VIRTUAL AUDIO LOCALIZATION WITH SIMULATED EARLY REFLECTIONS
AND GENERALIZED HEAD-RELATED TRANSFER FUNCTIONS

by

Darrin Kiysohi Reed

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Electrical Engineering

MONTANA STATE UNIVERSITY
Bozeman, Montana

November 2009
APPROVAL

of a thesis submitted by

Darrin Kiyoshi Reed

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citation, bibliographic style, and consistency, and is ready for submission to the Division of Graduate Education.

Dr. Rob C. Maher

Approved for the Department Electrical and Computer Engineering

Dr. Rob C. Maher

Approved for the Division of Graduate Education

Dr. Carl A. Fox
STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master’s degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library.

If I have indicated my intention to copyright this thesis by including a copyright notice page, copying is allowable only for scholarly purposes, consistent with “fair use” as prescribed in the U.S. Copyright Law. Requests for permission for extended quotation from or reproduction of this thesis in whole or in parts may be granted only by the copyright holder.

Darrin Kiyoshi Reed

November 2009
DEDICATION

To Mom and Dad... without your love, life’s teachings and unwavering support, none of the accomplishments of my education would have been realized.
ACKNOWLEDGEMENTS

Much thanks is owed to those directly responsible for the completion of my work at Montana State University. To Dr. Rob Maher I extend the utmost gratitude for allowing me to grow as a researcher under his guidance and to pursue my passion of audio technology. To Dr. Ross Snider and Dr. Bob Gunderson I owe sincere thanks for their time, teachings, and practical points of view. The assistance of these three mentors provided the crucial elements for a fulfilling passage. Other thanks to Dr. Kevin Repasky and Dr. Alan Leech is well deserved for their assistance with the completion of this research. Specific thanks are owed to Pearl Nixon and Muriel Holmquist for helping to maintain sanity amongst the endless details. Finally, I owe much thanks to all the experiment participants for this research, particularly Diego DeLeon for his help with preliminary testing and setup.

Other people that have made my journey possible must be recognized. Thank you to Jason Ringelstetter and Chris Parks for giving me the opportunity to be a part of your recording studio (and to “hang out with rock-stars”). You are a key reason for why I am here today. Special thanks to Dave Jensen for all of the practical mentoring in live-sound and general sound principles. To Dr. Elisa Barney-Smith and Dr. Jacob Baker for providing a model of education that I will always strive towards. Thank you to all my family and friends.
TABLE OF CONTENTS

1. INTRODUCTION .............................................................................................................1
   Motivation .....................................................................................................................2
   Thesis Organization ....................................................................................................3

2. BACKGROUND ...............................................................................................................5
   Lateralization vs. Localization ....................................................................................6
   Duplex Theory: ITD and ILD .......................................................................................7
   Head-Related Transfer Functions (HRTFs) ...............................................................9
      Generalized vs. Individualized HRTFs ...............................................................12
   Previous Related Work .............................................................................................13
      Supporting Work ....................................................................................................13
      Contradictory Work ...............................................................................................14
   Other Proposed Methods for Improved Localization ............................................15

3. METHODS ..................................................................................................................16
   Institutional Review Board .......................................................................................16
   Headphone Listening .................................................................................................17
      Headphone Selection .............................................................................................17
      Headphone Coupler ..............................................................................................19
      Headphone Calibration ..........................................................................................20
   Acquisition of Source Stimulus ..............................................................................20
   Virtual Source Placement .........................................................................................22
   Simulated Reflections ...............................................................................................24
   Experiment Design ...................................................................................................26
      Experiment Location ..............................................................................................26
      Participants ...........................................................................................................27
   Hearing Test .............................................................................................................27
   Experiment 1: Reflection Detection Experiment ...................................................28
      Task .......................................................................................................................28
      Design ...................................................................................................................29
      Procedure .............................................................................................................29
   Experiment 2: Anechoic Localization Experiment ...................................................29
      Task .......................................................................................................................29
      Design ...................................................................................................................30
      Procedure .............................................................................................................30
   Experiment 3: Non-Anechoic Localization Experiment .........................................30
      Task .......................................................................................................................30
      Design ...................................................................................................................30
TABLE OF CONTENTS – CONTINUED

Procedure ...........................................................................................................31
Hypothesis ..........................................................................................................32

4. RESULTS AND ANALYSIS..............................................................................33

Hearing Test Results ..........................................................................................33
Experiment 1: Reflection Detection Experiment Results ..................................34
Experiment 2: Anechoic Localization Experiment Results ...............................35
Experiment 3: Non-Anechoic Localization Experiment Results .........................38
More Data Analysis ............................................................................................43
Round Two Experiments .....................................................................................47
Statistical Discussion ..........................................................................................48
  Logistic Regression ..........................................................................................50
Another Consideration .........................................................................................52

5. ONGOING WORK.............................................................................................54

HRTF Interpolation Implementation ...................................................................54
  Processing Concerns .........................................................................................54
  Implementation Comments ...............................................................................56
Next Steps ............................................................................................................57
  Further Ideas with these Experiments ..............................................................57
  Future Areas of Research .................................................................................58
Conclusion ............................................................................................................59

REFERENCES CITED ..........................................................................................60

APPENDICES .......................................................................................................72

APPENDIX A: Human Participant Consent Form .............................................73
APPENDIX B: Headphone Coupler Drawings and Images .................................83
APPENDIX C: Pre-Test Calibration Data ............................................................86
APPENDIX D: Phrases for Sound Source Stimuli ...............................................88
APPENDIX E: Testing Room SPL Measurements .............................................90
APPENDIX F: Post-Experiment Questionnaire ...............................................93
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Reflection coordinates</td>
</tr>
<tr>
<td>3.2</td>
<td>Reflection factors</td>
</tr>
<tr>
<td>3.3</td>
<td>Hearing test initial SPLs</td>
</tr>
<tr>
<td>4.1</td>
<td>Reflection detection accuracy</td>
</tr>
<tr>
<td>4.2</td>
<td>Confidence interval values and p-values</td>
</tr>
<tr>
<td>4.3</td>
<td>Rear response percentages</td>
</tr>
</tbody>
</table>
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Experiment hypothesis</td>
<td>2</td>
</tr>
<tr>
<td>2.1</td>
<td>Cone of confusion</td>
<td>7</td>
</tr>
<tr>
<td>2.2</td>
<td>HRTF processing flow diagram</td>
<td>9</td>
</tr>
<tr>
<td>2.3</td>
<td>HRIR and HRTF examples</td>
<td>11</td>
</tr>
<tr>
<td>2.4</td>
<td>General HRTF coordinate system</td>
<td>11</td>
</tr>
<tr>
<td>3.1</td>
<td>Headphone frequency response curves</td>
<td>19</td>
</tr>
<tr>
<td>3.2</td>
<td>Frequency spectrum of a representative phrase sample</td>
<td>22</td>
</tr>
<tr>
<td>3.3</td>
<td>HRTF coordinate system used for this thesis</td>
<td>23</td>
</tr>
<tr>
<td>3.4</td>
<td>Eye-level quadrant locations</td>
<td>23</td>
</tr>
<tr>
<td>3.5</td>
<td>Reflection boundaries used for experiments</td>
<td>24</td>
</tr>
<tr>
<td>3.6</td>
<td>Signal processing block diagram</td>
<td>31</td>
</tr>
<tr>
<td>4.1</td>
<td>Hearing test results</td>
<td>34</td>
</tr>
<tr>
<td>4.2</td>
<td>Anechoic localization accuracy</td>
<td>36</td>
</tr>
<tr>
<td>4.3</td>
<td>Phrase analysis – Experiment 2</td>
<td>37</td>
</tr>
<tr>
<td>4.4</td>
<td>Specific experiment hypothesis</td>
<td>38</td>
</tr>
<tr>
<td>4.5</td>
<td>Percent difference from anechoic accuracy</td>
<td>39</td>
</tr>
<tr>
<td>4.6</td>
<td>Confidence intervals</td>
<td>40</td>
</tr>
<tr>
<td>4.7</td>
<td>Overall localization accuracy per reflection type</td>
<td>42</td>
</tr>
<tr>
<td>4.8</td>
<td>Phrase analysis – Experiment 3</td>
<td>43</td>
</tr>
<tr>
<td>4.9</td>
<td>Trial order analysis</td>
<td>44</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>----------</td>
<td>--------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>4.10</td>
<td>Rear hemisphere response bias</td>
<td>45</td>
</tr>
<tr>
<td>4.11</td>
<td>Front/back transition analysis</td>
<td>46</td>
</tr>
<tr>
<td>4.12</td>
<td>Round 1 &amp; 2 comparison</td>
<td>48</td>
</tr>
<tr>
<td>4.13</td>
<td>Relative operating characteristic (ROC) curves</td>
<td>51</td>
</tr>
<tr>
<td>4.14</td>
<td>Comb filtering examples</td>
<td>53</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>CIPIC</td>
<td>Center for Image Processing and Integrated Computing</td>
<td></td>
</tr>
<tr>
<td>HRIR</td>
<td>head-related impulse response</td>
<td></td>
</tr>
<tr>
<td>HRTF</td>
<td>head-related transfer function</td>
<td></td>
</tr>
<tr>
<td>ILD</td>
<td>interaural level difference</td>
<td></td>
</tr>
<tr>
<td>IRB</td>
<td>Institutional Review Board</td>
<td></td>
</tr>
<tr>
<td>ITD</td>
<td>interaural time difference</td>
<td></td>
</tr>
<tr>
<td>KEMAR</td>
<td>common manikin head made by G.R.A.S.</td>
<td></td>
</tr>
<tr>
<td>MATLAB</td>
<td>computing software by MathWorks™</td>
<td></td>
</tr>
<tr>
<td>NIH</td>
<td>National Institutes of Health</td>
<td></td>
</tr>
<tr>
<td>SLM</td>
<td>sound level meter</td>
<td></td>
</tr>
<tr>
<td>SPL</td>
<td>sound pressure level</td>
<td></td>
</tr>
<tr>
<td>TIMIT</td>
<td>corpus of read speech created by Texas Instruments (TI) and Massachusetts Institute of Technology (MIT)</td>
<td></td>
</tr>
</tbody>
</table>
In a natural sonic environment a listener is accustomed to hearing reflections and reverberation. It is conceived that early reflections could reduce front-back confusion in synthetic 3-D audio. This thesis describes experiments which seek to determine whether or not simulated reflections can reduce front-back confusions for audio presented with non-individualized head-related transfer functions (HRTFs) via headphones. To measure the contribution of the reflections, 13 human subjects participated in localization experiments which compared their localization ability with anechoic HRTF processing versus HRTF processing with a single early-reflection. The results were highly subject dependent; some showed improvement while others seemed to be inhibited by the reflections. Statistical analysis of the overall results concluded that a single reflection does not provide a significant difference in localization ability. Although this data rejects the hypothesis of this investigation, some suspicion regarding the contribution of lateral reflections in an auditory environment remains.
INTRODUCTION

People readily localize sound sources in daily life for communication, entertainment and maneuvering in hazardous situations. It is a sensory cue which is subconsciously learned, yet something that is generally completed by humans in a non-trivial fashion. It is desired to develop auditory displays for headphone audio which would provide listeners with the same spatial cues for localization.

For the past few decades the most widely used method for virtual/spatial placement of audio signals is the use of head-related transfer functions (HRTFs). These HRTFs model the spectral modification due to an individual’s pinna, head, and torso. One of the prominent difficulties with HRTFs is the need for individualized HRTF measurements. When using generalized (i.e. non-individualized) HRTFs the ability to properly distinguish sounds as being in either the front or the rear hemisphere becomes more difficult. This limitation has led to the investigation of methods for using generalized HRTF measurements that retain accurate spatial information for a larger population.

The presentation of sounds processed strictly through the HRTFs is an anechoic auditory display. Since natural listening environments generally produce reflections from nearby surfaces, spatial HRTF audio is unnatural and incomplete in presentation. It is possible that environment reflections are needed for the creation of a stable auditory space. Therefore, it is conceived that the addition of reflections could make a more realistic auditory display and provide additional sound cues for source localization. The experimental hypothesis shown in Figure 1.1 compares the localization performance of
virtual sound events for human subjects in HRTF auditory presentations with added reflections. The current investigation presents a method for evaluating the effect of adding a single early-reflection on source localization. Variations in reflection boundary orientations, reflection amplitudes, and reflection timings are all considered.

<table>
<thead>
<tr>
<th>Experiment Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0$: $\mu_{\text{anechoic}} = \mu_{\text{non-anechoic}}$</td>
</tr>
<tr>
<td>$H_A$: $\mu_{\text{anechoic}} &lt; \mu_{\text{non-anechoic}}$</td>
</tr>
</tbody>
</table>

Figure 1.1: Null ($H_0$) and alternative hypotheses ($H_A$) for the primary investigation of these experiments. The null hypothesis states that the average localization accuracy will be the same for anechoic and non-anechoic conditions. The alternative hypothesis considers that the average localization will improve with the addition of reflections.

