



Published in final edited form as:

J Acoust Soc Am. 2007 March ; 121(3): EL103–EL109.

Interaural time difference processing of broadband and narrow-band noise by inexperienced listeners

William A. Yost, Raymond H. Dye, and Stanley Sheft

Parmly Hearing Institute, Loyola University Chicago, 6525 North Sheridan Road, Chicago, Illinois 60626

Abstract

Recent functional magnetic resonance imaging (fMRI) data might be interpreted as being in disagreement with existing psychophysical data regarding the laterality of broadband noise stimuli presented with large interaural time differences (ITDs). This study investigated the possibility that lateral judgments made by inexperienced listeners who did not receive feedback might be different than those reported for experienced listeners, especially when the ITD is longer than that occurring in nature, and therefore data from inexperienced listeners presented unnaturally long ITDs for the first time might be more consistent with the possible interpretation of the fMRI results. The results from this study using inexperienced listeners were not basically different from those reported in the literature based on experienced listeners, suggesting a possible difference does exist between inferences drawn from fMRI data and human psychophysical results.

1. Introduction

Aspects of the classical coincidence-detection network originally proposed by Jeffress (1948) for processing interaural time differences (ITD) have recently been challenged based on observations of ITD processing in the inferior colliculus (IC). McAlpine and colleagues (see McAlpine *et al.*, 2001) suggest that a Jeffress-style coincidence network does not appear to exist in the IC of many animals (e.g., small mammals) and they have proposed an optimal population-based scheme (Harper and McAlpine, 2004) for processing ITD by all animals including humans. Thompson *et al.* (2006) recently published fMRI data recorded from regions of the IC in humans that appear consistent with observations of ITD processing in the IC of small mammals but could be interpreted as being inconsistent with published human psychophysical data, which can be accounted for by a Jeffress-style model. Thompson *et al.* (2006) suggest that models other than a Jeffress-style coincidence network might be required to account for ITD processing in humans and many other animals.

Thompson *et al.* (2006) used conditions in which relatively broadband noise (400-Hz wide) was presented with a whole-waveform ITD that was less than ($500 \mu\text{s}$) or greater than ($1500 \mu\text{s}$) a half period ($\pi=1000 \mu\text{s}$) of the noise's center frequency (CF=500 Hz). For $\text{ITD} < \pi$ of CF the IC on the side of the head opposite the ear receiving the first sound is activated, while for $\text{ITD} > \pi$ of CF the opposite IC is activated. If activation of the IC in each hemisphere is important in determining the laterality of images produced by varying ITD, this might suggest that broadband signals with small ITDs would be lateralized on one side of the head and those with large ITDs (relative to π at CF) on the other side. In human psychophysical experiments using similar broadband stimuli (see Trahiotis and Stern, 1989) the data suggest that broadband

signals are always lateralized on the side of the head of the ear receiving the leading sound even for very large ITDs. This observation can be accounted for by a Jeffress-style model designed to account for psychoacoustical data based on interaural cross correlation as the mathematical operation of a coincidence network and other assumptions (see Trahiotis *et al.*, 2001), but appears inconsistent with the functional interpretation of the IC recordings of Thompson *et al.* (2006) as explained above. Jeffress-style cross-correlation models and IC recordings and their functional interpretation are compatible with the psychophysical data showing that narrow-band stimuli with $ITD < \pi$ are located toward the leading ear, but for $ITD > \pi$ the sounds are lateralized on the opposite side of the head from the leading ear (see Trahiotis and Stern, 1989; Yost, 1981). Thus, for narrow-band noise, the interpretation of the recordings from the IC might be consistent with the psychophysical data, but there is a possible lack of consistency for broadband noise.

