AAMRL-TR-90-025

LOCALIZATION PERFORMANCE WITH SYNTHESIZED DIRECTIONAL AUDIO (U)

GERMAN VALENCIA
GLORIA L. CALHOUN
MARK A. ERICSON

ARMSTRONG AEROSPACE MEDICAL RESEARCH LABORATORY

JEFFREY R. AGNEW

MACAULAY-BROWN, INC.
DAYTON OH

JULY 1990

INTERIM REPORT FOR PERIOD JANUARY - SEPTEMBER 1989

Approved for public release; distribution is unlimited

ARMSTRONG AEROSPACE MEDICAL RESEARCH LABORATORY
HUMAN SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-6573

NOTICES

TECHNICAL REVIEW AND APPROVAL

AAMRL-TR-90-025

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

The voluntary informed consent of the subjects used in this research was obtained as required by Air Force Regulation 169-3.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

CHARLES BATES, JR.
Director, Human Engineering Division
Armstrong Aerospace Medical Research Laboratory

REPORT DOCUMENTATION PAGE

Form Approved OMB No 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE July 1990

3. REPORT TYPE AND DATES COVERED INTERIM, JANUARY-SEPTEMBER 1989

4. TITLE AND SUBTITLE
Localization Performance with Synthesized Directional Audio (U)

5. FUNDING NUMBERS
62202F
718426D3

6. AUTHOR(S)
Valencia, German
Calhoun, Gloria L.
Agnew, Jeffrey R.
Ericson, Mark A.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Armstrong Aerospace Medical Research Laboratory
AAMRL/HEA
Wright-Patterson AFB OH 45433-6573

8. PERFORMING ORGANIZATION REPORT NUMBER
AAMRL-TR-90-025

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

12b. DISTRIBUTION CODE
A

13. ABSTRACT *(Maximum 200 words)*
This report summarizes three studies designed to measure and compare the ability of subjects to localize sounds in azimuth, via headphones, generated by two prototype auditory localization cue synthesizers. In the first study, performance differences were found between the two synthesizers in certain areas of the azimuth plane. Additionally, the design of a synthesizer (e.g., resolution and interpolating between head-related transfer functions (HRTFs) can impact the perceived direction of the acoustic signals. Previous research with directional audio suggests that the verdicality of 3-D auditory displays could be optimized if individualized HRTFs are employed to synthesize the virtual sound sources, particularly in elevation. However, data from this experiment suggest that this design requirement can be relaxed, especially if only azimuth information is to be conveyed by the localization synthesizer. In the second study, two response methods for measuring localization performance were evaluated. No performance differences were found when subjects either verbally reported angular estimates or pointed to a circle to indicate the perceived direction of the target stimuli. In the third study, performance was impacted by manipulating the bandwidth of the acoustic signal and head movement.

14. SUBJECT TERMS
3-D Auditory Display
Auditory Cue Synthesizer
Acoustic Localization
Directional Audio
KEMAR Manikin
Psychoacoustics

15. NUMBER OF PAGES 61

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED

NSN 7540-01-280-5500

Standard Form 298 (Rev 2-89)
Prescribed by ANSI Std Z39-18
298-102

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow It is important to ***stay within the lines*** to meet ***optical scanning requirements***.

Block 1 Agency Use Only *(Leave blank)*.

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C -	Contract	**PR** -	Project
G -	Grant	**TA** -	Task
PE -	Program Element	**WU** -	Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *(If known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as Prepared in cooperation with ... ; Trans. of ...; To be published in ... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."
DOE - See authorities.
NASA - See Handbook NHB 2200.2.
NTIS - Leave blank.

Block 12b. Distribution Code.

DOD - Leave blank.
DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
NASA - Leave blank.
NTIS - Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*.

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

PREFACE

This report is the result of a joint research effort performed between the Biological Acoustics Branch (BBA) and the Visual Display Systems Branch at the Human Engineering (HE) Division, Armstrong Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio. This effort was accomplished under Work Unit Nos. 718426D3 (AAMRL/HEA) and 72313901 (AAMRL/BBA) between January and September 1989. Portions of the work reported herein were performed under Contract Number F33615-87-C-0534, MacAulay-Brown, Inc. The authors are grateful to Col Charles P. Hatsell, Laboratory Commander, Mr James W. Brinkley, Director, Biodynamics and Bioengineering Division and Mr Charles Bates, Director, Human Engineering Division for their support of this effort. Additionally, the authors wish to thank Dr Charles W. Nixon, Chief, Biological Acoustics Branch and Dr Wayne L. Martin, Chief, Visual Display Systems Branch for their cooperation during this effort.

The authors also wish to express their appreciation for the contributions made to this effort by Cadet Beth A. Siegrest, USAF Academy, CO and Ms Tracy R. Donohue, Wright State University, Dayton, OH.

TABLE OF CONTENTS

TABLE OF CONTENTS
(continued)

Section | Page

TABLE OF CONTENTS
(continued)

LIST OF FIGURES

LIST OF FIGURES
(continued)

LIST OF TABLES

1.0 INTRODUCTION

The Armstrong Aerospace Medical Research Laboratory (AAMRL) is developing three-dimensional (3-D) auditory displays for integration into flight simulators and advanced crew station systems. It is envisioned that 3-D auditory displays containing directional and distance information will enhance the pilot's ability to determine the location of impending threats. These displays will also enable the pilot to simultaneously monitor multiple communication channels.

The recent advances in digital signal processing technology have made it possible to develop auditory localization cue synthesizers capable of generating 3-D audio for headphone presentation. The basic approach to the design of a localization cue synthesizer is to reproduce the changes of a sound waveform as it enters the listener's ear canal from different directions in a natural environment. The auditory localization cue synthesizer incorporates the head-related transfer functions (HRTFs) and interaural time delays (ITDs), which are the primary parameters required to optimize a listener's ability to determine the apparent direction of sound sources.

As a step in AAMRL's research of 3-D audio for cockpit applications, two auditory localization cue synthesizers were developed. Each system, in real-time, adds directional information to a sound source for headphone presentation. One system, the Directional Audio Display (DIRAD) synthesizer, was developed in-house by the Biological Acoustics Branch (AAMRL/BBA). As currently implemented, the DIRAD synthesizer presents up to two sound sources that can vary in azimuth, coupled to the listener's head position. Software modifications are planned to add elevation and distance cues to provide a complete 3-D audio system. The second system was designed and fabricated by the Gehring Research Corporation (Toronto, Canada) under the sponsorship of the Human Engineering Division (AAMRL/HE). This auditory localization synthesizer, the AL-204, can present up to four sound sources that can vary in azimuth, elevation and distance and are coupled with the listener's head position.

1.1 Background

Several factors impact the design of an auditory localization cue synthesizer. One of these factors, the resolution of the head-related transfer functions (HRTFs), is related to the approach used to present the directional audio at specific locations in the azimuth plane. Within the synthesizers described herein, resolution refers to the specific number of locations for which HRTFs are measured on a Knowles Electronics Manikin for Acoustic Research (KEMAR). The DIRAD synthesizer uses 360 HRTF measurements to enable the directional audio to be presented at any one-degree location in azimuth. In contrast, the AL-204 uses 36 HRTFs to present the directional audio at every 10-degree location in azimuth. The AL-204 synthesizer employs interpolation routines to present directional sounds between any two HRTF 10-degree locations.

The psychoacoustical evaluation of the synthesizers is dependent, in part, upon the response method(s) employed in the evaluations. The response method is the procedure by which the subjects indicate the perceived direction of the target sounds. Several methods have been used in previous localization studies, for example, pointing to the sound with the hand (Oldfield and Parker, 1984), verbally reporting azimuth and elevation estimates of the sound position (Wightman and Kistler, 1989), and verbally reporting from which numbered loudspeaker the sound was perceived to emanate (Butler, 1969). In order to avoid response methods which involve a large motor component (e.g., pointing the hand) or those with potential problems with boundaries (e.g., between numbered loudspeakers), Wightman's verbal estimate procedure was evaluated. In addition, a new method in which the subjects point, on a circle, to the perceived direction of the sound was evaluated. The circle represents the azimuth plane with the subjects at the center. This research provides data for future investigations examining localization performance and the usability of these response methods.

Several factors related to the eventual integration of 3-D auditory displays into crew station designs also need to be examined. In particular, does localization performance, with directional audio, vary as a function of: 1) azimuth location of the sound source, 2) signal type (e.g., wide-band pink noise, tones, and human speech), and 3) head movement (e.g., listener moves head to determine sound direction during signal presentation)? These factors need to be viewed in terms of several cockpit-related questions. For example, if a 3-D auditory cue is used to alert the pilot of the direction of a threat, will the pilot's perception of the sound direction vary with the azimuth of that threat? In terms of the audio signals, which signal type(s) might be used to enhance the effectiveness of directional audio in the cockpit? Finally, will pilots be required to move their heads in order to accurately determine the direction of the impending signal?

1.2 Purpose

The purpose of this report is to document several localization performance research efforts conducted with the two auditory localization cue synthesizers. Since the DIRAD is not currently configured to present sounds varying in elevation, the investigations were limited to using auditory signals presented in the horizontal plane or azimuth (at ear level). The specific objectives of this report are as follows:

a. Provide a comprehensive description of the approaches used in the design of the DIRAD and AL-204 auditory localization cue synthesizers (Section 2.0).

b. Provide an overview of the auditory localization facility and methods employed to evaluate the performance of the synthesizers (Section 3.0).

c. Describe (Section 4.0) three pilot studies designed to:

1. Study I - Measure and compare localization performance with the two auditory localization cue synthesizers. The objectives of this study were to evaluate localization performance with the DIRAD and AL-204 synthesizers, and determine, with the AL-204 synthesizer, whether performance differed between

non-interpolated target sounds and target sounds generated by interpolating between HRTFs. Thus, these data provide an initial assessment of the two synthesizer design approaches.

2. Study II - Using the DIRAD synthesizer, examine localization performance obtained with each of the two response methods used by the subjects to indicate the perceived direction of the sound: verbal response (listeners report angular estimates of the perceived direction of a target stimulus) and circle pointing (listeners estimate the perceived direction of a target stimulus by pointing on a circle representing the azimuth plane).

3. Study III - Examine two conditions of head movement with the DIRAD synthesizer: fixed-head (listeners head movements are restricted to a straight-ahead or "zero-degree" position while determining the perceived direction of a target stimulus) and free-head (listeners head movements are unrestricted while determining the perceived direction of a target stimulus). Also, localization performance with four types of audio signals was compared.

2.0 SYSTEMS DESCRIPTION

2.1 Conceptual Overview

Both synthesizers employ HRTFs recorded on a Knowles Electronics Manikin for Acoustic Research (KEMAR) in a free-field environment. As previously mentioned, these systems differ in the resolution of the measured HRTFs recorded along the azimuth plane (0-degrees elevation): one degree for the DIRAD and 10 degrees for the AL-204. Thus, the AL-204 synthesizer contains 36 transfer functions per ear, one for each 10 degree increment and employs interpolation routines to present sounds between any 10 degree interval. In contrast, the DIRAD synthesizer contains 360 transfer functions per ear, one for each degree increment, and, no interpolation routines are employed.

2.2 DIRAD Synthesizer Description

The DIRAD system is a two channel, real-time, digital auditory localization cue synthesizer. It consists of the localization cue synthesizer (LCS) board, digital interface board (DIB), and analog interface board (AIB), all contained in a single box. The LCS board connects to the DIB to communicate with the host processor via a serial port (RS-232C) and a Polhemus 3-Space head tracker via a parallel port (see Figure 1). The synthesizer design is based on two Texas Instruments TMS- 320C25 processors which control the localization cue synthesis and a TMS-32020 processor that controls the input-output functions. The AIB performs 16-bit analog-to-digital (A/D) and digital-to-analog (D/A) conversions for the LCS board. The right and left audio signals are then amplified and displayed with headphones.

The synthesizer board connects with the DIB and AIB to communicate with the peripheral digital and analog equipment, respectively (see Figure 1). The host computer communicates with the DIB over a RS-232C serial data bus and the head tracker provides the real-time head angle information to the DIB via a 16-bit parallel bus. The digital board sends relative head angle indexes over two separate 16-bit parallel I/O boards to the right and left processors on the synthesizer board. The processors look-up the corresponding Finite Impulse Response (FIR) filter coefficients and encode the audio signal with the directional cues.

The presentation of directional audio information over headphones is based on two directional cues: HRTFs and ITDs (Ericson and McKinley, 1989). The HRTF and ITD cue information were collected on a KEMAR manikin at one degree increments in azimuth. Digital Finite Impulse Response (FIR) filters (179 taps) of the 360 HRTFs were generated, and the HRTF and ITD codes were implemented in look-up tables on the DIRAD's TMS-320C25 processors. The processors are clocked at 40 MHz in order to calculate the 179 tap FIR filter every 1 ms. The audio input is sampled at 40 kHz to provide a 10 kHz bandwidth at the output of the AIB. The update rate of the processor is limited by the update rate of the Polhemus 3-Space head tracker (60 Hz). The DIRAD synthesizer can update the angle positions for two input audio signals every 60 Hz. The processor has been designed to support and update angle positions at frequencies up to 1000 Hz, provided the head tracking technology is available to support this rate.

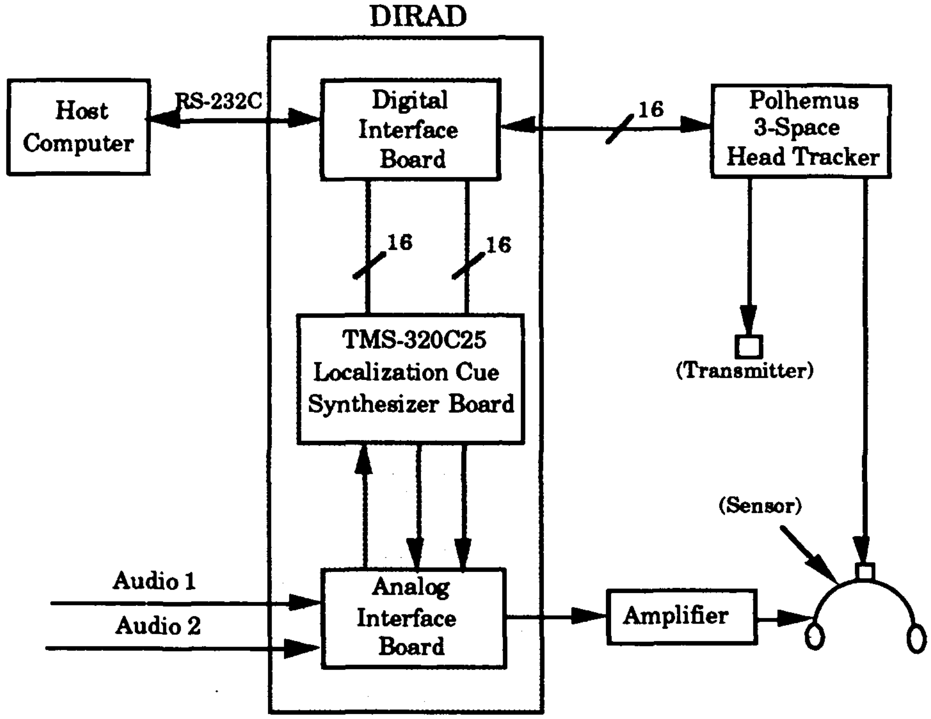

Figure 1. System Diagram of the Directional Audio Display (DIRAD) Synthesizer (from Ericson and McKinley, 1989)

2.2.1 DIRAD Host Computer Interface

The host computer communicates with the DIRAD synthesizer over an RS-232 serial interface. The serial port was configured for 9600 baud, no parity, 8 data bits and one stop bit. The command format consists of a byte string consisting of a send/receive (including the byte count), a one byte command code, zero to three data bytes and a cyclic redundancy check (CRC). The commands with data bytes are commands that send target angles. The first data byte is the higher order azimuth and the second byte is the lower order azimuth (0-359).