Motivation

The ability to produce a stable auditory image in both the front and rear hemispheres could lead to a new form of information theory. This advancement would provide many opportunities for a variety of commercial applications. The use of spatial auditory information has already been investigated for warning systems in aviation where the auditory cues could allow pilots and ground control to make quicker decisions on the location of potential dangers or conflicts. Similarly, auditory displays could provide improved underwater sonar displays. Mobile land-based positioning systems, such as GPS, could benefit by providing more intuitive directional guidance while reducing map interpretation errors in search-and-rescue or firefighting applications [2].

Auditory spatial information would enhance the realism in music, video game and movie applications, and would also find its way into the corporate office [3, 4]. As
companies become more reliant upon video conferencing for business operation, improvements in communication methods are certain to follow. Currently, multiple voices at a meeting site are conveyed to the video conferencing participant(s) through one or two audio channels causing a loss in discrimination between sound sources. However, by separating each of the voices at the meeting site and placing them in a spatial configuration agreeing with the video, the virtual participant(s) of the meeting could extract more information by focusing on specific voices or sound sources.

**Thesis Organization**

Investigations in this thesis will use human participants to determine if the problems with front-back confusions can be reduced while using generalized HRTFs. It is conceived that the addition of virtual reflections in headphone audio would create a better sense of realism and provide the listener with additional cues to localize the sound. Since the primary obstacle with localization is the ability to distinguish between front and rear hemispheres, the experiments presented in this thesis evaluate the influence of early-reflections on human subject’s front-back localization ability.

This thesis is organized as follows. Section two introduces foundational theories that will be applied to the experiments of this investigation. Other research that presented supporting or conflicting theories is also mentioned. Section three describes the experiment design and other completed work directed at the investigation of early-reflection’s effect on front-back localization. Section four presents a statistical analysis and discussion of the results. Experimental design considerations for future investigations
are also discussed in this section. The final section briefly describes other processing applications for HRTF data along with future avenues of spatial audio investigations.
BACKGROUND

Investigations in binaural hearing have extended from the nineteenth century in which scientists such as Giovanni Venturi, Ernst Chladni, and Gustav Fechner studied the interaction of sounds at both ears [5]. These original theories of varying sound arrival timings and levels were confirmed a century later by Lord Rayleigh [6]. Rayleigh conceived of two binaural cues: interaural time difference (ITD) and interaural intensity difference, more commonly referred to as interaural level difference (ILD). This binaural model proved to have limitations; primarily along equidistant curves around the head causing a localization ambiguity of front and back hemispheres [7]. This shortcoming has led most modern research to the use of head-related transfer functions (HRTFs) for sound spatialization techniques [8, 9, 10, 11, 12]. These HRTFs spectrally shape a free-field sound as if it were subjected to reflections and diffraction from one’s head, torso, and pinna.

It has been shown that individualized HRTFs allow for more accurate localization compared to generic HRTFs [13, 14]. However, the difficulty in rapidly obtaining individualized HRTFs forces many to use generalized HRTFs. The predominant degradation in localization performance with the use of these non-individualized transfer is a reduced ability in distinguishing between the front and rear hemispheres [13, 15].

Since listeners are accustomed to localizing sounds in reverberant environments, the strict presentation of a monaural signal processed solely by a set of anechoic HRTFs is an unnatural auditory display. It is possible that the addition of environment reflections could create a more realistic placement of sound events [12, 16, 17, 18, 19].
Lateralization vs. Localization

Prior to presenting the history and background of binaural audio it should be noted that there is a distinct difference between binaural and stereophonic audio presentations [20]. Binaural audio uses two separate channels, one for each ear, which contain time and spectral differences due to anthropometric shape and spacing. Ideally, these channel differences duplicate how a listener would naturally hear a sound-source and create the perception of position. Stereophonic audio is a two (or more) channel presentation of sound events which uses acoustic crosstalk among the sound channels to create the perception of position.

When describing the spatial positioning of sound sources there are two specific terminologies: lateralization and localization [21,22]. Lateralization refers to the left-right presentation of the sound and is sometimes referred to as inside-head localization. Stereophonic audio is intrinsically limited to a lateralized auditory display. On the other hand, localization is concerned with the perceived azimuth and distance of a sound appearing from an outside-head sound-source. Although the primary goal of spatial audio is localization, lateralization is what generally occurs since the source is not perceived to be externalized outside of the head. The experiments in this paper are not concerned with listener’s perception of externalization, thus the term localization will be loosely used for listeners’ perceived direction of sound.
Duplex Theory: ITD and ILD

Most modern binaural audio research references Lord Rayleigh’s Duplex Theory [6] as the origin of sound localization investigations. This theory defined the difference in timing (ITD) and level (ILD) of the signals at each ear, i.e. the ipsilateral ear (closest to sound source) will receive the sound event first and will have a larger amplitude signal than the contralateral ear (further from sound source). Studies subsequent to this theory revealed flaws in the strict use of ITD and ILD for sound localization. It was found that equidistant curves around the head resulted in the same ITD and ILD values as shown in Figure 2.1. These curves, termed “cones of confusion”, were shown to produce ambiguity between the front and rear hemispheres [7].

![Figure 2.1: Cone of confusion. Any point on a equidistant curve will result in the same ITD values which causes both front-back and up-down confusions.](image)

Other early localization research using pure tones revealed the frequency dependency of the ITD and ILD. Free-field experiments with tone bursts found that localization error rates were largest around 3kHz and reduced at very low and very high
frequencies [23]. The interpretation of this data indicated that there were two mechanisms for sound localization: one for low frequencies and one for high frequencies. These results were later confirmed by [24] where localization errors were mostly found in the 1.5 kHz to 5 kHz range with most uncertainty at 1.5 kHz.

The ambiguity at this frequency boundary agrees with the relationship of certain wavelengths and the geometry of an average human head [25]. Since the wavelengths of frequencies above the threshold are less than the distance between the two ears, multiple phase relationships could exist. This shows that ITD is a poor localization cue for high frequencies, however, high frequencies will be attenuated from the shadow of the head, thus, ILD is a good cue. Conversely, the amplitude of low frequencies will be relatively unchanged between the ears, yet the phase relationship would yield pertinent information for localization. When conflicting ITD and ILD cues are presented, the time difference dominates; however, when the low frequencies are removed, the ILD dominates the perceived direction [26].

In true monaural sound localization the ITD and ILD are non-existent, thus the Duplex Theory would not provide a means for localizing sound events. However, humans are able to localize events with only one ear using subtle diffraction and reflection cues from the folds of the pinna [27]. The need for spectral cues in addition to the ITD and ILD was presented as early as 1955 [28]. Other prominent early work and theories by Blauert are collected in a highly referenced book [29].
Head-Related Transfer Functions (HRTFs)

Limitations with the Duplex Theory guided researchers to find new methods for modeling accurate spatial audio. It was conceived that if sounds were recorded with the natural acoustical filtering due to the head and pinna, then the presentation of the sound would appear at the free-field position of the recorded sound [8, 9, 10, 22, 30]. Furthermore, if a known input was applied to an actual head/body system with a measurable response at the eardrum, an acoustic transfer function could be obtained for each ear which describes the filtering due to the interaction of sound waves with the body. A monaural sound-source could be filtered by these transfer functions and sent to the respective ear to create a perceived position for the original sound-source as shown in Figure 2.2.

![Figure 2.2](image.png)

Figure 2.2: Overview of HRTF processing for headphone audio. The HRTF filtering models spectral changes due to the reflection and diffraction of acoustic sound waves.

Some of the first binaural recordings were made in the 1970’s with microphones placed inside a manikin head [22] and in the ear-canal of a human listener [30]. These recordings allowed for a perceptual evaluation of virtually placed sources. Localization
and externalization were observed, but no comparison of binaural to free-filed localization ability was made. Subsequent work showed that binaural waveforms could duplicate free-field waveforms within a few dB [9], and localization ability using the personalized HRTFs was comparable to the localization of free-field sound events [10]. These documented improvements towards true localization laid the framework for modern HRTF processing.

HRTFs inherently encode the ITD and ILD parameters and account for the acoustic filtering that a free-field sound undergoes due to the head, pinna, and torso [31]. This can be seen in HRTF impulse responses plots, i.e. head-related impulse responses (HRIRs), for a source positioned in the left hemisphere as shown in Figure 2.3(a). Simple observation of the HRTFs in Figure 2.3(b) shows the overall reduction in gain for the contralateral ear. Other spectral HRTF characteristics, such as the minimal difference at low frequencies and the two prominent pinna notches, are common across HRTFs [32].

In principle, since a single HRTF measurement accounts for only one point in space, an infinite number of HRTF measurements would be necessary. To reduce measurement time, only a discrete set of points are measured around the head. The general convention for describing the horizontal and vertical location of these finite measurement points uses an azimuth (θ) and elevation (ϕ) coordinate system as shown in Figure 2.4. The measurements are usually taken at a fixed, far-field (r > 1m) radius although the spectral differences between near and far-field should be considered [33, 34, 35]. Methods have also been developed to modify far-field HRTFs for near-field presentations [36].
Figure 2.3: Left and right (a) HRIR and (b) HRTF for a sound source positioned in the left hemisphere.

Figure 2.4: HRTF coordinate systems commonly use two parameters $\theta$ and $\phi$ to represent azimuth and elevation, respectively. Figure taken from [37].

Using the discrete measurements it is possible to virtually place a sound at a particular azimuth and elevation, however, for many applications it desirable to create a moving source. The interpolation between HRTF measurements remains open to investigation as no single method has proven dominant [38]. A few methods of interpolation are described and attempted in the *Ongoing Work* section [33, 39, 40].
There are a few difficulties with acquiring accurate and repeatable HRTF measurements. For most researchers the ability to access an anechoic facility suitable for HRTF measurements is difficult. Some techniques have used time-windowed measurements to circumvent the need for an anechoic chamber [41]. Also, improper placement of the probe microphones and head movement [9, 42] can reduce fidelity of pinna effects when taking these measurements. The head movement can be remediated through the use of a bite bar or manikin head for measurements [9, 43]. Lack of protocol for something such as a standard measurement signal level [44] can also lead to discrepancies across various HRTF datasets.

Generalized vs. Individualized HRTFs

Due to difficulty in rapidly obtaining personalized sets of HRTFs, it is commonplace for researchers to use established HRTF libraries [43, 45]. The use of non-personalized measurements is a considerable field of interest since acquiring individualized measurement would be impractical for broader applications of spatial audio. However, generalized HRTFs have shown degradations primarily in elevation and front-back reversals [10, 12, 13, 15].

When the presentation of a sound in the front hemisphere is perceived as being in the rear, a front-back reversal has occurred. Conversely, if a sound is perceived in the front when its intended location was in the rear hemisphere, a back-front confusion has occurred. Throughout the remainder of this thesis the term front-back confusion will be used to reference both types of reversals unless a specific reversal direction is being addressed. Reduction of reversal rates is of great importance for realistic spatial audio.
Some of the research directed at front-back reversals will be discussed later in this section.

Attempts have been made to synthesize generalized HRTFs through principal component analysis [46, 47], acoustic raytracing techniques [48], spherical and ellipsoidal head models [49], head-torso models [50, 51], and pinna models [52, 53]. Although there is no certainty that combining spectral information from the individual components will result in a valid HRTF, it is useful for researchers to understand the isolated effects to help enhance generalized HRTFs to match a broader community of listeners.

**Previous Related Work**

Although this work seeks to better define the effects of reflections on HRTF localization, the concept of adding virtual environments and reflections to spatial HRTF audio is not novel. Due to the large number of factors in HRTF audio presentations, it is not surprising to find both supporting and contradictory work.

**Supporting Work**

Humans generally encounter very few real-world situations where sound localization in an anechoic environment is necessary. Therefore it seems reasonable to consider that the localization of a monaural signal processed strictly by a set of anechoic HRTFs would, also, be an unnatural auditory display. It is well-known that reflections provide a sense of distance and externalization to the listener in virtual environments [22, 54, 55] and create a heightened sense of realism [18, 19, 56]. These results would seem
to promote the theory that reflections and room-models could present a more natural display than with anechoic HRTFs alone and aid listeners in the localization of virtually placed sound events [16, 17, 19, 22, 57].

Begault compared a variety of elements for sound localization performance and applied early-reflections as a parameter of the investigation [11]. One experiment compared anechoic localization ability to instances when early reflections, reverberation or both were added. Some improvement in azimuth accuracy was reported, but no explicit claim was made regarding the effect of discrete early reflections on front-back reversals. In another investigation Zotkin, Duraiswami & Davis detail a virtual auditory system where environment modeling for audio scene rendering played a major component [12]. Though the system was well described with a variety of aspects pertaining to HRTF audio, minimal conclusions were drawn concerning the subjective changes in front-back reversal rates.

Contradictory Work

Research results that dispute the role of reflection information in improving localization ability has mostly considered the contribution of reverberant energy, not early reflections. Since late reverberation creates a lower correlation between the signals at the two ears compared to early-reflections [58], it is logical that late reverberation degrades speech intelligibility [59, 60] and localization accuracy [61, 62 63, 64] in free-field listening conditions. This degradation due to late-reverberation has been shown in some studies using binaurally processed audio to inhibit localization performance [65]
and in other studies to have no effect upon localization [66]. However, the binaural advantage is not completely destroyed for all source/listener positions within a room [67].

