Cues for localization in the horizontal plane

Jeppesen, Jakob; Møller, Henrik

Published in:
Proceedings of 118th AES Convention

Publication date:
2005

Audio Engineering Society

Convention Paper 6323

Presented at the 118th Convention
2005 May 28–31 Barcelona, Spain

Cues for Localization in the Horizontal Plane

Jakob Jeppesen[1], Henrik Møller[1]

[1]*Department of Acoustics, Aalborg University, 9220 Aalborg Ø, Denmark*

Correspondence should be addressed to Jakob Jeppesen (jaj@acoustics.aau.dk)

ABSTRACT

Spatial localization of sound is often described as unconscious evaluation of cues given by the interaural time difference (ITD) and the spectral information of the sound that reaches the two ears. Our present knowledge suggests the hypothesis that the ITD roughly determines the cone of the perceived position (i.e. the azimuth in a polar coordinate system with left-right poles), whereas the spectral information determines the position on the cone (i.e. the elevation in the same coordinate system). This hypothesis was evaluated in a series of listening tests, where the two cues were manipulated in HRTFs used for binaural synthesis of sound in the horizontal plane. The manipulation of cues resulted in HRTFs with cues ranging from correct combinations of spectral information and ITDs to combinations with severely conflicting cues. Both the ITD and the spectral information seems to be necessary for localization in the sense that sources are localized well when the two types of cues are correct. When the cues are severely conflicting the localization performance is highly degraded.

1. INTRODUCTION

When localizing a sound in space the use of timing and spectral characteristics of the sounds that reaches each of the ears is essential. The transmission to the two ears is given by the head-related transfer function (HRTF) for the given direction. The HRTF is defined as the pressure at the ear divided with the pressure in the middle of the head with the head absent [1, 2]. Head-related transfer functions exist in pairs in the sense that for a given direction there is a transfer function for the left ear and one for the right. Sometimes the term HRTF is used for such a pair of transfer functions, but in this paper the term is used for a single transfer function either to the left or the right ear.

The HRTF describes the filtering of the sound due to diffractions and reflections of the head, torso and especially the pinna of the involved listener [3]. This filtering depends on the angle of incidence, and the direction is often described by azimuth and elevation in a polar coordinate system with a vertical axis (poles above and under the listener). In this system the azimuth angle covers the entire circle from 0 to 360° (or 0 to ±180°), and the elevation angle is between -90° (down) to 90° (up). In the present work it was chosen to use a polar coordinate system with a horizontal axis and poles to the sides instead. In this the elevation angle lies within the range of ±180° with 0° marking horizontal directions in the frontal hemisphere, 90° directions in the upper median plane and ±180° horizontal directions in the rear hemisphere. The azimuth angle is in the range of -90° (left) to 90° (right) with 0° marking directions in the median plane. Since the axis of this system is similar to the interaural axis, directions with the same azimuth will (at least roughly) have the same ITD.

The characteristics in the HRTFs describing sound transmission from different directions are normally called cues. The literature generally describes them by the terms interaural time difference (ITD) and interaural level difference (ILD) [4, 5]. The ITD describes the difference in arrival time for a sound reaching our two ears. The term ILD could give the impression that the second cue consists of a simple level difference between the two ears. However, this is a much too simple description, since the transmission to each of the ears is highly frequency dependent, and for most directions the variation with frequency differs between the two ears, thus the ILD varies with frequency. The level is generally higher at the ipsilateral ear, but because of the variation with frequency for the transmission to each of the ears, frequencies may exist where the level is higher at the contralateral ear. Furthermore the ILD is zero in the median plane where the inputs to the two ears are identical (or virtually identical). Instead of "repairing" the idea of ILD, it will be more correct to consider - as the second cue in excess of ITD - the spectral characteristics in general that are present in the minimum-phase parts of the pair of HRTFs. The ITD and these spectral cues describe together the part of the transmission that is important for perception.

