8-2015

Effects of amplitude modulation on sound localization in reverberant environments.

Paul W. Anderson
University of Louisville

www.manaraa.com

EFFECTS OF AMPLITUDE MODULATION ON SOUND LOCALIZATION IN
REVERBERANT ENVIRONMENTS

By

Paul W. Anderson
B.S., Georgia College & State University, 2010
M.S., University of Louisville, 2012

A Dissertation
Submitted to the Faculty of the
College of Arts and Sciences of the University of Louisville
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy in Experimental Psychology

Department of Psychological and Brain Sciences
University of Louisville
Louisville, KY

August 2015

EFFECTS OF AMPLITUDE MODULATION ON SOUND LOCALIZATION IN
REVERBERANT ENVIRONMENTS

By

Paul W. Anderson
B.S., Georgia College & State University, 2010
M.S., University of Louisville, 2012

A Dissertation Approved on

May 27, 2015

by the following Dissertation Committee:

_____
Dissertation Director
Dr. Pavel Zahorik

_____
Dr. Jeffrey Weihing

_____
Dr. Christian Stilp

_____
Dr. Paul DeMarco

_____
Dr. Zijiang He

DEDICATION

This dissertation is dedicated to my mom, my dad, and sister Laura

for their constant love and support.

ACKNOWLEDGEMENTS

It is with great pleasure that I acknowledge and thank the following individuals for their contributions to this dissertation. I would first like to thank my mentor Pavel Zahorik for his patience and unwavering guidance throughout my time in Louisville. I would also like to thank my dissertation committee, Dr. Pavel Zahorik, Dr. Paul DeMarco, Dr. Zijiang He, Dr. Christian Stilp, and Dr. Jeffrey Weihing for their time and feedback. I also wish to thank my fellow lab mates, Dr. Eugene Brandewie, Greg Ellis, and Dr. Amanda O'Bryan, for their support, encouragement, and friendship. A special thanks to the listeners who participated in these experiments for their time and attention.

ABSTRACT

EFFECTS OF AMPLITUDE MODULATION ON SOUND LOCALIZATION IN
REVERBERANT ENVIRONMENTS

Paul W. Anderson

May 27, 2015

Auditory localization involves different cues depending on the spatial domain. Azimuth

localization cues include interaural time differences (ITDs), interaural level differences

(ILDs) and pinnae cues. Auditory distance perception (ADP) cues include intensity,

spectral cues, binaural cues, and the direct-to-reverberant energy ratio (D/R). While D/R

has been established as a primary ADP cue, it is unlikely that it is directly encoded in the

auditory system because it can be difficult to extract from ongoing signals. It is also

noteworthy that no neuronal population has been identified that specifically codes D/R. It

has therefore been proposed that D/R is indirectly encoded in the auditory system,

through sensitivity to other acoustic parameters that are correlated with D/R, such as

temporal cues (Zahorik, 2002b), spectral properties (Jetzt, 1979; Larsen, 2008), and

interaural correlation (Bronkhorst and Houtgast, 1999). An additional D/R correlate relies

on attenuation of amplitude modulation (AM) as a function of distance. Room

modulation transfer functions act as low-pass filters on AM signals, and therefore the

direct portion of a signal will have less modulation depth attenuation than the reverberant

portion. Although recent neural and behavioral work has demonstrated that this cue can

provide distance information monaurally, the extent to which the modulation attenuation

cue contributes to ADP relative to other ADP cues is not fully understood. It is also possible modulation attenuation by the room can provide additional directional localization information, perhaps through the resulting dynamic fluctuation of the ILD cue. The role of AM in directional sound localization has not been extensively studied, particularly in reverberant soundfields which can affect the modulation reaching the two ears in a directionally-dependent fashion. Three human psychophysical experiments assessed the role of AM signals in distance and directional auditory localization in reverberant soundfields. Experiment I focused on validating a graphical response method to be used in subsequent experiments. In Experiment II, an auditory distance estimation task was performed which yielded measures of the relative perceptual contributions of the modulation depth cue during ADP in a reverberant room. Experiment III investigated the effect of AM on binaural localization in the horizontal plane in a reverberant room.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

BACKGROUND

A. Distance Perception

Auditory distance perception (ADP) is known to be relatively inaccurate and highly variable (Zahorik, Brungart, and Bronkhorst, 2005; Anderson and Zahorik, 2014). It is thought to be governed by four acoustic cues: Intensity, spectral cues, binaural cues and the ratio of direct-to-reverberant sound energy (D/R). Binaural distance cues, like ITD and ILD, are most useful in the near field (within about 1 m of the head), while spectral distance cues are primarily effective for far distances where high frequency content is absorbed by the air between the source and the listener. D/R is a ratio of the amount of energy in the direct portion of the waveform (energy that is transmitted directly from the source to the listener without interacting with any surfaces of the environment) to the amount of reverberant energy (energy that interacts with surfaces in the environment) of the waveform that reaches the ears. As sound source increases the amount of direct energy decreases relative to the amount of reverberant energy which remains relatively constant as a function of distance. D/R information is important for making absolute distance judgments, and therefore a prerequisite for utilizing D/R is that the source signal be presented in a reverberant environment (Mershon and King, 1975; Bronkhorst and Houtgast, 1999). In anechoic space, distance estimation is dominated by intensity since no reverberant energy is available. Without D/R, only relative distance judgments can be made in anechoic space (Mershon and Bowers, 1979). The manner in

1

which D/R is encoded by the brain is uncertain, since no neuronal population has been found that directly codes it (For review see Kim, Zahorik, Carney, Bishop, & Kuwada, 2015). To address this issue, a number of theories have attempted to explain how the auditory system might code D/R via other correlated acoustical cues, including temporal cues (Zahorik, 2002b), spectral properties (Jetzt, 1979; Larsen, 2008), and interaural correlation (Bronkhorst and Houtgast, 1999). Our lab, in collaboration with colleagues at University of Connecticut, has proposed a new correlate of D/R specific to amplitude modulated (AM) sound sources.