Other Proposed Methods for Improved Localization

The problem with front-back reversals while using generalized HRTFs has attracted a large amount of research. The most documented method for reduction of front-back confusions is the listener’s ability to rotate their head while keeping the virtual sound source stationary [68, 69, 70]. Although this method has been shown to significantly reduce the reversal rates, it requires the use of additional head-tracking hardware which can be impractical for some applications. Similarly, sound-source movement has shown to reduce reversal rates, but the listener must have control over the source movement to achieve the improvement [68].

Since HRTF processing uses spectral shaping of sound sources to produce perception of direction, it is reasonable to attempt other manipulations of the spectrum. Some investigations have tried enhancing certain spectral peaks and notches [71, 72, 73, 74], while another has used spectral profiles of large protruding pinna for improvement [75]. Other research has attempted to “best fit” listeners from a set of HRTFs through tournament-style listening tests [76] or physical matching with anthropometric measurements [77]. Although the success of spectral manipulation is questionable, listening training sessions have shown to reduce front-back reversal rates [78].
METHODS

The presentation of sound events with HRTFs is an unnatural, anechoic auditory display and the use of generalized HRTFs causes an increase in front-back reversals. Experiments in this investigation will present human subjects with HRTF processed audio which contains a virtual early-reflection. Participants indicate if they perceive the sound as originating from either the front or the back hemisphere. Since this is a subjective test, the number of variables is limited to only those that can help determine the influence of a single reflection on front-back localization ability. This section details the preliminary procedures and experiment design necessary for determining the effects of early reflections on localization ability.

Institutional Review Board

Before any human subject experiments can be conducted it is necessary to obtain approval from the Institutional Review Board (IRB). The IRB is in place to ensure that no physical, psychological, social, legal and economic harm is caused due to participation in the experiments [79]. Since experiments of this nature had not previously been conducted by the investigators of this research, completion of a course designed to educate experiment supervisors on the protection of human research participants was necessary. The course was taken online through the [80] and provided a good review of participant rights along with less apparent pitfalls that researchers could easily overlook. The completed IRB application which includes the NIH certificate of completion and IRB approved Consent form for Participation is provided in Appendix A.
Headphone Listening

Since the acoustic crosstalk between the signals at the ears makes it extremely difficult to obtain a binaural image using loudspeaker arrangements, it is common to use headphones for binaural presentations. The headphone used in this investigation was carefully selected and calibrated to meet the needs of the experiments.

Headphone Selection

The choice of headphone was based on factors such as ease of placement on a subject’s head, good noise isolation, and a flat response curve. It was desired to have headphones which could be positioned repeatability by the listener so that spectral effects due to placement variability would be negligible. Although the experiment location was a quiet environment, it was beneficial to have the headphones attenuate exterior noise. The ideal headphone had minimal mechanical noise so that movements of the head would not produce sounds to interfere with localization. Finally, it was important that the headphone introduced minimal changes to the sound spectrum since no active headphone equalization was used to achieve a flat frequency response. Although the use of electronic headphone equalization has shown to enhance externalization of virtual sound images [81, 82], it has not shown to be a significant factor with localization accuracy [83].
There are five primary types of headphones:

1) Circum-aural headphones – completely cover the ear
2) Supra-aural headphones – rests on top of the ear
3) Intra-concha headphones – rests inside of the ear, but not inside the ear-canal
4) Ear-canal headphones – rest inside the ear-canal, and seal the canal opening

Under the employment of HeadRoom Corporation approximately 50 headphones of all four types were measured using an Audio Precision Cascade Sys-2522 system. Both the supra-aural and intra-concha types were not considered due to variability in placement. The ear-canal type achieved the best noise isolation characteristics and a flatter frequency response, however these properties were dependent on complete insertion of the transducer into the ear-canal which can cause discomfort to some listeners. Ear-canal type headphones also conduct more noise from cord movement since the headphone unit is more mechanically coupled to sensitive parts of the pinna.

There were two particular models of circum-aural headphones that produced reasonably flat responses: AKG K701 and Denon AH-D2000. The frequency responses of these headphones, along with an in-ear headphone and a headphone example with a poor frequency response, are shown in Figure 3.1. The more erratic response at high frequencies was observed across all types of headphones measured at HeadRoom; and was also found in an acoustic analysis of circum-aural headphones [84]. The Denon AH-D2000 response yielded the smoothest frequency response of the circumaural headphones, thus, it was chosen for the experiments conducted as part of this thesis.
Figure 3.1: Selected frequency response plots for headphones tested. The response for the headphones used in the experiments (Denon AH-D2000) is outlined in the top-left.

Headphone Coupler

To limit variability in the experiment responses, a regulated SPL at the headphones was desired. Since no device was readily available for this type of calibration, it was necessary to create a coupler for the sound level meter and headphones. A more sophisticated coupler following ANSI S3.7 standards [85] was unnecessary since only a known SPL at the headphone was desired. A flat-plate coupler similar to one described by [86, 87] was machined by the investigator of this research under the guidance of Dr. Kevin Repasky. Drawings and notes used for the machining of the coupler are provided in Appendix B.
Headphone Calibration

Using the coupler described in the previous section, the headphone was calibrated with a Brüel and Kjaer Type 2236 sound level meter (SLM). The SLM was calibrated before each use by the means of a Brüel and Kjaer Type 4231 acoustic calibrator. Before any subjective experiments were conducted the left channel of the headphone was calibrated to 75 dB at 1 kHz by adjusting the output of the HeadRoom Ultra Micro Amp. With the amplifier output unchanged, the right channel was measured and a negligible difference in SPL (< 1 dB) was found. This provided confidence that the headphone was balanced between left and right channels.

Prior to each subject’s participation in the experiments calibrations were conducted on the SLM and the headphone. The SLM was calibrated to 94 dB using the 4231 calibrator. The left channel of the headphones was calibrated to 75 dB at 1 kHz. A SPL level of 75 dB was chosen so that the acoustic reflex would not become a factor for the duration of the experiments [88]. Throughout the experiments no adjustment to the B&K meter or headphone amplifier was required. Calibration values for each participant are reported in Appendix C.

Acquisition of Source Stimulus

For the experiments in this investigation it was necessary to obtain source-audio samples which can be presented to participants for the localization evaluation. It is common for broadband noise to be used as the source stimulus so that the entire spectrum of the HRTF is excited, however, the usefulness of localizing noise stimuli has been
questioned since most applications for spatial audio would likely use realistic sounds, such as speech [11]. Listener’s familiarity with source stimuli has also been shown to affect localization ability [27, 29, 78, 89]. Since humans are more accustomed to localizing speech versus noise bursts, short phrases of human speech were chosen as the stimuli for these experiments.

It was desired to take the speech samples from the standard TIMIT database, which is a collection of phonetically variant sentences used for speech identification research developed jointly by Texas Instruments (TI) and the Massachusetts Institute of Technology (MIT) [90]. However, the phrases of this database were recorded at a sampling rate of 16 kHz. This limited bandwidth reduces the amount of HRTF information available for localization cues, so it was necessary to make special recordings for the localization sound samples.

Forty-one phrases with above average fricative content were chosen from the text transcript of the TIMIT sentences. A male speaker with a fundamental frequency below 100 Hz read the phrases. The short (1-2 sec) speech phrases were recorded at 48 kHz, 24 bit in a Model 802 Ray Proof Sound Shield rehearsal room using an AKG C414 microphone and Tascam HD-P2 digital-audio recorder. All forty-one phrases were amplitude normalized, and spectrally analyzed in 1/3 octave bands to find phrases containing significant energy in the upper frequency bands. The spectrum of a representative phrase is shown in Figure 3.2. All phrases read are listed in Appendix D with the ten phrases used in the experiments provided below.
Figure 3.2: Spectrum analysis of a representative phrase used in these experiments, ‘Outstanding Effects’. The frequency content energies are plotted on a 100dB scale.

Virtual Source Placement

To achieve perceptual positioning, the audio phrases are filtered by HRTFs corresponding to the virtual source placement at a particular location in space. The monaural audio sample is processed with both a left and right HRTF for each respective
ear to achieve a binaural presentation. Placement of the virtual sound sources was performed using the HRTF data from the KEMAR large pinnae dataset (Subj. 165) from the CIPIC database [45]. The coordinate system used for azimuth ($\theta$) and elevation ($\phi$) is shown in Figure 3.3. Speech samples were virtually placed in four, eye-level (zero elevation) quadrant locations around the head with locations 45° off the coronal and midsagittal planes as shown in Figure 3.4.

![Diagram of CIPIC coordinate system](image1)

Figure 3.3: CIPIC coordinate system which was used for the experiments of this thesis. Figure taken from [91].

![Diagram of eye-level quadrant locations](image2)

Figure 3.4: Four, eye-level (zero elevation) quadrant locations for the direct stimuli.
Simulated Reflections

To limit the number of experimental variables only a single reflection was added to the monaural source stimulus. Using the image-source method [92, 93], reflections were created for two boundary cases, floor and right-wall, as shown in Figure 3.5. Two variations of the right-wall reflection were used to simulate short (6ms) & long (20ms) delays. Only the 6ms delay was used for the floor since a 20ms delay would require an impractical observer-boundary relationship: 20ms would correspond to 6.9 meters of path length difference with a room temperature sound speed of 343 m/s. These boundary reflection delays are on an order of magnitude difference than the 0.34ms ITD of the HRTF dataset used in these experiments.

![Figure 3.5: Two types of reflections used in the experiments (a) floor and (b) right-wall.](image)

For these experiments it was desired to have no spectral modification from the raw HRTF measurements, so source and boundary distances were carefully chosen so
that no interpolation was necessary. Azimuth and elevation coordinates for each quadrant of each boundary case are provided in Table 3.1.

Table 3.1: Azimuth and elevation coordinates of reflections in the respective quadrant locations for three reflection types.

<table>
<thead>
<tr>
<th>Quadrant</th>
<th>Direct</th>
<th>Floor</th>
<th>Lateral Wall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Front-Left</td>
<td>(-45°,0°)</td>
<td>(-45°,-45°)</td>
<td>(65°,0°)</td>
</tr>
<tr>
<td>Front-Right</td>
<td>(+45°,0°)</td>
<td>(+45°,-45°)</td>
<td>(80°,0°)</td>
</tr>
<tr>
<td>Back-Left</td>
<td>(-45°,180°)</td>
<td>(-45°,-225°)</td>
<td>(65°,180°)</td>
</tr>
<tr>
<td>Back-Right</td>
<td>(+45°,180°)</td>
<td>(+45°,-225°)</td>
<td>(80°,180°)</td>
</tr>
</tbody>
</table>

A broad range of reflection levels were considered. These simulated reflections were simply variations in gain with no spectral change. Since the goal of this investigation was to determine if any type of first-order reflection would contribute to front-back localization ability, the range of relative levels include some simulated reflections not physically possible. Table 3.2 shows a list of reflection factors (ρ) for each of the boundaries investigated, where ρ = 1 corresponds to perfect reflection and ρ = 0 corresponds to perfect absorption. A ρ > 1 would indicate a reflection stronger than the incident sound, thus not reasonable in real-world environments.

It is well documented that certain reflections are inaudible or masked by the preceding waveform depending on the time and amplitude difference between the direct and reflected signal [29, 94, 95]. If the delay between the direct and reflected signals is too small and the amplitude of the reflection is below a certain threshold, the discrete reflection is inaudible and the two signals are perceived as one event. This psychoacoustic property is commonly termed the ‘precedence effect’. Spatial auralization
methods for video games use this principle to reduce the number of calculated reflections in complex environments with multiple sound sources [96]. Begault showed that the precedence effect is direction dependent and provided a few “rules of thumb” for determining perceptually significant reflections depending on the reflection amplitude and delay [97]. Since some of the reflections in Table 3.2 could be inaudible and insignificant to the results of the study, it would be desirable to reduce the number of trials by eliminating them from the localization experiments.

Table 3.2: Reflection factors (ρ) for various levels of reflections in right & left hemispheres.

<table>
<thead>
<tr>
<th>Floor Boundary</th>
<th>6ms delay</th>
<th>Right-wall Boundary</th>
<th>6ms delay</th>
<th>20ms delay</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Left</td>
<td>Right</td>
<td>Left</td>
</tr>
<tr>
<td>-6dB</td>
<td>0.76</td>
<td>0.85</td>
<td>1.53</td>
<td>2.22</td>
</tr>
<tr>
<td>-9dB</td>
<td>0.54</td>
<td>0.60</td>
<td>1.09</td>
<td>1.57</td>
</tr>
<tr>
<td>-12dB</td>
<td>0.38</td>
<td>0.42</td>
<td>0.77</td>
<td>1.11</td>
</tr>
<tr>
<td>-15dB</td>
<td>0.27</td>
<td>0.30</td>
<td>0.54</td>
<td>0.79</td>
</tr>
<tr>
<td>-18dB</td>
<td>0.19</td>
<td>0.21</td>
<td>0.38</td>
<td>0.56</td>
</tr>
<tr>
<td>-21dB</td>
<td>0.13</td>
<td>0.15</td>
<td>0.27</td>
<td>0.39</td>
</tr>
<tr>
<td>-24dB</td>
<td>0.10</td>
<td>0.11</td>
<td>0.19</td>
<td>0.28</td>
</tr>
<tr>
<td>-27dB</td>
<td>0.07</td>
<td>0.08</td>
<td>0.14</td>
<td>0.20</td>
</tr>
</tbody>
</table>

**Experiment Design**

**Experiment Location**

All experiments were conducted in a quiet, acoustically treated (non-anechoic) room on the MSU campus. The building location was greater than 40 meters to the nearest area of high-volume and large-vehicle traffic. Subjects were isolated in a room separate from the experiment controller. Experiment times were scheduled so that the
building was vacant of other tenants and did not conflict with the campus lawn mowing and maintenance schedule. Sound pressure level measurements of the experiment room were taken on two occasions when exterior noise levels were perceived to be high. The omni-directional microphone of the sound level meter was placed at the approximate location of the listeners’ head. Sound pressure levels in decibels were recorded with A-weighting (dBA) for two and three hour time windows. A summary of the results can be found in Appendix E. The SPL was less than 30 dBA for more than 95% of the time during each testing occasion.