2.3 AL-204 Synthesizer Description

The Auditory Localizer (AL-204) is a four channel, real-time, auditory localization cue synthesizer. The system consists of a digital controller, four high-speed digital signal processors with A/D and D/A converters, solid-state memory and operating system to form the AL-204, as shown in Figure 2. The synthesizer communicates with a host processor over a serial port (RS-232C). A Motorola 68000 processor residing on a VMEbus controls the AL-204 synthesizer input-output functions. Four high-speed digital signal processing boards (ICS-100V) interface directly to the VMEbus to provide the localization cue synthesis capability. The right and left side audio signals for each channel are mixed together, then amplified and displayed with headphones. The 68000 operating system is written in Assembler.

The presentation of directional audio information with headphones is based on the HRTFs and ITDs. The directional cue information was collected on a KEMAR manikin at 10 degree increments in azimuth and elevation. Digital finite impulse response filters (112 taps) of the HRTFs were generated, and the HRTF codes were implemented in look-up tables on the synthesizer's memory board. The analog audio input signals are converted to 16-bit digital data sampled at a programmable rate derived from the processor's 18 MHz clock. Typically, the sampling rate is approximately 44.1 KHz for each channel. The output D/A conversion is clocked by the input sampling rate, and the 12-bit A/D and D/A devices. Audio input and output is at line level through RCA phono jacks. Four audio source inputs and four binaural pair outputs (eight monoaural signals) are provided. A single binaural output corresponding to each input source can be provided and two or more output binaural pairs can be mixed simultaneously by specifying the input-output combination from the host computer. The audio output bandwidth for each channel is 22 KHz.

2.3.1 AL-204 Host Computer Interface

The AL-204 accepts commands for azimuth, elevation, and range from the host computer for each audio source. The command format is binary, and the communication protocol is software-configurable. The AL-204 is interfaced to the host computer via an RS-232C interface (selectable up to 9600 baud). The host command structure consists of an eight-byte binary incremental command. This data string consists of:

Byte	Command
1	Channel addressing
2-3	Azimuth position command
4-5	Elevation position command
6-7	Range mute/unmute command
8	Carriage return

The AL-204 receives the head tracker position data via the host processor (see Figure 2). The interpolation resolution of the system in azimuth and elevation is 10 bits (0.5 degrees). The range resolution is two bytes (0 - 255) and the mute command is equivalent to specifying an infinite distance. As currently implemented, the host can address a single or multiple combination of channels. If a single channel is invoked, the update rate is 60 Hz. However, this rate is effectively reduced by the number of channels invoked. For example, if three channels are invoked, the update rate is reduced to 20 Hz.

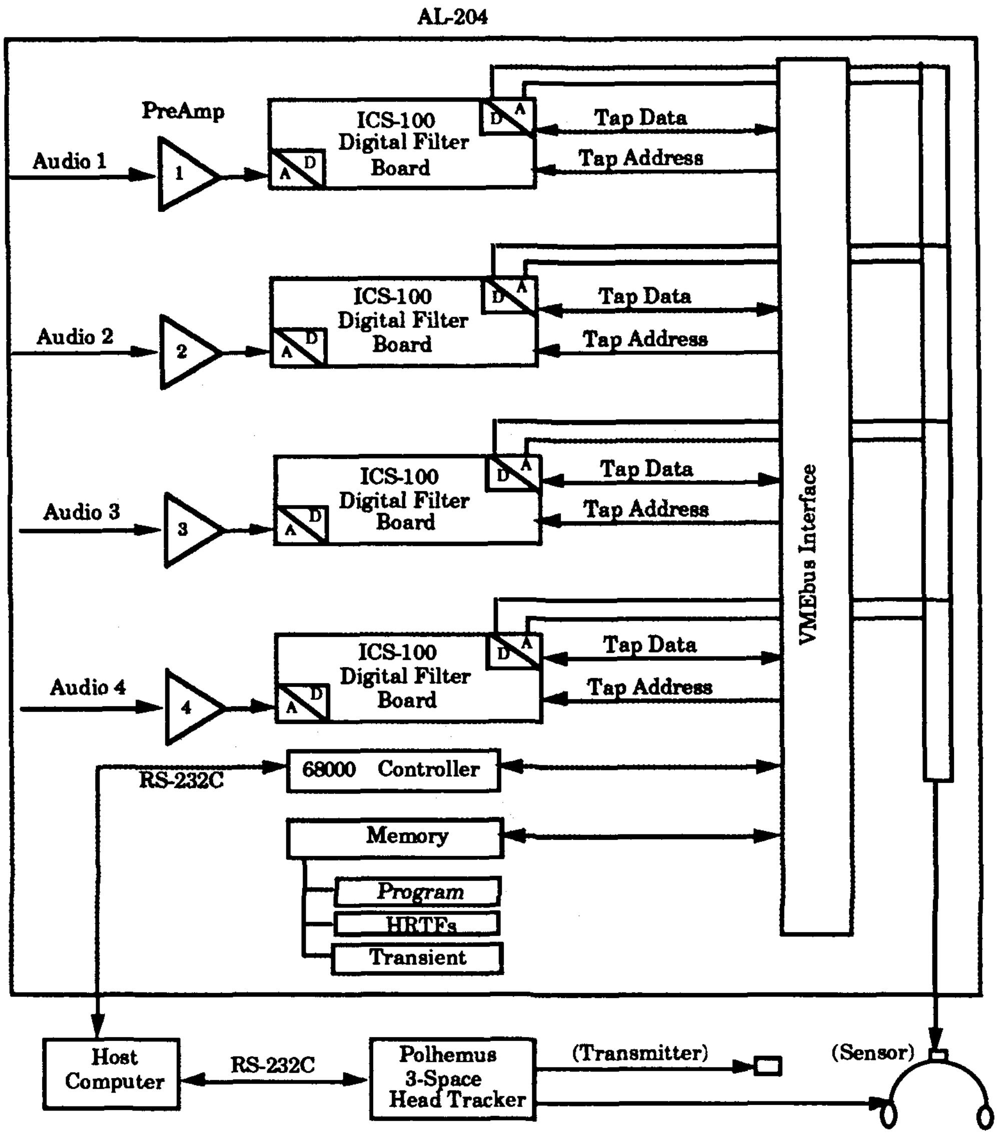

Figure 2. System Diagram of the AL-204 Synthesizer

3.0 AUDITORY LOCALIZATION FACILITY DESCRIPTION

3.1 Experimental Area

Since the experimental area employed in these studies was previously described in an AAMRL technical report (Valencia and Calhoun, 1988), it will only be briefly summarized herein. The area consisted of an approximately 4.7 sq m (50 sq ft) circular area (1.22 m diameter) enclosed by a curtain. The curtain was made of an opaque material which extended 2.75 m from the ceiling to the floor. The magnetic transmitter for the head tracker was attached to a plastic rod, measuring 16 inches from the ceiling. The subject stood directly underneath the transmitter in the center of the experimental area. The area was dimly lit (approximately 12.8 nits), and was not treated with any soundproof or acoustic material. The ambient noise level at the center of the experimental area was 56 dB(A) SPL.

3.2 Hardware Description

A detailed schematic of the hardware employed to evaluate the localization performance of the two auditory localization synthesizers is shown in Figure 3. A MicroVAX II was used to host both synthesizers over two RS-232C serial ports and collect the experimental data. The MicroVAX II was also employed to host the magnetic head tracker and a data tablet. The input azimuth commands were localized by referencing the appropriate HRTFs and ITDs in each synthesizer and displaying the output to a subject wearing headphones. The following sections provide additional information regarding the software and equipment employed in the studies.

3.2.1 Head Tracker Interface

A Polhemus 3-Space head tracker was utilized to boresight the sound's location to the straight-ahead (0-degree) position during the three studies. The system employs a low-frequency magnetic field to determine the orientation of a sensor relative to the transmitter. The head tracker also provided data on subjects' head movements which allowed the stimulus to be space-stabilized in the head-free condition in Study III. The angular resolution error in azimuth of the tracker, with the current experimental set-up, averaged 1.0 degree RMS.

For the studies involving the DIRAD synthesizer (Studies I, II and III), the head tracker was directly interfaced with the DIB over a parallel port, as previously described in Section 2.2. The update rate of the 3-Space tracker in this configuration was 56 Hz. For the AL-204 synthesizer (Study I), the head tracker was interfaced directly to the MicroVAX II through a RS-232C port, as shown in Figure 3.

3.2.2 Data Tablet Interface

In Study II, involving the circle pointing response method, a Summagraphics (MM-1201) tablet was employed which consists of two parts, the stylus and the tablet. The stylus (pen-like, hand-held device) emits a low-intensity magnetic field which is sensed by a grid inside the tablet whenever the stylus is located on the tablet's "active area".

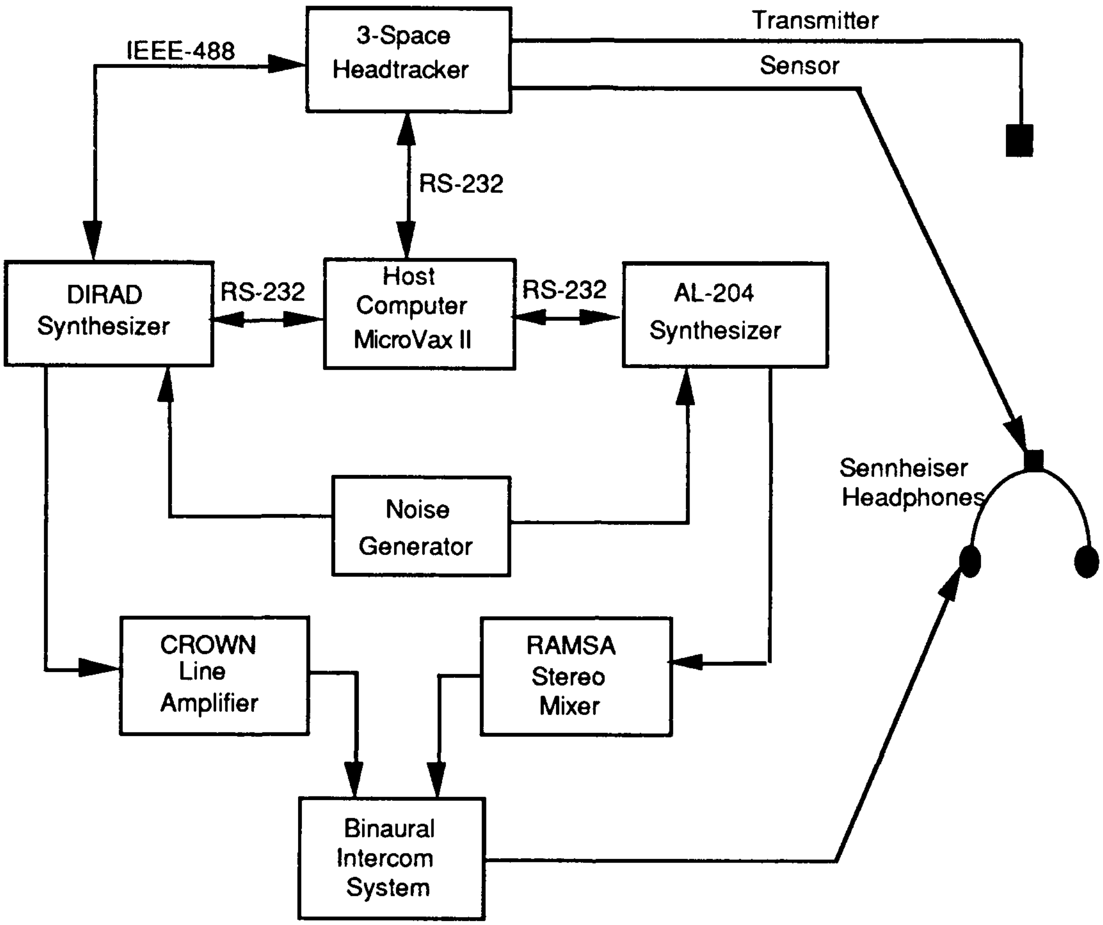

Figure 3. Layout of Experimental Set-up

The Summagraphics tablet translates the position of the stylus on the tablet into digital information and communicates it to the MicroVAX II via the data/ power cable. The MicroVAX II communicates with the tablet via a serial link (RS-232C) running at 9600 baud. The stylus position is expressed as an X and Y coordinate pair. For purposes of data collection, the active area was limited to a 100 square inch circle about the center of the tablet. The resolution of the device was set to 0.25 inches, which equates to one degree resolution in azimuth. The hardware components of the Summagraphics data tablet are illustrated in Figure 4.

3.3 Software Description

The major software components employed in the localization studies are the programs used to host the DIRAD and AL-204 synthesizers, the 3-Space Tracker, and those to monitor the experimental paradigm. These programs were written in FORTRAN and were part of a master test program used to initialize each synthesizer, enable experimenter control, calculate the perceived sound direction and accomplish data collection.

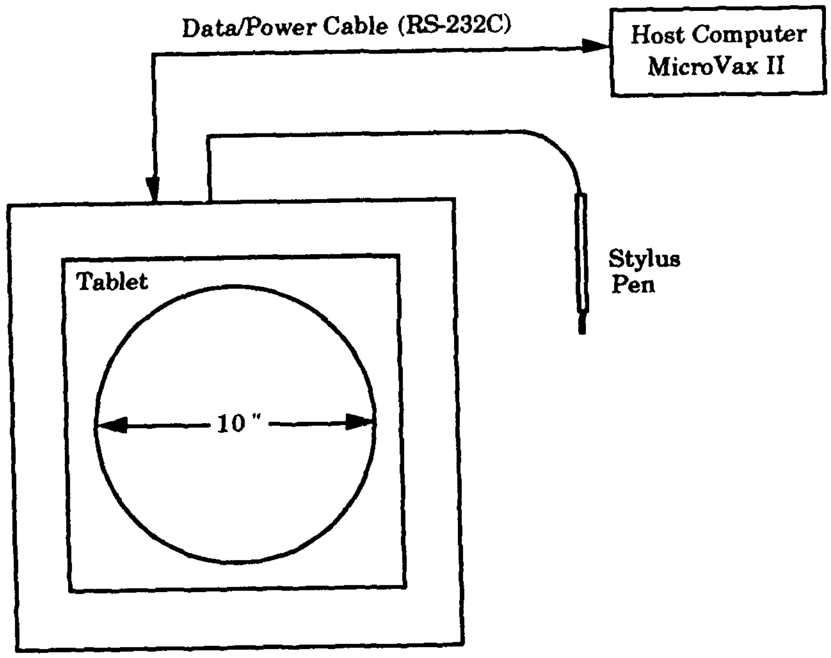

Figure 4. Summagraphics Data Tablet

4.0 LOCALIZATION PERFORMANCE EVALUATION

4.1 Objectives

Three pilot experiments were conducted to examine subjects' ability to perceive the direction of synthesized directional audio signals randomly presented in the azimuth plane. All studies measured localization performance in terms of localization accuracy, percentage of sound reversals and response time required to perceive the direction of the target sounds.

The specific objectives of each study were as follows:

a. Study I - Comparison of the AL-204 and DIRAD Synthesizers. This pilot study evaluated and compared localization performance with the two synthesizers. A second objective was to compare interpolated versus non-interpolated targets in the AL-204 to determine if interpolating between HRTFs results in decreased localization performance. The listeners' head movements were restricted in this study (fixed-head) and a verbal response method was employed (i.e., listeners provided angular estimates) to measure the perceived direction of the target stimulus.

b. Study II - Evaluation of Two Methods for Measuring Localization Performance: Verbal Response and Circle Pointing. In this study, a fixed-head paradigm was also used, but the audio cues were delivered only by the DIRAD synthesizer. Localization performance was compared with two response methods (subjects provided verbal estimates in degrees and subjects pointed to a position on a circle with a stylus pen).

c. Study III - Localization Performance as a Function of Signal Type and Head Movement. This study evaluated localization performance, within fixed-head and free-head conditions, with four types of audio signals delivered by the DIRAD synthesizer. The stimuli consisted of a wide-band pink noise signal, a 4 kHz low-pass pink noise signal, a 1 kHz pure tone and female speech.

4.2 Method

4.2.1 Apparatus

Section 3.0 provides a detailed description of the facility in which the studies were conducted. The following provides additional information, specific to each of the studies.

