head on a magnetic tape recorder to create large noise delays and asked listeners to reduce the ITD until the noise could be lateralized. Delays up to about 13 ms were perceived as directional for broadband noise. Blodgett et al. observed that this estimate of the maximum lateralizable delay was probably conservative, since the listeners were invariably correct in their lateralization judgment once they had reduced the delay to one that they perceived as lateralized. In a second experiment they measured a mean threshold of 14.0 ms for low-frequency bands of noise (106–212 Hz). Higher-frequency bands of noise (2400–4800 Hz) had large thresholds at large delays (mean 8.0 ms). At larger ITDs beyond 15–20 ms or so, the percept is of a diffuse noise, indistinguishable from uncorrelated noise, but at even longer ITDs (on the order of hundreds of milliseconds) the presence of an ITD is detectable once again as a subtle echo percept with no lateralization (Warren et al., 1981).

Some studies have reported just noticeable difference (jnd) measurements for ITDs against reference ITDs other than 0 μs. For 500-Hz pure-tone stimuli, Domnitz (1973) found consistently low jnd’s (9 μs) for 0-, 200-, and 400-μs reference ITDs, and higher jnd’s (26 μs) at 700-μs reference ITD. Domnitz and Colburn (1977) found similar trends with more intersubject variability in a virtually identical experiment which also included 1000-μs reference ITDs, at which jnd’s were, on average, 62 μs. Hershkowitz and Durlach (1969), also using a 500-Hz pure tone, found jnd’s which were less than 20 μs for reference ITDs of 200 and 400 μs, but that, above 600-μs reference ITD, subjects could not perform a jnd task reliably and reported changes in subjective cues. These changes included cue reversals, where stimuli with a larger time lead in the left ear produced a percept to the right of a stimulus that had a smaller time lead in the left ear. These cue reversals may have been associated with ambiguity about which ear receives the leading signal for pure tones with large ITDs. Hershkowitz and Durlach noted, however, that for three of four subjects the reversals occurred at a reference ITD of only 750 μs, somewhat less than the half-period of the signal (1000 μs). Kohenke et al. (1995) have reported jnd’s for 1/3-octave bandpass noise centered around 500 Hz and 4000 Hz at 0, 300, and 600 μs of ITD. They found that jnd’s for reference ITDs of noise with a center frequency of 500 Hz were between 20 and 40 μs for reference ITDs of 0 and ±300 μs and around 70 μs for ±600-μs reference ITDs. For stimuli with a center frequency of 4000 Hz, jnd’s were higher; around 60 μs for 0
and ±300 μs and near 250 μs for ±600-μs reference ITDs.

Using sinusoidal stimuli at large ITDs creates phase ambiguity when the ITD is equivalent to half a period of the sinusoid. This ambiguity can be removed by using broadband noise stimuli, rather than tones or narrow-band noise. However, no experiments to our knowledge have systematically studied the ability to perceive laterality of broadband noise at increasingly large ITDs, or have extensively examined jnd’s for ITDs greater than 0 μs. Here, we present two experiments designed to obtain objective measurements of the limits of human performance in discriminating differences between large ITDs.

Experiment 1 examines jnd’s in relation to reference ITDs from 0 to 3000 μs with the purpose of exploring the transition from ITDs experienced in everyday life (±600 μs) to larger, supra-ecological ITDs. Experiment 2 involves left/right discriminations of noise with even larger ITDs (from 0.5 to 10 ms) and a range of high-pass cutoffs (from broadband to 3000-Hz high-pass cutoff). The second experiment was designed to examine the extent to which a lateralizable perception is evoked by large ITDs, and also the effect of eliminating low-frequency components upon this lateralization ability. The high-pass filtered stimuli were used in order to explore the importance of low frequencies to the correct lateralization of large ITDs, by determining whether large ITDs could still be lateralized when the low frequencies were absent.

I. EXPERIMENT 1

The purpose of experiment 1 was to determine jnd’s for ITD at a range of ITDs from 0 to 3000 μs.

A. Subjects

Four adult subjects including the first author participated in experiment 1. Three had prior experience in binaural experiments. The subjects had no known hearing deficits.