Kim et al. (2015) recorded from AM sensitive neurons in the inferior colliculus (IC) of rabbits. Sounds were presented monaurally to the rabbits using virtual auditory space (VAS) techniques at varying sound source distances in both anechoic and reverberant environments. Stimulus level was equalized across distance to compensate for propagation loss by 6 dB per doubling of source distance. As modulation depth increased some of the AM sensitive neurons would increase their firing rate and others would decrease their firing rate. The AM sensitive neurons that decreased firing rates in response to decreased modulation depth also decreased their firing rates monotonically with source distance only if two conditions were true: 1) the sound was amplitude modulated and 2) the stimulus was presented in a reverberant environment. This supports the hypothesis that modulation depth attenuation as a function of distance can be used to code distance in IC neurons. Kim et al. also reported results from a behavioral experiment they performed where human listeners performed a distance judgment task under two conditions: modulated or unmodulated stimuli. As in the neural portion of their study, the stimuli were equalized for level. The human behavioral results where listeners performed

2

egocentric distance judgments showed that listeners could only discriminate source distance if the signal was amplitude modulated in a reverberant environment. For modulated sound sources perceived distance increased with physical distance, but for unmodulated signals perceived distance was independent of physical distance. The combination of these two studies provides evidence that AM can be used as a cue to distance in reverberant environments, and there is a neural basis for the cue.

Behavioral data from Zahorik & Anderson (2015) found more evidence that listeners can use AM as a monaural distance cue. Listeners performed an auditory distance perception task where stimuli were either modulated or unmodulated in both a reverberant and an anechoic environment. Headphone presentation was either monaural (to the contralateral ear) or binaural. Stimulus level was equalized across distance like in Kim et al. (2015). Distance judgments were only analyzed for sources less than 2 m from the listener. The results indicated that distance judgments were more accurate in a reverberant environment when stimuli were AM. This benefit from AM only occurred when stimuli were presented monaurally to the contralateral ear.   The basis for the hypothesis of using amplitude modulation as a distance cue comes from the modulation transfer function (MTF) of a room acting as a low-pass filter in the amplitude modulation domain (Houtgast and Steeneken, 1985), where higher modulation frequencies are progressively more attenuated with increasing physical distance. MTFs are computed so that they are independent of level. Attenuation of the modulation depth in the reverberant portion of the signal increases with source distance. The modulation depth of the direct energy portion, however, remains unattenuated with increasing source distance. Therefore, the direct portion of an AM signal will have modulation depth attenuated less

3

than the reverberant portion of the signal. The difference between the AM depth of the direct portion of the waveform and the reverberant portion of the waveform may then be used as a correlate of D/R.

A common AM signal is speech. Zahorik (2002a) investigated distance perception of speech stimuli and found no significant difference in distance estimates compared to noise stimuli. Distance judgments were also made at both 0° and 90°. No significant difference was found between the two orientations. However, the speech signal used was a single syllable, /da/, and thus contained little AM. Nielsen (1991) compared distance estimates of white noise and 5 s of speech stimuli (a female talker reading a sentence from a short story recorded in an anechoic space) finding no significant difference between the two. Any benefit from amplitude modulation on distance judgments may have been over powered by more reliable distance cues. This will be discussed in more detail below. Nielsen also collected distance judgments at azimuths other than midline using speech stimuli. Sources located off midline (45°, 90°, and 180°) were perceived as farther away than sources at 0°. However, four target distances were tested at 0° and 45°, and only one target distance was tested at 90° and 180°. When only one target source was present intensity was manipulated to create the percept of more sources. Positioning more sources along each azimuth and using a more reliably AM signal (like noise) would help elucidate results from these two studies.

Modulation depth is predicted to vary with distance as a function of azimuth. The description above of how modulation is attenuated in the direct and reverberant portions of the waveform is specific to midline where the signal reaching both ears is identical. When the source is moved away from midline the contralateral ear will have

4

more modulation attenuated by the room than the ipsilateral ear since the contralateral ear has more room exposure. At 90° the difference in modulation depth between ears will be greatest for a given distance. The direct portion will have similar AM depth regardless of arriving at the ipsilateral or contralateral ear; however, the AM depth of the reverberant portion will be more attenuated under contralateral stimulation since the far ear will have more exposure to the room. Figure 1.1 illustrates the difference of modulation depth of the reverberant portion of the signal as a function of distance at 90°. The left column shows modulation depth as a function of distance in a highly reverberant room, and the right column shows the same relationship in anechoic space. The figure demonstrates that modulation depth is attenuated more at greater distances in a room. The left column shows modulation depth at the contralateral ear (green) is more attenuated than the ipsilateral ear (blue) in a room. However, the right column shows that in anechoic space AM depth remains unchanged as a function of distance for both the ipsilateral and contralateral ear. Modulation depth changes similarly to how D/R changes as a function of distance. It is difficult to extend this relationship between modulation depth and source distance to the near field because the relationship between binaural cues (ITDs and ILDs) and physical distance is complex when a sound source is within approximately 1 m of the head, especially at more lateral azimuths (Zahorik, Brungart, & Bronkhorst, 2005).
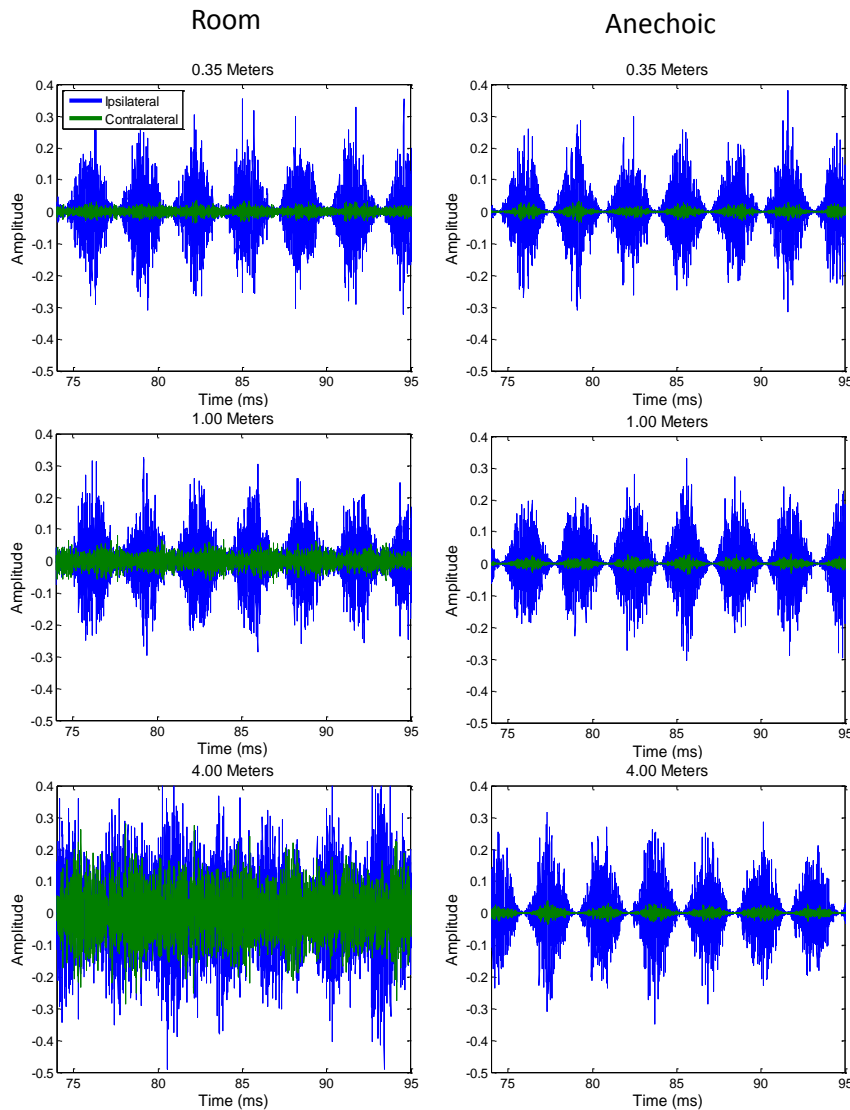
Figure 1.1: Example time waveforms that show the effect of distance on amplitude modulation at 90° azimuth under reverberant (left) and anechoic (right) conditions. The source signal is a sinusoidally amplitude-modulated 1-octave band of noise centered at 4 kHz, with modulation frequency of 32 Hz. Sound intensity was boosted by +6 dB of gain per doubling of distance to compensate for propagation loss with distance using 1.4 m as the reference distance. In reverberation, AM is attenuated as distance increases. Under anechoic conditions AM depth is relatively constant across distance. AM depth at the contralateral ear (green) in the room is more attenuated than the ipsilateral ear (blue).

