Acoustics 2008
Geelong, Victoria, Australia 24 to 26 November 2008

Acoustics and Sustainability:

How should acoustics adapt to meet future demands?

Human Sensitivity to Interaural Phase Difference for Very Low Frequency Sound

M. Irfan Jan Mohamed and Densil Cabrera

Faculty of Architecture, Design and Planning, University of Sydney, NSW 2006 Australia

ABSTRACT

Recent studies using subwoofers have provided consistent evidence that localisation along the left-right axis can occur for sound in the frequency range below 100 Hz, and even includes signals in the lowest octave of human hearing (however, front-back localisation fails for low frequency sound). If such left-right localisation is possible, the most likely explanation is a surprisingly acute sensitivity to interaural time or phase difference in the very low frequency range. The present study investigates this hypothesis using stimuli presented via headphones in a quiet anechoic room. Stimulus signals consisted of 1/3-octave noise bands centred on frequencies from 20 Hz -100 Hz with interaural time differences ranging between ±650 microseconds. The stimulus duration was 800 ms and was multiplied by a hanning window resulting in a smooth fade-in and fade-out (with the two channels faded together, regardless of the interaural time difference – hence this might be thought of as a frequency-dependent linear phase shift rather than a simple time difference). Tested on a head and torso simulator, the presentation sound pressure level was 40 dB(A), and distortion and background noise were both negligible. The subjects' task was to identify, on a scale from left to right, the location of the auditory image (i.e. the task was 'lateralisation' rather than localisation). Results show mild lateralisation for frequencies at and above 31.5 Hz with the lateralization of the image becoming clearer in the higher frequencies and the higher time delays across the frequency range tested, and so support the hypothesis.

INTRODUCTION

This study examines the hypothesis that interaural phase difference provides a viable cue for auditory localisation between left and right (more succinctly known as 'lateralisation') in the very low frequency range of human hearing (20 Hz – 100 Hz). The study is prompted by previous work on the localisation of sound from subwoofers, where results showed a surprising ability to localise – surprising because there appears to be a widespread belief that auditory localisation is very poor or indeed non-existent in the very low frequency range. For example, based on listening tests of a variable crossover stereophonic system with subwoofer, Borenius (1985) states that, "…it would seem that there is very little direction information contained in the sound signals below about 200 Hz, and none at all, practically speaking, below about 100 Hz." Similarly, others such as Feldman (1986) and Kügler and Thiele (1992) have advocated the use of a single low frequency channel or subwoofer for practical reasons, with little concern about trading off low frequency spatial reproduction. However, studies by Martens (1999), Braasch *et al.* (2004), Martens *et al.* (2004), Welti (2004), Subkey *et al.* (2005), Jan Mohamed (2007) and Raitio *et al.* (2007) show that multichannel audio in the very low frequency range can affect auditory spatial perception, both in terms of localisation and the width of the auditory image (through decorrelated signals in multiple channels). Localisation in such studies is restricted to left-right (lateralisation), with no ability to make front and back judgments.

With an ear on each side of the head, lateralisation can take advantage of comparisons of the signals arriving at the two ears. Interaural differences may consist of differences in sound pressure level (caused by differences in path length and by the shadowing, reflecting and diffracting effects of the head), and differences in arrival time (caused by differences in path length). By contrast, auditory localisation for angles around the interaural axis (known as 'polar' angles, or else confusingly as 'elevation' angles) relies on other cues such as spectral transformations due to the pinnae. Figure 1 illustrates the geometry of lateral angle (or azimuth) and polar angle, which are equivalent to lines of longitude and latitude respectively for a globe on its side. While listener head movements may allow interaural cues to be used more generally, the vast majority of studies of auditory localisation are concerned with fixed head listening, and furthermore, some listening scenarios (such as much audio reproduction) involve a fairly static head.

