and Dotson, 1993) suggest that for aircraft applications (i.e., presenting a warning message) that 3-D audio resolution may need to be as little as five degrees to be useful.

REFERENCES

1. Ericson, M. A., and McKinley, R. L. (1989). Auditory localization cue synthesis and human performance. In Proceedings of the IEEE National Aerospace Electronics Conference. New York: Institute for Electrical and Electronics Engineers. pp. 718-725.

2. Lee, M. D., Patterson, R. W., Folds, D. J., and Dotson, D. A. (May, 1993). Application of simulated three-dimensional audio displays to fighter cockpits: A user survey. Proceedings of the IEEE National Aerospace Electronics Conference. New York: Institute for Electrical and Electronics Engineers.

3. Valencia, G., and Agnew, J. R. (1990). Evaluation of a directional audio display synthesizer. Proceeding of the Human Factors Society 34th Annual Meeting. Human Factors Society: Santa Monica, California.

Model 656 provided sixteen-bit, two channel digital to analog and analog to digital conversion. Input from a Polhemus Isotrak headtracking device was used to correct the stereo output for head movements. High quality Sennheiser HD540 stereo headphones were used in the present study. The 3-D audio system combined to present the sounds to the subject, and to introduce localization cues. The SPARCstation was also used to collect subject responses. Subjects were seated facing a wall. Large black numbers (point size 360) and the word "level" printed on white, 8.5' by 11' inch sheets of paper were pasted on the wall at approximately eye level. The numbers 11, 1, 3, and 9 will be used to represent the 330, 30, 90, and 270 degree angles respectively. The words high level and low were also pasted on the wall, about 45 degrees above and below the subjects line of sight.

Procedure

Subjects read a set of written instructions and gave informed consent. The instructions explained the experimental protocol, and described how the head tracker was being used to correct the sound location for head movements. All subjects were encouraged to move their heads to help them localize the sounds. Subjects were seated in a chair and asked to wear the headphones. Two sample sounds were displayed, one to the subjects left (270 degrees azimuth, level elevation), and one to the subjects right (90 degrees azimuth, level elevation). Subjects were informed of the intended apparent location of the sample sounds. Subjects were permitted to adjust the sound level of the samples, and from that point on the level was fixed.

On each trial, the subject was presented with the sound, and then indicated it's apparent location in space by clicking a mouse on a button bar icon. The button bar icons were arranged in a circular pattern on the SPARCstation monitor, which approximated the layout of a clockface. An additional response button was located in the center of the circle, so that subjects could indicate the a sound appeared to come from directly above. Elevation buttons, indicating a high, level, or low response were available, arranged vertically in a column immediately to the right of the clockface. Subjects were instructed to use only the possible answers provided. After completion of the 44 trials, subjects were debriefed as to the nature, and purpose of the experiment. The duration of the entire study was approximately 30 minutes.

Design

The independent variables in this experiment is the intended apparent location of the sound source (within-subjects), and the signal bandwidth (between- subjects). Three signal bandwidths were used (10 kHz, 12kHz, and 16 kHz). There were 19 possible locations constructed from the six azimuths (30, 90, 150, 210, 270, and 330 degrees) and three elevations (eye level, 48 degrees above and 48 degrees below eye level). The final location was from 84 degrees in elevation, and zero degrees azimuth (directly in front). The sound stimuli were presented three times at each azimuth with the level elevation, two times at each azimuth for the high and low elevations, and twice from straight above for a total of 44 trials. The dependent measure was percent correct.

RESULTS

To date the data have only been summarized at the mean level. The average accuracy for each bandwidth is presented below in Table 1.

Table 1: Percent accuracy for each bandwidth group, for azimuth only, elevation only, or accuracy and elevation on the same trial.

Test Condition	Cor. Azimuth	Cor. Elevation	Cor. Total Position
10kHz stimulus	49%	37%	22%
12kHz stimulus	48%	38%	23%
16kHz stimulus	46%	42%	23%

DISCUSSION

Using current technology, it appears that localization accuracy is quite poor. Subjects only got about half of the trials correct in azimuth and slightly more than a third of the trials correct in elevation. This conflicts with previous studies that have used response terminated noise durations, and different accuracy measures. The present data also show that the addition of higher frequency sounds has little impact on azimuth accuracy, however there is a tendency for the addition of high frequencies to increase elevation accuracy. Further analyses will be directed toward assessing the statistical significance of the present finding, as well as assessing accuracy as a function of azimuth and elevation angle.

