Just noticeable differences of spatial cues in echoic and anechoic acoustical environments

Stefan Klockgether, and Steven van de Par

Citation: The Journal of the Acoustical Society of America **140**, EL352 (2016); doi: 10.1121/1.4964844
View online: https://doi.org/10.1121/1.4964844
View Table of Contents: https://asa.scitation.org/toc/jas/140/4
Published by the Acoustical Society of America

Just noticeable differences of spatial cues in echoic and anechoic acoustical environments

Stefan Klockgether and Steven van de Par
Department for Medical Physics and Acoustics, Cluster of Excellence "Hearing4All," Carl von Ossietzky University Oldenburg, Carl-von-Ossietzky-Strasse 9-11, 26129 Oldenburg, Germany
stefan.klockgether@uni-oldenburg.de, steven.van.de.par@uni-oldenburg.de

Abstract: The perceptual limits for detecting changes in binaural cues also define the boundaries for the perception of differences in spatial impression. This study reports just noticeable differences for interaural time delays (ITDs) and interaural level differences (ILDs) of the early part and for the interaural cross-correlation (IACC) of the early and diffuse part of binaural room impulse responses. The results show that ITDs only allow a high accuracy in localization in anechoic environments, whereas ILDs show a higher robustness against reverberation. Subjects are rather insensitive to changes in IACC, only changes that bring the IACC close to one are detectable.

[QJF]
Date Received: May 3, 2016 **Date Accepted:** September 9, 2016

1. Introduction

Humans are able to very accurately localize a sound source in the horizontal plane. The main cues for localization in the horizontal plane are the interaural time delays (ITDs) and the interaural level differences (ILDs) which both together allow for an unambiguous localization (Blauert, 1969). In reverberant environments, ITDs and ILDs are fluctuating across time due to the presence of delayed reflections from various directions. This results in a reduction of the interaural cross-correlation (IACC) measured on a longer time-scale.

For a fixed source and listener position, the essential acoustical properties of a room can be captured in a binaural room impulse response (BRIR). The perceived location of a sound source in a room is depending on the early part of that BRIR (Litovsky *et al.*, 1999). The IACC of the BRIR influences other spatial aspects of the perceived sound field. The IACC of the early part is directly linked to the perceived width of the source (Hidaka *et al.*, 1995). The IACC of the reverberant tail is linked to the perceived envelopment by a sound that a listener has in that room (Okano *et al.*, 1998). When listening to a continuous sound in a reverberant environment, the early part and the reverberant tail of the impulse response are concurrently present and listeners are not always able to distinguish between the early and late reverberations (Klockgether and van de Par, 2014).

Just noticeable differences (JNDs) of the ITDs and the ILDs have been well investigated for synthetic stimuli. Humans have a very small JND for ITDs which can be as low as 20 μs at low-frequencies (Klumpp and Eady, 1956). The JND for ILD is in the range of 0.5 to 1 dB (Mills, 1960). These small JNDs for ITDs and ILDs have typically been measured in head-phone experiments or in free-field conditions with artificial stimuli that are highly correlated between the left and right ear. Such measurements do not necessarily reveal the sensitivity to ITD and ILD changes in real environments, since the correlation between the two ears decreases due to the presence of reverberation in reverberant environments. In JND experiments of ITD and ILD cues the interaural cross-correlation of the stimulus has a significant influence on the thresholds. A study of Le Goff *et al.* (2013) shows an increase in ITD JNDs when the correlation is low.

The sensitivity to changes in IACC is depending strongly on the initial interaural cross-correlation a stimulus has before any manipulation to the correlation is applied. For narrow-band noise the sensitivity to changes in correlation is very high provided that the initial correlation is high (Durlach *et al.*, 1986). When the initial correlation is low, however, subjects are rather insensitive to changes in correlation (Gabriel and Colburn, 1981). This leads to smaller JNDs for correlation differences towards a higher IACC and larger JNDs towards uncorrelated signals (Walther and Faller, 2013).