In Study I, the directional signals were delivered by the DIRAD and AL-204 cue synthesizers. For Studies II and III, the directional signals were generated by the DIRAD synthesizer only. For all three studies, a General Radio 1382 - Random Noise Generator provided the wide-band pink noise stimuli. In Study III, the 4 kHz low-pass stimuli were generated by the noise generator which was connected to a Bruel & Kjaer band-pass filter set (Model 1612/S1A) and a summing panel (Model 1612/SP). The 1 kHz pure tone (sinewave) was provided by an HP-

8116 Function Generator. The female speech (bandwidth = 6 kHz recorded while reading from a technical report) was played back from an audio cassette tape.

The output of the noise generator was connected to a General Radio 1/3 Octave-band set to balance the signals to +/- 3 dB from 100 Hz to 20 kHz. The Sennheiser (HD-250) headphone signals were balanced to 70 dB(A) SPL, while the signals were presented at the straight ahead or "0-degree" azimuth position.

In Study II, a Summagraphics tablet (see Section 3.2.2) was employed in the circle pointing response method. In the free-head movement condition in Study III, the target stimulus position changed in real-time (54 Hz update rate) to provide the subject with a space-stabilized acoustic image over the headphones.

4.2.2 Subjects

The hearing thresholds of all the subjects tested in the three studies were not greater than 15 dB at any standard audiometric test frequency from 500 to 8000 Hz. In addition, the subjects' interaural threshold differences did not exceed 10 dB at any test frequency.

In Study I, six male subjects (mean age 25 years) were tested. Three were naive, paid volunteers, while the other three were in-house personnel experienced in listening to localization cues over headphones. In Study II, six paid subjects (3 males, 3 females; mean age 24 years) were tested. In Study III, six paid subjects (3 males, 3 females; mean age 25 years) were tested. No subjects participated in more than one of the three studies.

4.2.3 Procedure

In all three studies, the subjects' task was to indicate the perceived direction of the sound presented over the headphones. The following general procedure was used. The subjects positioned their heads on a chinrest and stood directly underneath the magnetic transmitter and on top of an "X" marked on the floor. Each subject wore a pair of Sennheiser headphones (HD-250), with a magnetic sensor centered on top, and a microphone to allow communication with the experimenter. To begin each trial, the subjects depressed the hand-held response switch to boresight. The host processor automatically boresighted the head tracking unit (to the straight ahead or "0-degree" position) and presented the target stimulus one second later. The auditory stimulus was delivered, via the headphones, from a target location along the azimuth plane. The subjects were instructed to depress the hand-held switch a second time, once they had perceived the direction of the stimulus.

To indicate the perceived direction of the stimulus, subjects used one of two response methods: verbal estimate (Studies I, II, III) and circle pointing (Study II). For the verbal response method, the subjects verbally indicated their best angular estimate of the sound's direction in terms of degrees (0 to 359). For example, if the sound appeared to be originating directly from the front, the subject responded "0 degrees" and if it originated directly at the right or left ears, the response was "90 degrees or 270 degrees", respectively. To aid the subjects in their estimates, a circle diagram, marked every ten degrees, was placed on a table in front of the subjects. In the circle-pointing response method (Study II), subjects were provided with a stylus pen with which to indicate on a circle diagram, placed on a table in front of them, the perceived direction of the sound (see Section 3.2.2 for description of Summagraphics data tablet). The circle diagram on the data tablet was marked at the 0, 90, 180 and 270 degree azimuth positions to aid the subjects while responding. (Note that this was a different diagram than the one used for the verbal estimate response method). The subjects' responses were based on the assumption that the circle represented the horizontal plane, with the subject at the center of the circle. The subject responded by placing the pen on a specific point on the circle which corresponded to the perceived direction of the stimulus. The perceived azimuth (AZ) was calculated by the host processor from the X and Y coordinates as follows: Perceived AZ = Arctangent (Y/X). The boresight procedure began two seconds later.

With the verbal response method, the subjects' responses were entered by the experimenter on the host processor and the next trial began three seconds later. This procedure was repeated 24 times per run. In all three studies, the trials were self-paced and the subjects were instructed to take as long as they desired to determine the perceived direction of the target stimulus. No performance feedback was provided to the subjects by the experimenter. A timeline of an experimental trial is depicted in Figure 5.

In Study III, localization performance was evaluated in a free-head condition, as well as a fixed-head condition. In the free-head condition, the procedure was identical to the fixed-head condition, with the exception that the subjects were free to move their heads for localizing during the stimulus presentation. Following their response, they faced forward for the next boresight. In the free-head condition, the sound was repositioned in real-time to provide a stable acoustic image while the subjects moved their heads for localizing.

4.2.4 Experimental Design

4.2.4.1 Independent Variables

Table 1 lists the independent variables used in the three studies. All studies examined localization performance as a function of the target sector (see Figure 6).

In addition, Study I examined localization performance as a function of synthesizer type and target type. An experimenter-controlled grouping was used to randomly present the stimulus and evaluate target type. A non-interpolated target was defined as any stimulus location which appeared at an angle evenly divisible by 10 degrees (e.g., 10, 20, 30, 40, ..., 350, etc.), while an interpolated target was defined by any other target angle (e.g., 5, 23, 46, 101, ..., 358, etc.).

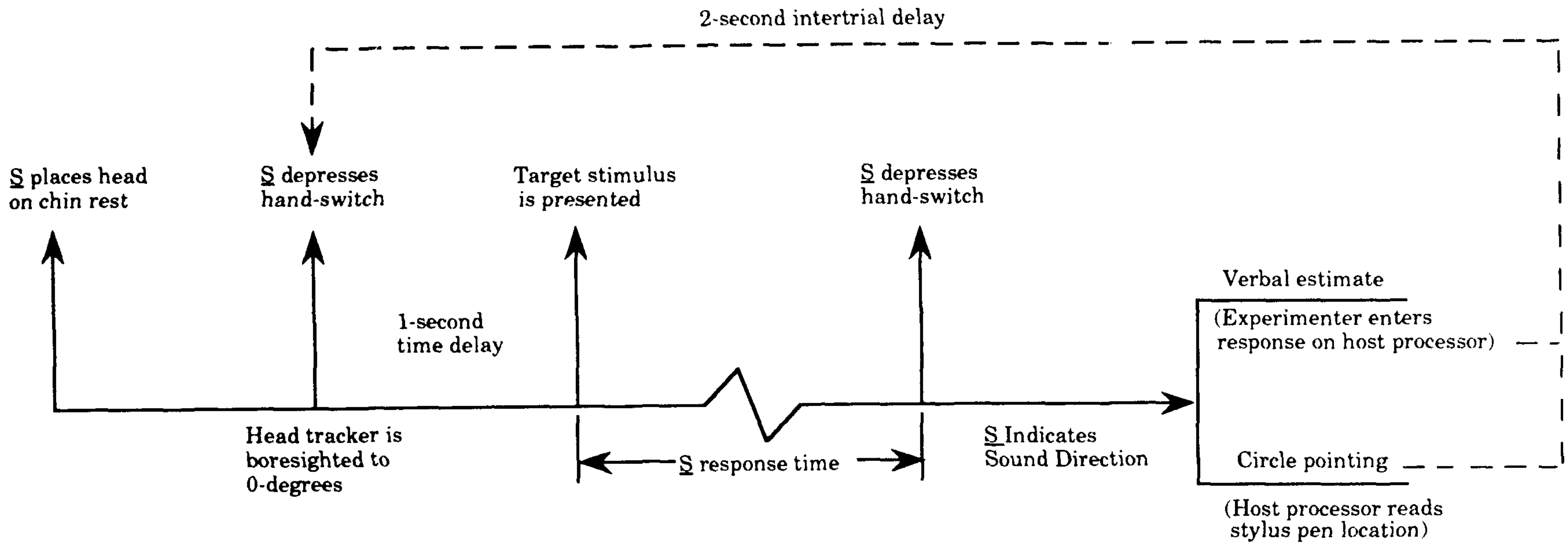

Figure 5. Timeline for an Experimental Trial

Table 1. Summary Table of Independent Variables for Three Studies

Study	Factor	Level	Description
I	Synthesizer Type	2	1. DIRAD 2. AL-204
I	Target Type	2	1. Non-interpolated (divisible by 10) 2. Interpolated (not divisible by 10)
II	Response Method	2	1. Verbal Estimate 2. Circle Pointing
III	Head Movement	2	1. Fixed 2. Free
III	Stimulus Type	4	1. Pink noise (PN) 2. 4 kHz Low-pass PN 3. Female speech 4. 1 kHz tone
I, II, III	Target Sector (Azimuth - degrees)	6	1. 0 - 59 2. 60 - 119 3. 120 - 179 4. 180 - 239 5. 240 - 299 6. 300 - 359

4.2.4.2 Dependent Variables

The data recorded in Studies I, II and III consisted of:

a. Magnitude Error (degrees): Absolute difference between the actual target azimuth and the perceived target azimuth, as indicated by the subject's response.

b. Response Time (seconds): The time from the stimulus onset to the subject's pressing the response switch.

c. Percentage of Front/Back Reversals: The occurrence of front/back reversals is a common phenomena in auditory localization studies, particularly when head movements are restricted. To resolve any front/back reversal, a technique previously used by Wightman and Kistler (1989) was employed in these studies. In this technique, any trial in which the perceived azimuth was reported in the opposite hemisphere (front vs. back, but on the same side) than the subject's response was defined as a front/back reversal. For example, a subject's perception of a 45-degree azimuth target as located at 130 degrees was treated as a front/back reversal. When a reversal occurred, the perceived azimuth was subtracted from 180 degrees to place the perceived target in the correct hemisphere. Then, the difference between the perceived azimuth and the target azimuth was recorded as the localization error. In the above example, the corrected perceived azimuth was 180 - 130 = 50 degrees. The localization error was then 50 - 45 = 5 degrees and the reversal count for the sector in which the target stimulus was presented was increased by one. The data submitted to analyses for the three dependent variables included those trials in which reversals occurred. Although this procedure results in improved error, it was employed here to facilitate comparison with data collected in earlier studies.

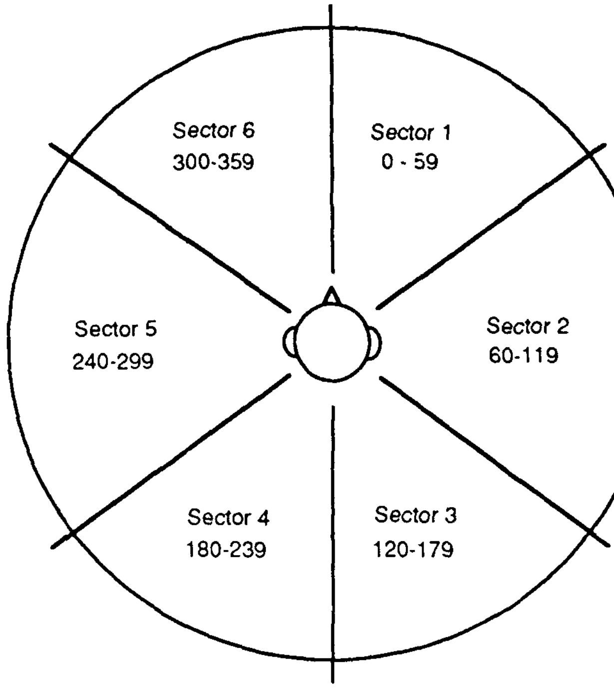

Figure 6. Grouping of Target Location into Six Azimuth Sectors

d. Subjective Measures: Following the completion of the studies, each naive subject responded to a questionnaire designed to elicit observations and comments pertaining to various aspects of the study (see Appendix B).

4.2.4.3 Target Selection Process

For Study I, the presentation order of the target sectors and the targets within each sector was determined in the following manner. The presentation order of the six angular sectors was randomized with the constraint that four targets, in each run of 24 trials, were presented in each sector. Furthermore, the four targets in each sector represented two targets for each target type: two targets randomly selected from six non-interpolated points common to both synthesizers and, two targets randomly selected from six interpolated locations.

The selections of the six interpolated locations were actually made from six intervals for each sector; each interval represented a nine-degree range of one degree increments. All totaled, 54 interpolated points were possible for each sector. Table 2 illustrates the locations for selecting targets in Sector #1, which spanned 0 - 59 degrees:

Table 2. Target Locations for Sector #1

Sector #1 Non-Interpolated	Sector #1 Interpolated (interval)
0	1 - 9
10	11 - 19
20	21 - 29
30	31 - 39
40	41 - 49
50	51 - 59

For each run, and within each sector, two intervals among the interpolated locations were randomly selected (without replacement). A target location was then randomly drawn from each selected interval. Similarly, two target locations were also randomly drawn (again without replacement) from the six non-interpolated locations within each sector. Thus, across a session of three runs, six unique locations were drawn separately, within each sector, for both the non-interpolated and interpolated target presentations. Table 3 lists one possible presentation for Sector #1.

Table 3. A Possible Target Presentation Order for Sector #1

	Non-Interpolated	Interpolated
Run 1	10	15
	40	26
Run 2	0	33
	30	8
Run 3	50	42
	20	54
Totals:	6/sector	6/sector

In summary, a total of 4 azimuth locations (2 interpolated and 2 non-interpolated) from each of the 6 sectors were tested in each run . For Studies II and III, the same process was employed.

It must be noted, however, that the DIRAD synthesizer does not perform interpolation routines to present targets between ten degree intervals. The use of the terms "interpolated" and "non-interpolated" target types is only meaningful when applied to the AL-204 synthesizer. The use of these terms with the DIRAD synthesizer was to ensure that the same target locations were examined with both synthesizers. The effect of interpolated and non-interpolated target locations on localization performance was investigated with AL-204 data only.

4.2.4.4 Data Collection

The number of days that subjects were tested in each study and the number of blocks of 24 trials conducted per day (or session) are shown in Table 4. Also shown are the factors examined in each study and the total number of trials. Additional information pertaining to the experimental design for each study is described below.

a. Study I. The presentation order of the synthesizers was counterbalanced by randomly dividing the subjects into two groups. In the first two sessions, the Group I subjects (inexperienced) were tested with the DIRAD synthesizer in the first three blocks and the AL-204 synthesizer in the last three blocks. For the last two sessions, Group I was tested with the AL-204 synthesizer in the first three blocks and the DIRAD synthesizer in the last three blocks. This order was reversed for the Group II subjects (experienced).

b. Study II. The presentation order for the response methods was counterbalanced by randomly dividing the subjects into two groups. In the first three sessions, the Group I subjects were tested with the circle pointing method during the first three blocks and the verbal response method during the last three blocks. For the last three sessions, the Group I subjects were tested with the verbal response method in the first three blocks and the circle pointing method in the last three blocks. This order was reversed for the Group II subjects (experienced).

c. Study III. Two sessions were conducted in the fixed-head condition, and two in the free-head condition. Each session consisted of 12 blocks of 24 trials, three blocks with each of the four stimulus types. The presentation order of the other stimulus types was determined by a 4x4 balanced Latin square, such that, across the four sessions, each stimulus was preceded equally often by each of the other stimuli. The head movement condition was balanced, such that half of the subjects were tested in the fixed-head condition in the first two sessions, followed by the free-head condition in the last two sessions. This order was reversed for the other half of the subjects.

Table 4. Breakdown of Number of Trials per Study

Study	# Days	#Blocks of 24 trials/day	# Blocks w/ each factor	Factors							Total Mean Scores / subject	Total # trials
				Subjects	Sectors	Head move.	Synth	Target type	Response method	Stimulus		
I	4	6	3 each synthesizer	6	6		2	2			24	3456
II	6	6	3 each response method	6	6				2		12	5184
III	4 head-fixed 2 head-free 2	12	3 each stimulus per day	6	6	2				4	48	6912

4.2.5 Data Analyses

Prior to any statistical analyses, the mean magnitude error, mean response time, and mean percentage of reversals were calculated by averaging performance across all the sessions in the study. Thus, the data matrix submitted for analyses was composed of one mean score for each factorial combination of the factors examined in the study.

For all three studies, the statistical analysis on the three dependent measures was conducted using an Analysis of Variance (ANOVA). For significant main effects, Tukey's test was employed, while analysis of simple effects/simple comparisons was employed to analyze significant interactions.