The large ITDs used in the Thompson *et al.* (2006) study and in many psychophysical experiments do not occur in nature as they are considerably larger than the ITD based on the diameter of a human head and the speed of sound (between 650 and 800 μ s; see Kuhn, 1987). It has been shown that ITD processing can be variable in that the lateral location of images based on ITD (e.g., Yost, 1981) and ITD discrimination thresholds (McFadden *et al.*, 1973) exhibit large individual differences, and for some conditions it takes practice before asymptotic ITD performance is reached (see Wright and Fitzgerald, 2001). In addition, one of the arguments concerning ITD processing in the IC is that it appears limited to ITDs near π at CF, that is, usually relatively small ITDs. Sounds presented with ITDs over headphones can vary in the perceived lateral location and in the diffuseness of the lateralized image (see Blauret, 1997). It is possible that lateral images produced with large ITDs are more diffuse than those produced with smaller ITDs and, as a consequence, untrained listeners might have more difficulty than experienced listeners in assigning a lateral location to images produced with large ITDs. Finally, the subjects in the fMRI experiments had limited experience with long ITDs. This paper investigates the possibility that untrained/inexperienced listeners performing without feedback might indicate a different pattern of results for unnaturally large ITDs than that typically obtained for trained listeners who are provided feedback. That is, data from untrained listeners experiencing unnaturally long ITDs for the first time might be more consistent with the interpretation of IC results from Thompson *et al.* (2006) than the data from studies involving highly trained listeners who receive feedback of one sort or another regarding ITD processing, especially regarding long ITDs.

2. Methods

Four groups of 19–25 undergraduate students were used. Each group performed a different set of conditions for about an hour without receiving any training or feedback, and several conditions were randomized within a block of trials. These procedures helped ensure that listeners were not trained to respond with any particular bias or pattern.

Three groups (200-, 400-, 800-Hz CF groups) participated in three tasks: left-right lateralization task-L-R, lateral movement task-MOVE, and diffuse/punctate judgment task-D-P. ITD and bandwidth were randomized within each block of trials (see Table 1 for conditions). In the L-R task, listeners indicated whether the noise appeared on the left or right side of the head. In the MOVE task, the first interval of each trial contained a stimulus with an ITD followed 200 ms later in the second interval with a stimulus with a larger ITD (except in one condition when the larger ITD was followed by a smaller ITD). In all cases the listener indicated if the sound image appeared to move left or right between the first and second interval. In the P-D task listeners were to indicate if the image was punctate or diffuse. At the beginning of each block of trials, the listeners were presented samples of diotic noises as examples of a punctate image

and interaurally uncorrelated noises as examples of a diffuse image, with the concept of a diffuse image explained before data collection. No feedback was provided for any task.

The fourth group (MIXED) performed a left-right lateralization task (L-R task) but with ITD, bandwidth, and CF (200, 400, and 800 Hz) all mixed within a block of trials (see Table 2 for conditions). Only two bandwidths were used (10 and 400 Hz, based on observations from the other three groups that performance appeared constant with fixed rather than proportional bandwidth). Again, no feedback was provided.

All listeners were also administered a screening test in which a 12 000-Hz low-pass noise was presented eight times monotically to either the left or right ear and they indicated if the sound was left or right. The data from listeners with reported hearing loss, recent significant exposure to loud sounds, or who did not perform with 87.5% or better accuracy in the screening task were not used (Tables 1 and 2 indicate the number of listeners out of 25 whose data were used). Every block began with a 1-min presentation of a 12 000-Hz low-pass noise and listeners were asked to adjust their headphones so that an image appeared in the middle of their heads, and to not move their headphones during a block of trials once a centered image was obtained.

All noises were generated in the frequency domain with Rayleigh distributed amplitudes and uniform and randomly determined phases, transformed to the time domain using a 32 768-point inverse FFT, presented at 32 768 points per second via a GINA sound card, and then low-pass filtered at 12 000 Hz. All stimuli were 200 ms in duration, shaped with 20-ms \cos^2 rise-fall times, and presented at an overall level of 65 dB SPL. All ITDs were waveform ITDs in that the right ear was always presented the signal first. The sounds were presented over Sennheiser headphones (HD 280 Pro) to listeners seated in a sound-proof room.

3. Results

Figures 1 and 2 indicate the data for the L-R and MOVE tasks for the 400-Hz CF group, while Fig. 3 shows the data from the MIXED group. None of the data from the P-D task are shown in that the only variable that affected average punctate/diffuse judgments for all three CF groups was noise bandwidth; the broader the bandwidth the higher the proportion of “diffuse” judgments. ITD had no influence on the average punctate/diffuse judgments for all bandwidths and CFs. The P-D protocol may be insensitive to evaluating the diffuseness of images produced with different ITDs. The data from the 200- and 800-Hz CF groups are very similar to those from the 400-Hz group in that all of the trends seen in Figs. 1 and 2 are found in the data for the 200- and 800-Hz CF groups.