In this study we investigate within a reverberant context by how much primary cues (ITD, ILD, IACC) must change for subjects to just detect the change. The perceived location of a sound source will only be influenced by the early part of the BRIR. Therefore, only the ITDs and ILDs of the early part of the BRIR are manipulated for determining the JNDs for these interaural cues. The later part of the BRIR will become dominated by diffuse reflections of the room and will carry no perceptible directional information. The sensitivity to interaural cross-correlation manipulations is investigated in both parts of the BRIR, which are linked to different aspects of the spatial impression of a sound in a room.

2. Method

The JNDs of the ITD, ILD, and IACC were measured with real sound signals which were generated by convolving anechoic music signals with BRIRs. The JNDs were determined with an alternative forced choice measurement (AFC) (Levitt, 1971), in which subjects had to listen to trials of three sound intervals. One of these intervals was a spatially manipulated version of the other two (reference) intervals and had to be detected by the subjects. The exact procedure is specified below.

2.1 BRIR-manipulation

For the psychoacoustic experiments, the BRIRs were divided in two parts, where the room-dependent transit between both parts is called the perceptual mixing time (PMT) (Lindau *et al.*, 2012). The splitting of the BRIR enabled to separately manipulate the early part and the diffuse tail. The crossover between both parts was realised with 4-ms long raised-cosine ramps which would create a perfect reconstruction of the BRIR in case no manipulation was performed.

In each experimental run, one of the three binaural cues was manipulated. In the case of the ITD, the timing of the first part of the right BRIR channel was delayed resulting in an ITD (see Fig. 1, left panel). The start value of the ITD was 300 μs, which was found to be clearly audible for the subjects in pilot experiments. For determining the ILD-JND, the early part of the left BRIR channel was increased by half of the value of the ILD while the corresponding early part of the right BRIR channel was reduced by an equal amount (see Fig. 1, right panel). The start value of the ILD was 10 dB. In both conditions the diffuse tail of the BRIR was not manipulated. The manipulation leads to a shift of the perceived location of the sound source to the left.

The manipulation of the IACC was achieved by cross-mixing the right and left channel (R and L) of the BRIR as shown in Eq. (1). The mixing parameter alpha is used frequency-dependent whereas the center frequencies and bandwidths were chosen according to the critical-band scale (Glasberg and Moore, 1990). The mixing parameter α is used to control the ratio of the mixture and it serves as the adaptive variable in the experiment to control the strength of the manipulation in the new right and left channels (R'_f and L'_f). After cross-mixing, a level normalization within each critical band (f) was made to avoid any coloration. Equation (1) shows the cross-mixing together with the normalization which is the multiplication of the ratio between the root mean square (RMS) of the initial left or right channel divided by the RMS of the respective new channel. This mixing method is similar to the method proposed by Xiang *et al.* (2015),

$$L' = \sum_f (L_f + \alpha R_f) \frac{\mathrm{RMS}(L_f)}{\mathrm{RMS}(L_f + \alpha R_f)}, \quad R' = \sum_f (R_f + \alpha L_f) \frac{\mathrm{RMS}(R_f)}{\mathrm{RMS}(R_f + \alpha L_f)}. \qquad (1)$$

Equation (1) shows that for $\alpha = 0$ both channels are not mixed at all and stay unaltered. $\alpha = 1$ would imply two identical signals in both channels. However, a very small difference ($\rho > 0.999$) might be left due to the independent level normalization in each critical band which preserves level differences between both channels to avoid

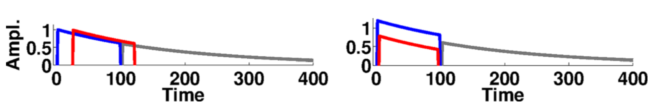

Fig. 1. (Color online) Schematic pictures of the ITD- and ILD-manipulations of the early part of the BRIRs. The left panel shows the addition of a time delay to the early part of the right channel. On the right panel an increase to the level of the early part of the left channel is shown, while the level of the early part of the right channel was decreased by the same amount. Note that for visual clarity the manipulations shown here are strongly exaggerated.

coloration. The impact of α on the IACC is shown in the data analysis of the experiment in Sec. 3.