The spectral cues associated with localisation around polar angles are restricted to high frequencies. Roffler and Butler (1968) found that veridical vertical localization in the median plane requires complex frequency content above 7 kHz, and that this is not affected by the presence or absence of low-frequency content. Related findings have been made by many others, generally showing that spectral cues above about 5 kHz may be influential for vertical localization in the median plane (e.g., Shaw and Teranishi, 1968; Blauert, 1969/70;

Hebrank and Wright, 1974; Mehrgardt and Mellert, 1977; Asano *et al.*, 1990). Spectral features related to accurate front-back localisation may be in a lower frequency range, but generally require spectral content above 2 kHz. While some weak spectral cues may exist below this (Carlile *et al.* 1999), they are unlikely to exist in any consistent way below 700 Hz (above which the spectral effect of the shoulder reflection may be observed, as shown by Algazi *et al.* (2001)). Hence there is very little prospect for body-related spectral cues to be used in very low frequency localisation, implying that localisation will be restricted to lateralisation in the absence of significant head movements.

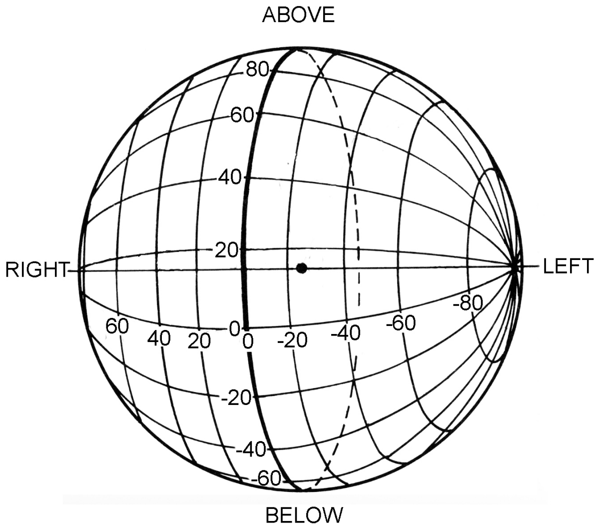

Figure 1. Illustration of a polar coordinate system attuned to the features of auditory localisation. The line from left to right is the *interaural axis*. The horizontal angles (between left and right, or -90 to 90 degrees) are known as *azimuth* angles, and are characterised by distinct interaural cues. The vertical angles (from -180 to 180 degrees at the back) are known as *polar* angles, and are not characterised by distinct interaural cues. The full range of polar angles for the 0 degree azimuth is known as the *median plane*, around which binaural differences are minimal.

The duplex theory of auditory localisation, proposed by Lord Rayleigh (Strutt 1907), is concerned with interaural cues, which are useful for lateralisation. Put briefly, low frequency tones are lateralised using interaural time differences, while high frequency tones are lateralised using interaural level differences. While considerable subtlety has been added to the understanding of auditory localisation since 1907 (c.f. McPherson and Middlebrooks 2002), the concept remains useful as a starting point for understanding lateralisation. We will first consider level difference. As the azimuth angle of a sound stimulus changes, the amount of sound received by each ear will change due in part to changes in the extent to which the contralateral ear is shadowed by the head. With the important exceptions of nearfield sources (Duda and Martens 1998) and reactive soundfields (e.g., in the vicinity of pressure nodes), substantial interaural level differences only occur in natural spatial hearing for relatively high frequency sound. Of course the interaural level differences due to low frequency room modes in small rooms do not provide a viable localisation cue, and so cannot contribute to the learnt ability to localise sound sources from the sound received at the ears. The interaural level differences in low frequencies due to nearfield sources do provide viable localisation cues, but are not relevant to the subwoofer localisation experiments mentioned at the start of this paper because subwoofers are not used in the nearfield. We can conclude that the reported ability of people to localise low frequency sound from subwoofers is unlikely to be related to interaural level differences.

Interaural time differences (ITD) occur due to the difference in path length between the source and each ear. Since the speed of sound is independent of frequency, in simple terms the interaural time difference for a given azimuth is the same across the entire frequency range. The range of interaural time differences for farfield natural hearing is approximately within ±650 µs. For steady state signals, the interaural time difference is analysed as an interaural phase difference (IPD, which is frequency dependent). However, the fixed distance between the two ears affects the possibility of exploiting this cue: for frequencies where a path length difference can be greater than half a wavelength, more than one azimuth angle is associated with a given interaural phase difference. This ambiguity limits the IPD cue to frequencies below 700 Hz. Also, as the wavelength becomes longer, the IPD becomes smaller, so in the very low frequency range the IPD might be too small to exploit. The extent to which this is the case is the question investigated in this paper: with IPD being the only viable non-dynamic cue for subwoofer lateralisation, the question is whether it is sufficient for the task. The hypothesis that IPD provides a viable very low frequency localisation cue is supported by the successful modelling of localisation in this frequency range by Braasch *et al.* (2004) using interaural cross-correlation.