B. Stimuli

The stimuli were broadband 8192-sample Gaussian noises, generated digitally (20-KHz sampling rate; 16-bit quantization) using a TDT II AP2 array processor, and duplicated for stereo channels. One channel was then transformed into the frequency domain, delayed by a given amount (by modifying the phase spectrum), and transformed back into the time domain. Application of a delay in the frequency domain meant that the delays did not change the onset time of the noise and that the delays could be specified precisely rather than being quantized to the nearest sampling period. Channels were gated on and off simultaneously with a 50-ms rise/fall time, for a total stimulus duration of 409.6 ms, including the rise/fall ramps. Thus each stimulus had an ongoing ITD, but not an onset or offset ITD. All ITDs had the left channel lagging the right. The stimuli were converted to analog waveforms and amplified (using the Tucker-Davis DD1 digital-to-analog converter, FT5-9 reconstruction filter with 10-KHz cutoff, PA4 digitally controlled attenuators and HB6 headphone amplifier), before being fed into a double-walled sound-attenuating chamber, and presented to a subject through Sennheiser HD414 headphones. Sound level output was calibrated using a Bruel & Kjaer artificial ear and 4130 half-inch microphone. Sounds were presented at equal A-weighted levels to each ear at 76 dB.

C. Procedure

A two-down/one-up adaptive method was used to determine jnd’s (Levitt, 1971). We employed a three-interval, two-alternative forced-choice task where the first of the three intervals was a noise with a set ITD (the ‘‘reference’’ stimulus). The second and third intervals were, in random order, the reference ITD again and noise with a larger ITD (the ‘‘test’’ stimulus). All noise bursts were generated from fresh noise samples.

Subjects indicated, by button press on a computer terminal whether the second or the third interval of the set appeared to be located in a different intracranial position than the first interval. The adaptive procedure adjusted the difference in ITD between the test stimulus and the reference stimuli on a logarithmic scale, as recommended by Saberi (1995), using a step size of 12.2% increase/decrease in ITD, until 14 reversals had taken place. The jnd values were determined from the geometric average of the microsecond values of the final ten reversals in direction. This procedure estimates the value of ITD yielding 70.7% correct performance.

Subjects attended a series of one-hour sessions during which they usually performed one adaptive run at each of 16 reference ITDs that ranged from 0 μs to 3000 μs. Subjects were trained until they were judged to be giving consistent performance across trials at each reference ITD value. This training took between 3 and 12 sessions depending on the subject. Sometimes during training, different subsets of the stimuli were interleaved to give subjects more practice on stimulus conditions that they found more difficult. At reference ITDs greater than about 800 μs, subjects required more training in order to achieve consistency of performance comparable to that obtained at lower reference ITDs. Once relatively consistent performance was achieved for all reference ITD magnitudes, 8 adaptive thresholds at each reference ITD were taken as the results.

D. Results

Figure 1 shows the mean jnd’s from 8 recorded thresholds with standard error bars for the 4 individual subjects at each of 16 reference ITDs. The reference ITDs tested were 0–800 μs in 100-μs steps, and 1000, 1200, 1400, 1600, 2000, 2500, and 3000 μs.

The magnitudes of the jnd values in Fig. 1 vary from subject to subject, with subject JM notably lower than the rest at large reference ITDs. However, the shape of each curve is similar across subjects. All curves begin, at 0-μs reference ITD, with low jnd values ranging from 12.3 μs (subject MT) to 62.2 μs (subject EH). From 0 to 700 μs, jnds became gradually larger for each subject, but remain generally low, usually below 100 μs. Between 700 μs and 1000 μs, all subjects show a sharp increase in jnd. Above about 1000 μs, the jnd’s become more variable, and do not con-
tener training, may account for the differences between the ITDs. They did not report listener training or feedback, but anticipated that there may be a shift in the available cues, as the authors discussed in II D. This is unsurprising considering listener's everyday experience.

The results for the nonecological ITD stimuli (>700 μs) presented in experiment 1 show that subjects can discriminate differences in the perceived position of sounds with supra-ecological ITDs to some extent. In this experiment, jnd’s increased by about an order of magnitude for ITDs greater than 700 μs.

II. EXPERIMENT 2

Experiment 2 was designed to test whether subjects could assign a sidedness (left or right) to broadband signals with various high-pass cutoff frequencies that had large ITDs of up to 10 ms.