As starting value in the adaptive procedure α was set to 0.8. The cross-mixing was applied either to the early part or to the diffuse tail of the BRIR.

2.2 Stimuli and subjects

Selected parts of longer anechoic music signals were convolved with the BRIRs and used as stimuli. For each trial, a 2 s long excerpt with randomly chosen starting point was selected for the convolution. This might have led to a slightly increased variance in the measured thresholds but the random timing was necessary to avoid that possible sound coloration cues were learned for distinguishing the intervals. Three different music signals were used, a guitar, a violin or a snare drum. The BRIRs were recorded with the FABIAN dummy head in a lecture hall ($T_{60} = 1.7$ s) and a seminar room ($T_{60} = 0.9$ s) of the TU Berlin. The PMT was set to 120 ms in the lecture hall and to 100 ms in the seminar room. These values were determined in a pre-experiment where subjects had to distinguish between a manipulated and a non-manipulated signal. The manipulation in the pre-experiment was done by interchanging the left and right channel of the reverberant tail of the BRIRs of the different rooms. The time at which the interchanging started was the adaptive variable in the pre-experiment and the resulting detection threshold was used as PMT of the particular room. The found thresholds corresponded well with the predictions of the PMT by a model proposed by Lindau *et al.* (2012). The anechoic condition was included in the ITD- and ILD-experiments to serve as a reference condition that can be compared against literature ITD- and ILD-JND data. Without any manipulations all sound sources were perceived in the front of the listener.

For determining the JNDs, an adaptive three-interval 2-AFC procedure was used. The first interval always contained the reference as well as either the second or the third interval, while the remaining interval contained the manipulated BRIR. The subjects had to determine which interval was manipulated. The thresholds were determined with an adaptive 1-up, 2-down staircase procedure (Levitt, 1971), which meant, that the strength of the manipulation was reduced, when the subjects correctly detected the manipulation in two consecutive trials. If the subject was not able to detect the manipulation, the strength of the manipulation was increased. Visual feed-back was given after each trial. After two and after four reversals the stepsize of the adaptive tracking procedure was reduced each time by a factor of 2. After that, the mean of the next six reversals of the tracking variable served as the threshold. The minimum stepsize was 0.05 in the IACC-conditions and about 10% of the current value of the ITD or ILD in the ITD- and ILD-conditions. Eight normal-hearing subjects participated in the experiment and each condition was repeated four times for each subject. Four of the subjects, including one of the authors, were considered as expert listeners, the other four as naive listeners. All signals were presented at a RMS-level of 65 dB sound pressure level. For the ILD-experiment, a random level off-set the range of ± 3 dB was added to each interval to avoid that monaural level cues in one ear could be used as a reliable cue to detect the ILD-manipulation.

3. Results

The results are shown in Figs. 2 and 3 as mean values with standard errors over all subjects for the different rooms and instruments. Results for the guitar excerpt are shown with purple (◇), the violin with green (▷), and the snare drum with orange (◁). Figure 2 shows the JNDs in ITD and ILD for the two echoic rooms and the anechoic condition. The JND for the ITD is about 20 μs for the anechoic condition and five to eight times higher in reverberant rooms for the violin and guitar. The ITD-JND for the snare drum is only two to three times higher in the reverberant conditions. This may be related to the steeper and also more isolated transients of the snare drum sound which may allow for a better localization due to the smaller influence of the reverberant tails. For all three instruments the variation of the results decreases in less reverberant conditions.