5.0 RESULTS/DISCUSSION

Tables showing the mean magnitude error, mean response time and mean percentage of reversals as a function of the independent variables for Studies I, II and II are presented in Appendix C. After a brief general discussion of the results, the results for each factor under investigation will be presented and discussed, in turn.

The overall mean magnitude localization error measured in the three studies (17.4 degrees) compared favorably with a previous study (Wightman and Kistler, 1990) in which subjects, using a verbal estimate paradigm, reported the perceived direction of stimuli synthesized with KEMAR HRTFs. The average magnitude error in Wightman and Kistler's study was 19.1 degrees for targets presented at 0 and 18 degrees elevation, across azimuth locations.

Fourteen of the 15 subjects perceived the sounds to be externalized (i.e., outside-the-head; see Appendix B - Subject's Questionnaire). This indicates that a realistic perception of 3-D sound was created by the synthesizers. Although all the target sounds were presented in only the azimuth plane (0-degree elevation), eight of the 15 subjects reported perceiving some of the sounds to be located above or below the azimuth plane. For one of these subjects, this perception occurred when targets were at the right and left side of the head. For two others, the frontal sounds were perceived as elevated above the horizontal plane.

5.1 Performance as a Function of Subjects' Localization Experience (Study I)

As mentioned in Section 4.2.2, three of the subjects used in the first experiment were inexperienced in listening to directional audio signals over headphones. The other three subjects, however, were in-house personnel experienced in listening to synthesized audio cues. An ANOVA was performed to determine if these data were sensitive to the subject's level of listening experience. The results showed that the experienced subjects had significantly less mean magnitude error (12.3 degrees) than the inexperienced subjects (16.7 degrees) ($F(1,4) = 18.3$, $p = 0.0129$). Significantly less reversals also occurred with experienced subjects (14.4%) compared to naive subjects (28.0%) ($F(1,4) = 9.77$, $p = 0.0353$). No significant differences were found for mean reaction time. Additionally, no interactions were significant.

5.1.1 Discussion of Performance as a Function of Subjects' Localization Experience

The level of prior listening experience to synthesized directional audio played a role in enhancing localization. Subjects with prior listening experience had a lower magnitude error and, most notably, a lower incidence of front/back reversals. Reversals tend to be related to the role that the monaural pinna (external ear) cues play in allowing listeners to resolve front/back confusions (Oldfield and Parker, 1984). The monaural pinna cues play a major role in allowing listeners to disambiguate front/back confusions. For these monaural cues to be an effective aid in localization, however, the listener must have a priori information concerning the spectrum generated by the source (Doll, Gerth, Engelman and Folds, 1986). This information permits the listener to judge whether a given spectral feature is a function of the source position, or an inherent property of the stimulus, regardless of source position. Repeated exposure to the source, or pre-experimental knowledge can provide such a priori information. Thus, the experienced subjects have apparently learned to resolve the front/back ambiguities through repeated exposure to the stimuli.

It is evident that the level of subject experience in listening to directional sound must be considered in future investigations. The amount of training required to achieve asymptotic performance has not been adequately investigated. The number of sessions in which a group of naive subjects would have to participate in to match the performance level of the experienced subjects is not known. Perhaps the unique experience of localizing over headphones, or localizing with HRTFs other than the listeners' own, also plays a significant role in the level of performance.

5.2 Performance as a Function of Target Sector (Studies I, II and III)

For Study I, the results of the ANOVA showed no significant differences between target sectors for mean magnitude error, $F(5,25) = 1.30$, $p = 0.2960$; mean response time, $F(5,25) = 2.14$, $p = 0.0942$; and mean percentage of reversals, $F(5,25) = 1.17$, $p = 0.3509$.

In Study II, the mean magnitude error and mean percentage of reversals (see Figures 7 and 8, respectively) varied significantly as a function of target sector ($F(5,25) = 8.48$, $p = 0.0001$ and $F(5,25) = 3.97$, $p = 0.0087$, respectively). Tukey's test showed that Sectors 2 (13.8 degrees) and 5 (13.6 degrees) had significantly lower mean magnitude error than the remaining sectors. The Tukey test also revealed that a significantly higher percentage of reversals occurred in Sector 1 (53.8%) than in Sectors 3 (15.5%) and 4 (15.9%). Mean response time did not vary as a function of target sector ($F(5,25) = 0.86$, $p = 0.5254$).

In Study III, the mean magnitude error, mean response time and mean percentage of reversals varied significantly as a function of target sector ($F(5,25) = 5.57$, $p = 0.0014$; $F(5,25) = 11.43$, $p = 0.0001$; and $F(5,25) = 5.80$, $p = 0.0011$, respectively; see Figures 9, 10, and 11). The Tukey test showed that the subjects' mean magnitude error for Sectors 3 (18.3 degrees), 4 (18.9 degrees) and 6 (18.0 degrees) was significantly greater than that for Sector 5 (11.3 degrees). The Tukey test also showed that the mean response time was significantly longer for Sector 4 (2.3 seconds) than for Sectors 2 (1.9 seconds) and 5 (1.9 seconds). Additionally, the Tukey test indicated that significantly more reversals occurred for Sectors 2 (23.7%) and 5 (25.2%) than for Sectors 1 (11.4%) and 4 (9.5%).

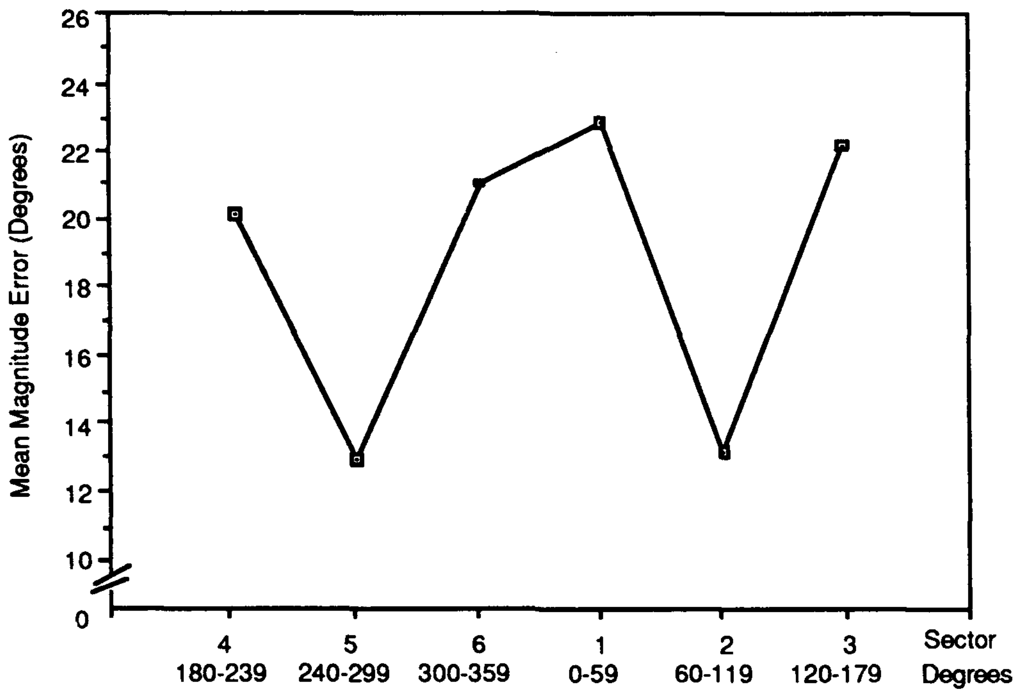

Figure 7. Mean Magnitude Error as a Function of Target Sector (Study II)

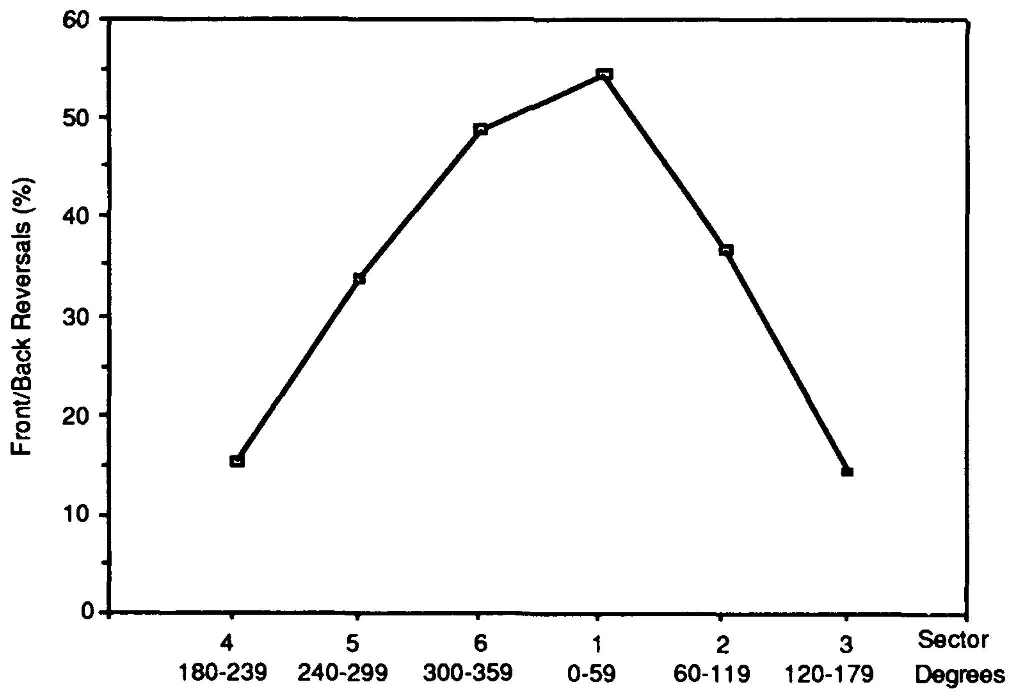

Figure 8. Mean Percentage of Front/Back Reversals as a Function of Target Sector (Study II)

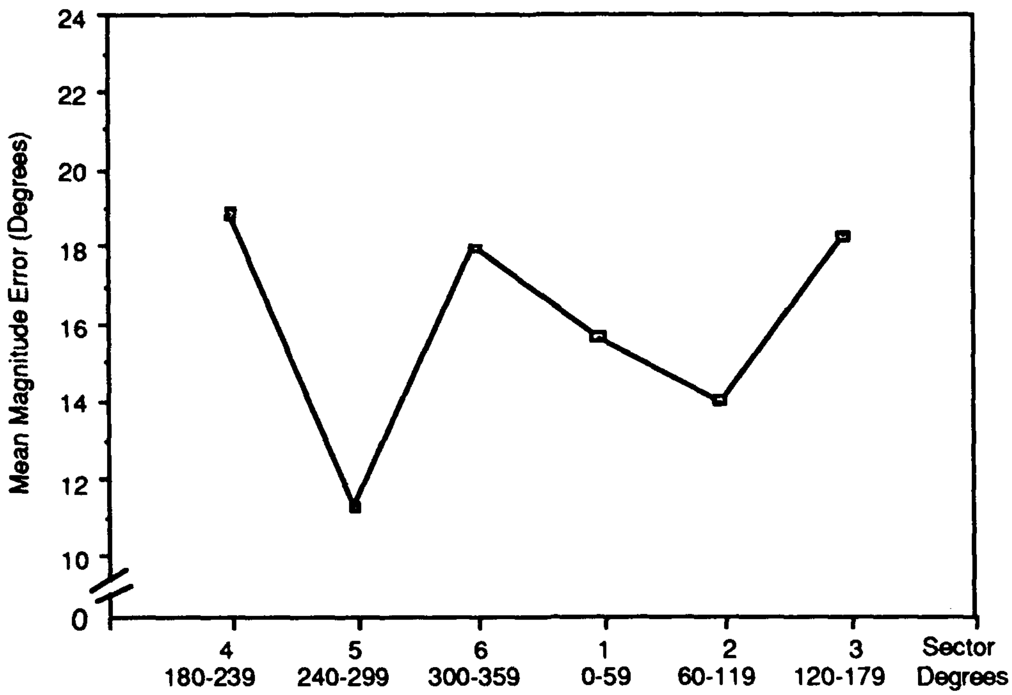

Figure 9. Mean Magnitude Error as a Function of Target Sector (Study III)

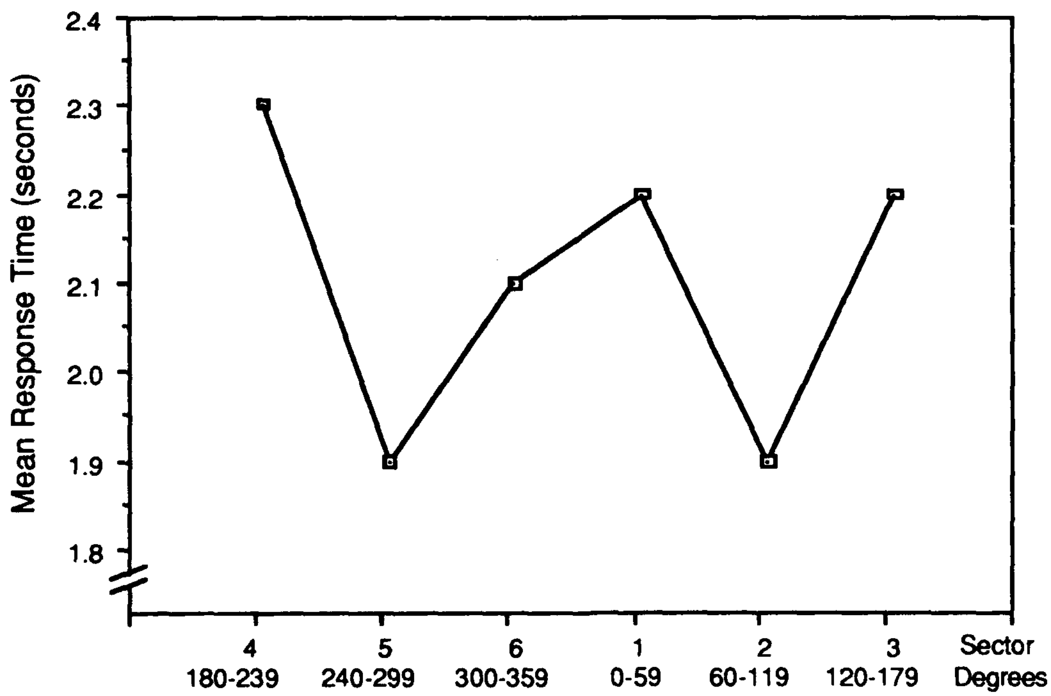

Figure 10. Mean Response Time as a Function of Target Sector (Study III)

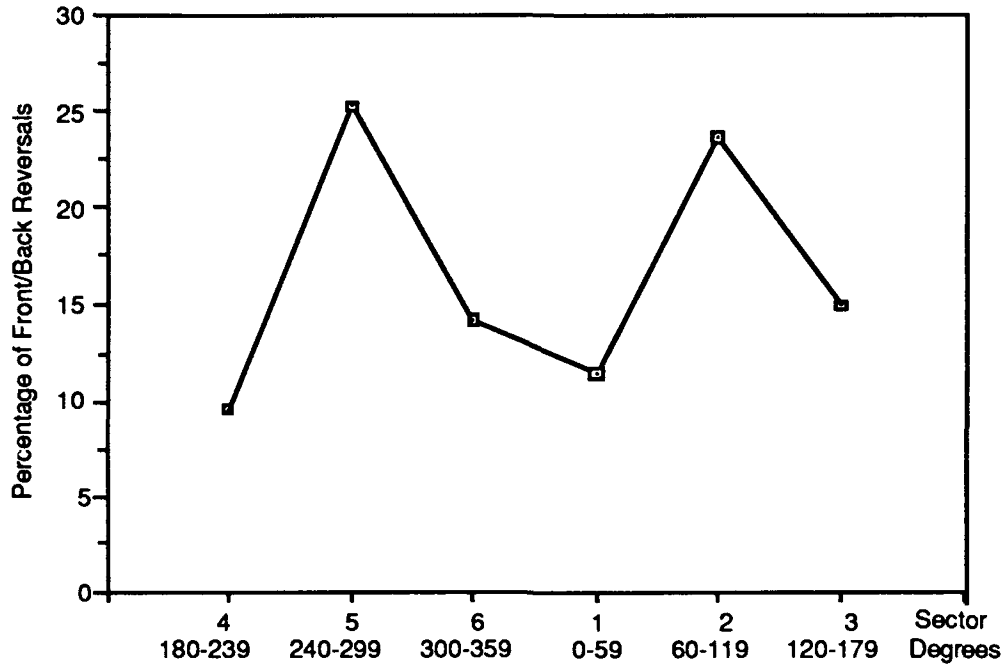

Figure 11. Mean Percentage of Front/Back Reversals as a Function of Target Sector (Study III)

5.2.1 Discussion of Performance as a Function of Target Sector

In Study I, target sector was not a significant factor in localization performance. However, in Studies II and III, the highest localization accuracy (e.g., lowest mean magnitude error) was found for the side targets, with performance at the front and back targets exhibiting comparable, but less accurate, responses. Data from the subjects' questionnaire (see Appendix B) closely match with the mean magnitude error findings. Subjects reported that the direction of sounds located at the sides were easier to perceive than sounds from the front or back. More than half of the subjects thought that the location of the front and back targets were difficult to perceive. In fact, two subjects reported that they never perceived sounds to be emanating from the front sectors.