Figure 1 displays the data for the L-R task as mean percent “Right” judgments as a function of ITD/IPD *re:CF* at π (IPD is the interaural phase difference computed based on CF). The bold vertical black line in the approximate middle of the figure indicates an ITD/IPD approximately equivalent to π . The results in Fig. 1 and for each CF group are similar in that for ITD $< \pi$, percent “Right” judgments for all bandwidths and CFs increased toward 100% as ITD increased. For broader bandwidths the proportion of “Right” judgments tended to be near 100% for large ITDs, whereas for narrow bandwidths large ITDs $> \pi$ produce more “Left” judgments (% “Right” < 0.5), with the proportion of “Left” judgments increasing (% “Right” decreasing) as ITD/IPD approached 2π .

Figure 2 shows the MOVE data plotted as the percent “Right Move” judgments versus the two ITD/IPD values presented in the two successive observation intervals of each trial. For all CF groups when the ITDs for both stimuli were less than π , the listeners indicated that the image moved right (high proportion of “Right Move” judgments) as the ITD increased from the first to the second observation interval. For the broadband noise the percent “Move Right” judgments were greater than 50% for all conditions in which the second stimulus contained

the larger ITD and less than 50% when the second stimulus contained the smaller ITD (far right data in Fig. 2). This is consistent with the data in Fig. 1 (and those for the other two CF groups) in that for the broadband stimuli the images were always on the side of the head to which ITD led (right side) and larger ITDs produce images more lateralized toward the leading ear than smaller ITDs.

A different pattern of results exists for the narrow-band noise for all CF groups. Whether the listener reported the image as moving right or left depended on the relationship of the ITD to the value obtained for π at CF. When the smaller ITD $< \pi$ and the larger ITD $> \pi$, listeners indicated that the image moved left (% “Right Move” $< 50\%$). This is also consistent with the data from Fig. 1 in that images for narrow-band stimuli for ITD $< \pi$ were on the right and those for ITD $> \pi$ were on the left. Thus, when a stimulus with an ITD $< \pi$ was followed by one with ITD $> \pi$, the image would be predicted to move left (not right) as the data in Fig. 2 suggest. The trends for both narrow-band and broadband noises for all CF groups are far less evident for the conditions in which the ITD is very large (the next to the far right data in Fig. 2). In these cases for both the narrow-band and broadband cases for all CF groups the percent “Move Right” judgments are near 50%.

In the MIXED condition, Fig. 3, the left-right judgments follow the same trends as shown in Fig. 1 (and the other CF groups). When the ITD was greater than π (the blue circled data) the images were always judged to be right when the bandwidth was broad, but left when it was narrow, even when the ITD was greater than 2π (data within green ellipse). Thus, randomizing CF within a block of trials (MIXED group) did not have a major effect on left-right judgments relative to keeping CF fixed within a block of trials (CF groups).

Two audio examples are provided for listening over stereo headphones:

Mm. 1. Audio 1 file (180 kB). This is a file of type “wav”.

Mm. 2. Audio 2 file (180 kB). This is a file of type “wav”.

4. Discussion

All of the main effects of the data from Figs. 1–3 and the other CF groups are consistent with past psychophysical data. That is, use of untrained listeners who were not provided feedback and were presented high uncertainty stimulus conditions produces lateralization results for processing narrow-band and broadband stimuli that are similar to the data in the literature obtained with trained listeners in minimal uncertainty tasks with feedback. As stated in the Introduction, if there is a switch in the activation of one IC to the other as ITD is increased beyond π at CF for broadband stimuli, then one might assume broadband stimuli would be perceived on different sides of the head, depending on the relative magnitude of the ITD. The current data along with those from other studies are not entirely consistent with such an interpretation of these IC results. Thus, listening experience, feedback, and condition complexity are probably not relevant to teasing out these relationships between physiology and perception.

There are, however, indications in the data suggesting that laterality is different when ITDs greatly exceed the biologically relevant ITD. The diffuseness of an image may be broader for large ITDs as compared to small ITDs. If the image is diffuse, then even though it might be perceived as being on one side of the head rather than the other, it might be more difficult to determine if such a diffuse image moved left or right. Thus, difficulty in making a left-right movement judgment might be an indirect measure of diffuseness. The MOVE data (Fig. 2 and for the other CF groups) for conditions in which both stimuli had ITDs much larger than the biologically possible ITDs indicate more difficulty in making motion judgments than when

the ITDs were smaller. This is consistent with the possibility that such conditions produce diffuse images (although the attempt to measure diffuseness in the P-D task did not produce such results). It is also possible that these large ITDs are already lateralized as far right as possible, although why this would be frequency dependent and apply to both narrow-band and wideband stimuli is not clear. Thus, it is possible that the arguments (based on IC data) that large ITDs may be processed differently than small ITDs are partially reflected in lateral judgments of inexperienced human listeners.