The term lateralisation can have more than one meaning in auditory localisation studies. It may be used to denote left-right pseudo-localisation that occurs for auditory images that are not externalised (but are experienced as inside the head). It could also be used more broadly to denote externalised localisation only with respect to azimuth. Since the present study is concerned with headphone-based production of auditory stimuli with interaural phase difference, the first definition of lateralisation applies. Nevertheless, there is little different between this and the sound received from subwoofers in an absorptive room – the main difference being that head-locking of the soundfield does not occur when loudspeakers are used.

METHOD

The subjective test in this study examines human sensitivity to lateralisation induced solely by interaural phase difference for very low frequency sounds.

Audio Signals and Equipment

Stimulus signals consisted of 1/3-octave noise bands centred on frequencies from 20 Hz to 100 Hz with interaural time differences within ±650 µs. The seven selected time delays are ±650 µs, ±501 µs, ±390 µs, 0 µs. The initial two-channel file with the relevant centre frequency was time shifted on either the left or right channel then cropped to yield a two channel steady state signal of 800 ms. A 10th order Chebychev low pass filter with a cut off frequency of 200 Hz was applied to all the stimuli to filter out any high frequency noise in the recordings. This reduced the remote possibility of coherent high frequency content in the stimuli providing a lateralisation cue. Each 800 ms stimulus was multiplied by a hanning window (i.e. a raised half-sine) resulting in a smooth fade in and fade out (with the two channels faded together, regardless of the interaural time difference – hence this might be thought of as a frequency-dependent linear phase shift rather than a simple time difference). A similar approach was used by Schiano *et al.* (1986), although for higher frequency signals (300 Hz and above). This process is illustrated in Figure 2.

Headphones (Sennheiser HD600) were used for reproducing the stimuli. The reproduction sound pressure level of the stimuli was 40 dB(A), as measured on each artificial ear of a Bruel & Kjaer type 4128C head and torso simulator. This level refers to the $L_{A,eq}$ of the 800 ms period. The gains of each channel contributing to the headphones were adjusted to have a negligible (<0.1 dB) difference between left and right. Although a sound pressure level of 40 dB(A) may appear weak, in the very low frequency range it represents substantial sound pressure levels (92 dB at 20 Hz, 84 dB at 25 Hz, down to 52 dB at 100 Hz). Hence this low A-weighted sound pressure level was chosen not only to avoid distortion in the headphones, but also to avoid exposing subjects to sound pressure levels that could be interpreted as being excessive. It must be noted, however, that subjectively the stimuli were all quiet to very quiet.

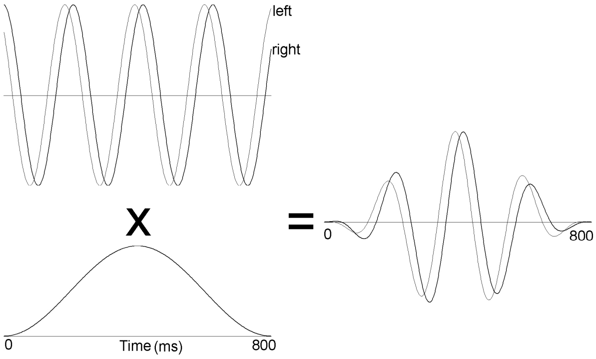

Figure 2. The process used to generate sound stimuli that have an interaural delay in the fine structure, but no delay in the envelope. A steady state wave with an interaural delay (top left) is multiplied by a hanning window function (bottom left) to produce the stimulus signal (right). In this example, the left channel is leading the right, and a cosine tone is shown instead of the noise band signals of the experiment.

The study was conducted in an anechoic room, and the main purpose of using this room was to provide a low background noise environment to avoid masking (or distraction from) the stimuli. Raitio *et al.* (2007) show that low frequency lateralisation can be particularly sensitive to interference from background noise. Background noise measurements were made with a Bruel & Kjaer type 2250 sound level meter at the listening position. Figure 3, which shows these measurements, together with the stimulus measurements, confirms that the background noise was likely to be inaudible, and far below the sound pressure levels of the stimuli. Note that while this anechoic room is by no means anechoic in the very low frequency range (it has an anechoic cut-off frequency of 200 Hz), its anechoic performance should not affect an experiment using headphone-produced stimuli.