Similar results were found by Wightman and Kistler (1989) using individualized (listener-specific) HRTFs and HRTFs obtained from a KEMAR manikin. Under both conditions, listeners demonstrated the highest degree of accuracy for the side targets, although the accuracy differences between the three regions (front, side and back) were smaller (2 - 4 degrees) than that found in Studies II and III (7 degrees).

How do data with synthesized sound compare with those studies which utilize real sound? There are conflicting data concerning the effect of target location on localization accuracy. Most auditory localization studies (e.g., Oldfield and Parker, 1984; Mills, 1958) have found greater accuracy for sounds presented in the front areas. Oldfield and Parker (1984) investigated localization performance using a white noise stimulus presented over a range of azimuth (0 - 180 degrees) and elevation (± 40 degrees) positions. Subjects, blindfolded and in a head-fixed paradigm (via a chin rest), pointed a hand-held "gun" to the perceived direction of the sound. Results showed that the azimuth error was lower in the front locations (approximately 5 degrees), in contrast to the back target locations (approximately 18 degrees). Mills (1958) found localization performance was better for positions directly in front of the subject. Mills measured auditory discrimination ability in terms of the minimum audible angle (MAA; defined as the smallest detectable difference in location between two stimuli). Subjects reported whether a second stimulus was located to the right or left of a first stimulus (reference stimulus). The task was conducted at 0, 30, 45, 60, 75, and 90 degrees in azimuth. Mills found optimum discrimination performance at 0-degrees azimuth (approximately 1-degree) with an increasingly poorer discrimination function as the azimuth increased to 90 degrees.

Oldfield and Parker (1984) suggest that poorer localization performance in the rear quadrants may be related to the structure of the external ear (pinna) and its orientation with respect to the head. These authors cite studies (e.g., Dirks and Gilman, 1979) which suggest that the binaural cues are more variable in the back quadrants.

Other studies, however, have not demonstrated better performance in the front areas. For instance, Wightman and Kistler (1989) evaluated localization accuracy in a free-field study in which subjects estimated the azimuth and elevation coordinates of a wide-band, noise-like signal. Subjects localized the side targets (±60, ±120 degrees azimuth, 0, ±18, ±36 and ±54 degrees elevation) with the lowest error (19.7 degrees), and the front and back targets with slightly higher errors (23.1 and 22.9 degrees, respectively).

The effect of target sector on the mean percentage of front/back reversals was different for each of the three studies reported herein. In Study I, there was no effect. However, in Studies II and III the highest percentage of reversals occurred in the front sectors (1, 6) and in the side sectors (2, 5) respectively. One might expect more reversals in the front sectors, as in Study II, since the subjects primarily confuse the front with the back targets, but usually not vice-versa (Wightman and Kistler, 1989). Also previous research (Weinrich, 1982 and Wightman and Kistler, 1989), indicates that listeners, with no head movement, have difficulty in perceiving sounds presented in the front areas. This is especially true when listening to sounds reproduced through manikin's ears, such as KEMAR.

Two other factors have influenced the reversal patterns in these studies. One, different subjects were used in all three studies. Wightman and Kistler (1989) found that subjects exhibited a wide range of front/back discrimination ability when localizing with KEMAR HRTFs, as well as individualized HRTFs. This ability is related to the function of the pinna cues (Oldfield and Parker, 1984). Wightman and Kistler (1989) also found large intersubject differences in the magnitude components of the HRTFs. These differences were greatest in the 5-10 kHz region, where the pinna cues are most salient.

Second, the manner in which the front/back reversal was defined in this study may have falsely elevated the metric in the side sectors (2 and 5). A front/back reversal was defined as a subject response in the opposite (front/back) hemisphere from the target location. In the 90 and 270 degree regions, however, a response could be classified as a reversal even though the actual error was only a few degrees. For example, a target sound presented at 89 degrees and a subject response of that sound as being located at 91 degrees was treated as a reversal. To determine whether this was a factor, these data were reexamined by calculating the percentage of total sector reversals in 10-degree intervals for Sectors 2 and 5 in all three studies. The analysis revealed that most reversals occurred close to 60 degrees and 300 degrees in Studies I and II, but close to 90 and 270 degrees in Study III. Thus, the high rate of reversals in Study III could have been due, in part, to our definition of reversal. Perhaps a more suitable method would be to add a constraint to the reversal definition such that a subject response would not be classified as a reversal if it were in the opposite hemisphere than the target but ,within ±5 or ±10 degrees of 90 or 270 degrees.

5.3 Performance as a Function of Synthesizer (Study I)

The ANOVA showed no significant differences between synthesizers for the mean magnitude error (AL-204 - 15.3 degrees; DIRAD - 13.7 degrees), $F(1,5) = 1.58$, $p = 0.2647$; mean response time (AL-204 - 1.6 seconds; DIRAD - 1.7 seconds), $F(1,5) = 0.38$, $p = 0.5667$ or mean percentage of reversals (AL-204 - 23.1%; DIRAD - 19.1%), $F(1,5) = 1.22$, $p = 0.3195$.

5.3.1 Performance with each Synthesizer as a Function of Target Sector

There was a significant interaction between synthesizer type and target sector for mean magnitude error and mean percentage of reversals ($F(5,25) = 4.46$, $p = 0.0044$; $F(5,25) = 4.05$, $p = 0.0079$, respectively, see Figures 12 and 13). The interaction was not significant for the mean response time ($F(5,25) = 1.36$, $p = 0.2725$). The analysis of simple effects showed that the mean magnitude error was similar with both synthesizers, except at Sector 1, where this error was significantly higher with the AL-204 ($F(5,25) = 24.34$, $p < 0.05$). With regard to reversals (Figure 13), the graph indicates similar trends for each synthesizer. However, significantly more front/back reversals occurred with the AL-204 synthesizer than with the DIRAD synthesizer in Sectors 1, 5, and 6 ($F(1,25) = 11.76$, 6.13 and 4.39, respectively; $p < 0.05$).

5.3.2 Discussion of Performance as a Function of Synthesizer (AL-204)

The significant interaction between synthesizer type and target sector indicates that performance differences did occur between the two synthesizers in certain sectors. It should be remembered that, in Study I, the synthesizer comparison was limited to targets presented in the azimuth plane, with the subjects' head movements restricted. Data from the subject's questionnaire indicate that some subjects felt that it was easier to perceive the direction of the frontal sounds with the DIRAD synthesizer. Additional research with the AL-204 synthesizer is needed to further verify and investigate why the magnitude error increased when the target sounds were presented in Sector 1 (right front sector).

5.4 Performance as a Function of Target Type

As stated previously, the two synthesizers differed in the resolution of the measured HRTFs (DIRAD: 1 degree, AL-204: 10 degrees.). Consequently, the AL-204 employed interpolation routines to present sounds between two 10-degree intervals. Thus, it was of interest to determine whether interpolation affected localization indices for sounds presented in the azimuth plane.

There were no significant differences between the AL-204 synthesizer target types for the mean magnitude error (non-interpolated - 15.8 degrees; interpolated targets - 14.8 degrees), $F(1,5) = 5.10$, $p = 0.0754$, or mean response time (non-interpolated - 1.6 seconds; interpolated - 1.6 seconds), $F(1,5) = 0.9000$, $p = 0.3685$. However, there was a significant decrease in the mean percentage of reversals when localizing the non-interpolated targets (21.1%), compared to interpolated targets (25.2%), $F(1,5) = 28.42$, $p = 0.0031$.

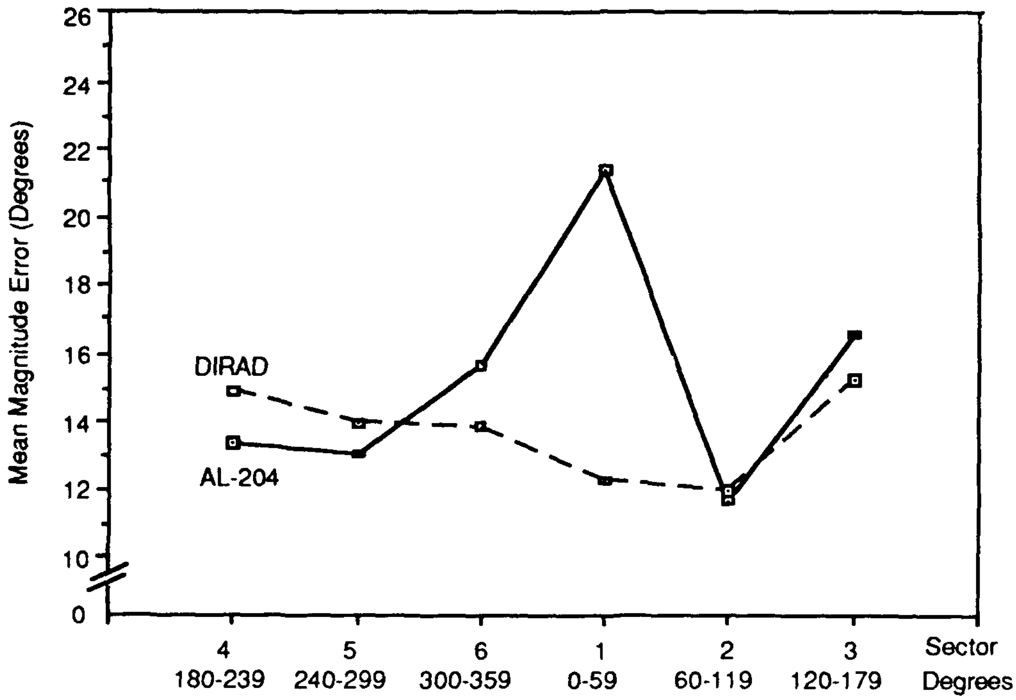

Figure 12. Mean Magnitude Error as a Function of Target Sector and Synthesizer (Study I)

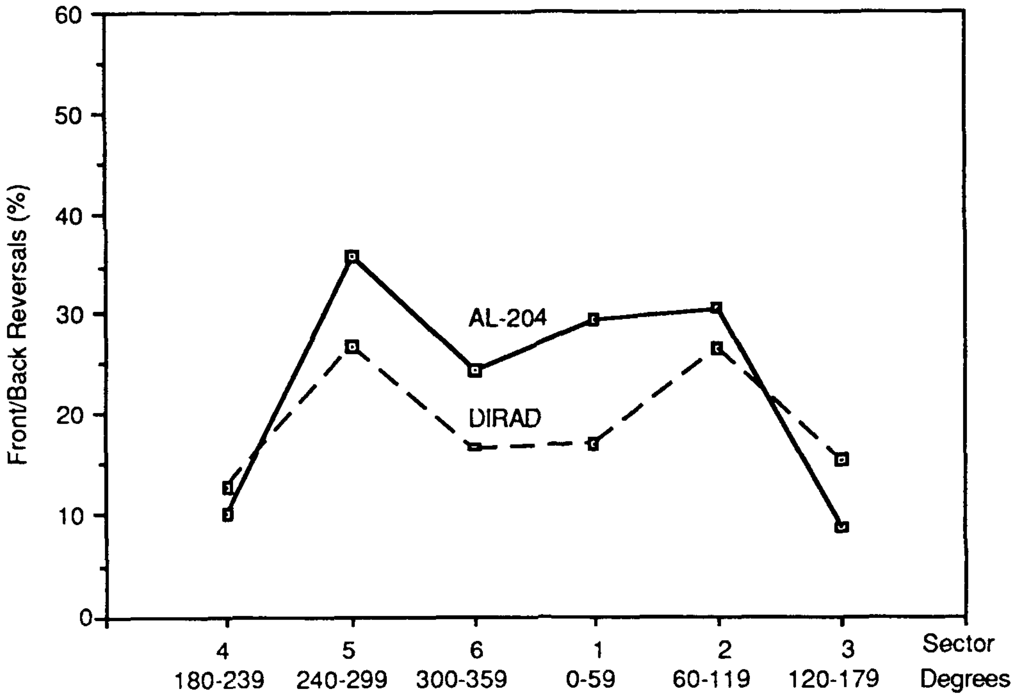

Figure 13. Mean Percentage of Front/Back Reversals as a Function of Target Sector and Synthesizer (Study I)

5.4.1 Performance with each Target Type as a Function of Target Sector

There was a significant interaction between target type and target sector for the mean magnitude error and mean percentage of reversals ($F(5,25) = 4.56$, $p=0.0043$, ($F(5,25) = 2.75$, $p = 0.0415$, respectively). This interaction was not significant for the mean response time. Referring to Figure 14, the mean magnitude error for interpolated and non-interpolated targets followed similar patterns as a function of target sector However, the mean magnitude error with the non-interpolated targets was significantly higher than the error with the interpolated targets at Sectors 3 and 6 ($F(1,25) = 13.94$ and 4.56, respectively; $p < 0.05$).

In addition, the percentage of front/back reversals for interpolated and non-interpolated targets, as a function of target sector, was similar in shape (see Figure 15). However, the number of reversals was significantly higher with the interpolated targets in Sectors 2 and 5 ($F(1,25) = 9.61$ and 14.17, respectively; $p < 0.05$).

5.4.2 Discussion of Performance as a Function of Target Type

In the AL-204 synthesizer, stored HRTFs are linearly interpolated between multiples of 10 degrees to present source positions. The results of this study indicate that interpolating between HRTFs did not result in increased magnitude error. In fact, in Sectors 3 (back-right) and 6 (front-left), the subjects' localization accuracy with the interpolated targets was significantly better than that with the non-interpolated targets. However, interpolating between HRTFs appeared to diminish the subjects' ability to resolve front/back confusions in Sectors 2 (right side) and 5 (left side). Thus, it is possible that employing algorithms to interpolate linearly between HRTFs may affect the localization indices in particular sectors. To date, no other research has been conducted to determine the effect of target type (i.e., interpolation) on localization performance. Additional investigations with this and other interpolating algorithms will help define the perceptual impact of interpolation on localization performance. If interpolation does not significantly affect localization performance, then the resources to measure HRTFs at every degree may not be required.