The conditions used in this set of experiments and in a great deal of the psychophysical and physiological literature are unnatural in several ways. First, the ITDs are often much larger than occur in nature. Second, the interaural level difference (ILD) is usually 0 dB while ITD is varied. This does not occur in nature as ITD and ILD covary. Third, the ITDs are either only ongoing ITDs or the rise times are long compared to the ITD, making the use of ITD-onset cues highly problematic (as was the case in the present experiments). For most real world sounds there are interaural onset cues for sound sources at different azimuths and these are known to affect sound localization judgments (see Buell *et al.*, 1991). Thus, attempts to model (psychophysically or physiologically) data based on such artificial conditions may not be appropriate for describing ITD processing in the real world. To the extent that existing models do reveal relevant processing schemes for ITD, the current possible disagreement in the literature between human psychophysical data and recordings in the IC suggest that either the IC is not *the* neural circuit that forms *the* basis for psychophysical judgments based on ITD *and/or* that current psychophysical models or new models need to consider the processing that takes place within IC.

Acknowledgments

This research was supported by NIDCD grants awarded to William A. Yost and to Stanley Sheft. We thank Joe Boomer and Jim Collier, the Parmlly Hearing Institute lab technicians, and Elizabeth Chrobak, an undergraduate student, for their assistance.

References and links

- Blauert, J. *Spatial Hearing*. (MIT Press; Cambridge: 1997).
- Buell TN, Trahiotis C, Bernstein LR. Lateralization of low-frequency tones: Relative potency of gating and ongoing interaural delays. *J. Acoust. Soc. Am* 1991;90:3077–3084. [PubMed: 1787246]
- Harper NS, McAlpine D. Optimal neural population coding of an auditory spatial cue. *Nature London* 2004;430:682–686. [PubMed: 15295602]
- Jeffress LA. A place theory of sound localization. *J. Comp. Physiol. Psychol* 1948;41:35–39. [PubMed: 18904764]
- Kuhn, G. Physical acoustics and measurements pertaining to directional hearing. In: Yost, WA.; Gourevitch, G., editors. *Directional Hearing*. Springer-Verlag; New York: 1987. p. 3-26.
- McAlpine D, Jiang D, Palmer AR. A neural code for low-frequency sound localization in mammals. *Nat. Neurosci* 2001;4:396–401. [PubMed: 11276230]
- McFadden D, Jeffress LA, Russell WE. Individual differences in sensitivity to interaural differences in time and level. *Percept. Mot. Skills* 1973;37:755–761. [PubMed: 4764506]
- Thompson SK, von Kriegstein K, Deane-Pratt A, Marquardt T, Deichman R, Griffiths TD, McAlpine D. Representation of interaural time delay in human auditory midbrain. *Nat. Neurosci* 2006;9:1096–1098. [PubMed: 16921369]
- Trahiotis C, Stern RM. Lateralization of bands of noise: Effects of bandwidth and differences of interaural time and phase. *J. Acoust. Soc. Am* 1989;86:1285–1293. [PubMed: 2808904]
- Trahiotis C, Bernstein LR, Akeroyd MA. Manipulating the ‘straightness’ and ‘curvature’ of patterns of interaural cross correlation affects listeners’ sensitivity to changes in interaural delay. *J. Acoust. Soc. Am* 2001;109:320–330.

- Wright BA, Fitzgerald MB. Different patterns of human discrimination learning for two interaural cues to sound-source location. *Proc. Natl. Acad. Sci. U.S.A* 2001;98:12307–12312. [PubMed: 11593048]
- Yost WA. Lateral position of sinusoids presented with interaural intensive and temporal differences. *J. Acoust. Soc. Am* 1981;70:397–409.