The ILD-JND shows similar results (Fig. 2, right). The less reverberant the rooms are, the lower the ILD-JND seems to be. In the anechoic condition the JND is about 0.6 dB and for more reverberant conditions it is only up to twice as high, suggesting that ILD-JNDs are affected less by reverberation than the ITD-JNDs. Similar to the ITD-JND (Fig. 2, left), the variations of the ILD thresholds are lower for the less reverberant conditions, at least for the violin and snare drum.

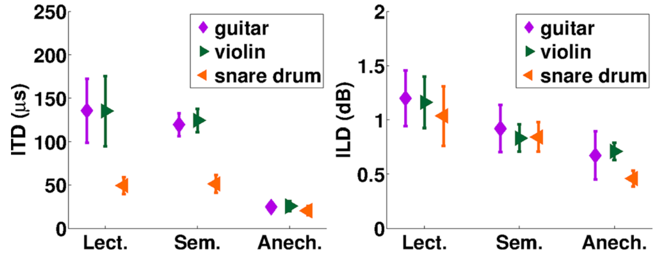

Fig. 2. (Color online) Just noticeable difference for the interaural differences (ITD, left panel; ILD, right panel) measured for three different rooms (lecture hall, seminar room, anechoic) and three different instruments.

Figure 3 shows the JND for the mixing parameter α for the two reverberant rooms and the three different instruments. The left panel shows the results for manipulating the IACC of the early part of the BRIR, the right panel for the manipulation of the diffuse tail. When manipulating only the early part, the mixing parameter needs to be about 0.4 to detect a change in interaural correlation for the lecture hall and a little smaller (about 0.3) in the seminar room. The thresholds for detecting changes in the IACC are a little lower for the guitar sound. The manipulation of the diffuse tail needs to be much stronger than the manipulation of the early part to be detected, both for the guitar (about 0.6) and the violin (about 0.7) sound. The snare drum sound allows one to detect the manipulation of the diffuse tail at a slightly lower mixing parameter as for the manipulation of the early part. For the guitar and the violin signal, some subjects were not always able to detect the interval with the manipulated diffuse tail, even when the mixing parameter α was at maximum. Since it made no sense to increase α beyond a value of 1, those thresholds were set equal to 1 for the analysis. The results show slightly lower thresholds for detecting manipulations to the interaural cross-correlation of the diffuse tail of a BRIR in the more reverberant room.

Figure 4 shows two graphs which should help to interpret the meaning of the JND of the mixing parameter in terms of changes in the interaural cross-correlation. The left panel shows the IACC of an exemplary stimulus from the experiment as a function of the mixing parameter for cross-mixing either the early part, the reverberant tail or the complete BRIR. It shows that the increase of the IACC due to cross-mixing is larger, when the IACC is low and that cross-mixing the early part of the BRIR has by far the stronger influence on the overall IACC than cross-mixing only the reverberant tail. All the stimuli used for the experiment also showed the same trend. The right panel of Fig. 4 shows the calculated JND for the IACC as a function of the initial correlation of a sound. The JND of the mixing parameter is used as a parameter in the calculation of the different curves labeled with 0.3–0.8. When the JND of the mixing parameter is used to determine the JND in interaural cross-correlation, the increase of the JND will depend on the initial correlation. Since in the stimulus this initial correlation is highly dependent on frequency we can only qualitatively characterize the extent to which the correlation needs to change to be detectable. It can be seen that whatever the initial correlation was, it needs to increase to a correlation near to 1 in order to be distinguishable from the initial correlation, suggesting that subjects are very insensitive to changes in IACC of the early or later part of the BRIR.