The question of which characteristics of the HRTFs that give rise to localization from specific directions has been addressed in a number of studies. Most of the studies have focused on the localization cues given by the monaural and interaural spectral information present in the HRTFs [6, 7, 8, 9, 10] whereas analysis of the influence of the ITD in localization is more seldom although some studies have been carried out, e.g. [11, 12]. Wightman et al. [11] compared wideband stimuli and highpass-filtered stimuli and found that if low frequencies are present, the ITD serves as the leading cue in localization when the ITD and the spectral information are conflicting. Kulkarni et al. [12] concluded that the ITD is necessary to provide reliable cues for localization. Furthermore they found that if the ITD is preserved the combination of a minimum-phase transfer function and a pure delay gives an adequate approximation to the empirical HRTF. Both studies used ITDs calculated as the maximum of the interaural cross-correlation in their experiments. This way of calculating the ITD was later found slightly inaccurate by Minnaar et al. [13] (differences were small, and the audible differences subtle).

1.1. Hypothesis

The present study seeks to evaluate a model that describes localization as a two-step procedure; the first step is to evaluate the ITD and thereby place the sound on a contour with constant ITD. The fact that the iso-ITD contour can be roughly estimated with a cone has lead to the expression cone-of-confusion which describes an eventual localization error on the iso-ITD contour.

The second step of the localization procedure is evaluation of the spectral information in the HRTFs to more precisely place the sound on the already determined contour.

The dominance and robustness of the ITD as the leading cue can be tested in the horizontal plane using combinations of ITDs and minimum-phase HRTFs from different directions and thereby creating conflicting cues for localization, see e.g. [12, 14]. It is expected that combinations of ITDs and minimum-phase transfer functions with slightly conflicting cues will lead to localization close to the direction indicated by the ITD rather than the direction indicated by the spectral information. When the cues are more severely conflicting the localization is expected to be less focused and might also happen to be in the area of the direction indicated by the spectral cues.

2. METHODS

To test the hypothesis a listening experiment was designed. In the experiment a number of stimuli with and without manipulated cues were presented to the listeners.

2.1. HRTF processing

From a database of HRTFs measured with a very high directional resolution using an artificial head [15, 16] a number of directions in the horizontal plane were selected for further processing.

The HRTFs were available as impulse response measurements with an initial pure delay corresponding to the propagation time from the sound source to the ears of the mannequin used for the measurements. The data could be more precisely named head-related impulse responses (HRIRs) and the transformation between the HRIRs and HRTFs can be done by means of the Fourier and the inverse Fourier transforms.

The minimum phase part of each selected HRTF was calculated using homomorphic filtering as described in [17]. The minimum-phase part of a transfer function contains all spectral information available in the transfer function. The remaining part of the transfer function is called the excess-phase and can be decomposed into a linear-phase part and an all-pass part as shown in (1) and (2) [12, 18].

$$HRTF = HRTF_{min-ph} \cdot HRTF_{ex-ph} \quad (1)$$

$$HRTF_{ex-ph} = HRTF_{lin-ph} \cdot HRTF_{all-pass} \quad (2)$$

The minimum-phase part contains the spectral information, whereas the ITD is contained in the linear-phase and all-pass-phase parts. Minnaar et al. [13] evaluated different ways of calculating the ITD and found the most correct way is to calculate the interaural group delay (IGD) at low frequencies of the excess-phase part of the HRTF. An easy way of calculating the IGD is also proposed and is based on calculating the centroids (centre of gravity) of the time representations of the HRTFs. When the ITD has been calculated this way there is no audible difference between the original HRTF and an HRTF consisting of the minimum-phase part and the ITD [18].

The ITD for a given direction is calculated as (3) where $\mathbb{C}$ represents the centroid function and l and r are the abbreviations of left and right.

$$\begin{aligned} ITD &= IGD_{ex-ph,\ 0\ Hz} \quad (3) \\ &= \Big(\mathbb{C}(HRTF_l) - \mathbb{C}(HRTF_{l,\ min-ph})\Big) - \\ &\quad \Big(\mathbb{C}(HRTF_r) - \mathbb{C}(HRTF_{r,\ min-ph})\Big) \end{aligned}$$

In order to be able to do this calculation properly it is important that the HRTFs are in their most correct form which implies the initial delay of the transfer function must consist of zeros and the remaining part must have a DC-component of unity [13]. The DC-component can be inspected as the first value in the Fourier transform of the impulse response.