The fine spectral content of the stimuli was measured using a 65536-pt fast Fourier transform (FFT) of the stimuli downsampled to 6 kHz, and the sound pressure levels measured from this process are shown in Figure 4. This shows that there is little spectral content beyond the intended frequency range of each stimulus.

Experiment control and graphical interface

The full stimulus set had two parameters: centre frequency and interaural time difference. There were eight centre frequencies (comprising 20, 25, 31.5, 40, 50, 63, 80 and 100 Hz) and seven interaural time delays, totaling 112 stimuli (8 x 2 x 7). The playback system utilised a Max/MSP patch that played each stimulus (in a random order) and recorded the listener's response (of the lateral position of the auditory image). Listeners could listen repeatedly to each stimulus in attempting to lateralise the auditory event. The response was given by adjusting on a slider on a computer touch screen. A response was required for each stimulus, but if the stimulus could not be heard, the subject could return a response using the 'Inaudible' button. The experiment interface is shown in Figure 5.

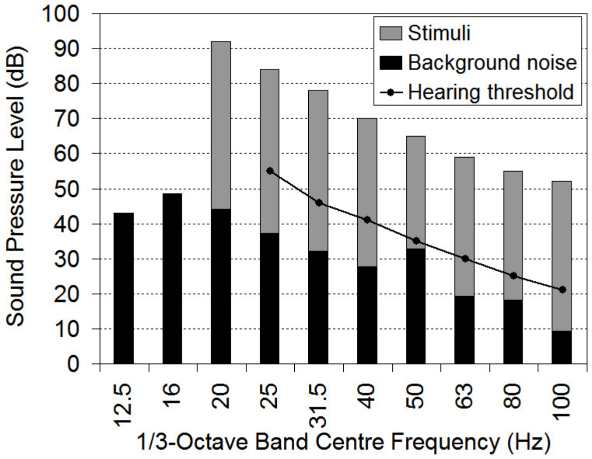

Figure 3. 1/3-octave band sound pressure levels of the background noise and stimuli in the anechoic room, also showing the hearing threshold curve from ANSI S12.2:1995.

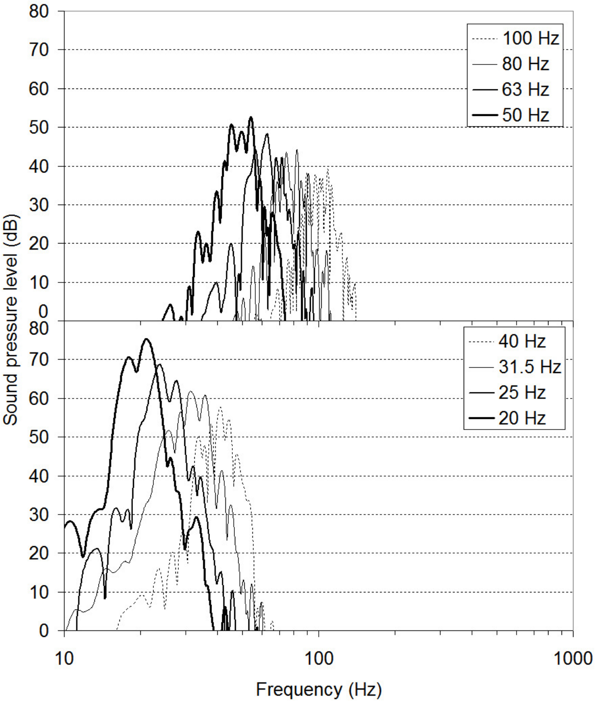

Figure 4. Fine spectral analysis of the stimuli: downsampled to a sampling rate of 6 kHz prior to 65536-pt FFT using a Blackman-Harris windowing function.