6

Auditory distance perception depends upon at least four cues: D/R, intensity, binaural cues, and spectral change. To form a stable distance estimate to a sound source, these cues must be combined by the perceptual processes underlying auditory distance perception. Zahorik (2002a) performed a study investigating how D/R and intensity were perceptually weighted during distance judgments under different stimulus conditions. Small amounts of independent random perturbation were applied to each parameter during stimulus presentation. Perceptual weights can then be estimated using a multiple regression model where the perturbed parameters are used to predict distance responses. It was found that perceptual weights assigned to D/R and intensity change as a function of source signal type and source direction. This approach of measuring perceptual weightings of auditory cues would be able to measure the relative contribution of modulation depth to distance judgments by perturbing modulation depth the same way Zahorik (2002a) perturbed intensity and D/R. While D/R is a cue specific to localization in the distance domain, differences of modulation depth as a function of azimuth may also aid directional localization ability in rooms.

B. Directional Perception

The primary directional localization cues are interaural level (ILDs) and interaural timing (ITDs) cues. According to Duplex theory these two cues derive from different frequency regions (Strutt, 1907). ILDs result from the head shadow effect where the signal reaching the contralateral ear is more attenuated than the signal reaching the ipsilateral ear. Head shadow is most effective for creating ILDs at frequencies higher than about 1000 Hz. For frequencies below about 1000 Hz ITDs can be used for localization where the signal arrives at the closer ear before the signal reaching the far ear

7

(Middlebrooks & Green, 1991). Directional sound localization in anechoic environments is generally known to be quite accurate: Minimum audible angles can be as small as 1° at midline for horizontally displaced sounds (Mills, 1958). Reverberation, on the other hand, has a complex influence on localization accuracy. While reliable distance perception requires the presence of a room, many directional localization studies have found rooms are detrimental to localization accuracy (Hartmann, 1983; Giguere and Abel, 1993; Ihlefeld and Shinn-Cunningham, 2011). Although localization is generally impacted negatively in rooms, there are mechanisms, such as the precedence effect (Wallach, Newman, and Rosenweig, 1949) which limit the detrimental effect of reverberation by placing importance on the timing of the arrival of the first wave front to reach the ears. Other studies have shown that localization ability in reverberant environments is equal or even improved relative to anechoic for certain types of stimuli, like high frequency noise (Begault, Wenzel, and Anderson, 2001; Ihlefeld and Shinn-Cunningham, 2011) and depending on the amount of exposure a listener has to the room (Shinn-Cunningham, 2001). The stimuli in all of these studies were unmodulated noise.

C. Directional Localization of AM Stimuli

Investigation of localization of AM stimuli has mostly focused on lateralization tasks or anechoic environments. Lateralization of AM tones (50 - 800 Hz mod rate; Bernstein and Trahiotis, 1985b) and noise (Trahiotis and Bernstein, 1986) have been investigated using a pointing task in which listeners adjusted the interaural level difference (ILD) of an auditory pointer to match the lateralization of a target tone. They found that lateral position could be coded by the ITD resulting from interaural differences in the amplitude envelopes. This effect was greatest for high-frequency tonal carrier

8

signals (2-4 kHz; Bernstein and Trahiotis, 1985) and for wide-band noise (Trahiotis and Bernstein, 1986). The auditory system is also sensitive to changes in dynamic ILDs in lateralization, especially at higher carrier frequencies (Grantham, 1984). While results from these studies indicate the importance of envelope ITDs and dynamic ILD cues on lateralization, it is not clear what the impact of AM would be on localization when pinna cues are present and stimuli are presented in a real environment with reverberation.

Directional localization of speech, which can be considered an AM signal, has been studied, and localization errors with speech are similar to those with noise stimuli (Begault and Wenzel, 1993). Begault and Wenzel used spoken speech samples that were one or two syllable words 0.7- 1.3 s in duration. A complication with speech localization is that the tested speech signals are often low-pass filtered such that they do not contain frequencies in the range of 8-16 kHz, which have been shown to be critically important for localization of speech compared to broadband noise (Best, Carlile, Jin, and van Schaik, 2005). Eberle, McAnally, Martin, and Flanagan (2000) investigated localization ability using more reliable AM stimuli in an anechoic environment at modulation frequencies of 20, 80, and 320 Hz. They found localization was more accurate only when the signal was modulated at the highest frequency; however, they believe the effect was due to the broadening of the signal's spectral energy from side bands produced by modulation. Wagenaars (1990) performed a localization experiment with sinusoidal stimuli in a room at very high modulation rates (500 and 2000 Hz) and found that the sinusoidal stimuli could be localized well if there was an abrupt onset or offset which is similar to how the precedence effect functions (Wallach, Newman, and Rosenweig, 1949).