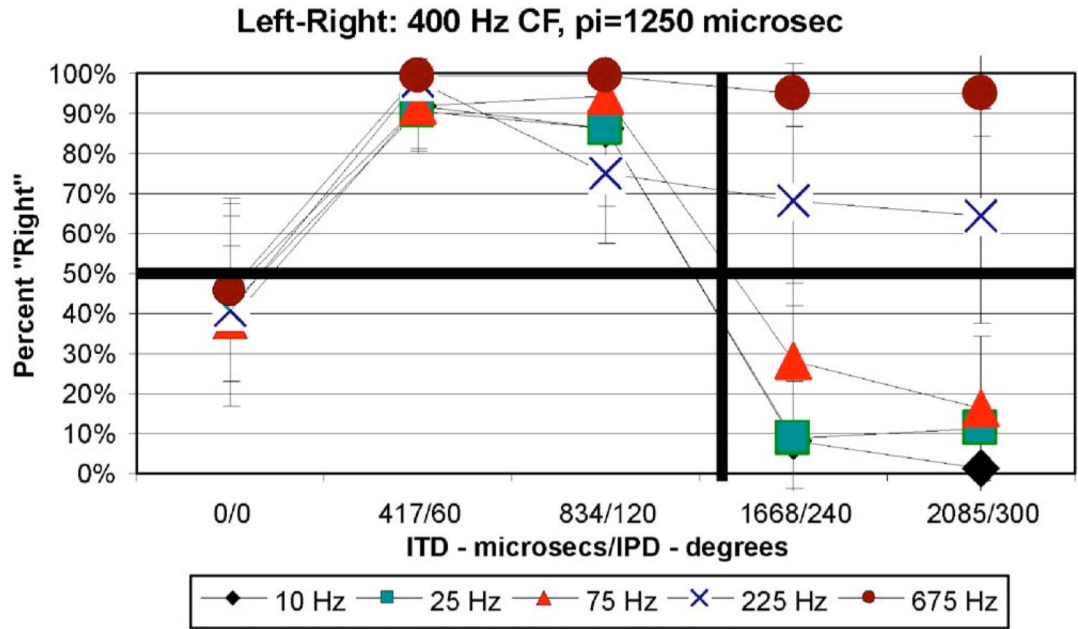


Fig. 1. (Color online) Average data (with standard error of the mean) for the 400-Hz CF group for the L-R task indicating percent “Right” judgments as a function of ITD (μs)/IPD (relative to π at CF). Different curves represent different bandwidth conditions indicated in the legend. The thick vertical line in the approximate middle of the figure represents an ITD at π of CF.

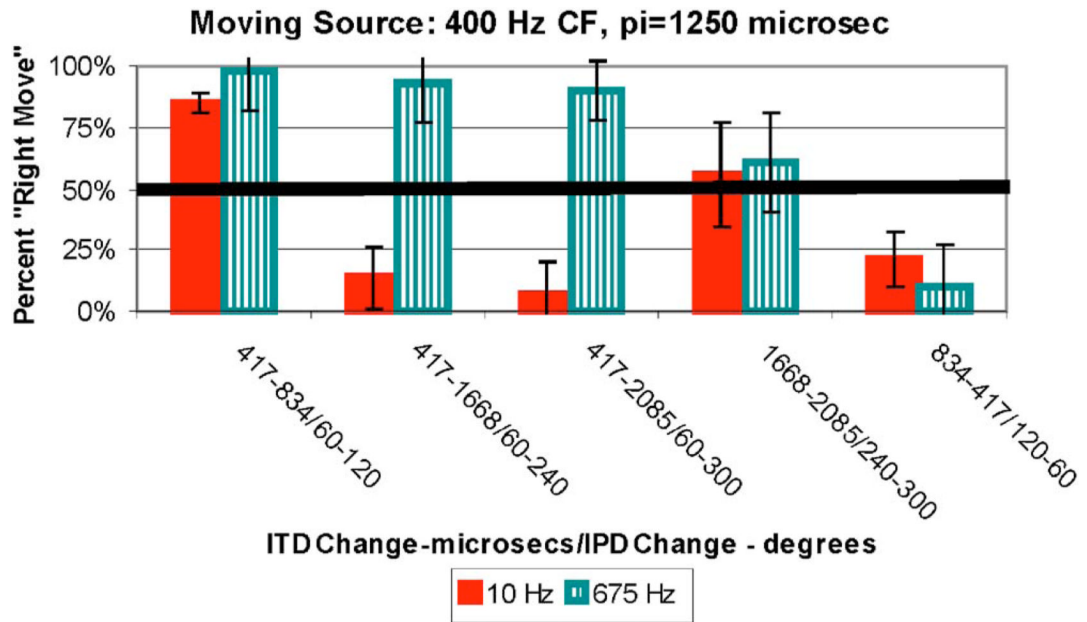


Fig. 2. (Color online) Average data (with standard error of the mean) for the 400-Hz CF group for the MOVE task indicating percent “Right Move” judgments as a function of the ITD (μs)/IPD (relative to π at CF). The first ITD/IPD is that for the first stimulus presented and the second ITD/IPD is for the second stimulus. Different bars represent different bandwidth conditions indicated in the legend.