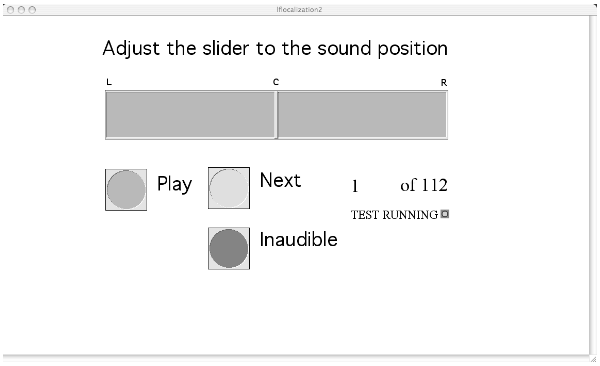
Adjust the slider to the sound position

Figure 5. The graphical user interface used for by the experiment subjects.

Thirteen subjects participated in the test (1 female), with ages between 20 and 40. All listeners were briefed prior to beginning the experiment on the placement of headphones, using the Max/MSP patch and the concept of locating auditory images on the left-right axis.

RESULTS

Figure 6 shows the mean response data for each of the stimuli. A positive slope across the range of ITDs is an indicator of some localisation ability for a given frequency band. The only band that does not have an overall positive slope is the 25 Hz band. However, the slopes in the other lower frequency bands are much shallower than that of 100 Hz. This indicates that ITD has less effect on lateralisation, or at least a less consistent effect on lateralisation, in the lower part of the frequency range tested.

Although it might be hoped that an ITD of ±650 µs would yield a highly lateralised response, the limited response scale means that a spread of responses will average out to a value closer to centre, because it is impossible to record a response beyond the left or right extremes of the scale. Since this 'ceiling effect' produces a central tendency in scattered responses as well as limiting the spread of responses, the greater slope for 100 Hz compared to lower frequency bands could reflect a mixture of greater certainty in response and greater lateralisation effect. One way of examining the difference between these influences is to consider the range of responses (e.g., standard deviation or standard error) for the 0 µs ITD stimuli, which, with the exception of 20 Hz and 80 Hz, exhibits a decline with frequency. This confirms that the results reflect greater certainty in response as frequency increases.

The strength and significance of the effect of ITD was assessed using factorial analysis of variance (ANOVA) for each frequency band. The F-values (which represent the absolute strength of the effect), P-values (the probability that the results confirm the null hypothesis, for which values less than or equal to 0.05 are conventionally considered to signify significance) and the proportion of response variance related to the effect of ITD are shown in Table 1, This confirms that significant results are found for all of the frequency bands apart from 25 Hz, although the effect is weak for the lower bands.

Another way of looking at the extent of the effect is to examine the significance of mean differences between pairs of stimuli using a *post hoc* test. With seven stimuli in each 1/3-octave band, there are 21 possible pair combinations. The Tukey-Kramer test was used, with results shown in Table 2.

Table 1. Strength and significance of the effect of ITD derived from ANOVA (df =6) for each 1/3-octave band**.**

	F-Value	P-Value	Proportion of Variance
20 Hz	6.649	<.0001	0.33
25 Hz	0.768	0.768	0.04
31.5 Hz	2.463	0.0305	0.15
40 Hz	4.495	0.0005	0.24
50 Hz	4.468	0.0006	0.24
63 Hz	4.161	0.001	0.23
80 Hz	9.33	<.0001	0.40
100 Hz	19.969	<.0001	0.59