2.2. Selection of directions and manipulation of HRTFs

In total seven different ITDs were chosen. Each ITD exist for two directions in the horizontal plane (front and back) thus 14 different directions were available. Minimum-phase transfer functions from all 14 directions were calculated. Combining minimum-phase transfer functions from all different directions with all different ITDs resulted in an HRTF data-set with a total of 98 new pairs of HRTFs ranging from correct combinations of minimum-phase transfer functions and ITDs to combinations with severely conflicting cues. The ITDs were chosen in order to have values presented in the range from $0\mu s$ to close to maximum which - for the present mannequin - was a little higher than $\pm 600\ \mu s$. The chosen ITDs were $0\ \mu s$, $\pm 200\ \mu s$, $\pm 400\ \mu s$ and $\pm 600\ \mu s$. The directions involved are shown in Fig. 1.

In addition to this HRTF data-set a second data-set with only correct HRTFs was used for reference. The second data-set contained HRTFs from the left hemisphere in a resolution of 4° from 0° to -88° azimuth with elevation 0° and 180° - in total 46 directions. All HRTFs in the second data-set were processed as described above and stored with correct combinations of ITD and minimum-phase transfer functions only. As the HRTFs did not contain conflicting cues the localization was expected to be close to the presented direction although it is well known

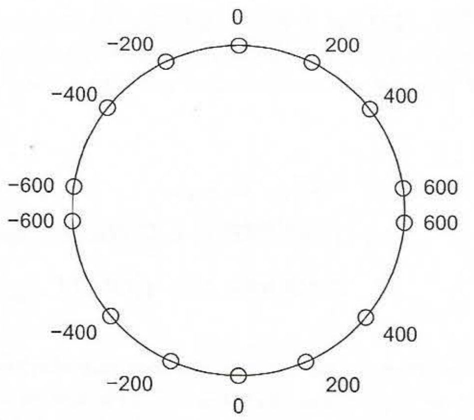

Fig. 1: Horizontal plane with the selected directions from the HRTF database and the ITDs in μs.

that localization with non-personal HRTFs is somewhat degraded compared to real life localization.

2.3. Stimuli

All stimuli were prepared by convolving the HRTFs (whether original or manipulated) with a pink noise sequence. The noise sequence was band-limited from 200 Hz to 15 kHz. The stimuli had 100 ms onset and offset ramps. The resulting sequence was 1 second long. The sampling frequency was 48 kHz. In total 98 different stimuli were presented from the first HRTF data-set and 46 different stimuli were presented from the second HRTF data-set. Each stimulus was repeated six times to each subject, thus each subject gave 864 responses. The stimuli were presented to the subjects in six sessions each with a duration of 10-12 minutes. The six sessions were divided on two days with three sessions each day. A break of 10 minutes was introduced between the sessions. Each of the 144 stimuli were presented once in each session in a random order.

2.4. Response method

The stimuli were presented to the subjects via headphones in a static environment hence eventual head turns did not provide any help to the subjects which was also explained to the subjects prior to the experiment. The subjects were asked from which direction they heard the sound and had to mark their response on a digitizer pad using an electronic pen. The response pad held a circle of 8 cm in diameter illustrating the horizontal plane with an indication of the listener being in the middle of the circle. It was possible to indicate all directions in the horizontal plane without discretization. The response was marked with the pen on the circle. Fig. 2 shows a downscaled version of the response pad.

During the experiment the subjects were guided with a small "traffic light" that indicated when to listen and when to give their response. There was no limit on the response time.

All subjects were presented with the same written instruction. It was specified that they were expected to evaluate the direction of the sound in the horizontal plane only. If they found the sounds not being in the horizontal plane they were supposed to respond corresponding to the projection of the direction to the horizontal plane. Furthermore they were not expected to evaluate distance to the sound source. In interviews with the subjects after the experiment, none of the participants expressed any particular problems with neither elevated sound sources nor confusing distance perception - e.g. the feeling of the sound being inside the head. The subjects were encouraged to be fully concentrated during the experiment.