Because these studies on localization of AM stimuli were all performed in anechoic space, the effect of reverberation on localization of AM stimuli is currently unknown. It is possible that reverberation may provide additional information which could aid localization. Figure 1.2 shows how AM depth is attenuated as a function of azimuth in both a highly reverberant room (right column) and in anechoic space (left column). In the room, AM at the contralateral ear (green) is more attenuated as the source is moved away from midline while the AM depth at the ipsilateral ear (blue) remains relatively constant. This causes the interaural difference in modulation depth to increase as the source moves away from midline. In anechoic space the AM depth at both ears is constant across all azimuths. Since the modulation depth reaching each ear would differ in reverberation, and the amplitude of the signal would be fluctuating, this may cause dynamic ILD cues to be used as well. While localization is already highly accurate, it is uncertain whether the addition of modulation will provide any benefit to localization ability.
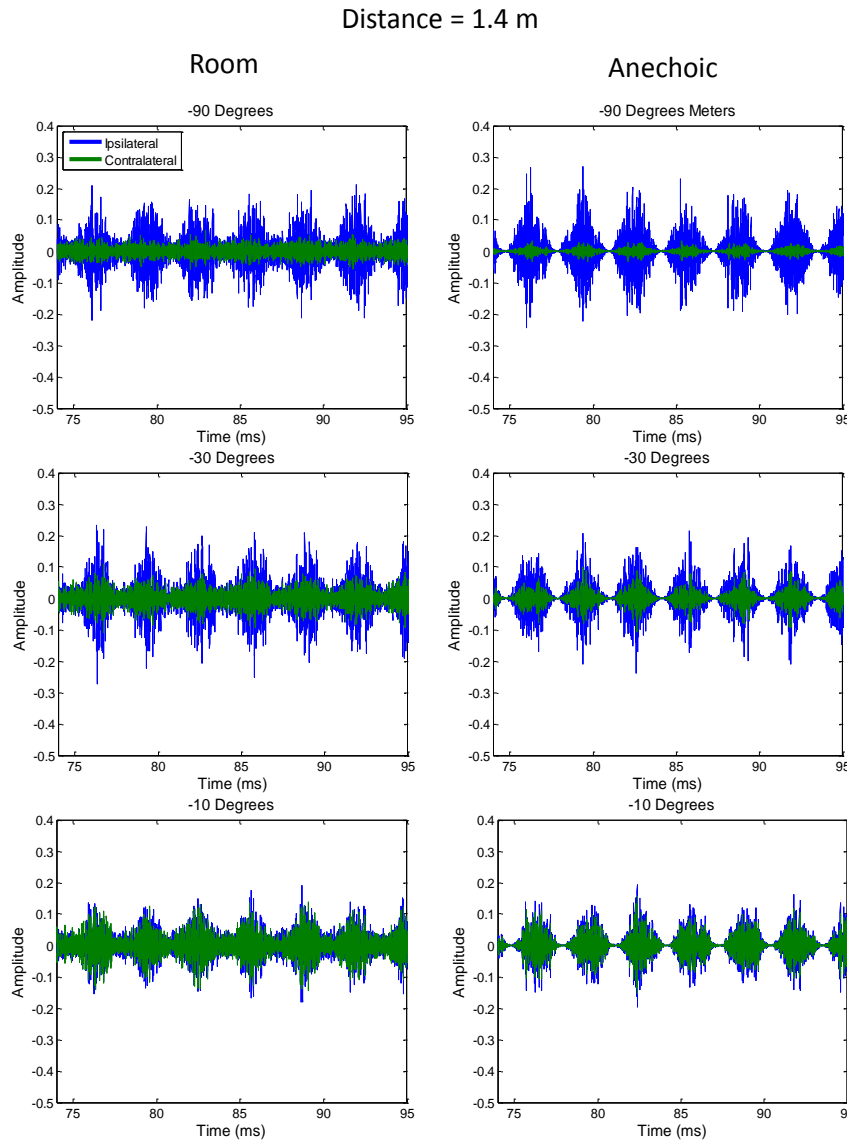
Figure 1.2: Example time waveforms that show the effect of azimuth on amplitude modulation at 1.4 m under reverberant (left) and anechoic (right) conditions. The same waveform parameters from figure 1 were used here. The amplitude modulation depth of the contralateral signal (green) in the room is progressively less attenuated as the source is moved toward midline while the modulation depth of the contralateral signal (blue) maintains approximately unchanged. In anechoic space the AM depth of the contralateral and ipsilateral (blue) signals remains relatively unchanged as a function of azimuth and no difference in AM depth between ears is noticeable.

11

D. Response Methods in Distance and Directional Localization Studies

Auditory distance perception and directional localization studies use a variety of response methods. While some studies employed forced-choice procedures (Hartmann, 1983) for localization, where speakers are visible during the task, this is not ideal because of the visual bias and restricted response space inherent in the design (Perrett and Noble, 1995). Studies using verbal reports for both direction (Wightman and Kistler, 1989b) and distance (Anderson and Zahorik, 2011; Anderson and Zahorik, 2014) provide reliable responses from participants when no visual cues are available, but the subjects in certain cases must learn a coordinate system for responding, and the responses are often highly variable. In an effort to streamline data collection in the current study, the use of a response method similar to Nielsen (1991) was considered. Nielsen employed a polar plot graphical user interface (GUI) that allowed the listener to respond using the cursor by clicking inside of a circle that represents a two dimensional top-down view of the listener and surrounding space. Distance from the center of the circle represented egocentric distance while the azimuth around the center represented directional of the sound source. This method allows for simultaneous directional and distance responses. Nielsen found response times averaged around 5.8 seconds per trial, which is shorter than observed for distance judgments in Anderson and Zahorik (2014). Shinn Cunningham, Santarelli, and Kopco (2000) used a similar polar response GUI to collect distance judgments from listeners without a noticeable effect on the results. Ihlefeld and Shinn-Cunningham (2011) had success using a similar response method with a GUI of a frontal hemi-field rather than the full 360° space surrounding the listener. A similar, but more complex, GUI method was used by Begault, Wenzel, and Anderson (2001) which allowed listeners

12

to provide localization estimates that included azimuth, distance, and elevation estimates. In addition to the response time benefit, judgments may also be more accurate since listeners would be able to spatially view their judgments before responding.