The just noticeable difference (jnd) at zero reference ITD are higher than Klumpp and Eady’s (1956) measurement of 10 μs for 75% correct detection of ITD of broadband noise. The present study used a relatively small amount of training, involved 409.6-ms stimuli which had no interaural onset-time differences, employed no response feedback, and used a 70.7% jnd criterion (Levitt, 1971). Klumpp and Eady used longer duration stimuli (about 1700 ms) which had onset as well as ongoing ITDs. They did not report listener training or feedback, but the differences between the stimuli used, and perhaps in listener training, may account for the differences between the present jnd’s at 0 μs and Klumpp and Eady’s values. In experiment 1, the smallest thresholds obtained were more typical of naive, than of well-trained listeners (Colburn, 1996).

The reference ITD at which jnd’s increase for all subjects (about 700 μs) corresponds approximately to the ITD at which sounds exceed the ecological range (about ±600 μs). This is unsurprising considering listener’s everyday experience.

Listeners’ performance at supra-ecological ITDs was also examined here. In designing this experiment, it was anticipated that there may be a shift in the available cues, as the ITD increased. At some point beyond the ecological range, it was anticipated that the lateral position cue may be superseded by a diffuseness cue. As ITD increases from zero, the image would be perceived further and further to the right, but with even larger ITDs the image would also become more diffuse, and perhaps spread back towards the midline. Thus for large reference ITDs, the target interval may be discriminable on the basis of having a more diffuse image than the two reference intervals. However, as noted in Sec. II D, only one listener occasionally reported using cues which seemed consistent with this expectation. All four listeners described the dominant percept as being intracranial position.

The experimental design does not distinguish directly between the cues (intracranial position or diffuseness) that the listeners were actually using. Although the listener’s comments were obtained informally, they give some insight into the possible cues that they were using. The dominance of an intracranial location cue over a diffuseness cue has some precedent in the work of Jeffress et al. (1962) and of Pollack and Trittipoe (1959). Jeffress et al. measured the precision with which listeners could adjust the ITD of a partially correlated noise in order to center it within the head (i.e., find zero ITD using purely auditory feedback). They found that listeners could center the noise almost as accurately when it had a correlation of only 0.2 as when it was perfectly correlated. Thus it appears that the binaural system is very sensitive to small movements in an auditory image even when that image is very diffuse. On the other hand, Pollack and Trittipoe (1959) showed that listeners are poor at discriminating the interaural correlation of two noises when both have small interaural correlations (see Durlach et al., 1986, Table III). To correctly discriminate (75% accuracy) a second noise from a reference noise with a cross correlation of 0.2, their listeners required an increment in interaural correlation of 0.38. For higher reference correlations listeners were much more sensitive. For a reference correlation of 0.93 listeners only needed an increment of 0.06 to achieve 75% accuracy. Hence, listeners have relatively coarse abilities to discriminate differences between diffuse noises. At the large interaural delays used in experiment 1, the image of the reference noise is already somewhat diffuse, perhaps explaining why listeners reported continued use of intracranial position rather than diffuseness of the sound image at large ITDs.

The results for the nonecological ITD stimuli (>700 μs) presented in experiment 1 show that subjects can discriminate differences in the perceived position of sounds with supra-ecological ITDs to some extent. In this experiment, jnd’s increased by about an order of magnitude for ITDs greater than 700 μs.
A. Subjects

Five subjects participated; the four subjects from experiment 1 and one other subject with normal hearing and previous experience in binaural experiments.

B. Stimuli

The stimuli were Gaussian noises generated digitally using WAVE software (Culling, 1996). As in experiment 1, the noises were filtered in the frequency domain. In this experiment, the amplitude spectrum was also manipulated to create a range of high-pass cutoffs. The high-pass cutoff frequencies used were 0 (broadband), 500, 1000, 1500, 2000, 2500, and 3000 Hz. The noises were then duplicated for two channels. One channel was interaurally delayed or advanced by between 500 µs and 10 ms (in 500-µs steps) in order to create an ongoing ITD. The two channels were gated simultaneously with a rise/fall time of 50 ms. Each stimulus was 409.6 ms in total duration.

The delayed and filtered noise was added to a 24-dB-less-intense interaurally uncorrelated broadband noise (0–10 kHz) which was used to prevent listeners from using any low-frequency lateralization cues that may be reintroduced by distortion products or by spread of excitation.

The complete set of 280 individual stimuli consisted of one exemplar of each of 140 conditions delayed to the left and right sides. The conditions were 7 high-pass cutoff frequencies×20 ITDs. The A-weighted stimuli were presented at between 76 and 79 dB, depending on the particular high-pass cutoff. The stimulus set was periodically regenerated in order to prevent the characteristics of individual noise bursts from influencing the results.