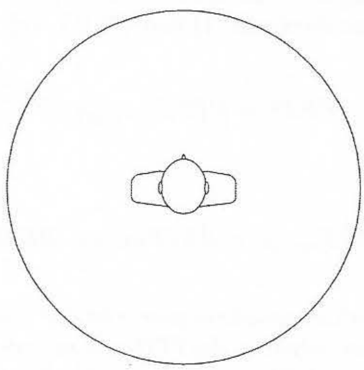

Fig. 2: Response pad used in the experiment.

2.5. Subjects

Ten subjects participated in the listening experiment. The subjects were paid for their participation and were all young people between 20 and 28 years of age. Prior to the experiment an audiometric test in the octave bands from 125 Hz to 4 kHz was conducted on each test subject. All test subjects were of normal hearing since none had greater hearing loss than 15 dB at the measured frequencies. Also the differences in hearing level between left and right ear did not exceed 5 dB at the measured frequencies for any of the subjects. All subjects were naive regarding localization experiments.

Prior to the experiment the subjects were trained for approximately 10 minutes. The purpose of the training session was primarily to help the subjects to get familiar with the response method. As a secondary purpose they also had a possibility to adapt to the presented stimuli. Although the presented stimuli were of relatively simple nature it can often feel rather unnatural to listen to noise sequences if one never did before.

2.6. Headphone equalization

When reproducing binaural signals to a listener it is important to ensure that exactly the same signal as was recorded (or synthesized) is presented to the listener [2]. Therefore it is necessary to minimize the influence of the playback system and by equalization make sure that the reproduction chain has a flat frequency response.

The reproduction chain consisted of three units, namely a pc with a soundcard (RME DIGI96/8), an amplifier (Pioneer A616) with a fixed gain of 0 dB and a headphone (Beyer Dynamic DT990 Pro). By measurements it was verified that the soundcard and amplifier both had a flat frequency response within the audible range. However the headphone did not have a flat frequency response and equalization was therefore necessary.

Since the used HRTFs are measured at a blocked entrance to the ear canal the equalization must also be done at the entrance of a blocked ear canal to ensure correct reproduction of the binaural signals [2]. The electro-acoustical transfer function of the headphone from the terminals to the blocked entrance of the ear canal on a human being when the headphone is correctly fitted to the head was measured. The measurement procedure was adopted from [19] and was as follows. A miniature electret microphone from Sennheiser (KE4-211-2) was mounted in an expandable EAR earplug. The earplug was fitted in the ear canal of the test subject in such way that the microphone was placed precisely at the entrance to the ear canal. The ear canal was thereby blocked by the

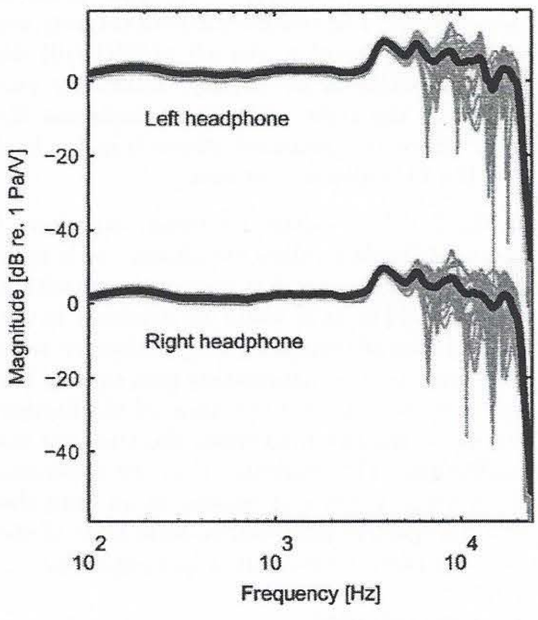

Fig. 3: Measured left and right headphone responses (gray) and averaged left and right headphone responses.