E. Goals of Present Research

The present investigation studied the effect of amplitude modulation in reverberant environments on distance and directional localization in the horizontal plane. The purpose was to 1) Find a response method that requires less response time while obtaining more accurate localization judgments, 2) investigate a recently identified auditory distance cue and compare its perceptual salience to established distance cues, and 3) determine whether the auditory system can exploit characteristics of degraded signals in reverberant environments to improve directional localization. Three experiments were conducted. The first experiment, involving validation of the polar plot GUI, informed how subsequent experiments were designed. Experiment II investigated the contribution of modulation depth during distance judgments. Experiment III investigated whether directional localization estimates are influenced by modulation depth changes in a room.

13

CHAPTER II

EXPERIMENT I

The graphical response method for collecting localization judgments is
potentially more desirable than direct estimates because data collection can be performed
more rapidly, potentially with less response variability, and does not require the
participant to learn a coordinate system for responding. For these reasons, a graphical
response method using a polar plot of space surrounding the listener was developed that
was similar to those used by Shinn-Cunningham, Santarelli, and Kopco (2000) and
Nielsen (1991). Because these studies did not provide data collected with other response
methods, it is not possible to validate the methods. Experiment I therefore sought to
validate the polar plot response method by comparing results to a traditional direct
estimate method in which listeners directly estimated sound source location using an
internally represented coordinate system.

A. Methods

1. Participants

In this experiment, 16 (15 female) listeners participated ranging in age from 18 to
34 ($M = 22.00$). Listeners were recruited through University of Louisville's subject pool
and received course credit for their participation. All listeners had normal hearing based
on self-reports because the amount of time available for running participants was limited
due to subject pool. Informed consent was obtained from all listeners prior to data

14

collection. All procedures involving human subject participants were approved by the University of Louisville Institutional Review Board (IRB).

2. Stimuli

Materials. Binaural room impulse responses (BRIRs) were measured in a large lecture hall (Bigelow Hall, University of Louisville). The hall was the shape of a trapezoidal box with a raised stage at one end of the room. All BRIR measurements were recorded on the floor in front of the stage with chairs removed. The hall's total volume was approximately 11074 m^3 (L = 14.0208; H = 5.6388; A = 25.908 x H; B = 23.7744 * H; L x H x (A + B)/2). The broadband reverberation time (T60) was 2.3 s (ISO-3382, 1997). BRIRs were measured using a KEMAR mannequin (G.R.A.S. Type 45BM) with IEC711 ear-canal simulators (G.R.A.S. RA0045) and large pinnae (G.R.A.S. KB1060/1) positioned at a fixed location in the hall. The sound source was a high quality 2-way co-axial loudspeaker (Bag End PM-6). The loudspeaker was moved away from KEMAR toward the stage to manipulate distance. To manipulate azimuth KEMAR was rotated on a turntable.

A total of 11 BRIR measurement locations were used for Experiment I. BRIR measurement locations for distance were made 0.3, 1.22, and 4.88 m from the measurement microphone at 0°. For azimuth locations BRIR measurements were recorded .3 m and 4.88 m from the measurement microphone at 0, -30, -60, -90, and -150° azimuths. BRIRs were estimated using Maximum Length Sequence (MLS) system identification techniques (Rife & Vanderkooy, 1989). The MLS signal was 2.73 s in duration (17th order MLS), sampled at 48000 kHz with 24-bit resolution. 16 repetitions of this signal were presented and averaged to improve signal-to-noise ratio (SNR), which

15

was 47dB (25 Hz – 12.5 kHz) at 1 m after averaging. All BRIR measurements were post-processed to compensate for response characteristics of the measurement loudspeaker and the presentation headphones (Beyerdynamic DT-990 Pro). The source signal for virtual synthesis was a 1 s sample of Gaussian shaped wide-band noise.

3. Design

Listeners were tested in a within-subjects design using two response methods. In the direct estimate condition listeners responded by using a computer keyboard to input both a distance judgment and azimuth judgment for each stimulus. Figure 2.1 shows a screenshot of the direct estimate graphical user interface (GUI). The GUI has two dialog boxes, one for the distance response and one for the azimuth response. A diagram below the text boxes reminded listeners that azimuths to the left were negative and to the right were positive. The second response method, the polar plot GUI, required listeners to use only the computer mouse to make their response. This method used a GUI in which a two-dimensional top down view of the space surrounding the listener is displayed (see figure 2.2). As the mouse moved around the figure the current location (distance and azimuth) of the cursor was displayed at the top of the GUI. Concentric circles represented the distance from the center, while lines radiating from the center represented azimuth. A zoomed-in version of the GUI is shown in Figure 2.3. When the user responded, a red 'x' appeared at the chosen location and he/she could click 'confirm' to enter the response. The 'play' button allowed listeners to listen to the stimulus as many times as desired during a trial. The polar plot GUI was made up of a circle with concentric circles at 1.52, 3.05 and 4.57 m and with lines radiating from the center starting at 0° and increasing in 30° increments. The distance units displayed in the GUI depended on whether the listener