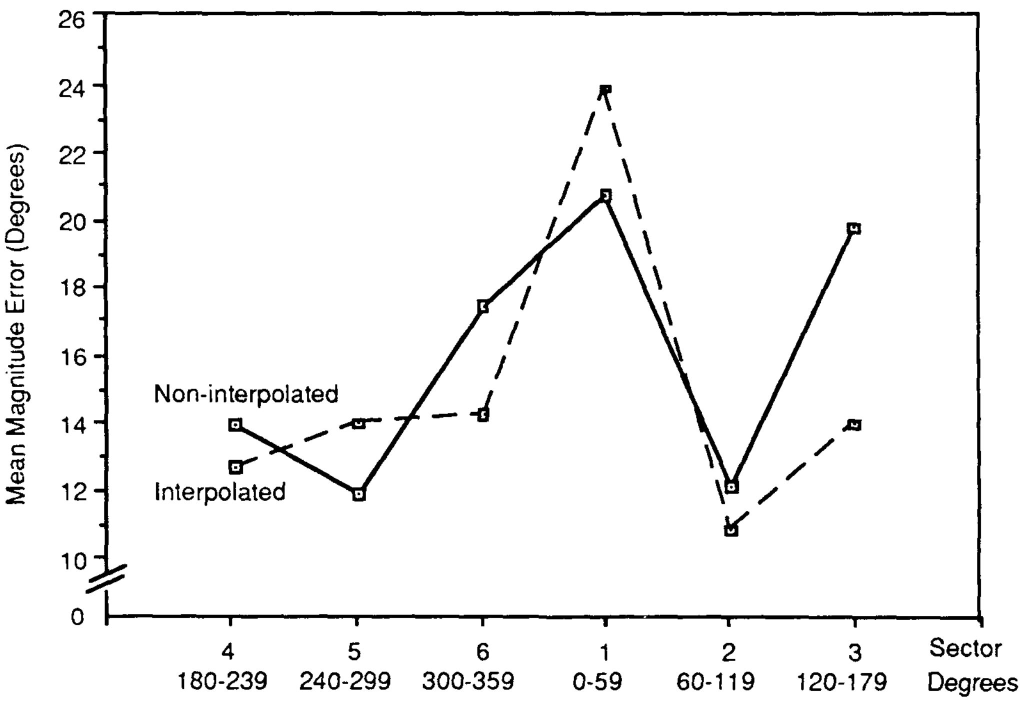

Figure 14. Mean Magnitude Error as a Function of Target Sector and Target Type for the AL-204 Synthesizer (Study I)

5.5 Performance as a Function of Head Movement Condition (Study III)

The ANOVA showed no significant difference in the mean magnitude error between the two head movement conditions, although the two differed by four degrees (fixed-head = 18.1 degrees; free-head = 13.9 degrees), $F(1,5) = 1.13$, $p = 0.3371$. Conversely, the mean response time and percentage of reversals did vary significantly as a function of head movement condition ($F(1,5) = 7.07$, $p = 0.0449$ and $F(1,5) = 7.59$, $p = 0.0401$, respectively). The mean response time was longer in the free-head condition than in the fixed-head condition (2.7 and 1.5 seconds, respectively). Additionally, the percentage of reversals was significantly greater in the fixed-head condition than in the free-head condition (means = 25% and 8%, respectively). No significant interactions were found.

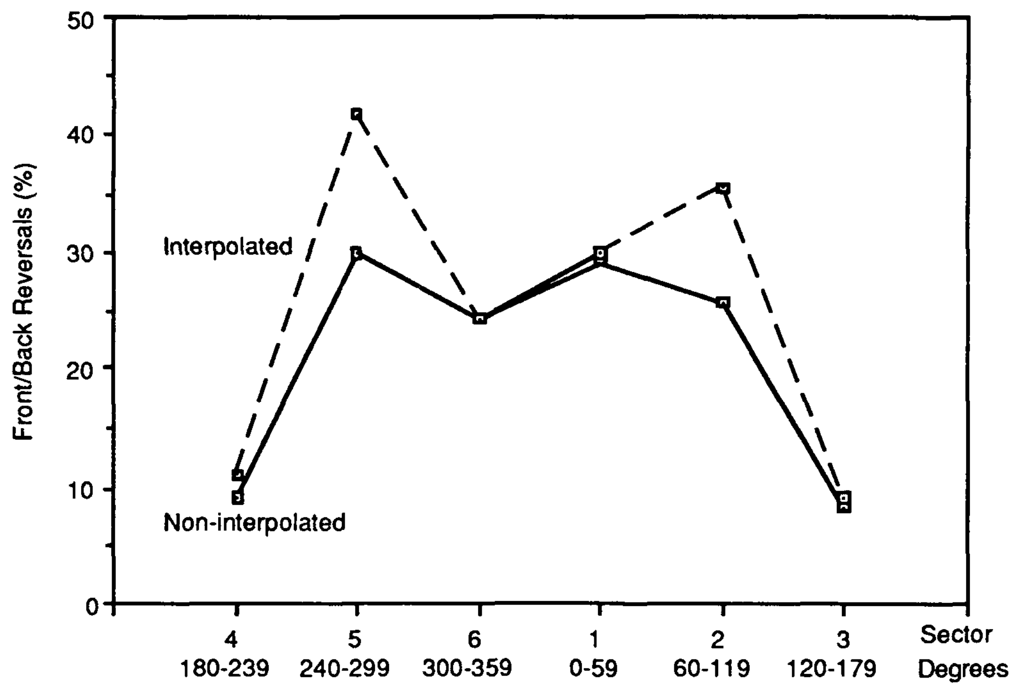

Figure 15. Mean Percentage of Front/Back Reversals as a Function of Target Sector and Target Type for the AL-204 Synthesizer (Study I)

5.5.1 Discussion of Performance as a Function of Head Movement

The results indicated that, when head movements were not restricted, localization performance improved slightly. The fact that the mean magnitude error was not significantly lower in the free-head condition, as one would expect, may reflect the subject's individual varying localization strategies. Data from the post-experimental questionnaire indicated that some subjects used a great deal of head movement (e.g., moved their heads and faced the sound), while others appeared to utilize smaller head movements. Also, the subjects reported that localization appeared easier when they were allowed to move their heads. In terms of the target locations, subjects reported using very little head movement for sounds in front, moderate amounts for sounds on the side and a great deal of movement when localizing the back targets. If the subjects had been instructed to face the sounds, the free-head condition may have produced significantly lower mean magnitude error. Facing the sound would have allowed the subjects to "zero-in" on the perceived direction by balancing out the interaural time and intensity differences, thus resulting in a more accurate perception of the stimulus direction.

The subjects tested in the free-head condition had significantly longer mean response times than in the fixed-head condition. This reflects the different "zeroing" response strategies employed by some of the subjects, which would elicit longer response times. The response times in the free-head condition also reflect the inherent latency in initiating a motor response, such as head movement.

The greatest impact of the head movement condition during localization was manifested in the occurrence of front/back reversals. A decrease from 25% to 8% occurred when the subjects were allowed head movement. One explanation (Wallach, 1940) may be that moving the head modulates the binaural acoustic cues, thus enhancing localization. Wallach suggested that these head movements are especially helpful in resolving front/back ambiguities.

5.6 Performance as a Function of Stimulus Type (Study III)

The mean magnitude error, mean response time and percentage of reversals varied significantly as a function of stimulus type ($F(3,15) = 4.28$, $p = 0.0227$; $F(3,15) = 16.21$, $p = 0.001$; and $F(3,15) = 5.91$, $p = 0.0072$, respectively; see Figures 16, 17, and 18). Tukey's test showed that subjects had significantly lower mean magnitude errors when localizing the wide-band pink noise (15.1 degrees) than the 1 kHz pure tone (17.5 degrees). The Tukey test also showed that subjects took a significantly longer time to determine the direction of the speech stimulus (2.5 seconds) compared to the wide-band pink noise (2.0 seconds), the 4 kHz low-pass noise (2.1 seconds) and the 1 kHz pure tone (1.8 seconds). In terms of reversals, significantly more occurred with the 1 kHz pure tone (19.7%) than with the wide-band pink noise (15.0%) or the 4 kHz low-pass signal (14.9%). No significant interactions were found.

5.6.1 Discussion of Performance as a Function of Stimulus Type

The stimulus type significantly affected localization performance. The pink noise signal was localized with significantly greater accuracy than the 1 kHz pure tone. This finding was consistent with questionnaire responses where the subjects reported the pink noise signal to be much easier to localize than the pure tone. Half of the subjects found it difficult to localize the tone and the female speech, while the broad-band signals (e.g., pink noise and 4 kHz low-pass noise) were rated moderately easy to localize. It would appear that the multiple range of frequencies in the broad-band signals generate a greater number of localization cues compared to the limited bandwidth of the pure tone (Doll et al., 1986). The pure tone caused significantly more front/back reversals to occur, compared to the pink noise and 4 kHz low-pass signal. The pure tone does not permit adequate spectral manipulation by the pinna to generate salient pinna cues. These cues have been shown to be important in resolving front/back localization ambiguities (Oldfield and Parker, 1984). The female speech signal required significantly longer time to localize than the other three stimuli. This may have been due to the speaker pauses between sentences which, according to several subjects' comments, inflated their response times and diminished their front/back discrimination abilities.

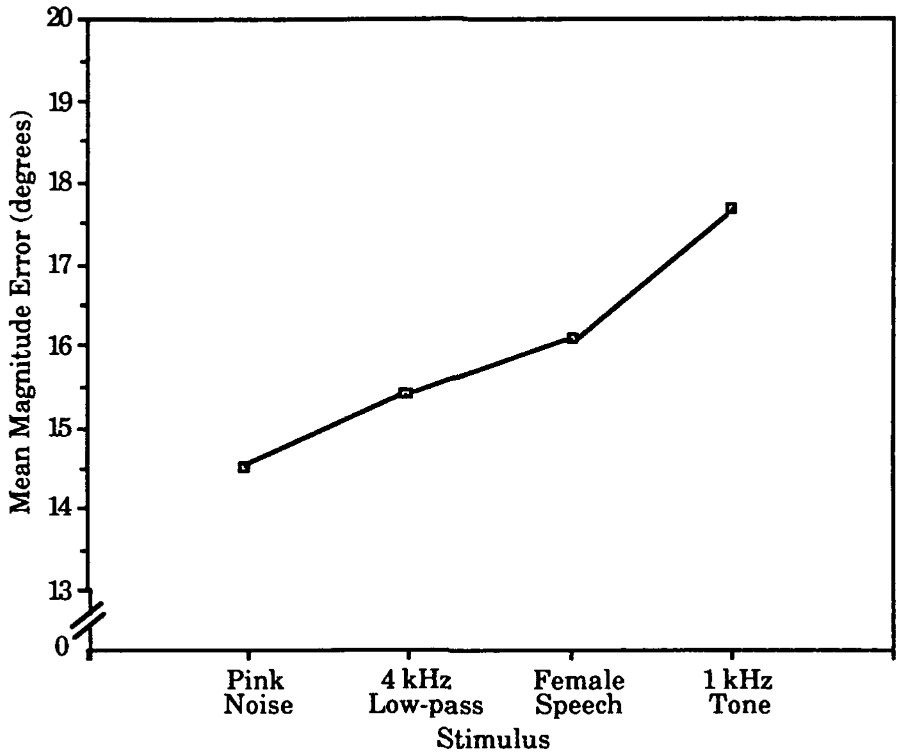

Figure 16. Mean Magnitude Error as a Function of Stimulus (Study III)

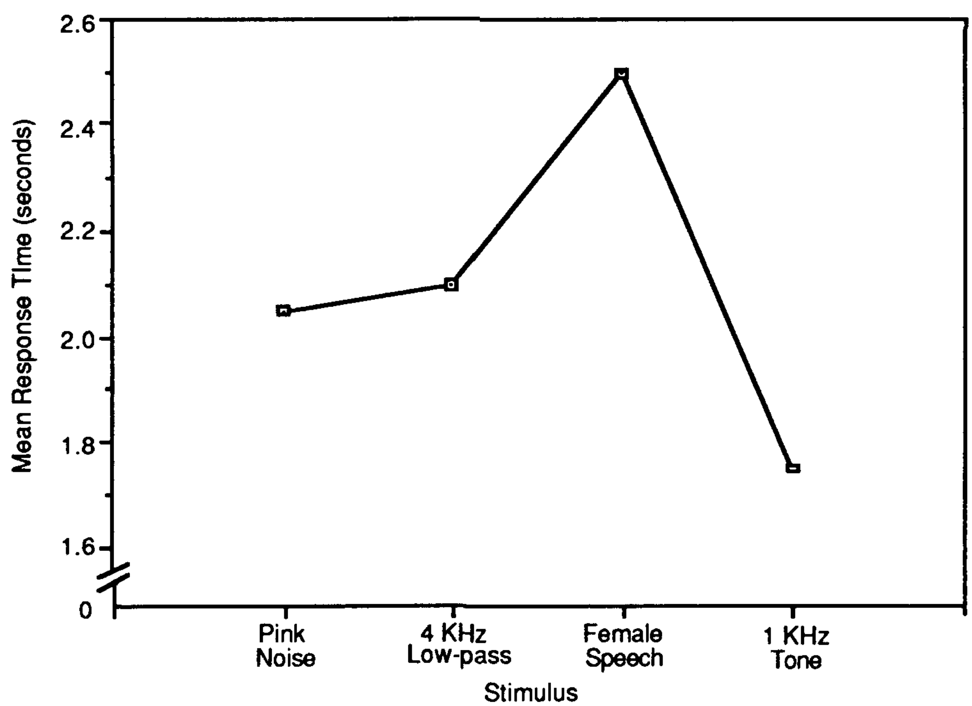

Figure 17. Mean Response Time as a Function of Stimulus (Study III)

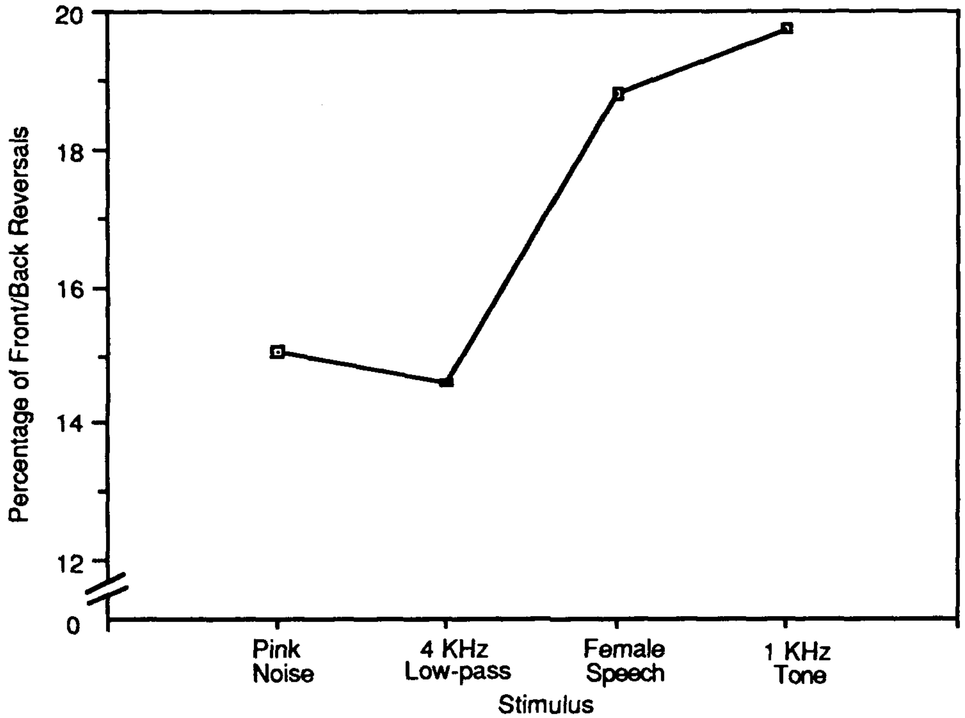

Figure 18. Mean Percentage of Front/Back Reversals as a Function of Stimulus (Study III)

5.7 Performance as a Function of Response Method (Study II)

The ANOVA revealed no significant difference in the mean magnitude error between the circle pointing (18.0 degrees) and verbal response (19.8 degrees) methods ($F(1,5) = 3.93$, $p = 0.1043$). The mean response time was also not significant between the circle pointing (1.8 seconds) and verbal response (1.6 seconds) methods ($F(1,5) = 1.23$, $p = 0.3179$). However, the percentage of reversals was sensitive to the effect of response method, with significantly more reversals occurring with the verbal response method (36.0%) than with the circle pointing response method (31.8%), ($F(1,5) = 6.99$, $p = 0.0458$). The interaction of response method and target sector was not significant.