The present data have both practical and theoretical significance. Three-dimensional audio devices are designed based on theories about which cues are important for accurate auditory localization. While we have reasons to believe the hardware and software are providing subjects with the cues we ask, it appears that these cues as used in current systems are insufficient to provide good spatial resolution. This indicates the need for more basic research into auditory localization cues.

From a practical perspective, the data suggest that high frequencies contribute little to elevation accuracy. This information is important because the usable range of sound frequencies can limits the design of aural signals. In addition, this issue has serious implications for the engineering of 3-D audio systems. Of the many characteristics of digital sound that may impact the engineering of 3-D audio systems, required signal bandwidth is one of the most critical issues in determining system sample rate. Of course, as the accuracy of 3-D audio systems improves, this issue will need to be revisited because the very cues that improve the quality of 3-D audio systems, may involve the use of high frequency sound components.

The present results also suggest that 3-D audio technology may not be accurate enough for cockpit applications at the present time. Interviews with pilots (Lee, Patterson, Folds,

The Perception of Location Using Synthetic Auditory Localization Cues: Accuracy and the Effects of Stimulus Bandwidth

Mark D. Lee and David A. Burgess

Georgia Institute of Technology
mlee@psy.gatech.edu, burgess@cc.gatech.edu

ABSTRACT

The present experiment was designed to examine the accuracy with which people can localize sounds generated by three-dimensional or 3-D audio devices. Sixty subjects listened to two second pulse trains of white noise, using bandwidths of either 10 kHz, 12 kHz, or 16 kHz. A forced choice paradigm was used in which subject indicated that sound came from one of six azimuth locations (30, 90, 150, 210, 270, or 330 degrees) and one of three elevation locations (-48, zero or level, and 48 degrees). Stimuli could also come from directly overhead. In all nineteen possible locations were used (six azimuth by three elevation, plus straight up). Results indicated that with current technology and theoretical knowledge about auditory localization, performance is quite poor. Results are discussed in terms of the application of 3-D audio technology to aircraft cockpits.

KEYWORDS

Auditory localization, Auditory Displays, Stimulus bandwidth.

INTRODUCTION

Devices which alter sound signals to make them appear as if they originate from various locations, despite the fact that the sounds are actually delivered over headphones, are referred to as three-dimensional or 3-D audio devices. Most 3-D audio systems are actually two-dimensional, allowing the user to manipulate only elevation and azimuth of the sound source, and assuming a constant distance. 3-D audio devices typically produce the perception of location by introducing interaural time and intensity differences in the sound which arrives at each ear. More sophisticated systems transform sounds to more closely reflect the effects of pinna, shoulder, and torso reflections.

The application of such technology to fighter aircraft could be very beneficial. Some obvious applications for 3-D audio technology include threat warning messages, and communications tasks (Lee, Patterson, Folds, and Dotson, 1991). Adding the location dimension to a sound may provide additional information without increasing a pilot's cognitive load, enhance a pilot's "situational awareness", and increase survivability. In advance of designing a 3-D audio subsystem, a myriad of issues need to be resolved.

The purpose of the present experiment is to examine two of these issues. First, we need to determine how well people can use simulated auditory localization cues, as provided by current technology, to localize sound sources presented over headphones. This issue has only been examined in paradigms that use response terminated stimuli and/or mean accuracy difference measures, which may change the nature of the localization task (Ericson and Mckinley, 1989; Valencia and Agnew, 1990). In addition, the perception of elevation may be critically dependent upon the high frequency components of the sound stimulus (Butler and Helwig, 1983). Therefore the effects of signal bandwidth were examined to determine the importance of adding high frequency bands to artificially localized stimuli.

The present experiment asked subjects to localize short duration white noise sound sources in azimuth and in elevation. Bandwidth of the white noise was manipulated as a between-groups variable.

METHOD

Subjects.

Sixty volunteers from the Georgia Institute of Technology's, School of Psychology undergraduate subject pool served as subjects. Subjects were given experimental credit for their participation. Sample characteristics such as age and sex have not yet been tabulated.

Stimulus

Two second long pulse trains of uniformly distributed white noise served as stimuli in the present experiment. Pulse trains were constructed from either 10kHz, 12 kHz, or 16 kHz bandwidth white noise, with an interruption rate of 4 Hz, and an 80% duty cycle.The white noise was turned on for 200 ms, then turned off for 50 ms. This on-off sequence was repeated eight times, so the total duration of the stimulus was two seconds. The sound pressure level of the noise stimulus was adjusted to a comfortable level by the subject.

Apparatus

The 3-D audio system for the present experiment was constructed from a SUN Microsystems SPARCstation IPX with a Ariel S-56X DSP coprocessor board. An Ariel ProPort