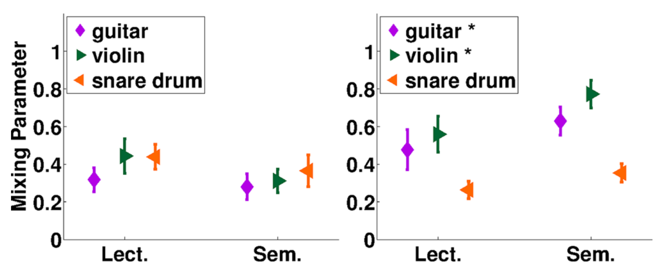

Fig. 3. (Color online) JND for the indirect manipulation of the IACC of either the early part (left panel) or the reverberant tail (right panel) of the BRIRs, measured for two different rooms (lecture hall, seminar room) and three different instruments, using the mixing parameter α to alter the IACC. The asterisks in the right picture indicate that some JNDs of these conditions were set to $\alpha = 1$, when the subjects were not able to detect the manipulation at all.

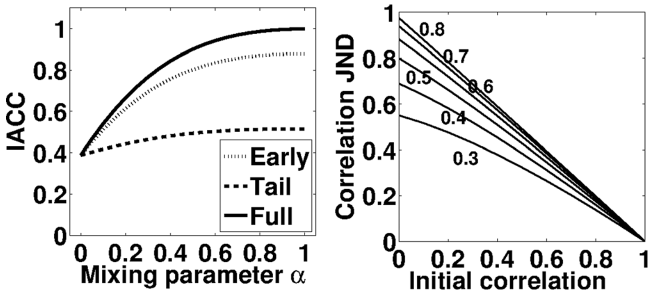

Fig. 4. Relation between interaural cross-correlation and mixing parameter. The left panel shows the normalized IACC of an exemplary stimulus after cross-mixing the left and right channel of either the early part, the reverberant tail or the complete BRIR as a function of the mixing parameter α. The right panel translates the JNDs of the mixing parameter into JNDs of the IACC. The graphic shows the JND for the IACC as a function of the initial correlation of a stimulus for different JNDs of the mixing parameter (0.3–0.8).

4. Discussion

For this study a psychoacoustic experiment was conducted that measured JNDs for ITD, ILD, and IACC in anechoic and reverberant conditions. For the measurement ITDs or ILDs were imposed on the early part of a BRIR, for the IACC, the early or late part of the BRIR was increased in correlation by cross-mixing. The results show that the JNDs of the ITD and the ILD for the anechoic condition match quite well with literature data (Klumpp and Eady, 1956; Mills, 1960). Just noticeable differences for ITD and ILD were found to increase with increasing reverberation times. The ITD-JND increases much stronger than the ILD-JND. This indicates that the detection of an ILD is more robust against reverberation than the detection of an ITD,[1] which may indicate that ILD cues are more reliable in reverberant environments than ITD cues. This may be due to the low baseline ITD-JND in anechoic conditions, and to the fact that ITD-JNDs primarily depend on the low frequency fine structure below 1.5 kHz (Klumpp and Eady, 1956). The JNDs of the ITD for the percussive sound (snare drum) seem to be more robust against reverberation than more tonal sounds. This might be related to the occurrence of more temporally isolated transients which have less temporal overlap with the reverberant tails of previous signal components.

The JNDs for the manipulations of the IACC are very large. If, for example, the reference stimulus has an IACC of 0.6 the IACC of the target stimulus has to be close to one to be detected. In some conditions, some subjects were even unable to detect the manipulation at all. From literature it is also known that the human auditory system is rather insensitive to correlation changes when the initial correlation is close to zero (Robinson and Jeffress, 1963). Note that it can be expected that the reverberant tail will typically have a lower IACC than the early part of the BRIR. Therefore it is expected that subjects are less sensitive to correlation changes of the tail of the BRIR, which can also be seen in the data (see Fig. 3).

The large correlation-JNDs, that we found in our experiments, have important implications for perceptual parameters like apparent source width and listener envelopment that have been argued to be dependent on IACC and which are important factors in the appreciation of concert halls. The low sensitivity to correlation changes indicates that small differences in the IACC in real rooms will not be distinguishable for listeners.