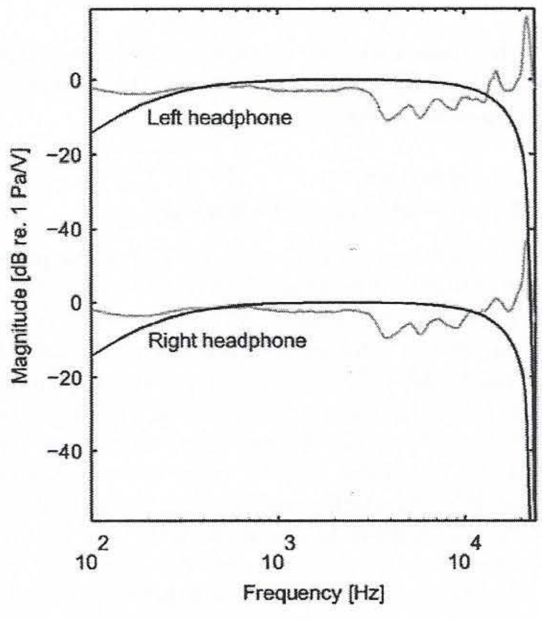

Fig. 4: Inverse of averaged left and right headphone responses (gray) and equalized left and right headphone responses.

expandable earplug. The transfer function was measured using a maximum length sequence (MLS) system. On seven subjects five repeated measurements of the left and the right channel of the headphone were carried out with the headphone refitted every time. In total 35 transfer function measurements were done for each channel.

The measured transfer functions for each subject were inspected for outliers (none was found) and adjusted with respect to the sensitivity of the microphone. The averaging was done in two steps; first by averaging the five transfer functions for each subject on a sound level basis and thereafter finding the median transfer function across subjects. The differences in the transfer functions within subjects were in general very small and the mean value could be considered as an appropriate measure whereas the differences across subjects were larger, thus the median were considered better than averaging.

The left and the right channel were averaged independently and the averaged transfer functions were inverted using deconvolution with regularization as described by [20]. Furthermore the equalization filters were band limited with a second order band pass filter from 200 Hz to 15 kHz.

The measured and averaged headphone responses can be seen in Fig. 3 and the inverted and equalized responses in Fig. 4.

3. RESULTS

3.1. Localization with original cues

The data contains the responses from the presented 46 different directions in the left hemisphere. The responses were collected with a high directional resolution and subsequently discretized in steps of 4° for visualization. The responses are shown with respect to azimuth only, thus front/back confusions cannot be seen. Front/back confusions were present in 20% of the responses, but the value varied substantially between directions ranging from around 10% at directions slightly to the sides, 20% near the median plane and 50% at the sides. For geometric reasons the high value at the sides comprise also directional inaccuracies, since these may very likely go into the opposite hemisphere. The mean and standard deviation of the responses for a given presented direction has been calculated as well as the median value and although the localization performance for the participating subjects is generally quite good it must be noted that there is a tendency to overestimate the azimuth angle except for large azimuth angles where the azimuth is underestimated. The data is presented in Fig. 5 and in Table 1.

3.2. Localization with manipulated cues

The HRTFs in this part of the experiment ranges from combinations of spectral cues and temporal cues being correct to combinations where the two types of cues are severely conflicting. The complete data-set is presented in Fig. 6. Assuming to symmetry the data has been pooled for positive and negative ITDs (with corresponding mirroring of responses and minimum-phase HRTFs). The responses are presented as the azimuth directions only, thus it is not possible to see if a response was given in the frontal or in the rear hemisphere. This, off course, also eliminates the possibility to see eventual cone-of-confusions errors but since the interest in this experiment is focused on the ITD such errors do - by definition - not exist.

Each column holds the responses of the presented ITD combined with the minimum-phase transfer functions corresponding to the ITD given in the parentheses. The combinations of severely conflicting cues are shown to the left of HRTF(0) whereas the combinations of slightly conflicting cues are shown to the right. The arrow indicates the column where the minimum-phase transfer functions and the ITD give correct cues.

In Table 2 the calculated mean, standard deviation and median values are shown. It is seen that the responded azimuth is not consistent within each presented ITD as it could be expected if the ITD was an overall dominant cue. However when the ITD and spectral information give slightly conflicting cues the standard deviation of the localized direction is smaller than when the cues are severely conflicting. This indicates that the localization in the horizontal plane is dependent on both the ITD and the spectral information, since none of the cues provides correct localization independently.