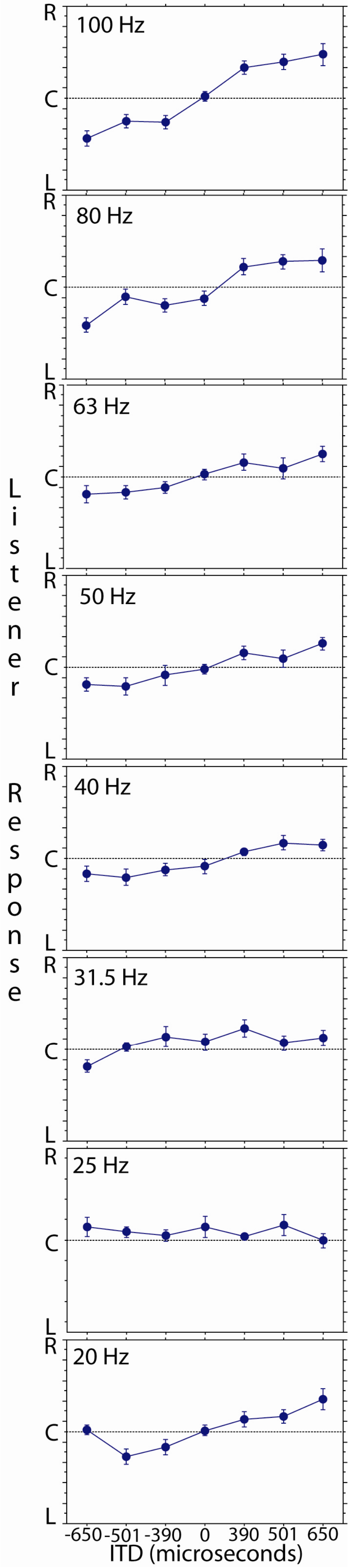

Figure 6. Mean subjective responses (±1 standard error).

Table 2. Significant differences between lateralisation responses for pairs of stimuli within each frequency band. An 'S' denotes a significant difference (95% or greater confidence) using the Tukey-Kramer Honestly Significant Difference *post hoc* test.

100 Hz	-650	-501	-390	0	390	501
-501						
-390						
0	S					
390	S	S	S			
501	S	S	S	S		
650	S	S	S	S		

80 Hz	-650	-501	-390	0	390	501
-501						
-390						
0						
390	S		S			
501	S		S	S		
650	S	S	S	S		

63 Hz	-650	-501	-390	0	390	501
-501						
-390						
0						
390						
501						
650	S	S	S			

50 Hz	-650	-501	-390	0	390	501
-501						
-390						
0						
390		S				
501						
650	S	S				

40 Hz	-650	-501	-390	0	390	501
-501						
-390						
0						
390						
501	S	S				
650	S	S				

31.5 Hz	-650	-501	-390	0	390	501
-501						
-390						
0						
390	S					
501						
650						

25 Hz	-650	-501	-390	0	390	501
-501						
-390						
0						
390						
501						
650						

20 Hz	-650	-501	-390	0	390	501
-501						
-390						
0						
390		S				
501		S				
650		S	S			

DISCUSSION AND CONCLUSION

A common sceptical response to experiment results showing lateralisation at very low frequencies is that perhaps the signal contained some unintended high frequency content. Such content might be due to the onset and offset of the stimulus envelope, or due to non-linear distortion of audio components (such as the loudspeaker driver, or 'chuffing' due to air turbulence in loudspeaker ports). It is relatively simple to avoid these problems by proper design of stimulus envelopes and the appropriate selection and use of audio components. Nevertheless, it is difficult to prove to a sceptic that such problems were entirely avoided. An advantage of the present experiment's use of headphone reproduction is that should any non-linearities exist, they may reveal nothing about the interaural phase difference (so long as the non-linearities do not possess corresponding phase differences). Needless to say, we ensured that distortion was not audible. The potential introduction of high frequency content due to the signal envelope was well contained by using a very smooth envelope, namely a hanning window function, over a relatively long duration signal. The data in Figure 4 describe the actual (rather than nominal) spectral content of the stimuli.

The experiment results confirm that lateralisation based on ITD or IPD can occur in the very low frequency range, although the strength of the effect weakens as the frequency descends. No effect was observed at 25 Hz, but a surprisingly strong effect was observed at 20 Hz. It is not clear whether the 20 Hz response is due to an experiment artefact (which perhaps is suggested by the fact that it is so different to 25 Hz), and until this can be checked through further experimentation we do not conclude that lateralisation occurs for 1/3-octave band stimuli at 20 Hz. However, results at and above 31.5 Hz form a straightforward pattern of increasing lateralisation sensitivity as frequency increases, and so we do conclude that lateralisation is possible at frequencies as low as 31.5 Hz (albeit marginal) and 40 Hz.

The commonly held view, that localisation acuity is poor in the very low frequency range, is confirmed (at least for IPD as a cue), but this study confirms that at least some lateralisation is possible in this frequency range. The findings from this study support the conclusion by Braasch et al. that interaural phase difference is the cue or mechanism for lateralisation of low frequencies. It also reinforces conclusions of previous subwoofer-based experiments of potential benefits in spatial audio rendering from multiple subwoofers, or more specifically, a pair of subwoofers on the left and right of the listener.