Participants

Thirteen untrained subjects (nine male, four female) between the ages of 20 and 28 participated in the localization experiments. Subjects were not paid but were provided basic refreshments for their participation. Since the duration of the subjective experiments can adversely affect responses [98], the overall duration of the experimental procedure was carefully considered. Five participants were asked to complete all three experiments while the remaining eight participated in only the anechoic and non-anechoic localization experiments. At the end of the testing session subjects were asked to complete a post-experiment survey to gather relevant feedback about the experiment. A blank survey can be found in Appendix F.

Hearing Test

Although symmetrical hearing loss has not been shown to degrade localization performance [29], it was desired to obtain the subjects’ hearing thresholds to eliminate
any unexpected hearing impairments. Subjects were screened on six octave-band center frequencies at 250, 500, 1k, 2k, 4k, and 8kHz using an implementation of the classic adaptive maximum-likelihood procedure [99, 100]. Since no headphone equalization was applied, the actual SPL of the tones varied due to the frequency response of the headphones. Table 3.3 shows the initial SPL for each of the frequencies tested.

Table 3.3: Initial SPLs for each of the frequencies tested in the hearing screening.

<table>
<thead>
<tr>
<th>Freq. (Hz)</th>
<th>SPL (dBA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>250</td>
<td>53.8</td>
</tr>
<tr>
<td>500</td>
<td>57.7</td>
</tr>
<tr>
<td>1k</td>
<td>63.2</td>
</tr>
<tr>
<td>2k</td>
<td>58.1</td>
</tr>
<tr>
<td>4k</td>
<td>60.9</td>
</tr>
<tr>
<td>8k</td>
<td>52.8</td>
</tr>
</tbody>
</table>

Experiment 1: Reflection Detection Experiment

The first experiment was used to determine thresholds of audibility for reflections since the testing of non-audible reflections would provide no insight on their contribution to front-back localization.

Task

Participants listened to pairs of audio samples and made a same or different forced-choice response.
Design

Five subjects (all from Group A) were presented with two instances of the same virtually placed speech sample. One of the samples was an anechoic presentation and the other potentially contained a reflection. One trial per quadrant for each permutation of boundary case (floor 6ms, wall 6ms & wall 20ms) and reflection amplitude relative to the direct signal (-6dB, -9dB, -12dB, -15dB, -18dB, -21dB, -24dB & -27dB) was tested. This resulted in 100 trials per experiment and took participants approximately 15 minutes to complete.

Procedure

Participants were limited to a two interval forced-choice (same/different) response and indicated their choice via keyboard through a MATLAB interface. Prior to beginning this experiment, participants listened to instructions read by the same voice as the test phrases.

Experiment 2: Anechoic Localization Experiment

The second experiment established subjects’ baseline localization ability in anechoic conditions.

Task

Participants listened to a single audio sample and made a front or back forced-choice response.
Design

All thirteen subjects were presented with speech samples virtually placed at one of four, eye-level quadrant locations. Each subject was presented with fifteen trials in each of the quadrant locations. This resulted in 60 trials per experiment and took participants less than 10 minutes to complete.

Procedure

Participants were limited to a binary (front/back) response and indicated their choice via keyboard through a MATLAB interface. Prior to beginning this experiment, participants listened to instructions and a single virtually placed example per quadrant. Both the instructions and example were read by the same voice as the test samples.

Experiment 3: Non-Anechoic Localization Experiment

The final experiment investigated listeners’ localization ability when a single, audible early-reflection was added to the direct anechoic phrase.

Task

Participants listened to a single audio sample and made a front or back forced-choice response.

Design

All thirteen subjects were presented a sequence of audio samples containing a direct signal virtually placed at one of four, eye-level quadrant locations. Based on the results of Experiment 1, a reflection type considered perceptually detectable was added to
the direct signal. The block diagram in Figure 3.6 describes the signal flow and processing for the simulated reflections.

Figure 3.6: Block diagram of speech sample processing with an added reflection.

Two trials for each permutation of quadrant and boundary condition were tested for the -6dB case; three trials were used for all other relative amplitude cases (-9dB to -18dB). To supplement the anechoic baseline experiment, three trials without reflections were also presented per quadrant. This resulted in a total of 144 trials per experiment and took participants approximately 15 minutes to complete.

Procedure

Participants were limited to a binary (front/back) response and indicated their choice via keyboard through a MATLAB interface. Prior to beginning this experiment, participants listened to instructions and a single virtually placed example per quadrant. Both the instructions and examples were read by the same voice as the test samples.
Hypothesis

Consideration of existing literature developed a few theories regarding the effect of reflections. Since floor and lateral reflection geometries are symmetric between the front and rear hemispheres, there is nothing particularly unique about the reflection to provide localization cues. This did question the usefulness of the added reflection, but it was reasoned that providing two HRTF processed signals from the same front/back hemisphere would help to discriminate between the front and rear hemispheres. This directional support from the reflection is easily conceived for any case of floor reflection and any occasion when the sound source is on the same side as the boundary. However, when the source is placed on the side opposite the boundary, it seems possible that a confounding auditory display could be presented. This investigator suspects that sources placed in the left hemisphere with a reflection due to a right-wall will be more poorly localized than a source placed in the right hemisphere with a reflection from the same boundary. In addition, since there is a greater sensitivity to lateral reflections, it is believed that the floor reflection will likely provide little assistance whereas the reflection from the right-wall could show some contribution to localization accuracy.
RESULTS AND ANALYSIS

The experiments described in the previous section used responses from human participants to determine if the addition of a single early-reflection had an effect on localization ability when using generalized HRTFs. This section provides the results and analysis of the experiments conducted.

Hearing Test Results

All subjects participated in an informal hearing screening as described in the previous section. Of the thirteen subjects, two participants reported clinically-classified hearing problems due to work environment conditions. These problems were considered symmetric between both ears. Two other participants mentioned a suspected decline in hearing though not medically determined. It can be seen in Figure 4.1 that the results of the hearing screening show a trend in higher detection thresholds for these four individuals (Subj. 10-13). The subjects' overall threshold trends followed the expected pattern: best acuity in the 3-4 kHz range, with less sensitivity at lower and higher frequencies. The thresholds in Figure 4.1 are with respect to the initial SPLs of Table 3.3.

As will be discussed later in this section, the four participants reporting possible hearing limitations showed comparable localization accuracy to the nine “normal hearing” participants. If it is assumed that the change in thresholds was symmetric for both ears, previous findings support that localization should be unaffected based on hearing thresholds for these four individuals [29]. This assumption allows the following
data analysis to be conducted across all thirteen subjects, thus maintaining a larger sample size.

![Hearing Test Results](image)

Figure 4.1: Hearing screening results for the thirteen participants. Thicker hearing curves indicate the four subjects (10-13) with suspected hearing problems. Threshold values are reported with respect to the initial SPL levels of Table 3.3

**Experiment 1: Reflection Detection Experiment Results**

The first experiment aided in determining which reflections were useful for the localization experiments. To reduce the number of trials used for the investigation, only reflections that showed reasonable detectability were used. The percentage of detectability for all five subjects who participated in Experiment 1 was calculated for each boundary and relative amplitude condition. The overall results are presented in Table 4.1 where the shaded reflections were those chosen for the localization experiments. The choice of reflection types was reasonable given the results; however, it was conceived to eliminate the -9dB floor (6ms) and -18dB wall (6ms) reflections. These
two cases were retained to provide additional trials for floor reflections and to maintain symmetry between the wall reflection trials. Table 4.2 divides the subjective detectability for the two wall-reflection types into left and right hemispheres. These left-right hemisphere results agree with [97] that larger lateral differences between the direct and reflected signals result in a lower threshold of detection.

Table 4.1: (a) Reflection detection accuracy for the three boundary cases at varying relative amplitudes. Shaded reflection categories indicate cases used in localization experiments. (b) Left and right hemisphere reflection detection accuracies for both (6ms & 20ms) lateral reflections due to a right-wall boundary.

<table>
<thead>
<tr>
<th>Relative dB</th>
<th>Floor 6ms</th>
<th>Wall 6ms</th>
<th>Wall 20ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>-6dB</td>
<td>55%</td>
<td>95%</td>
<td>100%</td>
</tr>
<tr>
<td>-9dB</td>
<td>25%</td>
<td>95%</td>
<td>95%</td>
</tr>
<tr>
<td>-12dB</td>
<td>0%</td>
<td>50%</td>
<td>100%</td>
</tr>
<tr>
<td>-15dB</td>
<td>0%</td>
<td>45%</td>
<td>50%</td>
</tr>
<tr>
<td>-18dB</td>
<td>15%</td>
<td>35%</td>
<td>50%</td>
</tr>
<tr>
<td>-21dB</td>
<td>0%</td>
<td>20%</td>
<td>15%</td>
</tr>
<tr>
<td>-24dB</td>
<td>10%</td>
<td>10%</td>
<td>5%</td>
</tr>
<tr>
<td>-27dB</td>
<td>10%</td>
<td>15%</td>
<td>10%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Relative dB</th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>-6dB</td>
<td>100%</td>
<td>95%</td>
</tr>
<tr>
<td>-9dB</td>
<td>100%</td>
<td>90%</td>
</tr>
<tr>
<td>-12dB</td>
<td>90%</td>
<td>60%</td>
</tr>
<tr>
<td>-15dB</td>
<td>75%</td>
<td>20%</td>
</tr>
<tr>
<td>-18dB</td>
<td>70%</td>
<td>15%</td>
</tr>
<tr>
<td>-21dB</td>
<td>25%</td>
<td>10%</td>
</tr>
<tr>
<td>-24dB</td>
<td>15%</td>
<td>0%</td>
</tr>
<tr>
<td>-27dB</td>
<td>15%</td>
<td>10%</td>
</tr>
</tbody>
</table>

**Experiment 2: Anechoic Localization Experiment Results**

The second experiment provided a baseline localization accuracy in anechoic conditions for which to compare localization accuracy with added reflections. Figure 4.2 plots the results for each participant in the four quadrant locations. The average localization accuracy for each participant is indicated in the respective quadrant locations. Although no direct reasoning was determined, it is interesting to note that the localization accuracy was about 10% better in the front-left and rear-right quadrants versus the front-right and rear-left quadrants.
Of the two participants with clinically classified hearing problems, one reported the lowest localization accuracy in the front-left quadrant whereas the other participant recorded perfect localization accuracy in that same quadrant. These two subjects showed average localization accuracy in the other quadrants. One of the other two subjects with suspected hearing loss showed below average localization ability in all quadrants, however did not record the poorest localization ability in any of the quadrants. The other participant with suspected hearing impairment showed average localization ability. These results provide satisfactory evidence that the data for all participants can be grouped together without introducing a bias due to hearing loss. Again, the assumption that localization performance is not degraded by individuals with moderate symmetrical hearing loss is supported by previous findings [29].
Since ten different phrases were used, it was conceived that some phrases could potentially be better or worse for localization. Figure 4.3 does show that the phrase “porch steps” was the most incorrectly localized phrase by nearly 8%. It is possible that this phrase was too short and did not allow sufficient time to localize as specified by [101], however, the phrase, ‘these shoes’, which was only milliseconds longer, recorded much better localization accuracy. In addition, the longest phrase, ‘instructions are crucial’, was the second-worst localized phrase. Other reasoning for the poor localization ability could point to spectral content; however, ‘porch steps’ contained energy in the 6 to 10 kHz bands comparable to the best localized phrases, ‘offensive as innocence’ and ‘outstanding effects’. Without conclusive evidence to discard this single phrase, the experiment was not modified to exclude the phrase, ‘porch steps’.

Figure 4.3: Response accuracy across all subjects for each phrase used in Experiment 2.
Experiment 3: Non-Anechoic Localization Experiment Results

The final experiment sought to answer the primary question in this investigation: does a single early-reflection assist in the front-back localization of stationary sound sources when using generalized HRTFs. To determine if there was an effect on localization performance, the difference between anechoic (Experiment 2) and non-anechoic (Experiment 3) localization accuracy was calculated for every subject in each of the four quadrants.

The null hypothesis ($H_0$) for this investigation stated that the average response accuracy of Experiment 3 should be equal to Experiment 2. Since it is believed that reflections potentially assist response accuracy, the alternative hypothesis ($H_A$) stated that the average response accuracy of Experiment 3 should be greater than Experiment 2. This experimental statement is summarized in Figure 4.4. For this experiment a 90% significance level ($\alpha = 0.1$) was chosen.