4. DISCUSSION

It was hypothesized that the ITD is the leading cue for localization in the horizontal plane even when the spectral information in the minimum-phase transfer

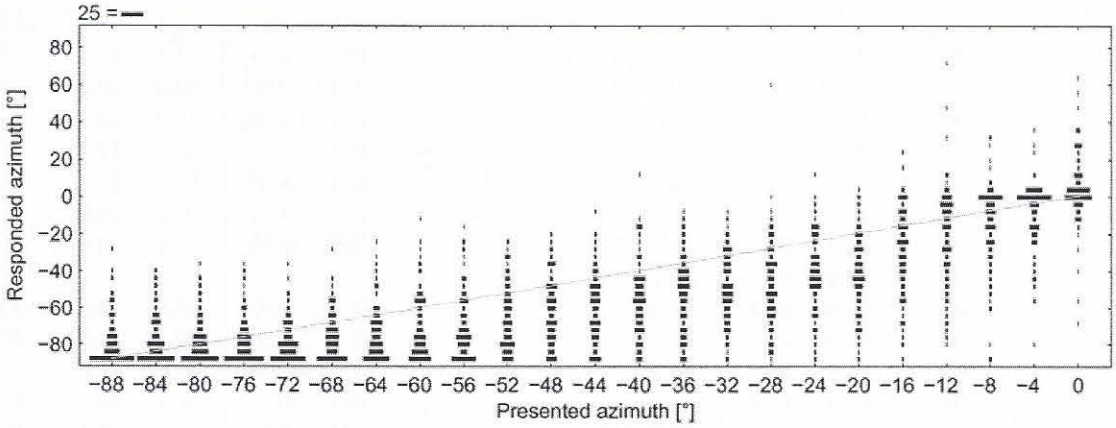

Fig. 5: Data for presented stimuli with original cues. Responses are discretized in 4° steps.

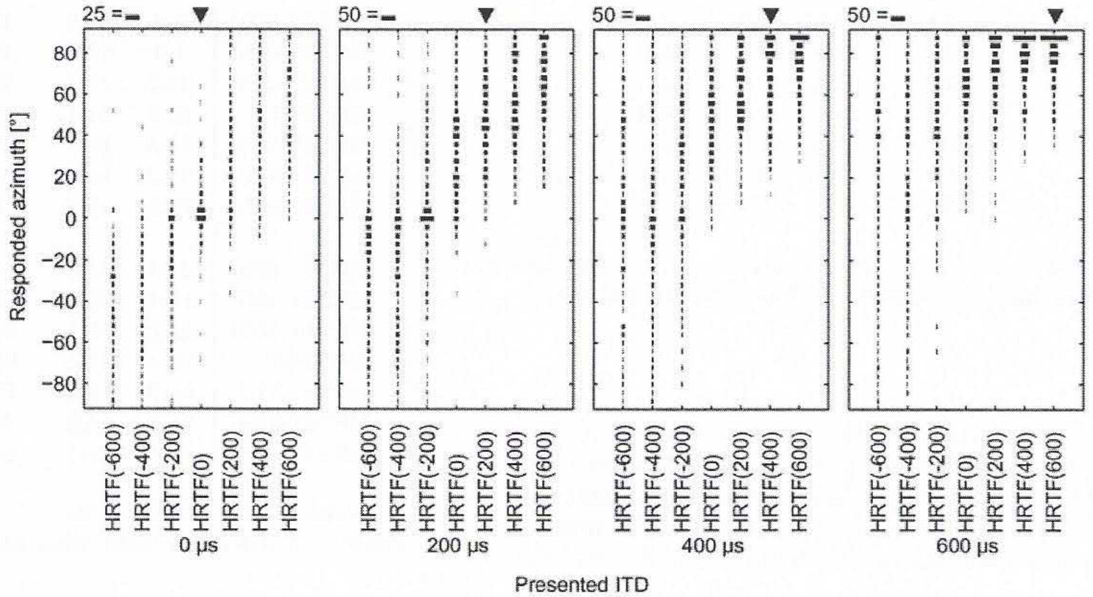

Fig. 6: Data for presented stimuli with manipulated cues. Responses are discretized in steps of 4° . Each of the four subfigures corresponds to all stimuli presented with the specified ITD.

functions gives conflicting cues. It was expected that combinations of ITDs and minimum-phase transfer functions with slightly conflicting cues would lead to localization close to the direction indicated by the ITD rather than the direction indicated by the spectral cues. This is to some extent found in the data from the present experiment, especially when the presented ITD is large. The responded azimuths for an ITD of e.g. 600 μs with an HRTF(400) is not substantially different from those of an ITD of 600 μs with an HRTF(600) although the mean value is sligthly lower and the standard deviation slightly