5.7.1 Discussion of Performance as a Function of Response Method

The results obtained with each response method, circle pointing and verbal report estimate, were essentially the same, except for the percentage of front/back reversals. There appears to be no obvious explanation why significantly more reversals occurred with the verbal estimate than with the circle pointing method. The similar performance results can be attributed, in part, to the commonality of the procedures used with the two methods. A circle diagram, marked every degree and numbered every 10 degrees (see Figure A-1, Appendix A), was placed on the response table to assist the subjects in their angular estimate with the verbal response method. For the circle pointing method, a circle diagram marked and numbered at the 0, 90, 180 and 270 degree locations was used. Thus, each response method required the subject to complete a mental conversion by shifting from an "egocentric" perspective, when localizing the target sounds, to a "God's eye-view" perspective while reporting the direction of the sound. Perhaps if the subjects had not been provided with a reference card with the marked azimuths when using the verbal report method, significant differences may have occurred. Eight of the nine subjects reported that the reference card helped them in making verbal estimates of the sound direction. In addition, the subjects felt that they reported more accurately using the verbal response method.

Previous research (Wightman and Kistler, 1989) has found that the verbal report method, also called the absolute judgment method, tends to produce very stable judgments and is relatively easy for the subjects to learn. In the present studies, both methods were easy to use. Additional research is needed to determine the best or most suitable method of indicating the direction of target sounds which vary in azimuth, as well as in elevation. The inherent error in the response methods should be investigated as well.

6.0 CONCLUSIONS

The major findings from the three pilot localization studies are as follows:

1. Localization performance was fairly consistent across studies.

In order to compare localization performance across the three pilot studies, data collected under similar conditions (e.g., naive subjects and wide-band pink noise stimuli presented by the DIRAD synthesizer; using the verbal report method) are summarized in Table 5.

Table 5. Summary of Comparable Sessions from Localization Studies

	Study		
	I (n = 36)	II (n = 36)	III (n = 36)
Mean Magnitude Error (degrees)	16.3	19.8	16.5
Mean Response Time (seconds)	2.2	1.6	1.3
Mean Percentage of Reversals	28.4	36.1	24.0

The consistent effect of target sector on the localization indices across the three studies is also of particular interest. Table 6 shows the sectors, ranked from highest (top) to lowest (bottom) in terms of the means for each dependent variable. The same data used to generate Table 5 was used here. Notice that, for the most part, the lowest magnitude errors and response times occur in the side sectors (2,5) with the front (1,6) and back sectors (3,4) ranked higher. Also, the lowest incidence of front/back reversals generally occurred at the back sectors (3,4).

Table 6. Localization Performance Ranking by Dependent Variable and Target Sector

Mean Magnitude Error			Mean Response Time			Mean Percentage of Reversals		
Study			Study			Study		
I	II	III	I	II	III	I	II	III
1	1	4	1	6	4	5	1	5
3	3	3	6	1	3	2	6	2
6	6	6	3	4	1	1	2	3
4	4	1	5	5	6	6	5	6
5	2	2	2	3	2	3	4	1
2	5	5	4	2	5	4	3	4

2. The effect of target sector on localization performance with synthesized directional sound may differ from that observed with real sound sources.

A different effect of target sector on magnitude error was observed in the three studies compared to some past studies employing real sound sources. The overall localization error at the side sectors (2,5) was lower than that for the front (1,6) or back sectors (3,4). The majority of studies employing real sound sources (e.g., Oldfield and Parker, 1984; Mills, 1958) have found the lowest magnitude errors in the frontal regions.

3. Prior experience in listening to synthesized directional acoustic stimuli enhances localization performance.

Experienced subjects performed the localization task with lower magnitude errors and with a lower incidence of front/back reversals. The largest impact of experience was manifested in the number of reversals.

4. A wide-range of front/back discrimination ability was demonstrated between the subjects.

Subjects, localizing the wide-band pink noise stimulus with their heads fixed, for example, showed incidences of reversals ranging from 8% to 38%. This variability could be possibly attributed to individual perceptual abilities, as well as localization cues generated via KEMAR. It may be that, when localizing with KEMAR-modeled stimuli, individual front/back discrimination abilities may be a function of how close the subject's pinna match those of KEMAR. In essence, the listeners' ability to localize through KEMAR's ears may have played a role in performance.

5. Performance with the AL-204 synthesizer differs from that with the DIRAD synthesizer at certain sectors.

Mean magnitude error was similar with both synthesizers, except at Sector 1, where this error was significantly higher with the AL-204. With regard to reversals, similar trends were found for each synthesizer. However, significantly more front/back reversals occurred with the AL-204 synthesizer than with the DIRAD synthesizer in Sectors 1, 5, and 6.

6. Interpolating between HRTFs affects front/back discrimination.

With the AL-204 synthesizer, utilizing interpolation to present targets between two HRTFs did not affect the magnitude of the localization error, but it did decrease the subjects' ability to discern front targets from back targets.

7. Head movement enhances localization performance.

When subjects are free to move their heads during the localization task, the number of front/back reversals is significantly reduced, since head movements provide enhanced binaural acoustic cues. The magnitude of the localization error was reduced, although not significantly.

8. Stimulus characteristics affect localization performance.

The subjects localized the wide-band stimulus (pink noise) with greater accuracy and less front/back reversals than with the narrow-band 1 kHz pure tone. It may be that the wide range of frequencies present in the pink noise signal generates a greater number of localization cues compared to the limited bandwidth of the tone, thus enhancing localization performance.

9. The circle pointing and verbal response methods (both utilizing cards depicting degrees on a circle) elicit comparable localization performance, in terms of magnitude error and response time. However, significantly more reversals occurred with the verbal response method than with the circle pointing method.

7.0 SUMMARY

The research described in this report summarized three pilot experiments designed to measure and compare the ability of subjects to localize sounds in the azimuth, via headphones, generated by two prototype auditory localization cue synthesizers. In the first study, localization performance differences were found between the two synthesizers in certain areas of the azimuth plane (e.g., front area). The experience or familiarity in listening to synthesized directional sound also significantly affected localization performance. Additionally, the design of a synthesizer (e.g., resolution and use of interpolation between KEMAR HRTFs) can impact the perceived direction of the acoustic signals. Wightman and Kistler (1990) suggest that the veridicality of 3-D auditory displays could be optimized if individualized HRTFs are employed to synthesize the virtual sound sources, particularly in the elevation plane. However, data from the present experiment suggest that this design requirement can be relaxed, especially if only azimuth information is to be conveyed by the auditory localization cue synthesizer.

In the second study, two response methods for measuring localization performance were evaluated. The results indicate that no performance differences were found when subjects either verbally reported angular estimates or pointed to a circle to indicate the perceived direction of the target stimuli. In the third study, localization performance was impacted by manipulating the bandwidth of the acoustic signal and head movement. Performance was enhanced when a broad-band signal (e.g., pink noise) was used as the test signal. By allowing the subjects to move their heads freely while determining the direction of target stimuli, localization performance was enhanced.

Results from the data collected in these studies indicate that subjects are able to discriminate and externalize synthesized directional acoustic signals which are presented over headphones by these prototype devices. Since these auditory cue synthesizers are portable and can be easily integrated into flight simulators, they promise to be valuable research tools for studying the utility of an auditory display on increasing threat assessment, enhancing communication discriminability and alleviating operator workload.

8.0 BIBLIOGRAPHY

Butler, R.A. (1969). On the relative usefulness of monoaural and binaural cues in locating sound in space. Psychonomic Science, 17, 245-246.

Dirks, D.D. and Gilman, S. (1979). Exploring azimuth effects with an Anthropometric Manikin. Journal of the Acoustical Society of America, 66, 696-701.

Doll, T.J., Gerth, J.M., Engelman, W.R. and Folds, D.J. (1986). Development of simulated directional audio for cockpit applications (AAMRL-TR-86-014). Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory.

Ericson, M.A. and McKinley, R.L. (1989). Auditory localization cue synthesis and human performance. In Proceedings of National Aerospace Electronics Conference, Vol 2, pp. 718-725.

Friedman, S.J. and Fisher, H.G. (1968). The role of the pinna in auditory localization. In S.J. Freedman (Ed.), The Neuropsychology of Spatially Oriented Behavior (pp. 135-152). Homewood, IL: Dorsey Press.

Mills, A.W. (1958). On the minimum audible angle. Journal of the Acoustical Society of America, 30, 237-246.

Oldfield, S.R. and Parker, S.P. (1984). Acuity of sound localization: a topography of auditory space. Perception, 13, 581-617. I. Normal hearing conditions: 581-600, 2. Pinna cues absent: 601-617.

Valencia, G. and Calhoun, G. L. (1988). Headphone localization with a single sound transducer and movable manikin head (AAMRL-TR-89-005). Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory.

Wallach, H. (1940). The role of head movements and vestibular and visual cues in sound localization. Journal of Experimental Psychology, 27, 339-368.

Weinrich, L. (1982). The problem of front-back localization in binaural hearing. In O.J. Pedersen and T. Poulsen (Eds.), Binaural Effects in Normal and Impaired Hearing (Scandinavian Audiological Supplement No. 15, pp. 135-145).

Wightman, F. L., Kistler, D. J., and Perkins, T. (1988). In W.A. Yost and W. Gourenvitch (Eds.). Directional Hearing, Springer Verlag Press.

Wightman, F. L. and Kistler, D. J. (1989). Headphone simulation of free-field listening - II: Psychophysical Validation. Journal of the Acoustical Society of America, 85, 868-878.

Wightman, F. L. and Kistler, D. J. (1990). Field Measurement of Head Related Transfer Functions (AAMRL-TR-90-019). Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory.

9.0 APPENDICES

APPENDIX A

SUBJECT'S INSTRUCTIONS FOR GENERAL AUDITORY LOCALIZATION STUDY

Introduction

Thank-you for serving as a subject in the auditory localization experiment. The purpose of this experiment is to examine your ability to determine the direction of audio signals presented to you over headphones. That is, we want to examine how well you can determine the direction of sounds presented around you. Evaluation of your performance as well as your responses to a questionnaire will help indicate the feasibility of using 3-D sound to enhance a pilot's situation awareness in advanced crew station design concepts.

Procedure Set Up

You will be asked to stand on top of an "X" at the center of the curtain area and place a pair of stereo headsets on your head. You will be provided with a hand-switch to start/stop the signals. You will also be asked to rest your head on the chinrest and look straight ahead during the trials. Once you are positioned correctly, the experimenter will close the curtain. The experimenter will always be in communication with you via a microphone attached to the headsets.

Experimental Task

In this experiment, you will be asked to determine the direction of a sound presented (through the headsets) around you. Two methods of indicating the perceived locations of the sounds may be evaluated in this study. These are described below. With the first method, verbal report, you will be facing a circle diagram (see Figure A-1) and verbally reporting the location of the sound in terms of degrees (0 to 359). For example, if the sound appears to be originating directly in front of you, you would respond "0 degrees" and if it originated directly at your left ear, you would respond "270 degrees". You are asked to provide your best estimate of the sound direction to the nearest degree. The circle diagram is an aid to help you report a numerical estimate.

In the second method, circle pointing, you will use a special pen to depress one area of the circle where you estimate the sound to originate. Your response should be based on the assumption that the circle represents the area around you, with you at the center of the circle. Only point to one location and make sure you point on the circle, not inside or outside of the circle. A computer will automatically record where you depressed the pen on the circle.

Regardless of the particular response method used, the procedure will be as follows. You will place your chin on the chinrest and look straight ahead. For every trial, you will be presented one sound initially; this is a boresight sound for which you do not need to determine its direction. Press the hand-switch and the boresight sound will stop and, one second later, the test sound will be presented. Once you have determined the direction of the sound, press the hand-switch a second time. Next, you are to verbally report or indicate on the circle where you perceived the direction of the sound.

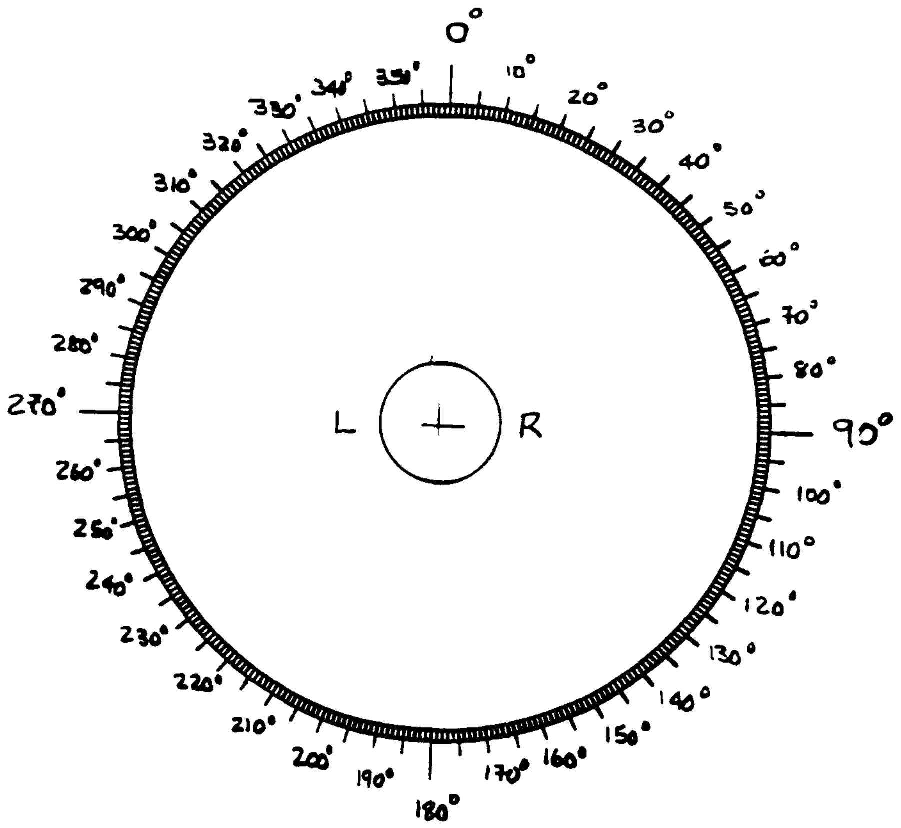

Figure A-1. Example of Degree Estimates for Experiment

If you are using the circle pointing method, place the pen on the point on the circle representing the direction and push down on the pen. Keep the pen on the circle until you hear the boresight sound again, which starts the next trial.

It is very important that you press the hand-switch when you determine the direction and then respond by one of the two methods. This insures an accurate response time measurement.

In addition, two different sound generating systems may be used in this study. System #1 uses a pistol hand-switch and System #2 uses a hand-held push-button. Try to remember any differences in sound quality, ability to determine sound location, etc., when using the two systems/response switches. Finally, you may be localizing different types of auditory signals (tones, static noise, speech, etc.) in a fixed-head and free-head condition. For the fixed-head condition, please keep your head positioned on the chin rest and do not move your head during a stimulus presentation. For the free-head condition, you are free to move your head to determine the direction of the stimulus. For both response methods, the next trial will begin two seconds after your response.

Precautions. If an emergency should occur and evacuation of the laboratory is required, perform the following steps, in order:

a) remove headset.

b) open velcro-attached curtain.

c) exit through one of the two doors in the lab.

Do you have any questions?

APPENDIX B

AUDITORY LOCALIZATION STUDY QUESTIONNAIRE

Immediately following the final data collection session of each study, each subject completed a questionnaire concerning various aspects of the study. The subjects' responses to the questions are presented below. Following each question, in parenthesis, is the study or studies pertinent to that question. Note: Only the naive subjects in Study I completed this questionnaire.

1. During each session, the sounds were presented at different locations around you. Rate the difficulty in perceiving the sounds for the following locations (Studies I, II, III).

	Never Perceived	Very Difficult	Moderately Difficult	No Opinion	Moderately Easy	Very Easy
In Front	2	0	6	0	3	4
In Back	0	2	6	0	3	4
Right Side	0	0	0	0	8	7
Left Side	0	0	1	0	8	6

Was this true with both synthesizers? (Study I) Yes: 2 No: 1

Comments:

- The AL-204 was louder than the DIRAD synthesizer.
- With the AL-204, it was harder to distinguish 60-90 and 270-290 degrees sounds.
- The DIRAD seemed to have clearer sound (especially perceiving frontal noise).
- I never perceived sounds directly at 90 or 270 degrees.