![Experiment Hypothesis](image)

Figure 4.4: Null ($H_0$) and alternative ($H_A$) hypotheses for the primary investigation of these experiments.

The individual differences in response accuracy from anechoic conditions are plotted in Figure 4.5(a). The data is divided into the three reflection types and plotted
together in the respective quadrant location. Maximum, minimum, and average values for difference in localization ability are provided in Figure 4.5(b).

Figure 4.5: Percentage difference in response accuracy between non-anechoic (Exp. 3) and anechoic (Exp. 2). Plot of individual results (a) and values for maximum, minimum, and average differences (b).

The data reveal large variability for all reflection types and indicates that the results were highly subject dependent. Some individuals localized with considerable improvement while others seemed to be inhibited by the additional reflection. Given these data it would be difficult to suggest that, in general, reflections contribute significant information for front-back sound localization using generalized HRTFs. The large p-values shown in Table 4.3 indicate that there is not much evidence against the null hypothesis \( \text{H}_0 \) stated in Figure 4.4. Likewise, since the 90% confidence intervals contain 0% mean difference in localization ability, the evidence would argue against the
alternative hypothesis \((H_A)\). A graphical view of the confidence intervals is shown in Figure 4.6 for easier comparison.

Table 4.2: Statistical analysis results. 90% confidence intervals and p-values at the 90% significance level are reported for the twelve experimental categories: three types of reflections in four quadrants.

<table>
<thead>
<tr>
<th></th>
<th>90% Confidence Interval</th>
<th>P-Value (\alpha = 0.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td><strong>Front Left</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Floor 6ms</td>
<td>-20%</td>
<td>3%</td>
</tr>
<tr>
<td>Wall 6ms</td>
<td>-1%</td>
<td>15%</td>
</tr>
<tr>
<td>Wall 20ms</td>
<td>-4%</td>
<td>17%</td>
</tr>
<tr>
<td><strong>Front Right</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Floor 6ms</td>
<td>-6%</td>
<td>15%</td>
</tr>
<tr>
<td>Wall 6ms</td>
<td>-18%</td>
<td>4%</td>
</tr>
<tr>
<td>Wall 20ms</td>
<td>-19%</td>
<td>4%</td>
</tr>
<tr>
<td><strong>Rear Left</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Floor 6ms</td>
<td>-13%</td>
<td>16%</td>
</tr>
<tr>
<td>Wall 6ms</td>
<td>-10%</td>
<td>12%</td>
</tr>
<tr>
<td>Wall 20ms</td>
<td>-5%</td>
<td>12%</td>
</tr>
<tr>
<td><strong>Rear Right</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Floor 6ms</td>
<td>-12%</td>
<td>11%</td>
</tr>
<tr>
<td>Wall 6ms</td>
<td>-13%</td>
<td>-2%</td>
</tr>
<tr>
<td>Wall 20ms</td>
<td>-22%</td>
<td>2%</td>
</tr>
</tbody>
</table>

Figure 4.6: Confidence intervals for the twelve experimental categories. Intervals containing 0% indicate that there is potentially no average improvement.
It is easily seen that nearly all confidence intervals contain 0%, i.e. no change. This would support the conclusion that reflections do not in general assist front-back localization ability for every listener. The one confidence interval that did not contain zero yielded the largest p-value. Although this result does not indicate statistical significance for inhibiting localization performance, it did direct attention toward a future investigation of that nature.

Analysis of the p-values show that sources located in the front-left quadrant with an added 6ms wall reflection yielded significance at the 90% level. However, since the same dataset was used to create multiple tests (i.e. twelve p-value tests for each reflection type and quadrant location), the initial alpha \( \alpha_{\text{initial}} = 0.1 \) must be corrected to account for multi-testing. It is necessary to place a stricter bound by correcting \( \alpha_{\text{initial}} \) so that instances of p-values less than \( \alpha_{\text{initial}} \) are statistically based rather than being attributed to pure chance. Dividing \( \alpha_{\text{initial}} \) by 12, i.e. the number of data divisions, yields

\[
\alpha_{\text{corrected}} = 0.0083.
\]

Given this new threshold for significance, the p-value of 0.08405 for a front-left sound source combined with a 6ms wall reflection is no longer statistically significant on the 90% level. Due to the concern with multi-testing on the same dataset, statistical analysis beyond answering the initial hypothesis was discouraged [102].

Although the primary hypothesis was shown to have no statistical basis, additional data analysis was conducted to determine any other information that could be gleaned from this investigation. It was observed that the effect of the floor boundary yielded a greater variance for difference in response accuracy than did the wall
reflections. This supports the investigator’s theory that floor reflections are less helpful than lateral reflections.

Strict comparison of localization accuracy difference in Figure 4.5(a) shows a very slight decrease in localization ability for sources in the right hemisphere when a reflection from a right-wall boundary is added. Conversely, a slight improvement in localization ability is seen for sound sources located in the left-hemisphere when a reflection from a right-wall boundary is added. If these relationships are assumed to be relevant, they would contradict the investigators theory that sound sources would be more easily located if both the sound source and reflection originated from the same side. A different look at the overall responses for each of the reflection boundary and amplitude types is shown in Figure 4.7. The anechoic mean for all subjects is indicated in each quadrant. No conclusions were drawn from this visualization; the plot is included here for completeness.

Figure 4.7: Overall localization accuracy for different categories of reflection type. Lines in each quadrant represent the average localization accuracy in anechoic conditions.
As in Experiment 2, analysis of localization accuracy per phrase was conducted for Experiment 3. Participants’ responses appeared evenly distributed across the ten phrases as seen in Figure 4.8. The poorest localized phrase in Experiment 2, ‘porch steps’, was the second poorest localized phrase in Experiment 3. However, the phrase yielded nearly the same response accuracy to ‘outstanding effects’ which was one of the best localized phrases in Experiment 2. This random distribution of localization accuracy for each of the ten phrases may provide an indication that certain phrases did not significantly affect localization performance.

![Phrase Analysis – Experiment 3](image)

Figure 4.8: Response accuracy across all subjects for each phrase used in Experiment 3.

More Data Analysis

One of the questions from the post-experiment survey (see Appendix F) asked participants if they felt their ability to correctly localize sound events improved over the course of the experiments. Subjects unanimously reported that they perceived an
improvement. Improvements of this nature would agree with [78] that training and exposure to localization tasks can improve localization ability. However, subjects’ response accuracy for Experiment 3 actually showed little evidence for improvement throughout the duration of the experiments. Localization performance appeared random with respect to time as shown by a representative subject’s responses in Figure 4.9.

Figure 4.9: Response accuracy versus trial number for a representative participant. No improvement in accuracy was seen with time.

Another question from the post-experiment survey asked participants if they perceived a bias towards a particular front/back hemisphere when the location of the sound source was questionable. A majority of the subjects answered yes, however, the answers were evenly split between the front and rear hemispheres. The percentage of subject’s ‘rear’ responses for both Experiments 2 and 3 are shown in Figure 4.10. The average across all participants (~50%) show there was no overall bias for either experiment. Table 4.4 shows the percentage of rear responses for each individual. Only two subjects showed a noticeable bias; both towards the rear hemisphere. Ironically, on the post-experiment survey neither of these subjects reported a preference in hemisphere.
Figure 4.10: Plot of individual rear response percentages for Experiment 2 (left) and Experiment 3(right). No bias to front/back hemispheres is indicated by 50%.

Table 4.3: Percentage of rear responses for each individual. No bias to front/back hemispheres is indicated by 50%.

<table>
<thead>
<tr>
<th>Subject #</th>
<th>Exp. 2</th>
<th>Exp. 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>070109_001</td>
<td>52%</td>
<td>51%</td>
</tr>
<tr>
<td>070209_001</td>
<td>53%</td>
<td>59%</td>
</tr>
<tr>
<td>070709_001</td>
<td>53%</td>
<td>45%</td>
</tr>
<tr>
<td>070909_001</td>
<td>53%</td>
<td>42%</td>
</tr>
<tr>
<td>071009_001</td>
<td>45%</td>
<td>42%</td>
</tr>
<tr>
<td>071309_001</td>
<td>52%</td>
<td>53%</td>
</tr>
<tr>
<td>071409_001</td>
<td>70%</td>
<td>64%</td>
</tr>
<tr>
<td>071609_001</td>
<td>55%</td>
<td>53%</td>
</tr>
<tr>
<td>072209_001</td>
<td>48%</td>
<td>38%</td>
</tr>
<tr>
<td>072209_002</td>
<td>32%</td>
<td>32%</td>
</tr>
<tr>
<td>072309_001</td>
<td>57%</td>
<td>62%</td>
</tr>
<tr>
<td>072309_002</td>
<td>47%</td>
<td>47%</td>
</tr>
<tr>
<td>072409_001</td>
<td>42%</td>
<td>56%</td>
</tr>
</tbody>
</table>

Since head and/or source movement has been reported to assist in the resolution of front-back localization [68, 69, 70], it was suspected that subjects would show
improved localization performance when the location of the sound source transitioned from front-to- back or back-to-front in adjacent trials. Informal post-experiment discussions with many of the participants indicated general agreement with this suspicion. Since the presentation order of the trials was random, the number and order of transitions was different for all participants. This required a comparison based on the localization accuracy percentages between transition and non-transition trials. Figure 4.11 shows a slight average improvement in front-back localization accuracy for transition trials in the two localization experiments. However, the modest improvement does not strongly support the theory that the transitions are highly useful events for subsequent localization trials.

![Front/Back Transition Analysis](image)

Figure 4.11: Response accuracy for trials where the sound source was presented in the opposite front/back hemisphere on the previous trial.
Round Two Experiments

Approximately a month and a half after the initial round of experiments, Experiment 2 and 3 were repeated with eleven of the same thirteen participants from the first round. Since the first round indicated that floor reflections had no influence on the localization performance, they were eliminated from the round two set of trials. This shortened the experiment to approximately 20 minutes in length. Data analysis similar to that described for the first round of experiments was conducted on the second round data. Comparable results with large variability were found and nothing of statistical significance was revealed. However, a minimal decrease in localization performance was observed for sound sources located in the right hemisphere and slightly improved performance was found in the left hemisphere. Though these differences were very slight, they were similar to the overall analysis of round one as mentioned previously in this section.

A comparison of the trial responses was conducted to determine how consistent the subjects’ responses were between the two rounds. Since the same samples were used for both rounds (same phrase, quadrant location, reflection type, and relative reflection amplitude), the data allowed a direct comparison. The only two differences between experimental rounds were: 1) the presentation order of trials was random for both rounds, and 2) the trial count was reduced for the second round. It was assumed that these two differences would not directly influence response consistency. Figure 4.12 shows the percentage of consistent (correct and wrong) responses for the anechoic, 6ms wall reflection, and 20ms wall reflection conditions. Both the front-left and rear-right
quadrants show that participants chose the same response for both rounds two-thirds of the time. However, the other two quadrants only achieved consistent responses slightly better than half of the time. It is interesting to note the similar relationship between these “cross” quadrants as noted in the Experiment 2 discussion. No reasoning for this tendency was determined, but one possibility could attribute this characteristic to nuances of the HRTF dataset used to virtually place the sounds for these experiments.

Figure 4.12: Comparison of responses between round 1 and round 2. The percentages refer to the amount of trials were answered the same for both rounds (regardless if wrong or correct).

**Statistical Discussion**

Given the data collected and the statistical significance of the various parameters, one would conclude, in general, that *a single early-reflection does not assist or inhibit listeners when localizing sound events while using generalized HRTFs*. However, these
experiments report looked at a large number of variables: the reflection boundary type (3 cases) and relative amplitude of the reflection (5 cases). This resulted in an experiment design which likely fell into the pitfall of trying to extract too much information from the limited amount of time and number of participants. It may have been better to take a particular element of the initial hypothesis rather than trying to answer if a variety of single, early-reflections contribute to localization performance. Following the initial theory that lateral reflections from a wall would assist more than floor reflections, it would have been possible to focus on the benefit of a single, lateral reflection compared to anechoic localization conditions. Another test (independent of the first) could have then compared the front-back localization performance between left and right hemispheres to answer the investigator’s other theory that the lateral orientation of the listener, sound source, and reflective boundary affects localization performance. Obviously many other independent experiments can be conceived, but it is important to keep fundamental elements of statistical analysis in mind.

Several recommendations for experimental procedures come from this project. First, choose a specific hypothesis which can be answered within the both time and financial budgets. Second, choose a level of statistical significance, i.e. $\alpha$, that is appropriate for the investigation. For these experiments where human health will not be compromised from the outcome, it is possible to seek a looser indication of statistical significance by choosing a larger alpha. Finally, acquire the necessary number of participants which will allow the hypothesis to be answered with the desired level of statistical significance. These participants should be chosen so that they ideally represent
the entire population of interest thus providing a normal distribution and a means to approximate the true population average.

**Logistic Regression**

It was initially of interest to analyze the data using a classical Receiver Operating Characteristic (ROC) analysis developed by Swets [103]. Discussion with two statisticians provided supplementary information on the methods to applying ROC to the data of this investigation, however, the previously mentioned size limitations and other shortcomings of the subjective experiments limited the potential benefits of the formal statistical analysis. A basic description of the methodology is provided for future reference.