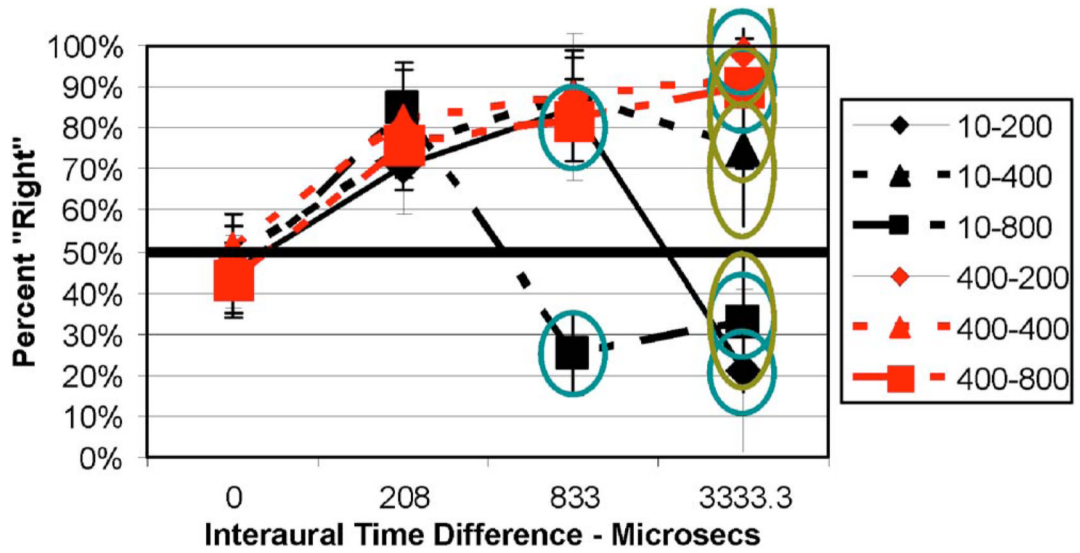


Fig. 3.
 (Color online) Average data (with standard error of the mean) for the MIXED group for the L-R task indicating percent “Right” judgments as a function of ITD (μs). Red curves for the 400-Hz-wide noise; black curves for 10-Hz-wide noise. Different curve and symbol types for different CF conditions as indicated in the legend. Data within blue circles are for ITD $> \pi$ at CF. Data within green ellipses are those for ITD $> 2\pi$. See Table 2.

Table 1
Stimulus and condition values for the three CF groups.

Conditions	200-Hz CF group	400-Hz CF group	800-Hz CF group
ITDs (μ s)	0, 417,834,1688, 3333, 4176	0, 417,834,1688,2085	0,209,417,834,1034
ITD μ s for π	2500	1250	625
Bandwidths (Hz)	195–205 (10), 187.5–212.5(25), 162.5–237.5 (75), 112.5–287.5 (175), 12.5–387.5 (375)	395–405 (10), 387.5–412.5 (25), 362.5–437.5 (75), 287.5–512.5 (225), 62.5–737.5 (675)	795–805 (10) 775–825 (50) 725–875 (150) 462.5–1137.5 (675) 125–1475 (1350)
No. of trials/condition/listener	20; 40; 20	20; 40; 20	20; 40; 20
L-R; MOVE; P-D tasks			
No. of subjects	19	22	21

Table 2

Stimulus and condition values for the MIXED group.

Conditions	MIXED group
ITDs (μs)	0, 208, 833, 3333.3
CFs (Hz)	200, 400, 800
IPD (deg)	0 208 833 3333.3
	CF ITD (μs)=
	200 Hz 0 0 60 240
	400 Hz 0 0 120 480
	800 Hz 0 0 240 (120)
(IPD $re:2\pi$)	
(IPD $re:4\pi$)	
Bandwidths (Hz)	10-Hz wide and 400-Hz wide, see Table 1
No. of trials/condition/listener, L-R task	50
No. of subjects	25