	[°]		
azimuth	μ	σ	median
0	4.3	17.4	2.1
-4	-8.8	17.3	-1.5
-8	-15.2	22.8	-8.1
-12	-22.0	25.6	-17.6
-16	-29.2	22.3	-26.0
-20	-40.5	20.9	-38.9
-24	-41.5	20.3	-41.5
-28	-47.2	22.0	-49.2
-32	-49.2	18.4	-50.0
-36	-51.9	19.5	-49.5
-40	-55.8	19.8	-56.2
-44	-60.3	18.3	-60.2
-48	-62.3	17.7	-62.7
-52	-66.0	17.1	-68.9
-56	-69.9	17.8	-75.7
-60	-71.2	16.9	-76.8
-64	-73.9	14.8	-77.0
-68	-75.9	12.5	-78.0
-72	-78.6	10.9	-81.5
-76	-78.1	12.3	-80.9
-80	-78.6	11.6	-82.1
-84	-77.0	13.3	-81.1
-88	-77.0	12.7	-84.7

Table 1: Mean (μ), standard deviation (σ) and median values calculated for data with original cues.

higher. For small ITDs, e.g. 0 μs combined with an HRTF from any other direction than the median plane the importance of the spectral cues being correct is clearly seen. Here even a slight change in the spectral cues leads to a completely different responded azimuth as seen in the left subfigure of Fig 6.

When the cues were severely conflicting the localization was expected to be less focused. This is indeed the case as it can be seen in e.g. the left columns in the subfigures with ITDs of 200 μs, 400 μs and 600 μs in Fig. 6. The results do not indicate that the spectral information is the leading cue either and it seems reasonable to conclude that both correct ITD and correct spectral cues are necessary although the importance of the spectral cues is dependent on direction.

can be seen in e.g. the left columns in the subfigures

		[°]		
		μ	σ	median
ITD 0	HRTF(-600)	-41.1	25.1	-42.7
	HRTF(-400)	-38.3	26.2	-40.8
	HRTF(-200)	-17.7	25.1	-15.0
	HRTF(0)	4.3	17.4	2.1
	HRTF(200)	29.9	25.3	28.0
	HRTF(400)	46.2	25.0	46.8
	HRTF(600)	51.0	22.8	51.0
ITD 200	HRTF(-600)	-17.9	33.8	-15.7
	HRTF(-400)	-20.7	32.9	-17.1
	HRTF(-200)	3.4	32.5	1.8
	HRTF(0)	32.8	24.4	33.3
	HRTF(200)	47.3	21.1	46.7
	HRTF(400)	55.2	19.5	55.3
	HRTF(600)	62.7	19.6	64.9
ITD 400	HRTF(-600)	9.7	42.6	8.1
	HRTF(-400)	1.3	43.2	0.3
	HRTF(-200)	26.2	32.7	30.5
	HRTF(0)	49.2	21.1	49.6
	HRTF(200)	61.3	17.9	61.3
	HRTF(400)	65.2	18.6	69.3
	HRTF(600)	72.5	15.2	75.7
ITD 600	HRTF(-600)	22.4	47.8	31.8
	HRTF(-400)	17.4	45.0	19.5
	HRTF(-200)	47.0	30.1	50.5
	HRTF(0)	62.4	18.4	65.0
	HRTF(200)	69.9	17.2	73.5
	HRTF(400)	75.1	14.5	78.9
	HRTF(600)	77.3	13.2	81.3

Table 2: Mean (μ), standard deviation (σ) and median values calculated for data with manipulated cues.

with ITDs of 200 μs, 400 μs and 600 μs in Fig. 6. The results do not indicate that the spectral information is the leading cue either and it seems reasonable to conclude that both correct ITD and correct spectral cues are necessary for localization in the horizontal plane although the importance of the spectral cues is dependent on direction.

A two-step model for localization was described and the results from the experiments give no reason to reject this model. Although the temporal cue might not be as robust as expected for all ITDs, the localized direction is fairly accurate when the two types of cues are supporting each other. It is still believed that the spectral characteristics of the minimum-phase transfer functions contain information about localization on the iso-ITD contour.