had selected feet or meters as their preferred measure. At the center of the circle was a drawing representing the listener's head facing forward toward 0°. Azimuths in the left hemisphere were labeled with negative degrees while azimuths to the right were positive. Above the circle the current location of the mouse inside the circle was displayed, both in azimuth and distance from the center. The listener could use the scroll wheel on the mouse to zoom in on the polar plot to make more accurate judgments if they chose (see figure 2.3). Both response conditions were self-paced, so the listener would press 'Play' to start the trial and press 'Confirm' to enter the response and move to the next trial
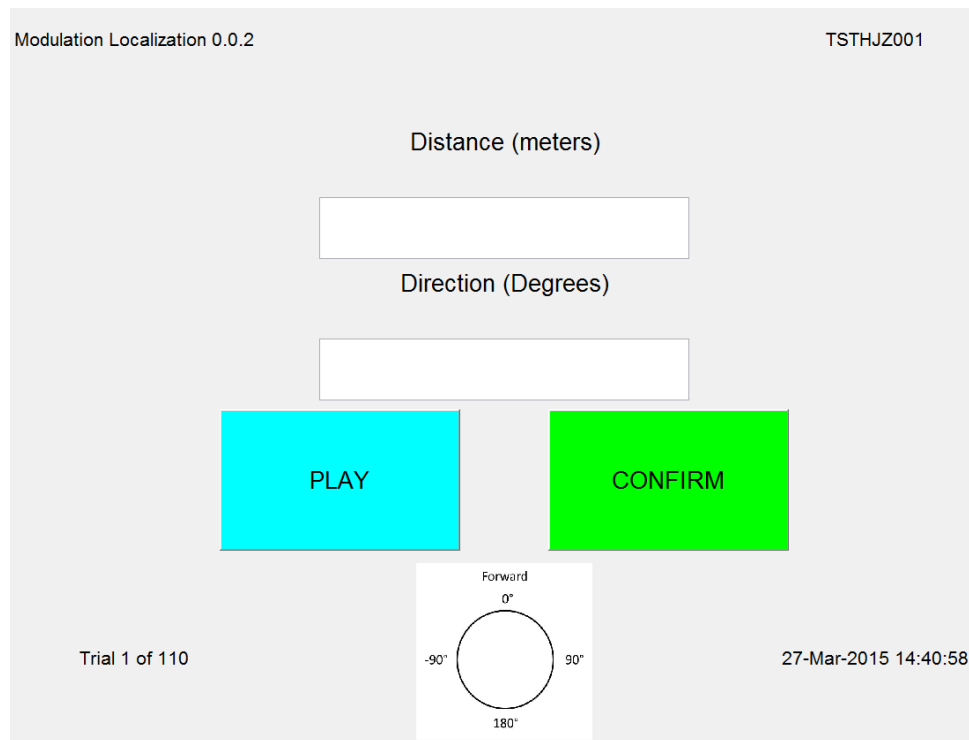
Figure 2.1: Direct estimate GUI. Listeners responded with distance judgments using the 'Distance' text box and azimuth judgments in the 'Direction' text box. Listeners clicked 'Play' to listen to the stimulus and 'Confirm' to enter a response. A diagram at the bottom reminded the listener that judgments to the left were negative, judgments to the right were positive, and directly behind the listener was 180°.
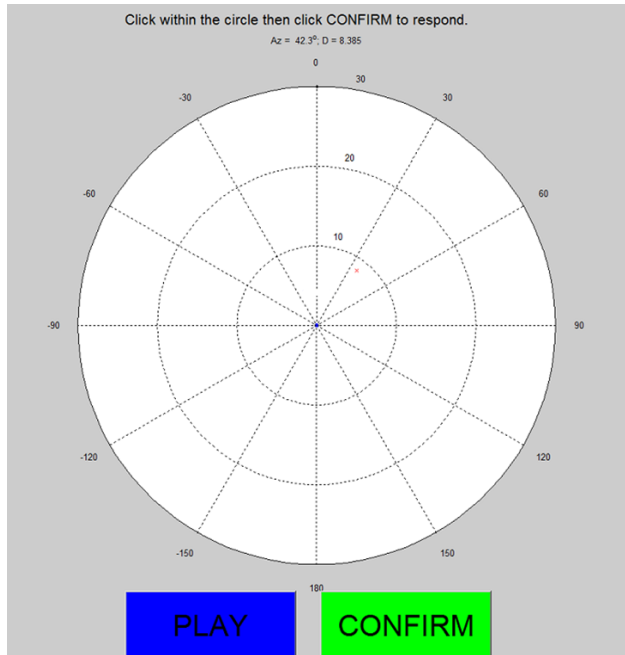
Figure 2.2: Proposed polar plot GUI. In the center of the polar plot lies a circle representing the listener's head facing 0°. As the mouse moves around the figure to click a location the azimuth and distance of the current location of the mouse was displayed at the top of the GUI. Concentric circles represented the distance from the center, while line radiating from the center represented azimuth. The scroll wheel on the mouse allowed the user to zoom in or out to adjust the resolution of their response. When the user made a response a red 'x' appeared at their chosen location and they can clicked 'confirm' to enter their response. The 'play' button allowed listeners to listen to the stimulus multiple times during a trial.
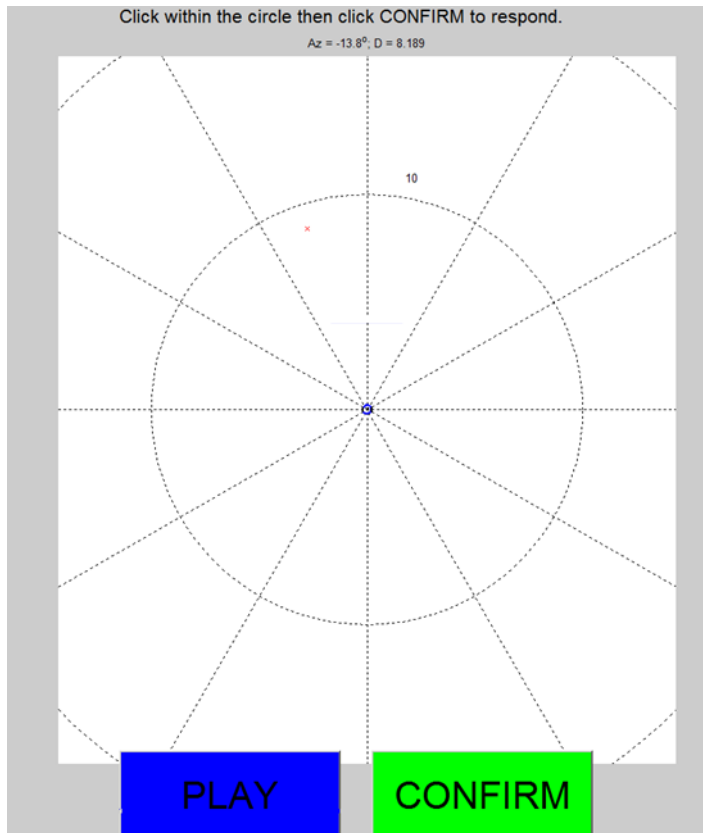
Figure 2.3: Zoomed-in version of figure 2.2.

4. Procedure

Listeners first responded to a GUI asking for their preferred units of

measurements (either feet or meters) and the response GUIs were updated to reflect the

listener's choice. Stimulus location varied in both azimuth and distance, and listeners

were asked to simultaneously respond in both dimensions. Response method was

blocked, and the order of which response method came first was counterbalanced

between listeners. Each response condition had 110 trials (3 distances at 0° x 10 reps + 4

azimuths at 0.3 m x 10 reps + 4 azimuths at 4.88 m x 10 reps). Trial order was pseudo-

randomized for each condition prior to running participants (each condition had a

different order of trials) and the order of trials was kept the same for all participants. For

the polar response condition listeners were instructed to click where the sound source was

19

located within the circle and were shown that they could zoom-in on the GUI. The direct estimate condition required additional instructions for the listener. The listener was instructed to only respond with a distance of 'zero' if the sound was located inside-the-head (Blauert, 1997, p. 132), and to use two decimal-point precision for distance judgments to allow for higher resolution. Both response methods had positive degrees in the right hemisphere and negative degrees to the left. The experiment took place inside a custom-double walled sound attenuating booth. No feedback was provided to the listener. Stimulus presentation and data collection were carried out using custom MATLAB (Mathworks Inc., Natick, MA) software.