To utilize ROC analysis it is first necessary to develop a curve which predicts the probability of a correct answer given the variables in the experiment. It was suggested by [102] to use logistic regression and an equation of the form of (4.1) to develop this curve.

\[
\log \left( \frac{P}{1-P} \right) = \beta_0 + \beta_1 x_{\text{Amplitude}} + \beta_2 x_{\text{Type}} + \gamma_{12} x_{\text{Interaction}} \quad (4.1)
\]

The y-intercept (\( \beta_0 \)), reflection amplitude (\( \beta_1 \)), reflection type (\( \beta_2 \)), and reflection type/amplitude interaction (\( \gamma_{12} \)) coefficients are adjusted so that equation (4.1) achieves a best fit to the experimental data. The area under this describing equation is then compared to the area under an ROC curve similar to those of Figure 4.13. Depending on the degree of discrimination for the analysis, an ROC curve with the appropriate level of accuracy can be chosen. If Equation (4.1) results in a greater area
than the ROC curve, the experimental hypothesis has validity based on the degree of accuracy chosen for analysis.

Figure 4.13: Example ROC curves for discriminating against different levels of accuracy. Figure taken from [103].

A supplementary comment by two statistics consultants from the Montana State University Mathematics Department [102] suggested that it was also necessary to deflate the standard error using a quasi-binomial probability mass function so that the coefficients would not overestimate the probability of a correct response. If no correction is applied, it is possible that level of significance determined is not appropriate for the experimental dataset.
Another Consideration

When a delayed version of a sound sample is combined with itself, a phenomenon called “comb-filtering” produces notches and peaks in the frequency spectrum due to destructive and constructive interference. The notches are equally spaced at \((1/\text{delay})\) Hz with the first notch located at the half-wavelength of the delay. The relative amplitudes of the direct and delayed signal determine the degree of interference. Figure 4.14 shows four comb-filter examples (top two correspond to a 6ms delay and bottom two correspond to a 20ms delay) all with different relative amplitudes for the delayed signal. As expected, as the two sound sources become closer in amplitude, the greater the degree of interference is observed. A comb filter with the delayed signal only 6 dB down from the direct signal produces a 3.5 dB increase in gain at each of the peaks whereas a delayed signal 18 dB down from the direct signal produces only a 1 dB increase in gain at the peaks. It has been shown that the peaks, not the notches, are more perceptually relevant [104].

Since the entire basis for this investigation involved the addition of a delayed version to the primary signal, it is possible that spectral shaping from the comb-filter influences the audible cues for localization. The perceptual effect of these spectral notches and peaks is more noticeable at lower frequencies where the spacing is more distinct [105], however, the spectral information used for localization is contained in the mid to upper frequency regions. As seen in Figure 4.14, the spacing becomes very small near 2 kHz for the 6ms delay and is smaller for the 20ms delay. Although shorter delays might provide more perceptual effects in the mid to upper frequency regions, the comb-
filtering due to the delay choices used in these experiments did not likely have a significant effect on localization.

Figure 4.14: Comb-filtering examples for two delays, 6ms (top) and 20ms (bottom), and four relative amplitudes of the delayed signal, -6, -9, -15, and -19dB.
ONGOING WORK

At the onset of this project there were numerous potential ideas to be investigated. Limited time did not allow for many of these topics to be addressed. A primary topic of interest was HRTF interpolation methods. Other potential avenues for future research are also summarized.

HRTF Interpolation Implementation

Since many applications require HRTF locations not corresponding to any of the measured points in an HRTF dataset, interpolation among the available measurements is useful. For the experiments described above, interpolation could have been used to simulate exact azimuth/elevation coordinates for the reflections, however, since the orientation of the source, listener, and reflective boundary were chosen to minimize this error below 1.5° for the reflections, no interpolation was applied. As a pilot investigation, three documented methods for HRTF interpolation were implemented [22, 39]: closest point choice, a weighted linear interpolator based on the closest point, and spectral shaping using a spherical head model.

Processing Concerns

For all interpolation implementations there were two questions necessary to address: 1) how to process lengthy sound samples and 2) what length of fast-Fourier transform (FFT) should be used. The overlap-add method is a common time-domain method for decomposing a long audio file into segments with a time window (Hamming,
Bartlett, etc) and then reconstructing the segments back into the original audio file. The need for overlapping the segments is due to additional sample length resulting from the convolution of the window and audio segment. For example, if the segment is N samples long and the window is of length M samples, the resulting convolution would yield an output of N+M-1 samples. Since the beginning of adjacent audio segments are separated by N samples, it is necessary to overlap the segments by M-1 samples to reconstruct the original audio file [106]. Experimentation with the amount of overlap found popping artifacts between segments and perceived amplitude modulation of the source file.

Frequency-domain processing of the audio segments leads to the question of what length of FFT to use. Although it is fairly common to blindly apply the FFT to avoid working with time-domain convolution, care should be taken when selecting the length of FFT to avoid circular convolution. For example, consider a filter with impulse response of length P samples which is convolved with a sound segment of length Q samples. The resulting convolution yields an output of P+Q-1 samples. Instead, if it was desired to avoid the convolution computation by taking an FFT of length L on both the filter and sound segment, the frequency domain multiplication would return an FFT of length L. If L < (P+Q-1) then an insufficient amount of frequency bins were used to represent the actual result. Since the FFT is periodic on L, the additional samples are pushed into the adjacent period causing frequency aliasing which would return an incorrect spectrum for the output [107].
Implementation Comments

The first interpolation method simply chose the nearest measurement for a given set of azimuth and elevation coordinates. Therefore, the resolution of the source movement was limited to the resolution of the discrete HRTF measurements. The sound source was swept across the horizontal plane for evaluation. As expected, there were obvious jumps in the sound when the position transitioned from one HRTF measurement to another.

The second interpolation implementation was a weighted linear interpolator and was taken from [39]. This method selected the three nearest measurements to form a triangle around the desired position coordinate. The weights of the three HRTF measurements, based on geometric relationships, were applied to its respective measurement. These three weighted HRTF measurements were then added together to obtain the interpolated HRTF. Simple evaluation of this method showed significant improvement from the nearest measurement method.

The third method applied the spherical head model from [33] to create an approximated HRTF for exact coordinates of interest. The magnitudes of this approximated HRTF were then averaged with the magnitudes of the HRTF measurement closest to the desired coordinates. Since no smoothing between HRTF measurements was implemented the perceptual evaluation of this method observed transitions between the HRTFs. It would be preferred to implement a weighting system similar to the second interpolation method.
Next Steps

The educational opportunity to research the vast corridors of spatial audio theory and technologies has revealed exciting avenues towards the development of novel ideas. This section presents some additional theories for furthering the initial hypothesis of these experiments. Other indirectly related research conceived during this investigation is mentioned as well.

Further Ideas with these Experiments

Although the statistical analysis of these data does not support the primary hypothesis that reflections assist with localization, it is still of interest to make a more focused investigation of specific elements from this investigation. An experiment to address the investigator’s theory that lateral reflections play some role in the perception of virtual auditory environments could easily be developed using the methods from this thesis. It would also be interesting to see if a different HRTF dataset yields similar response accuracy averages in the “cross” quadrants (i.e. front-left and rear-right versus front-right and rear-left). A separate idea involved with resolving front-back confusions considers the use of an “auditory anchor” to aid listeners in the front-back localization of sound events. It is conceived that if a listener is provided prior knowledge regarding the location of a reference signal, then front-back localization will be improved if the reference is presented with the sound event of interest. Implementation of the interpolation algorithms onto hardware such as a FPGA would allow for a fuller
perceptual evaluation. This could also allow for experimentation to test listener’s control of the sound within an audio environment.

Future Areas of Research

The evaluation of much literature helped to reveal different angles on the perception of spatial audio. Many recent papers indicate that the research of neural mechanisms underlying sound localization is a relevant area of ongoing research. Although only a shallow investigation regarding the neural aspects of hearing was appropriate in the limited time-frame of this project, a few selected works which tie into the theories of this investigator are briefly described.

A significant issue to address for creating more effective spatial auditory displays is the conflict between multi-sensory information when localizing sounds [108, 109, 110, 111]. There is evidence that the visual dominance over auditory perception must be treated before a sense of audio immersion can be completed. However, if a visual display is not possible for certain applications, an enhancement of listener’s attention within a spatial auditory environment must be developed to achieve a heightened cerebral response in the auditory cortex [112, 113]. Another problem with creating a realistic auditory display is the need for more complex environments. This investigator believes that localization within a dense display of auditory information has a neural basis as shown by [114], but also is greatly reliant upon principles from auditory scene analysis (ASA) [115, 116]. If interested in other psychoacoustic fundamentals, Moore’s book [25] would be a highly recommended place to start.
The design of realistic auditory displays is a complex task which must account for individual nuances as shown by Mrsic-Flogel et al. [117] that personal experiences calibrate the neural mechanisms for auditory space perception. This extends individualization far beyond the issue between generalized and individualized HRTFs. The diversity of potential listeners opens the door to a variety of other questions to be asked: How does the multi-sensory dominance of vision change for visually impaired people? How does the emotional status or level of stress change a person’s perception of auditory space? Would manipulation of aspects of the inner-ear such as the cochlear “bark bands” [74] help hearing impaired listeners with spatial sound environments?

**Conclusion**

The research in this thesis looked to address the issues with using non-individualized HRTFs by evaluating the effect of a single, early-reflection on front-back localization performance. Although additional experimentation with lateral reflections is suggested, the data from these experiments provide no statistically significant results to confirm that a single early-reflection assists with resolving front-back confusions. Nonetheless, the investigation has been an intriguing journey on the road towards understanding the complexities of sound spatialization with headphone audio.

After 30+ years of applying HRTFs to the spatialization of sound events, successful implementation for general applications have yet to be realized. Is it possible that individualized HRTFs are truly necessary or is there a fundamental aspect of sensory input that is being overlooked? Either way, the methods for achieving a realistic auditory display will continue to be investigated and developed.


[79] IRB website: [http://www2.montana.edu/irb/](http://www2.montana.edu/irb/)


APPENDICES
APPENDIX A

HUMAN PARTICIPANT CONSENT FORM
MONTANA STATE UNIVERSITY
Institutional Review Board Application for Review
(revised 11/17/08)

[Include copies of PI's and Co-PI's "Completion Certificate(s)" as proof that all have received the education and instructions for researchers using human subjects. The preferred instruction and education is that from the National Cancer Institute. http://cancer.gov - Human Participant Protections


Beginning January 1, 2006, University policy requires that all protocols submitted from individuals NOT employed by or students of Montana State University be charged a $500 review fee per application. Renewals for those proposals will be at no charge.

THIS AREA IS FOR INSTITUTIONAL REVIEW BOARD USE ONLY. DO NOT WRITE IN THIS AREA.

Application Number: Approval Date: IRB Chair's Signature:
Disapproved: 

Submit 14 copies of all application (including the signature copy), along with 14 copies of the subject consent form and 14 copies of all other relevant materials, to Institutional Review Board, 960 Technology Blvd., Room 127, Montana State University, Bozeman, MT 59717-3610. (Please staple, bind or clip together the application form, surveys, etc. as 14 individual packets; one complete packet for each board member.) Submit one copy of grant contract proposal for the office file. For assistance, call 994-6783 or contact the Institutional Review Board Chair, Mark Quinn at 994-5721.

Date: 

I. Investigators and Associates (list all investigators involved; application will be filed under name of first person listed)

NAME: Darrin K. Reed
DEPT: Electrical and Computer Engineering
ADDRESS: 102 Grant Chamberlain Apt #2D
E-MAIL ADDRESS: darrinreed@hotmail.com
DATE TRAINING COMPLETED: 4/19/09
TITLE: Research Assistant
PHONE #: (208)-880-1859
Bozeman, MT 59715

NAME: Rob Maher
DEPT: Electrical and Computer Engineering
ADDRESS: P.O. Box 173780 Bozeman, MT 59717-3780
E-MAIL ADDRESS: rmaher@ece.montana.edu
DATE TRAINING COMPLETED: 9/3/09
TITLE: Department Head
PHONE #: (406)-994-7759

Do you as PI, any family member or any of the involved researchers or their family members have consulting agreements, management responsibilities or substantial equity (greater than $10,000 in value or greater than 5% total equity) in the sponsor, subcontractor or in the technology, or serve on the board of the Sponsor? YES X NO
If you answered YES, you will need to contact the Director of the Technology Transfer Office, Dr. Rebecca Mahurin at 406-994-7888.

II. Title of Proposal: Evaluation of early reflection contribution to front-back resolution in binaural, headphone audio.

III. Beginning Date for Use of Human Subjects: ASAP

IV. Type of Grant and/or Project (if applicable) Thesis Project: Internal funding

VI. Signatures
Submitted by Investigator
Typed Name: Darrin K. Reed
Signature: 

Faculty sponsor (for student)
Typed Name: Rob Maher
Signature: 

Date: 5/12/09
VII. Summary of Activity.

Provide answers to each section and add space as needed. Do not refer to an accompanying grant or contract proposal.

A. RATIONALE AND PURPOSE OF RESEARCH

Can front-back localization errors in binaural, headphone audio be reduced through the addition of 1st order reflections?