As for the ITD JND, the JND for IACC manipulation of the snare drum sound differs for the tail manipulation from the other two instruments. The snare drum, being a percussion instrument, also has transients with a shorter decay than the other instruments, which might reduce the overlap of the reverberant tails from the transients. This may lead to a clear perception of the reverberant tail and therefore to smaller JNDs for the IACC of the reverberant tail.

The results of this study indicate that JNDs in the basic spatial cues (ITDs, ILDs, and IACC) depend highly on the room acoustics and also on the nature of the used sound source. This study also shows that the JNDs of the basic spatial cues measured with natural sounds in anechoic environments match quite well with the results and the expectations of previous studies which used artificial stimuli to measure the JNDs.

Acknowledgments

The authors would like to thank the Audio communication group of the TU Berlin, Germany for the BRIR recordings and the Deutsche Forschungsgemeinschaft for funding this research in the context of the research unit SEACEN. The authors also would like to thank the reviewers for their helpful comments.

[1]Also converted to a linear amplitude scale the ILD-JND increases less than the ITD-JND with reverberation.

References and links

Blauert, J. (**1969**). "Sound localization in the median plane," Acustica **22**, 205–213.

Durlach, N. I., Gabriel, K. J., Colburn, H. S., and Trahiotis, C. (**1986**). "Interaural correlation discrimination: II. Relation to binaural unmasking," J. Acoust. Soc. Am. **79**, 1548–1557.

Gabriel, K. J., and Colburn, H. S. (**1981**). "Interaural correlation discrimination: I. Bandwidth and level dependence," J. Acoust. Soc. Am. **69**, 1394–1401.

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

Hidaka, T., Beranek, L. L., and Okano, T. (**1995**). "Interaural cross-correlation, lateral fraction, and low- and high-frequency sound levels as measures of acoustical quality in concert halls," J. Acoust. Soc. Am. **98**, 988–1007.

Klockgether, S., and van de Par, S. (**2014**). "A model for the prediction of room acoustical perception based on the just noticeable differences of spatial perception," Acta Acust. Acust. **100**, 964–971.

Klumpp, R. G., and Eady, H. R. (**1956**). "Some measurements of interaural time difference thresholds," J. Acoust. Soc. Am. **28**, 859–860.

Le Goff, N., Buchholz, J., and Dau, T. (**2013**). "Spectral integration of interaural time differences in auditory localization," Proc. Mtgs. Acoust. **19**, 050160.

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

Lindau, A., Kosanke, L., and Weinzierl, S. (**2012**). "Perceptual evaluation of model- and signal-based predictors of the mixing time in binaural room impulse responses," J. Audio Eng. Soc. **60**, 887–898, available at http://www.aes.org/e-lib/browse.cfm?elib=16633.

Litovsky, R. Y., Colburn, H. S., Yost, W. A., and Guzman, S. J. (**1999**). "The precedence effect," J. Acoust. Soc. Am. **106**, 1633–1654.

Mills, A. W. (**1960**). "Laterization of high-frequency tones," J. Acoust. Soc. Am. **32**, 132–134.

Okano, T., Beranek, L. L., and Hidaka, T. (**1998**). "Relations among interaural cross-correlation coefficient ($IACC_E$), lateral fraction (LF_E), and apparent source width (ASW) in concert halls," J. Acoust. Soc. Am. **104**, 255–265.

Robinson, D. E., and Jeffress, L. A. (**1963**). "Effect of varying the interaural noise correlation on the detectability of tonal signals," J. Acoust. Soc. Am. **35**, 1947–1952.

Walther, A., and Faller, C. (**2013**). "Interaural correlation discrimination from diffuse field reference correlations," J. Acoust. Soc. Am. **133**, 1496–1502.

Xiang, N., Bosun, X., and Cox, T. J. (**2015**). "Recent applications of number-theoretic sequences in audio and acoustics," in *Acoustics, Information and Communication*, edited by N. Xiang and G. M. Sessler (Springer International Publishing, Cham, Switzerland), pp. 93–111.