   Data Analysis. Distance and directional data were analyzed separately. Distance analyses were performed only on stimulus locations at 0° azimuth because the other azimuths off midline only had sources located at two distances (near and far). The geometric means of distance judgments were fit to power functions of the form $\hat{y}_r = k\Phi_r^a$ ($\hat{y}_r$ = perceived distance, k = constant, a = power-law exponent, Φ_r = target source distance) for each condition, as used in previous studies (Anderson and Zahorik, 2014; Zahorik, Brungart, & Bronkhorst, 2005). Fit parameters, a and k, were used to describe the amount of linear and non-linear compression, respectively. R^2, derived from the power function fit, was used as a measure of within-subject variability. For azimuth data, unsigned angular errors at each target azimuth were computed after front-back reversals were recorded and resolved as performed in Wightman & Kistler (1989b). Angular error was defined as the unsigned error between the judged vector and the vector from the origin (the listener) to the target position. Additionally, the amount of time to complete each condition was recorded. Analyses were used to compare how localization judgments

20

varied between the two response methods. Fit parameters, R^2 values, angular errors, and time to complete were compared between conditions using paired *t*-tests. Independent *t*-tests were used to analyze order effects for constants, exponents, R^2 values, and angular errors between conditions. All analyses were performed using custom MATLAB software.

B. Results

1. Distance Analyses

Distance and azimuth analyses were performed separately. For distance analyses, only judgments to target positions located on the front midline were analyzed. Distance judgments from the polar plot GUI were measured as the distance of the response from the origin of the circle. Before any distance analyses were performed all distance judgments were converted to meters and log transformed. Figure 2.4 displays distance judgments of representative listener (Subject ID: QES) for the polar plot (left panel) and direct estimate (right panel) conditions. Data from each condition were fit with a power function ($\hat{y}$; solid line) of the form described above. Dots represent raw distance judgments (y): 10 replications/target distance. Open circles indicate geometric means ($\overline{y}$) for each target distance. The dashed line represents a perfectly accurate relationship between judged and target distance (i.e., $a = 1$, $k = 1$). Each panel includes the R^2 value, exponent, and constant values.
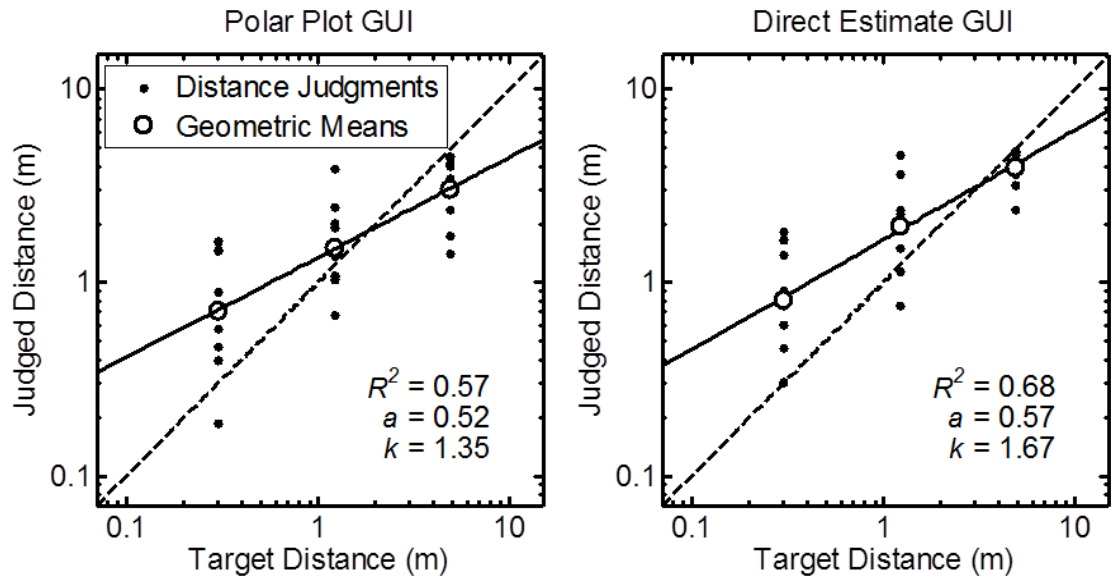
Figure 2.4: Data from a single representative listener (Subject ID: QES) for the Polar Plot GUI (left), and direct estimate GUI (right) conditions plotted on logarithmic axes. Dots show raw distance judgments (y): 10 replications/target distance. Open circles indicate geometric means ($\overline{y}$) for each target distance. Data from each condition were fit with a power function ($\hat{y}$; solid line) of the form $\hat{y}_r = k\Phi_r^a$ ($\hat{y}_r$ = perceived distance, k = constant, a = power-law exponent, Φ_r = target source distance). Fit parameters and the proportion of variability accounted for by the fit (R^2) are shown in each panel. Perfectly accurate performance is indicated by the dotted line in each panel.

The R^2 values from the power function fits for the polar plot ($M = 0.516$, $SD = 0.099$) and direct estimate ($M = .625$, $SD = 0.251$) were compared across listeners using paired t-tests. There was no significant difference between R^2 values in the two conditions, $t(15) = -1.648$, $p = 0.120$. The high R^2 values in each condition indicate that the power functions were good fits to the data.

22

Constant values, which measure the amount of linear compression of distance

judgments, were compared between conditions using a Wilcoxon Matched-Pairs test

because the constant values in the Direct Estimate condition were positively skewed.

There was no significant difference between the constants in the Polar Plot (Mdn = 1.605,

IQR = 1.248 - 2.068) and Direct Estimate (Mdn = 1.661, IQR = 1.214 - 2.373) conditions,

Z = -0.879, p = 0.379. This suggests that the amount of linear compression was not

influenced by response method type.

Exponents, based on power functions fit to distance judgments in both conditions

were compared using paired t-tests. The exponents in the polar plot response (M = 0.468,

SD = 0.162) method were not significantly different from those in the direct estimate (M

= 0.468, SD = 0.191) method, t(15) = -0.005, p = 0.996. The response type did not affect

amount of non-linear compression.