B. RESEARCH PROCEDURES INVOLVED. Provide a short description of sequence and methods of procedures that will be performed with human subjects. Include details of painful or uncomfortable procedures, frequency of procedures, time involved, names of psychological tests, questionnaires, restrictions on usual life patterns, and follow up procedures.

The procedure will be broken into two separate testing periods which are preferentially conducted on two separate days. Upon inquiry of participation subjects will be informed of the normal hearing requirement for participation per ANSI S1.6-2004. Prospective subjects will be asked the following questions to help them consider if they are capable of participation:

1) Do you feel handicapped by a hearing problem?
2) Do you have difficulty when someone speaks in a whisper?
3) Does a hearing problem cause you difficulty when listening to TV or radio?
4) Do you feel that any difficulty with your hearing limits or hampers your personal/social life?

BACKGROUND: The 3D spatialization of sounds for this experiment will be created from generic Head-Related Transfer Functions (HRTF) provided by the UC Davis (CIPIC) database. Noise exposure in all tests will follow the Occupational Safety and Health Administration (OSHA) guidelines 1910.95(b)(1)) to ensure that none of the participants will be subject to noise induced hearing loss. Subjects will be required to sign an informed consent waiver which complies with HHS regulations 45 CFR 46.116.

The first of the two separate testing periods will be composed of a hearing screening, baseline front-back localization test, and reflection detection experiment. The hearing screening follows basic Pure-tone Audiometry methods. An approximate threshold will be determined for seven pure tones at octave-band center frequencies (250, 500, 1K, 2K, 4K, 8K, 16KHz) for both ears. All tones will be presented via Matlab GUI using a psychoacoustic Maximum Likelihood procedure.

The baseline front-back localization test will present the subject with 40 trials of a single binaurally processed speech sample (~1sec) to the participant. The subject will be asked to identify if the sound appeared to be in the front or rear hemisphere. This experiment will take approximately 10 minutes.

The reflection detection experiment will present subjects with (~1 sec) speech samples. These samples will be modified with the presence (yes/no) of first order reflections from boundaries (floor/side); the samples also have three (3) relative direct-reflection amplitudes and two (2) delay times between direct and reflected signals. The sound will also be presented in both front and rear hemispheres. This yields 25 permutations presented to the subjects. Subjects will be presented with two binaural speech signals: 1) reference and 2) test; subjects report same or different. If ten subjects complete the experiment, then six trials per permutation would be desired. Considering each trial is approximately eight seconds, the experiment is expected to take 20-25 minutes.

The second testing period, composed of three separate experiments, will test the subject's ability to locate a speech sample (~1 sec) in either the front or rear hemisphere. The first experiment will test the effect of a first order floor reflection on the subject's ability to correctly localize a sound as being presented in the front or back. Similar to the reflection detection experiment in the first period, three (3) relative direct-reflection amplitudes and two (2) delay times between direct and reflected signals will be presented. Subjects will be presented with a single speech sample where they will be asked if the speech was presented in the front or rear.
hemispheres. If ten subjects complete the experiment, then six trials would be desired. Considering each trial is approximately five seconds, the experiment is expected to take approximately 15 minutes. The second experiment is identical to the first experiment with the exception that the reflection is due to the side boundary not the floor.

The third experiment is an optional test that seeks to test if an auditory reference (spatial "anchor") assists in front-back localization. The subject will be first presented a broadband noise sample or non-competitive speech signal with known location (front or rear hemisphere). Subsequently the test speech sample will be presented and the subject will be asked to identify if the test signal is in the front or back. With 40 iterations this experiment is expected to take 15 minutes.

C. DECEPTION - if any deception (withholding of complete information) is required for the validity of this activity, explain why this is necessary and attach debriefing statement. NA

D. SUBJECTS
1. Approximate number and ages
   How Many Subjects: 20 individuals with normal hearing
   Age Range of Subjects: 18 – 35 years

2. Criteria for selection:
   Male or female, ideally in the 18-35 age range, who report normal hearing

3. Criteria for exclusion:
   A preliminary screening for normal hearing per ANSI S3.6-2004 will be necessary. Eight standard audiometric frequencies (125, 250, 500, 1k, 2k, 4k, 8k, 16kHz) will be tested.

4. Source of Subjects (including patients):
   Volunteer subjects from Bozeman community
   Montana State University departments: Engineering, Music and Psychology

5. Who will approach subjects and how? Explain steps taken to avoid coercion.
   Primary Investigator: Darrin K. Reed
   Request for participants through various MSU departments

6. Will subjects receive payments, service without charge, or extra course credit? Yes or NO
   (If yes, what amount and how? Are there other ways to receive similar benefits?)

7. Location(s) where procedures will be carried out.
   Digital Audio Signal Processing Laboratory, Unit 21, Faculty Court.

E. RISKS AND BENEFITS (ADVERSE EFFECTS)
1. Describe nature and amount of risk and/or adverse effects (including side effects), substantial stress, discomfort, or invasion of privacy involved.
   Noise exposure in all tests will follow the Occupational Standards and Health Administration regulations 1910.95(b)(1) to ensure that none of the participants will be subject to noise induced hearing loss.

2. Will this study preclude standard procedures (e.g., medical or psychological care, school attendance, etc.)? If yes, explain. NO

3. Describe the expected benefits for individual subjects and/or society.
   There will be no direct benefit to individual subjects.
F. ADVERSE EFFECTS
1. How will possible adverse effects be handled? NA
   By investigator(s):
   Referred by investigator(s) to appropriate care:
   Other (explain):

2. Are facilities/equipment adequate to handle possible adverse effects? Yes or No
   (If no, explain.)

3. Describe arrangements for financial responsibility for any possible adverse effects.
   MSU compensation (explain):
   Sponsoring agency insurance:
   Subject is responsible:
   Other (explain):

G. CONFIDENTIALITY OF RESEARCH DATA
1. Will data be coded? Yes or No
2. Will master code be kept separate from data? Yes or No
3. Will any other agency have access to identifiable data? Yes or No
4. How will documents, data be stored and protected?
   Locked file: Computer with restricted password: Coded data on Lab/Office PCs
   Other (explain): ECE Main Office

III. Checklist to be completed by Investigator(s)
A. Will any group, agency, or organization be involved? Yes or No
   (If yes, please confirm that appropriate permissions have been obtained.)

B. Will materials with potential radiation risk be used (e.g. x-rays, radiotopes)? Yes or No

C. Will human blood be utilized in your proposal? Yes or No

D. Will non-investigational drugs or other substances be used for purposes of the research? Yes or No

E. Will any investigational new drug or other investigational substance be used? Yes or No

F. Will an investigational device be used? Yes or No

G. Will academic records be used? Yes or No

H. Will this research involve the use of:
   Medical, psychiatric and/or psychological records Yes or No
   Health insurance records Yes or No
   Any other records containing information regarding personal health and illness Yes or No
   If you answered "Yes" to any of the items under "H," you must complete the HIPAA worksheet.

I. Will audio-visual or tape recordings or photographs be made? Yes or No

J. Will written consent form(s) be used? Yes or No
SUBJECT CONSENT FORM FOR PARTICIPATION IN HUMAN RESEARCH
AT MONTANA STATE UNIVERSITY
SPATIAL AUDITORY PERCEPTION EVALUATION- STUDY #1

You are being asked to participate in a study of spatial auditory perception. The purpose of this study is to establish a basis for auditory localization and echo detection ability. You were identified as a possible subject because you are between 18 and 35 years old and have reported normal hearing ability.

Preliminary Screening
Participants in the following experiments are required to possess normal hearing. Your hearing will be screened through a series of yes/no questions regarding the presence of various tones. NOTE: This screening is not an absolute measure of your hearing threshold. Please contact an audiologist if you have concerns regarding your hearing ability.

Procedures Involved
If you agree to participate, the test will consist of two experiments and should take approximately 1 hour to complete. There is no payment for participation and this study has no direct benefit to you. This first experiment will record your perceived location of a short speech sample presented through headphones. Your response is limited to either the front or rear hemispheres. The second experiment will determine the audibility of a reflection, i.e. echo, added to a short speech sample. You will be presented with two audio samples where your objective is to judge if the samples are same or different.

Risks
The level of auditory stimulus presented follows Occupational Safety and Health Administration (OSHA) regulations 1910.95(b)(1). None of the procedures will cause discomfort. The audio samples are phrases read by male and female speakers, and contain non-invasive statements/questions.

Alternatives Available
You may withdraw from the study at any time without penalty or prejudice.

Confidentiality of Records
Your experimenter will treat your identity with professional standards of confidentiality. All records from the experiment will be stored under a subject number and your name will not be directly linked with the data. The information obtained in this study may be presented in future publications, but your identity will not be revealed.

Injury and Compensation
You should understand that in the very unlikely event of an injury occurring as a subject in this study, there is no compensation available from Montana State University, and any necessary medical care will be your own responsibility.

If you have any additional questions about the experiment please contact the principal investigator, Darrin Reed at 208-880-1859. If you have questions concerning the rights of human subjects contact Mark Quinn, Human Subjects Committee Chair at 406-994-4707.

APPROVED
MSU IRB
05/21/2009
05/20/2010
Expiration date
SUBJECT CONSENT FORM FOR PARTICIPATION IN HUMAN RESEARCH
MONTANA STATE UNIVERSITY
SPATIAL AUDITORY PERCEPTION EVALUATION- STUDY #2

You are being asked to participate in a study of spatial auditory perception. The purpose of this study is to determine if additional information from sound reflections or from a known auditory event location can assist a listener in the location of a speech sample. You were identified as a possible subject because you are between 18 and 35 years old and have reported normal hearing ability.

Preliminary Screening
Participants in the following experiments are required to possess normal hearing. Your hearing will be screened through a series of yes/no questions regarding the presence of various tones. NOTE: This screening is not an absolute measure of your hearing threshold. Please contact an audiologist if you have concerns regarding your hearing ability.

Procedures Involved
If you agree to participate, the test will consist of two experiments and will take approximately 1 hour to complete. There is no payment for participation and this study has no direct benefit to you. This first experiment will record your perceived location of a short speech sample presented through headphones. Your response is limited to either the front or rear hemispheres. The second experiment will determine if the presence of an auditory event with known location assists in the front-back localization of a different auditory event.

Risks
The level of auditory stimulus presented follows Occupational Safety and Health Administration (OSHA) regulations 1910.95(b)(1). None of the procedures will cause discomfort. The audio samples are phrases read by male and female speakers, and contain non-invasive statements/questions.

Alternatives Available
You may withdraw from the study at any time without penalty or prejudice.

Confidentiality of Records
Your experimenter will treat your identity with professional standards of confidentiality. All records from the experiment will be stored under a subject number and your name will not be directly linked with the data. The information obtained in this study may be presented in future publications, but your identity will not be revealed.

Injury and Compensation
You should understand that in the very unlikely event of an injury occurring as a subject in this study, there is no compensation available from Montana State University, and any necessary medical care will be your own responsibility.

If you have any additional questions about the experiment please contact the principal investigator, Darrin Reed at 208-880-1859. If you have questions concerning the rights of human subjects contact Mark Quinn, Human Subjects Committee Chair at 406-994-4707.

APPROVED
MSU IRB
05/21/2009
Date approved
05/20/2010
Expiration date
AUTHORIZATION

I have read the above and understand the discomforts, inconvenience and risk of this study. I,
__________________________ (name of subject), agree to participate in this research. I understand that I may later refuse to participate, and that I may withdraw from the study at any time. I have received a copy of this consent form for my own records.

Signed:________________________________________
Witness:________________________________________
Investigator:____________________________________
Date:__________________________________________

I have read the above and understand the discomforts, inconveniences and risks of this study. I,  
__________________________ (name of parent or guardian), related to the subject as ______________________ (relationship), agree to the participation of  
__________________________ (name of subject) in this research. I understand that the subject or I may later refuse participation in this research and that the subject, through his/her own action or mine, may withdraw from the research at any time. I have received a copy of this consent form for my own records.

Signed:________________________________________
Witness:________________________________________
Investigator:____________________________________
Date:__________________________________________
Certificate of Completion

The National Institutes of Health (NIH) Office of Extramural Research certifies that Darrin Reed successfully completed the NIH Web-based training course "Protecting Human Research Participants".

Date of completion: 04/10/2009
Certification Number: 217875
Certificate of Completion

The National Institutes of Health (NIH) Office of Extramural Research certifies that Rob Maher successfully completed the NIH Web-based training course "Protecting Human Research Participants".