2. Did you perceive the sounds to be located outside your head? (Studies I, II, III)

Yes: 14 No: 1

If yes, where were the sounds located?

Close to your head	4
Inside the curtain area	5
Close to the curtain	3
Outside the curtain area	2

Was this true for all locations? Yes: 7 No: 2

Was this true for both synthesizers? (Study I) Yes: 3 No: 0

Comments:

- Frontal sounds seemed to be closer than the back sounds.
- The sounds seemed very close to my head when the sound was at 90 or 270 degrees.
- Some sounds seemed close, some further away.

3. Did you perceive any of the sounds to be above or below the horizontal plane through your ears? (Studies I, II, III)

Yes: 8 No: 7

If yes, was this true for all locations? Yes: 3 No: 5

Was this true for both synthesizers? (Study I) Yes: 3 No: 0

Comments:

- Just a few of the 180-degrees signals and most, if not all, of the frontal area signals were not at the horizontal plane.
- It sounded as if some sounds were higher and others lower than the horizontal plane.
- Sometimes the sounds seemed to be located evenly with the left or right ear, along the horizontal plane.
- Sounds above or below the horizontal plane seemed to occur at the right and left side of the head.
- One of the "jet noises", [wide-band pink noise or 4 KHz low-pass pink noise] the sounds in front (270-90 degrees) sounded like it was coming from about a 45-degree angle above the horizontal plane. In the back (90-270 degrees), it sounded like it was horizontal.

4. In regards to each block of 24 trials, the difficulty level in accomplishing the localization task tended to be: (Studies I, II, III)

Constant across blocks	10
More difficult in the first block	0
Easier in the first block	0

Varied across the blocks	5
Can't remember/no opinion	0

Was this true for both synthesizers? (Study I) Yes: 3 No: 0

5. In regards to both the sessions in this experiment, the difficulty level in accomplishing the localization task tended to be (Studies I, II, III):

Constant across days	7
More difficult the first day	5
Easier in the first day	0
Varied across the days	3
Can't remember/no opinion	0

Was this true for both synthesizers? Yes: 3 No: 0

6. Were you able to maintain your concentration during each block of trials? (Studies I, II, III)

Yes: 14 No: 1

7. When required to keep your head steady, did you find it difficult to do so? (Studies I, II, III)

Yes: 1 No: 14

Was this true for all sound locations? (Study I) Yes: 13 No: 2

8. Rate how comfortable you found the following: (Studies I, II, III)

	Very Un-comfort.	Slightly Uncomfort.	Not Noticeable/ No Opinion	Slightly Comfort.	Very Comfort.
headset	1	1	3	2	8
response equip	0	2	2	5	6
*static sounds	0	1	3	4	1

*Subjects were not asked to rate the static sound in Study III. Thus, only 9 responses are recorded.

Comments:

- The response switch for the DIRAD synthesizer was more comfortable than the switch for the AL-204.
- After the 6th run of trials, the sound got on my nerves.

- Didn't like the response switch we had to push.
- Standing was the only thing that was uncomfortable.
- Became uncomfortable half-way through the trials.

9. Did the circle diagram help you in making a verbal estimate of the sound location? (Studies I, II)

Yes: 8 No: 1

Was this true for all locations? Yes: 4 No: 5

10. Rate the difficulty in transferring your perception of the locations of the sound to numerical degree estimates. (Studies I, II)

Very Easy	2
Easy	2
Satisfactory	3
Hard	2
Very Hard	0

Was this true for all locations? Yes: 4 No: 5

Was this true for both synthesizers? (Study I) Yes: 3 No: 0

Comments:

- The sounds in the back were the most difficult. Zero-degrees was the easiest, because it sounded just like the static (boresight) that preceded it.
- 110-180 degrees and 180-240 degrees were hard to tell apart.
- 170-23 was very easy, 90-170 and 23[illegible]70 was easy and 270-290 was hard.
- As the sounds got closer to 90, 180, and 270, they became easier to identify as coming from a specific direction.
- The sound occasionally seemed as though it was located directly above you. In these instances, it was difficult to determine whether to choose the zero-degree or 180-degree area.

11. Rate the difficulty in transferring your perception of the sound locations to a position on the circle located on the table. (Studies II, III)

Very Easy	2
Easy	5
Satisfactory	1
Hard	4
Very Hard	0

Was this true for all locations? Yes: 11 No: 1

Comments:

- The sound occasionally seemed as though it was located above. This made it difficult sometimes.
- I think it would be easier to mark the point on the circle if there were more numbers on the circle and also points marked around the curtain so they could be matched up.
- It was hard to be very accurate. I knew where the sound came from, but it was hard to find the spot on the paper sometimes.

12. Which method did you feel you were more accurate with in reporting the sound location? (Study II)

Circle pointing: 2 Verbal report: 4

13. Were there any noises outside the curtain which disturbed your concentration/performance? (Study III)

Yes: 1 No: 5

Comments:

- Although noises, such as footsteps, could be heard at times, they weren't much of a distraction.
- I could hear people talking in other sections, computers printing, but not all days.

14. Rate the difficulty of determining the direction of the following sounds: (Study III)

	Very Difficult	Moderately Difficult	No Opinion	Moderately Easy	Very Easy
Female Speech	0	3	1	2	0
Static I (pink noise)	0	1	0	4	1
Tone (1 kHz sine)	2	1	0	3	0
Static II (low-pass)	0	2	0	4	0

Comments:

- Speech was more difficult, because it was not a constant stimulus like the other three.

- The tone seemed louder on beginning days and very annoying to listen to, let alone determine the direction.

15. Did you find determining the location of the sound was easier when you were allowed to move your head? (Study III)

Yes: 6 No: 0

16. When you were free to move your head, rate how much head movement you used in determining the direction of the sound for the following locations: (Study III)

	None	Very little	Moderate amount	A great deal
Sounds in front	0	6	0	0
Sounds on side	1	0	5	0
Sounds in back	0	0	0	6

Did this amount of head movement vary according to the type of sound you were listening to?

Yes: 2 No: 4

Comments:

- I had to move my head more with the speech and tone.
- Speech required the most head movement, then tones, then static.

17. Do you have a preference concerning the different stimuli (sounds) presented in terms of distinguishability or accuracy? (Study III)

Yes: 4 No: 2

Comments:

- The louder static [wide-band pink noise] seemed to be much easier to localize than the other three.
- The louder static was easier to tell where it came from.
- Static and speech were at comfortable levels, the tones were not.
- Static sounds are easiest. The softer static was slightly better than the louder static. The louder static sound was "harsher" and seemed to cover a wider band of degrees.

18. Was the light pen easy to operate, that is, did it consistently respond to your first input? (Study III)

Yes: 1 No: 5

Comments:

- Four to five times a run it did not register, due to my error in pushing the button almost simultaneously with placing the pen.
- It did not always work at first, but after moving it a little bit it worked.
- Most of the time I needed to use it twice, sometimes three times, to continue.

19. Do you feel you were given adequate instructions and training prior to the start of the experiment? (Study III)

Yes: 6 No: 0

20. Did you know what was expected of you while participating in the experiment? (Study III)

Yes: 6 No: 0

Comment: What I didn't know, I understood after the practice trials.

21. Were the sessions too long? (Study III)

Yes: 0 No: 6

22. Have you participated in a similar experiment before? (Study III)

Yes: 6 No: 0

Comments:

- I took part in another localization experiment in which I had to turn my entire body toward the sound, and the difference between my body angle to the boresight and the given sound angle was taken as data.
- I am doing an experiment similar to this at this time. I started it the same day this experiment started.
- I took part in a localization experiment involving a ball divided into segments.
- Presently in a study where you position your body towards the sound.
- In an experiment where you face the sound.
- Experiment where you move your body in the direction of the stimulus.

23. General comments:

- I definitely would not recommend tone stimuli for localization. (Study I)
- It might be more accurate if a tone was used instead of static. A tone would be a "cleaner" noise and subjects could arrive at a more exact heading. (Study II)
- Sessions were too long. (Study II)
- The sessions didn't last too long, the equipment worked great, and I was adequately briefed about the experiments. (Study II)
- You might cut this experiment down to two or three days instead of four. I don't think that would be too much for the subjects. (Study I)
- It seemed there could have been more testing in the 270 to 360 degree directions. (Study I)

APPENDIX C

TABLES OF RESULTS FOR STUDIES

Table C-1 (Study I)

Mean Magnitude Error (degrees) as a Function of Synthesizer, Target Type and Target Sector

Synthesizer	Target Type	Target Sector (degrees) 180-239	240-299	300-359	0-59	60-119	120-179	Group Means
AL-204	Non-interpolated	13.9	12.1	17.1	20.1	12.3	19.2	15.8
	Interpolated	12.8	14.0	14.2	22.9	11.2	14.0	14.9
	Mean:	13.4	13.1	15.7	21.5	11.8	16.6	15.4
DIRAD	Non-interpolated	15.2	14.2	15.0	11.3	11.1	16.1	13.8
	Interpolated	14.6	13.7	12.7	13.3	12.9	14.5	13.6
	Mean:	14.9	14.0	13.9	12.3	12.0	15.3	13.7
Overall Means:		14.2	13.6	14.6	16.9	11.9	16.0	14.6

Table C-2 (Study I)

Mean Response Time (seconds) as a Function of Synthesizer, Target Type and Target Sector

		Target Sector (degrees)						Group
Synthesizer	Target Type	180-239	240-299	300-359	0-59	60-119	120-179	Means
AL-204	Non-interpolated	1.6	1.6	1.7	1.7	1.6	1.6	1.6
	Interpolated	1.6	1.6	1.6	1.6	1.6	1.5	1.6
	Mean:	1.6	1.6	1.7	1.6	1.6	1.6	1.6
DIRAD	Non-interpolated	1.6	1.6	1.8	1.9	1.6	1.8	1.7
	Interpolated	1.6	1.8	1.7	1.9	1.7	1.7	1.7
	Mean:	1.6	1.6	1.7	1.9	1.6	1.7	1.7
Overall Means:		1.6	1.7	1.6	1.8	1.6	1.7	1.7

Table C-3 (Study I)

Percentage of Front/Back Reversals as a Function of
Synthesizer, Target Type and Target Sector

Synthesizer	Target Type	Target Sector (degrees) 180-239	240-299	300-359	0-59	60-119	120-179	Group Means
AL-204	Non-interpolated	9.0	29.9	24.3	29.2	25.7	8.3	21.1
	Interpolated	11.1	41.7	24.3	29.9	35.4	9.0	25.2
	Mean:	10.1	35.8	24.3	29.5	30.6	8.7	23.1
DIRAD	Non-interpolated	13.2	23.6	16.0	18.1	23.6	13.9	18.1
	Interpolated	12.5	29.9	17.4	16.0	29.2	16.7	20.3
	Mean:	12.8	26.7	16.7	17.0	26.4	15.3	19.2
Overall Means:		11.5	31.2	20.5	23.3	28.5	12.0	21.2

Table C-4 (Study II)

Mean Magnitude Error (degrees) as a Function of
Response Method and Target Sector

Response Method	Target Sector (degrees) 180-239	240-299	300-359	0-59	60-119	120-179	Group Means
Verbal Estimates	20.7	13.7	22.2	25.5	13.7	23.0	19.8
Circle Pointing	19.6	13.6	19.9	19.9	13.9	21.1	18.0
Overall Means:	20.2	13.7	21.1	22.7	13.8	22.1	18.9

Table C-5 (Study II)

Mean Response Time (seconds) as a Function of
Response Method and Target Sector

Response Method	Target Sector (degrees) 180-239	240-299	300-359	0-59	60-119	120-179	Group Means
Verbal Estimates	1.6	1.6	1.8	1.7	1.5	1.6	1.6
Circle Pointing	1.8	1.7	1.8	1.8	1.7	1.8	1.8
Overall Means	1.7	1.7	1.8	1.8	1.6	1.7	1.7

Table C-6 (Study II)

Percentage of Front/Back Reversals as a Function of Response Method and Target Sector

	Target Sector (degrees)						
Response Method	180-239	240-299	300-359	0-59	60-119	120-179	Group Means
Verbal Estimates	15.3	36.3	51.6	57.6	40.3	15.3	36.1
Circle Pointing	16.4	31.5	44.9	50.0	32.9	14.8	31.8
Overall Means:	15.9	33.9	48.3	53.8	36.6	15.0	34.0

Table C-7 (Study III)

Mean Magnitude Error (degrees) as a Function of
Head Condition, Stimulus and Target Sector

Head Condition	Stimulus	Target Sector (degrees)						Group Means
		180-239	240-299	300-359	0-59	60-119	120-179	
Fixed	Pink Noise (PN)	18.5	10.7	21.9	15.5	14.3	18.7	16.6
	4 kHz Low-pass PN	19.0	11.1	20.1	18.2	13.2	18.6	16.7
	Female Speech	25.0	11.1	22.0	18.5	14.7	21.6	18.8
	1 kHz Tone	25.8	14.3	25.7	19.5	16.3	20.6	20.3
	Mean:	22.1	11.8	22.4	17.9	14.6	19.9	18.1
Free	Pink Noise (PN)	14.9	10.8	13.4	13.5	12.2	16.6	13.6
	4 kHz Low-Pass PN	17.0	10.3	13.6	12.4	14.2	16.8	14.0
	Female Speech	15.1	10.1	14.0	13.3	12.7	16.3	13.6
	1 kHz Tone	15.7	12.2	13.5	14.6	14.1	17.6	14.6
	Mean:	15.7	10.8	13.6	13.5	13.3	16.8	14.0
Overall Means:		18.9	11.3	18.0	15.7	14.0	18.3	16.1

Table C-8 (Study III)

Mean Response Time (seconds) as a Function of
Head Condition, Stimulus and Target Sector

Head Condition	Stimulus	Target Sector (degrees) 180-239	240-299	300-359	0-59	60-119	120-179	Group Means
Fixed	Pink Noise (PN)	1.4	1.2	1.4	1.4	1.2	1.4	1.3
	4 kHz Low-Pass PN	1.7	1.5	1.6	1.6	1.5	1.7	1.6
	Female Speech	2.1	1.5	1.8	1.7	1.6	2.0	1.8
	1 kHz Tone	1.4	1.2	1.3	1.4	1.2	1.4	1.3
	Mean:	1.6	1.3	1.5	1.6	1.4	1.6	1.5
Free	Pink Noise (PN)	3.0	2.6	2.7	2.8	2.6	3.0	2.8
	4 kHz Low-Pass PN	2.8	2.4	2.6	2.7	2.3	2.7	2.6
	Female Speech	3.6	2.9	2.9	3.4	3.0	3.4	3.2
	1 kHz Tone	2.5	2.0	2.1	2.3	2.0	2.4	2.2
	Mean:	3.0	2.4	2.6	2.8	2.5	2.9	2.7
Overall Means:		2.3	1.9	2.1	2.2	1.9	2.2	2.1

Table C-9 (Study III)

Mean Percentage of Front/Back Reversals as a Function of Head Condition, Stimulus and Target Sector

Head Condition	Stimulus	Target Sector (degrees) 180-239	240-299	300-359	0-59	60-119	120-179	Group Means
Fixed	Pink Noise (PN)	14.6	29.9	26.4	20.8	26.4	25.7	24.0
	4 kHz Low-Pass PN	6.3	34.0	34.7	19.4	23.6	17.4	22.6
	Female Speech	20.1	38.2	12.5	18.8	28.5	27.8	24.3
	1 kHz Tone	33.3	34.7	20.1	20.1	35.4	31.9	29.3
	Mean:	18.6	34.2	23.4	19.8	28.5	25.7	25.0
Free	Pink Noise (PN)	0.0	13.2	2.1	0.7	16.7	4.2	6.1
	4 kHz Low-Pass PN	0.7	14.6	2.8	1.4	20.2	3.5	7.2
	Female Speech	0.7	17.4	6.3	3.5	18.8	4.2	8.4
	1 kHz Tone	0.7	20.1	9.0	6.3	20.1	4.9	10.2
	Mean:	0.5	16.3	5.0	3.0	18.9	4.2	8.0
Overall Means:		9.5	25.3	14.2	11.4	23.7	14.9	

☆ U.S.Government Printing Office: 1990—548-018/20217