C. Procedure

The task was single-interval forced choice, where the subjects heard each of the 280 stimuli in a random sequence. After each stimulus, they were asked to choose by button press on a computer terminal whether the sound was located on the left or the right of the midline. No feedback was provided.

Subjects attended eighteen half-hour sessions in which they did four blocks of the complete stimulus set, for a total of 72 responses for each stimulus condition. Subjects received no training in this task. Their responses at the first of 18 sessions were the very similar to the last.

D. Results

Figure 2 shows results matrices for each of the five subjects from experiment 2 in the form of a grid, where ITD is represented on the ordinate axis and high-pass cutoff frequency on the abscissa. The gray scale shade at each vertex represents percent correct left/right discrimination, where each vertex is one high-pass cutoff frequency and ITD combination. Black is 92%–100% correct, dark gray 82%–92%, medium gray 72%–82%, light gray 62%–72% and white below 62% (not significantly above chance).

The most accurate subject (JM) was able to correctly discriminate left from right above chance as indicated by the lightest gray band, for all ITDs tested, up to and including 10 ms. At higher cutoff frequencies, JM discriminated correctly to at least 5.0 ms. Subjects EH, AM, and GH performed significantly above chance for broadband stimuli to 10 ms. MT was able to correctly discriminate up to 8 ms for broadband and 500-Hz cutoff stimuli.

All subjects were able to perform above chance at all cutoff frequencies to at least 3.5-ms delay.

E. Discussion

The data from experiment 2 confirm and extend those of Blodgett et al. (1952). Listeners are able to lateralize broadband noises at very large ITDs, at least 15 times larger than those which are experienced in free-field listening. The results also support Blodgett et al.’s finding that a greater range of delays can be lateralized when low frequencies are
The function has the general form of a reciprocal relationship, which may be explained in at least three different ways.

The maximum lateralizable ITD might be related to the period of the cutoff frequency. However, at each cutoff frequency the maximum lateralizable ITD is many times this period of the cutoff frequency. However, listeners are still able to make location-based judgments at supra-ecological ITDs, but have much higher thresholds, indicating that there is at least some degree of perceptible spatial information in these large ITDs. Experiment 2 shows that even larger ITDs up to at least 10 ms, especially with the presence of low-frequency information, are perceived as having some spatial location (left versus right).

These data, showing that listeners are able to lateralize large ITDs, may have implications for centrally weighted models of sound lateralization like those of Stern and Colburn (1978), Shackleton et al. (1992), and Stern et al. (1988), which all emphasize ITDs near zero. However, the fact that listeners are able to lateralize large ITDs and discriminate one large ITD from another does not directly imply that the system contains delay lines which apply delays of magnitudes comparable to the range over which lateralization and discrimination can be measured, since a number of other factors come into play. These factors include peripheral nonlinearities, the specificity with which coincidence detectors respond to stimuli with one particular ITD and no other (especially stimuli with longer ITDs), the way in which lateralization information is integrated across frequency channels, and the way that the information is then transformed into a lateral position percept. It is probably the case that any single set of data may be explained by manipulating different parameters of the same model. The present study provides two new forms of data which may provide useful constraints on such models.

ACKNOWLEDGMENTS

We would like to thank Dennis Phillips, whose supervisory role in a previous project led directly to this one. Steve Colburn and Trevor Shackleton and two anonymous reviewers provided helpful comments on an earlier version of this manuscript. This work was supported by a Natural Sciences and Engineering Research Council (Canada) postgraduate scholarship, an Oxford McDonnell-Pew award, and an Overseas Research Studentship to JFC, and a Medical Research Council postdoctoral fellowship to JFC.

1 Since the analysis buffer was completely filled with noise, the waveform wrapped around when a delay was applied, so that the tail end of the waveform appeared at the front, preventing a delay in the onset time. Generating noise in the time domain and then filtering it in the frequency domain is considerably more computationally efficient than generating arrays of Rayleigh-distributed amplitude values and rectangularly distributed phase values and then generating and summing sine waves with those parameters. In a sense, a Fourier transform of Gaussian noise is an efficient means of generating an appropriate set of amplitude and phase values for Gaussian noise, while an inverse Fourier transform is an efficient way of generating and summing sine waves.