Date of completion: 05/13/2009

Certification Number: 229928
APPENDIX B

HEADPHONE COUPLER DRAWINGS AND IMAGES
MACHINING STEPS FOR HEADPHONE COUPLER

Cut segment of bar stock
Centered piece on the rotational stage using dial indicator
Surfaced outside with endmill to obtain proper circumference

Working on top-side
Faced top surface with 1” endmill
Created inner circle with 1” endmill
Beveled inner ring
FLIP OVER

Working on bottom-side
Faced top surface down to desired thickness with 1” endmill
Drilled main hole
Cut outer circle by rotating with the 1” endmill
Cut inner circle with 1” endmill (using side-cut methods)
“Artistically” cut slots for plastic piece of SPL calibrator
Drilled holes for supports
Used tap to create threads
Attached ¼” posts

MISCELLANEOUS NOTES
Take no more than 20th of an inch per pass
Keep cutting tool properly oiled
When “side cutting”, go down first and then slide in to cut.
When tapping threads, rotate the tool in and out to prevent tap from breaking
APPENDIX C

PRE-TEST CALIBRATION DATA
<table>
<thead>
<tr>
<th>DATE</th>
<th>TIME</th>
<th>SUBJ. #</th>
<th>METER SPL</th>
<th>HEADPHONE SPL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Round 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7/1/2009</td>
<td>1:15pm</td>
<td>070109_001</td>
<td>94.2</td>
<td>75.2</td>
</tr>
<tr>
<td>7/2/2009</td>
<td>9:15am</td>
<td>070209_001</td>
<td>94.2</td>
<td>75.2</td>
</tr>
<tr>
<td>7/7/2009</td>
<td>1:40pm</td>
<td>070709_001</td>
<td>94.2</td>
<td>74.8</td>
</tr>
<tr>
<td>7/9/2009</td>
<td>3:20pm</td>
<td>070909_001</td>
<td>94.2</td>
<td>74.8</td>
</tr>
<tr>
<td>7/10/2009</td>
<td>10:05am</td>
<td>071009_001</td>
<td>94.2</td>
<td>74.6</td>
</tr>
<tr>
<td>7/13/2009</td>
<td>1:40pm</td>
<td>071309_001</td>
<td>94.2</td>
<td>74.8</td>
</tr>
<tr>
<td>7/14/2009</td>
<td>2:30pm</td>
<td>071409_001</td>
<td>94.2</td>
<td>74.8</td>
</tr>
<tr>
<td>7/16/2009</td>
<td>9:30am</td>
<td>071609_001</td>
<td>94.2</td>
<td>74.8</td>
</tr>
<tr>
<td>7/22/2009</td>
<td>3:40pm</td>
<td>072209_001</td>
<td>94.2</td>
<td>74.7</td>
</tr>
<tr>
<td>7/22/2009</td>
<td>4:35pm</td>
<td>072209_002</td>
<td>94.2</td>
<td>74.8</td>
</tr>
<tr>
<td>7/23/2009</td>
<td>9:20am</td>
<td>072309_001</td>
<td>94.2</td>
<td>74.8</td>
</tr>
<tr>
<td>7/23/2009</td>
<td>1:25pm</td>
<td>072309_002</td>
<td>94.2</td>
<td>74.6</td>
</tr>
<tr>
<td>7/24/2009</td>
<td>9:00am</td>
<td>072409_001</td>
<td>94.2</td>
<td>74.8</td>
</tr>
<tr>
<td><strong>Round 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9/11/2009</td>
<td>1:50pm</td>
<td>091109_001</td>
<td>94.1</td>
<td>75.2</td>
</tr>
<tr>
<td>9/15/2009</td>
<td>8:40am</td>
<td>091509_001</td>
<td>94.2</td>
<td>75.3</td>
</tr>
<tr>
<td>9/15/2009</td>
<td>9:50am</td>
<td>091509_002</td>
<td>94.2</td>
<td>75.3</td>
</tr>
<tr>
<td>9/17/2009</td>
<td>8:00am</td>
<td>091709_001</td>
<td>94.2</td>
<td>75.0</td>
</tr>
<tr>
<td>9/17/2009</td>
<td>11:00am</td>
<td>091709_002</td>
<td>94.2</td>
<td>75.3</td>
</tr>
<tr>
<td>9/22/2009</td>
<td>8:00pm</td>
<td>092209_001</td>
<td>94.1</td>
<td>75.4</td>
</tr>
<tr>
<td>9/23/2009</td>
<td>11:30am</td>
<td>092309_001</td>
<td>94.1</td>
<td>75.1</td>
</tr>
<tr>
<td>9/24/2009</td>
<td>2:20pm</td>
<td>092409_001</td>
<td>94.1</td>
<td>75.3</td>
</tr>
<tr>
<td>9/29/2009</td>
<td>5:55pm</td>
<td>092909_001</td>
<td>94.2</td>
<td>75.2</td>
</tr>
<tr>
<td>9/29/2009</td>
<td>6:20pm</td>
<td>092909_002</td>
<td>94.2</td>
<td>75.2</td>
</tr>
</tbody>
</table>
APPENDIX D

PHRASES FOR SOUND SOURCE STIMULI
<table>
<thead>
<tr>
<th>offensive as innocence</th>
<th>expense account</th>
</tr>
</thead>
<tbody>
<tr>
<td>calico cat</td>
<td>Straw hats</td>
</tr>
<tr>
<td>Stimulating discussions</td>
<td>sleigh bells</td>
</tr>
<tr>
<td>Regular attendance</td>
<td>subway escalator</td>
</tr>
<tr>
<td>gas shortage</td>
<td>speech recognition</td>
</tr>
<tr>
<td>porch steps</td>
<td>identical twins</td>
</tr>
<tr>
<td>musical instruments</td>
<td>tiny icicles</td>
</tr>
<tr>
<td>Guess the question</td>
<td>cashmere sweater</td>
</tr>
<tr>
<td>Birthday parties</td>
<td>rose corsage</td>
</tr>
<tr>
<td>steep slope</td>
<td>beach balls</td>
</tr>
<tr>
<td>Artificial intelligence</td>
<td>these shoes</td>
</tr>
<tr>
<td>eggs for breakfast</td>
<td>swiss cheese</td>
</tr>
<tr>
<td>Scientific progress</td>
<td>dogs chase cats</td>
</tr>
<tr>
<td>fascinating techniques</td>
<td>suitable gift</td>
</tr>
<tr>
<td>Too much curiosity</td>
<td>speech symposium</td>
</tr>
<tr>
<td>instructions are crucial</td>
<td>pressure changes</td>
</tr>
<tr>
<td>standardized tests</td>
<td>Strong teeth</td>
</tr>
<tr>
<td>algebraic expression</td>
<td>school districts</td>
</tr>
<tr>
<td>ancient ruins</td>
<td>family requests</td>
</tr>
<tr>
<td>outstanding effects</td>
<td>case basis</td>
</tr>
<tr>
<td>snow flurries</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX E

TESTING ROOM SPL MEASUREMENTS
### Bruel & Kjaer
**SLM Type 2236**

**SETTINGS:**

<table>
<thead>
<tr>
<th>F</th>
<th>20-100 dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMS: A</td>
<td>Peak: C</td>
</tr>
</tbody>
</table>

**LEVEL DISTRIBUTION:**

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Elapsed Time</th>
<th>Pauses</th>
<th>Overload</th>
<th>MaxP</th>
<th>Lav5</th>
</tr>
</thead>
<tbody>
<tr>
<td>23 Jun 2009</td>
<td>08:50:21</td>
<td>00:03:43:38</td>
<td>0</td>
<td>0.0 %</td>
<td>103.0 dB</td>
<td>---:-- dB</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level Range</th>
<th>Percentage</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>66.0 - 66.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>65.0 - 65.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>64.0 - 64.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>63.0 - 63.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>62.0 - 62.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>61.0 - 61.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>60.0 - 60.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>59.0 - 59.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>58.0 - 58.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>57.0 - 57.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>56.0 - 56.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>55.0 - 55.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>54.0 - 54.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>53.0 - 53.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>52.0 - 52.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>51.0 - 51.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>50.0 - 50.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>49.0 - 49.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>48.0 - 48.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>47.0 - 47.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>46.0 - 46.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>45.0 - 45.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>44.0 - 44.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>43.0 - 43.9 dB</td>
<td>0.0 %</td>
<td></td>
</tr>
<tr>
<td>42.0 - 42.9 dB</td>
<td>0.1 %</td>
<td></td>
</tr>
<tr>
<td>41.0 - 41.9 dB</td>
<td>0.1 %</td>
<td></td>
</tr>
<tr>
<td>40.0 - 40.9 dB</td>
<td>0.2 %</td>
<td></td>
</tr>
<tr>
<td>39.0 - 39.9 dB</td>
<td>0.2 %</td>
<td></td>
</tr>
<tr>
<td>38.0 - 38.9 dB</td>
<td>0.2 %</td>
<td></td>
</tr>
<tr>
<td>37.0 - 37.9 dB</td>
<td>0.1 %</td>
<td></td>
</tr>
<tr>
<td>36.0 - 36.9 dB</td>
<td>0.2 %</td>
<td></td>
</tr>
<tr>
<td>35.0 - 35.9 dB</td>
<td>0.2 %</td>
<td></td>
</tr>
<tr>
<td>34.0 - 34.9 dB</td>
<td>0.2 %</td>
<td></td>
</tr>
<tr>
<td>33.0 - 33.9 dB</td>
<td>0.2 %</td>
<td></td>
</tr>
<tr>
<td>32.0 - 32.9 dB</td>
<td>0.3 %</td>
<td></td>
</tr>
<tr>
<td>31.0 - 31.9 dB</td>
<td>0.3 %</td>
<td></td>
</tr>
<tr>
<td>30.0 - 30.9 dB</td>
<td>0.4 %</td>
<td></td>
</tr>
<tr>
<td>29.0 - 29.9 dB</td>
<td>0.4 %</td>
<td></td>
</tr>
<tr>
<td>28.0 - 28.9 dB</td>
<td>0.5 %</td>
<td></td>
</tr>
<tr>
<td>27.0 - 27.9 dB</td>
<td>0.7 %</td>
<td></td>
</tr>
<tr>
<td>26.0 - 26.9 dB</td>
<td>1.0 %</td>
<td></td>
</tr>
<tr>
<td>25.0 - 25.9 dB</td>
<td>1.1 %</td>
<td></td>
</tr>
<tr>
<td>24.0 - 24.9 dB</td>
<td>1.1 %</td>
<td></td>
</tr>
<tr>
<td>23.0 - 23.9 dB</td>
<td>1.4 %</td>
<td></td>
</tr>
<tr>
<td>22.0 - 22.9 dB</td>
<td>2.2 %</td>
<td></td>
</tr>
<tr>
<td>21.0 - 21.9 dB</td>
<td>4.3 %</td>
<td></td>
</tr>
<tr>
<td>20.0 - 20.9 dB</td>
<td>26.1 %</td>
<td>Underrange</td>
</tr>
</tbody>
</table>
Brul & Kjaer
SLM Type 2236

SETTINGS:

F 20-100 dB
RMS: A  Peak: C

LEVEL DISTRIBUTION:

25 Jun 2009 09:54:49
Elapsed Time 00:00:44:35
Pauses 0
Overload 0.0 %

MaxP 100.2 dB
Lav5 ---- dB

61.0 - 61.9 dB 0.0 %
60.0 - 60.9 dB 0.0 %
59.0 - 59.9 dB 0.0 %
58.0 - 58.9 dB 0.0 %
57.0 - 57.9 dB 0.0 %
56.0 - 56.9 dB 0.0 %
55.0 - 55.9 dB 0.0 %
54.0 - 54.9 dB 0.0 %
53.0 - 53.9 dB 0.0 %
52.0 - 52.9 dB 0.0 %
51.0 - 51.9 dB 0.0 %
50.0 - 50.9 dB 0.0 %

49.0 - 49.9 dB 0.0 %
48.0 - 48.9 dB 0.0 %
47.0 - 47.9 dB 0.0 %
46.0 - 46.9 dB 0.0 %
45.0 - 45.9 dB 0.0 %
44.0 - 44.9 dB 0.0 %
43.0 - 43.9 dB 0.0 %
42.0 - 42.9 dB 0.0 %
41.0 - 41.9 dB 0.0 %
40.0 - 40.9 dB 0.0 %
39.0 - 39.9 dB 0.0 %
38.0 - 38.9 dB 0.0 %
37.0 - 37.9 dB 0.1 %
36.0 - 36.9 dB 0.1 %
35.0 - 35.9 dB 0.1 %
34.0 - 34.9 dB 0.1 %
33.0 - 33.9 dB 0.1 %
32.0 - 32.9 dB 0.1 %
31.0 - 31.9 dB 0.1 %
30.0 - 30.9 dB 0.2 %
29.0 - 29.9 dB 0.2 %
28.0 - 28.9 dB 0.3 %
27.0 - 27.9 dB 0.5 %
26.0 - 26.9 dB 0.8 %
25.0 - 25.9 dB 1.0 %
24.0 - 24.9 dB 1.6 %
23.0 - 23.9 dB 2.8 %
22.0 - 22.9 dB 4.8 %
21.0 - 21.9 dB 8.3 %
20.0 - 20.9 dB 30.9 %

Underrange 47.7 %
APPENDIX F

POST-EXPERIMENT QUESTIONNAIRE
CLOSING QUESTIONS

Spatial Auditory Perception Evaluation

1) Have you ever participated in any other audio/psychoacoustic experiments? If yes, please describe.

2) Was the loudness of the sound samples lower/equivalent/louder than your normal listening levels for headphone audio?

3) Did you tend to preferentially select front or back if you were undecided on the location of the sound? If so, which one?

4) Did you feel your ability to hear the front/back and reflections improved after listening to more samples?

5) Was the length of the test satisfactory, or did you find yourself having attention lapses?

6) Was there anything confusing or awkward about the experimental procedure or the instructions? If so, please explain.

7) Would you be interested in participating in similar experiments in the future? If so, may we contact you regarding future experiments (please provide preferred method(s) of contact)?