5. ACKNOWLEDGMENTS

The authors would like to thank Søren Krarup Olesen for his help with the software used to control the presentation of stimuli and collection of responses and the Danish Technical Research Council for economic support.

6. REFERENCES

[1] J. Blauert. Räumlishes hören. *S. Hirzel Verlag*, 1974.

[2] H. Møller. Fundamentals of binaural technology. *Applied Acoustics*, 36:171–218, 1992.

[3] H. Møller, M. F. Sørensen, D. Hammershøi, and C. B. Jensen. Head-related transfer functions of human subjects. *J. Audio Eng. Soc.*, 43(5):300–321, 1995.

[4] E. M. Wenzel, M. Arruda, D. J. Kistler, and F. L. Wightman. Localization using nonindividualized head-related transfer functions. *J. Acoust. Soc. Am.*, 94(1):111–123, 1993.

[5] J. Blauert. Spatial hearing. *MIT Press*, 1997.

[6] R. A. Butler and K. Belendiuk. Spectral cues utilized in the localization of sound in the median sagittal plane. *J. Acoust. Soc. Am.*, 61(5):1264–1269, 1977.

[7] F. Asano, Y. Suzuki, and T. Sone. Role of spectral cues in median plane localization. *J. Acoust. Soc. Am.*, 88(1):159–168, 1990.

[8] J. C. Middlebrooks. Narrow-band sound localization related to external ear acoustics. *J. Acoust. Soc. Am.*, 92(5):2607–2624, 1992.

[9] E. H. A. Langendijk and A. W. Bronkhorst. Contribution of spectral cues to human sound localization. *J. Acoust. Soc. Am.*, 112(4):1583–1596, 2002.

[10] C. Jin, A. Corderoy, S. Carlile, and A. van Schaik. Contrasting monuaral and interaural spectral cues for humans sound localization. *J. Acoust. Soc. Am.*, 115(6):3124–3141, 2004.

[11] F. L. Wightman and D. J. Kistler. The dominant role of low-frequency interaural time differences in sound localization. *J. Acoust. Soc. Am.*, 91(3):1648–1661, 1992.

[12] A. Kulkarni, S. K. Isabelle, and H. S. Colburn. Sensitivity of human subjects to head-related transfer-function phase spectra. *J. Acoust. Soc. Am.*, 105(5):2821–2840, 1999.

[13] P. Minnaar, J. Plogsties, S. K. Olesen, F. Christensen, and H. Møller. The interaural time difference in binaural synthesis. *Audio Engineering Society Preprint*, AES 108th Convention, Paris(5133), 2000.

[14] D. J. Kistler and F. L. Wightman. A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. *J. Acoust. Soc. Am.*, 91(3):1637–1647, 1992.

[15] B. P. Bovbjerg, F. Christensen, P. Minnaar, and X. Chen. Measuring the head-related transfer functions of an artificial head with a high directional resolution. *Audio Engineering Society Preprint*, AES 109th Convention, Los Angeles(5264), 2000.

[16] P. Minnaar, F. Christensen, S. K. Olesen, B. P. Bovbjerg, and H. Møller. Head-related transfer functions measured with a high directional resolution. *J. Audio Eng. Soc.*, submitted.

[17] A. V. Oppenheim and R. W. Schafer. Discrete-time signal processing - international edition. *Prentice Hall*, 1989.

[18] P. Minnaar, F. Christensen, H. Møller, S. K. Olesen, and J. Plogsties. Audibility off all-pass components in binaural synthesis. *Audio Engineering Society Preprint*, AES 106th Convention, Munich(4911), 1999.

[19] H. Møller, D. Hammershøi, C. B. Jensen, and M. F. Sørensen. Transfer characteristics of headphones measured on human ears. *J. Audio Eng. Soc.*, 43(4):203–217, 1995.

[20] O. Kirkeby, P. A. Nelson, H. Hamada, and F. Orduna-Bustamante. Fast deconvolution of multi-channel systems using regularisation. *IEEE Transactions on Speech and Audio Processing*, 6(2):189–194, 1998.