ACKNOWLEDGEMENTS

The authors thank the experiment volunteers, as well as those who helped with other aspects of this study including Kenneth Stewart, Luis Miranda, Mark Fulgado, Matt Storey and Trishay Trada for their assistance and counsel.

REFERENCES

Algazi, V. R., Avendano, C. and Duda, R. O. 2001, "Elevation localization and head-related transfer function analysis at low-frequencies," *J. Acoust. Soc. Am.* 109, 1110–1122

Asano, F., Suzuki, Y. and Sone, T. 1990, "Role of spectral cues in median plane localization," *J. Acoust. Soc. Am. 88* 159-168

Blauert, J. 1969/70, "Sound localization in the median plane," *Acustica* 22, 205-213

Borenius, J. 1985, "Perceptibility of direction and time delay errors in subwoofer reproduction," *79th Audio Eng. Soc. Conv.*, New York, USA

Braasch, J., Martens, W.L. and Woszczyk, W. 2004, "Modelling auditory localization of subwoofer signals in multichannel loudspeaker arrays," *117th Audio Eng. Soc. Conv.*, San Francisco, USA

Carlile, S., Delaney, S., and Corderoy, A. 1999, "The localization of spectrally restricted sounds by human listeners," *Hearing Research* 128, 175-189

Duda, R.O. and Martens, W.L. 1998 "Range dependence of the response of an ideal rigid sphere to a point sound source," *J. Acoust. Soc. Am.* 105(5), 3048-3058

Feldman, S. 1986, "Preparation of master tapes: a producer's and mixer's guide to mastering for analog disks, cassettes and compact discs," *J. Audio Eng. Soc.* 34, 914-940.

Hebrank, J. and Wright, D. 1974, "Spectral cues used in the localization of sound sources on the median plane," *J. Acoust. Soc. Am.* 56, 1829-1834

Jan Mohamed, M. I. 2007, "Low frequency localization on the horizontal plane," unpublished report, Faculty of Architecture, Design and Planning, University of Sydney.

Kügler, C. and G. Thiele, G. 1992, "Loudspeaker reproduction: study of the subwoofer concept," *92nd Audio Eng. Soc. Conv.*, Vienna, Austria

Martens, W.L, Braasch, J., and Woszczyk, W. 2004, "Identification and discrimination of listener envelopment percepts associate with multiple low-frequency signals in multichannel sound reproduction," *117th Audio Eng. Soc. Conv.*, San Francisco, USA

Martens, W.L. 1999 "The impact of decorrelated low-frequency reproduction on auditory spatial imagery: are two subwoofers better than one?" *16th Audio Eng. Soc. Int. Conf.*, Rovaniemi, Finland

McPherson, E.A. and Middlebrooks, J.C. 2002, "Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited," *J. Acoust. Soc. Am.* 111(5), 2219-2236.

Mehrgardt S, Mellert V. 1977, "Transformation characteristics of the external human ear," *J. Acoust. Soc. Am.* 61, 1567-1576

Raitio, T., Lehtonen, H.-M., Laine, P. and Pulkki, V. 2007, "Detection and lateralization of sinusoidal signals in the presence of dichotic pink noise," *122nd Audio Eng. Soc. Conv.*, Vienna, Austria.

Roffler, S. K., and Butler, R. A. 1968, "Factors that influence the localization of sound in the vertical plane," *J. Acoust. Soc. Am.* 43, 1255-1259

Schiano, J.L., Trahiotis, C. and Bernstein, L.R. 1986, "Lateralization of low frequency tones and narrow bands of noise," *J. Acoust. Soc. Am.*, 79, 1563-1570

Shaw, E.A.G. and Teranishi, R. 1968, "Sound pressure generated in an external-ear replica and real human ears by a nearby point source," *J. Acoust. Soc. Am.* 44, 240-249

Strutt, J. W. 1907 ‘‘On our perception of sound direction,’’ Philos. Mag. 13, 214–232.

Welti, T.S. 2004, "Subjective comparison of single channel versus two channel subwoofer reproduction," *117th Audio Eng. Soc. Conv.*, San Francisco, USA