The Impact of Signal Bandwidth and Generic Head-Related Transfer Functions for Localization on the Horizontal Plane in Virtual Acoustic Space

G. Robert Arrabito, R. Brian Crabtree, Sharon M. McFadden and Sherry Van Blyderveen
DCIEM, 1133 Sheppard Ave. W., P.O. Box 2000, Toronto, ON M3M 3B9

INTRODUCTION

Auditory localization refers to the ability to determine the location of a source of sound solely on the presence of acoustic information. The localization of sound sources depends on the encoding of interaural differences in the level and time of arrival of the sound at the two ears in addition to spectral and temporal information contributed by the pinnae, head, and upper torso. A common phenomenon in auditory localization is for the listener to make front/back reversals, i.e., perceiving the mirror image of the sound source. In order to reduce front/back reversals, auditory signals must be broadband and contain spectral energy above approximately 3 kHz when presented in the free-field.

The use of directional auditory cuing via a three-dimensional (3-d) audio display to pilots in military aircraft could potentially increase pilot performance by off-loading the visual modality. In order to synthesize the location of a sound in virtual auditory space over headphones, digital filters are used to implement head-related transfer functions (HRTFs) which have been measured from humans or acoustic mannequins for many sound source positions in the free-field. The physical differences between individuals cause HRTFs to differ from person to person. Upon playback of a spatial signal over headphones, the HRTFs could be either those measured from the listener's head ("individualized") or those measured from another head ("generic").

Due to the limited bandwidth of aircraft communication systems, one would expect that front/back reversals would be increased. Such confusions could have potentially serious consequences in practical applications. An experiment in progress is examining auditory localization on the horizontal plane in a 3-d audio display using different signal bandwidths and generic HRTFs. A subsequent auditory localization experiment in the free-field will assist to psychoacoustically validate the results obtained in virtual auditory space.

METHOD

Participants

To date, three women and two men have been tested in a pilot study. The mean age was 33 years. A Békésy audiometric test confirmed that all participants had no more than a 20 dB bilateral hearing loss at any frequency between 125 Hz and 8 kHz.

Stimuli and Apparatus

The stimuli were low pass 3 kHz, high pass 3 kHz, and low pass 14 kHz white noise. These were chosen to allow an assessment of the effectiveness of binaural and spectral cues. Seven different generic HRTFs were used to filter the stimuli in order to synthesize the location of the sound in virtual auditory space.

Participants were tested individually in an IAC sound-attenuating listening booth. The Tucker-Davis Technologies (TDT) PD1 was used to spatialize the stimulus in real-time in conjunction with custom software. The TDT system is a suite of audio equipment (e.g., filters, analog-to-digital converters, digital-to-analogue converters, attenuators, convolvers), under computer control, which is responsible for creating and presenting sound to observers, and for collecting and storing their responses. The stimulus was presented over Stax electrostatic headphones (model SR-Λ Signature) at a sound level of approximately 71 dB(A). A localization judgement was made by pressing a button on a response box whose buttons were arranged in the same configuration as the virtual speaker array.

Procedure

The participant's task was to identify the perceived location of the stimulus that had a 300 ms duration. The stimulus was presented at one of eight static virtual positions on the horizontal plane at 45 degree intervals with the 0 degree azimuth position located in front of the participant. Each block consisted of 8 practice trials followed by 40 experimental trials, with each position used once in the practice trials and five times in the experimental trials. Each trial began by flashing a 0.5 second warning light on the wall in front of the participant followed by a 0.5 second delay prior to the presentation of the stimulus. A maximum of ten seconds was given to make a response corresponding to the virtual speaker that had emitted the stimulus. Each block was comprised of one of the three stimuli and one of seven HRTF conditions. The order of trials and blocks was presented according to a balanced Latin square design. A session contained 21 blocks. Participants completed four sessions each on separate days. No feedback was given to the participants as to the correctness of the localization judgements.

RESULTS AND DISCUSSION

A preliminary analysis of the pilot data suggests that localization performance measured by percent correct, type of reversal (e.g., front/back, left/right, and diagonal), and response time was not significantly affected by the choice of HRTF or stimulus. These findings were not expected. One might expect that performance would be significantly affected by the process of measuring the HRTFs, as measurement techniques differ across laboratories and are motivated by the different goals of the investigators (see references 2, 3, 5-27 of (Moller, Sorensen, Hammershoi & Jensen, 1995)). Some of the parameters that vary significantly in the measurement of HRTFs are type of test stimulus (e.g., sinusoidal tones or noise bursts), the point in relation to the ear canal where the measurement is made (e.g., at the blocked ear canal or a point somewhere along the ear canal), and the number of source positions. For example, Wightman and Kistler (1989) measured 144 source positions while Bronkhorst (1995) measured 967 source positions. A smaller number of measured sound source positions requires that the between sources interpolation procedure be more precise. The seven HRTFs used in this study were each measured differently. As for the choice of stimulus, one would expect that the low pass 3 kHz and 14 kHz stimulus would be significantly different from the high pass 3 kHz stimulus due to the dominant role of interaural time differences (ITD) (Wightman & Kistler, 1992). In that study Wightman and Kistler (1992) reported that when lower frequency ITD cues are present in the stimulus they override interaural level difference (ILD) and spectral shape cues present in low and high frequencies.

Similar results to those reported in this study were obtained by Bronkhorst (1995) who examined auditory localization in the free-field and under virtual conditions as a function of the cutoff frequency of the stimulus. He reported that front/back reversals under virtual presentation on the horizontal plane did not vary significantly as a function of the cutoff frequency regardless of the choice of HRTFs, (individualized or generic). However, under free-field presentation on the horizontal plane, Bronkhorst (1995) found that the number of front/back reversals was statistically affected by the cutoff frequency. This latter finding was also reported by King and Oldfield (1997). The implications of our findings on auditory localization performance in the free-field will be reported as further data become available.

REFERENCES

Bronkhorst, A.W. (1995). Localization of real and virtual sound sources. J Acoust Soc Am, 98(5), 2542–53.

King, R.B. and Oldfield, S.R. (1997). The impact of signal bandwidth on auditory localization: Implications for the design of 3-dimensional audio displays. Hum Fact, 39(2), 287–95.

Moller, H., Sorensen, M.F., Hammershoi, D. and Jensen, C.B. (1995). Head-related transfer functions of human subjects J Aud Eng Soc,43(5), 300–21.

Wightman, F.L. and Kistler, D.J. (1992). The dominant role of low-frequency interaural time differences in sound localization. J Acoust Soc Am, 91(3), 1648–61.

Wightman, F.L. and Kistler, D.J. (1989). Headphone simulation of free-field listening part I: Stimulus synthesis. J Acoust Soc Am, 85(2), 858–67.