Range effects were a concern for distance judgments because the polar plot GUI

had an inherently restricted range within which listeners could respond. Additionally,

listeners may quantize azimuth responses in the polar plot GUI along the dotted lines

radiating from the center in 30° increments. The distribution of distance and azimuth

responses were pooled across listeners and plotted to visualize possible range effects in

the two GUIs. Figure 2.5a-b displays distributions of log transformed judged distances, in

meters, pooled across listeners in the polar plot GUI and direct estimate GUI

respectively. Each plot includes the mean and standard deviation of the distribution. The

distribution of the Polar Plot GUI responses is negatively skewed with responses stopping

abruptly at approximately 4 m which closely aligns with the maximum response of 4.572

m (0.6601 $\log_{10}$(m)) allowed by the polar response GUI. Comparatively, the distribution

23

of responses from the direct estimate GUI more closely approximates a Gaussian distribution with many responses exceeding the maximum response limit in the polar plot GUI. The standard deviation of the distribution of judgments from the Direct estimate GUI is larger than the standard deviation of responses from the polar plot GUI condition. It is important to allow for variability of listeners' distance judgments given that large variability of responses is an inherent characteristic of auditory distance perception (Zahorik et al, 2005; Anderson & Zahorik, 2014). By restricting the distance judgment range the amount of variability of distance judgments is artificially reduced.
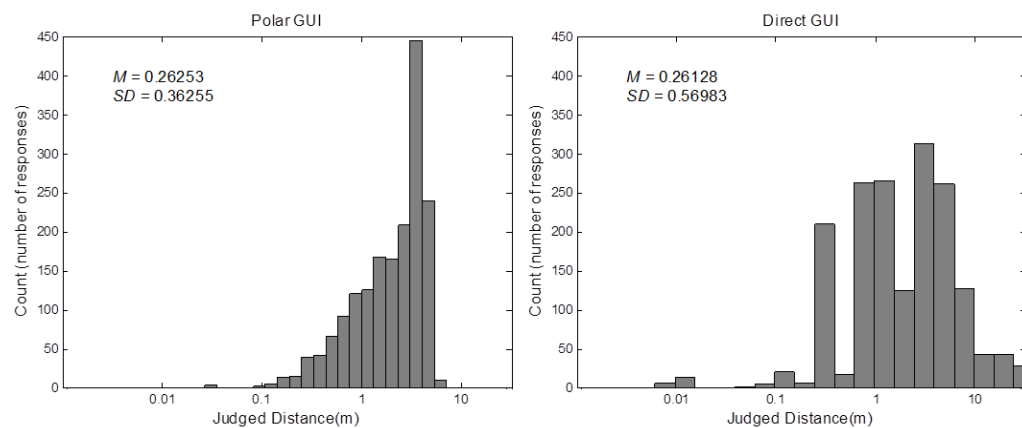
Figure 2.5: Distribution of log transformed distance judgments pooled across listeners from the polar plot GUI (left) and direct estimate GUI (right). Each panel includes the mean and standard deviation of the distribution. Note that the x-axis is plotted in log space.

2. Azimuth Analysis

Front back reversals were resolved if the azimuth error of the judgment was reduced by reflecting the judgment across the interaural axis. Table 1.1 lists the number of front-back reversals across all subjects for both the polar plot GUI and direct estimate GUI conditions. Chi-square tests of independence (degrees of freedom = 1) were

performed to statistically compare the number of reversals collapsed across target

azimuth between conditions. The table shows that there were significantly fewer reversals

in the direct estimate condition than in the polar plot condition for seven of the 16

listeners. The remaining nine listeners did not have a significant difference between

conditions. When the number of reversals was summed across listeners in each condition,

it was found that there were significantly fewer reversals in the direct estimate GUI

condition ($N = 431$) than in the polar plot GUI condition ($N = 717$); $\chi^2(1) = 71.251$, $=$

0.001)

Table 1.1. Displays the percentage of responses that were reversals in each condition for all listeners. The bottom displays the means across all listeners. On the right the χ^2 statistic and associated p-values are displayed comparing the number of reversals in each condition for all listeners. Listeners with superscript '1' performed the localization task using the polar plot GUI first.

Subject ID	Polar Plot GUI %-Reversals	Direct Est GUI %-Reversals	χ^2	p
QEM[1]	36.7	36.0	0.0	0.92
QEN	39.3	26.0	4.1	0.04
QEO[1]	20.7	22.7	0.1	0.71
QEQ[1]	38.7	17.3	12.2	0.00
QER	33.3	33.3	0.0	1.00
QES[1]	32.7	28.0	0.5	0.46
QET	38.0	11.3	21.6	0.00
QEU[1]	14.7	10.7	0.9	0.33
QEV	26.0	16.0	3.6	0.06
QEW[1]	27.3	6.7	18.8	0.00
QEY[1]	28.7	4.0	27.9	0.00
QEZ	35.3	6.0	31.2	0.00
QFB[1]	26.7	7.3	16.5	0.00
QFC	26.7	18.7	2.1	0.15
QFE[1]	14.7	12.0	0.4	0.53
QFF	38.7	31.3	1.2	0.28
Mean	**29.9**	**18.0**		

Azimuth judgments were plotted using the double pole plotting method, adapted from Kistler & Wightman (1992), which separates azimuth judgments into two domains based on the extent of laterality (right-left angle) and on extent from the interaural axis (front-back angle). Figure 2.6 shows azimuth judgments for left-right judgments (top panel) and front-back judgments (bottom panel) for a representative listener (Subject ID: QFC). For right-left angle the largest possible judgments are -90° (left) and 90° (right).

Using this measure, 15° and 165° both have the right-left angle of 15°. Red dots represent judgments from the polar plot condition and blue dots represent judgments from the direct estimate condition. In the top panel, the solid horizontal bar represents midline, so judgments below are to the left and judgments above are to the right. The diagonal dashed line represents a perfect relationship between judged and target position. Because of the transformation to double pole coordinates, the target position of -30° in the figure represents locations at both -30° and -150°. This figure shows that for this listener almost all judgments were in the correct left-right hemisphere. Figure 2.6b (bottom) displays azimuth judgments for front-back judgments. For front-back angle the extremes are -90° directly behind the listener and 90° directly in front of the listener. Judgments of 30° and -30° have a front-back angle of 75°. The horizontal solid black line at 0° represents the interaural axis so judgments plotted above are in the front hemisphere and judgments plotted below are in the back hemisphere.
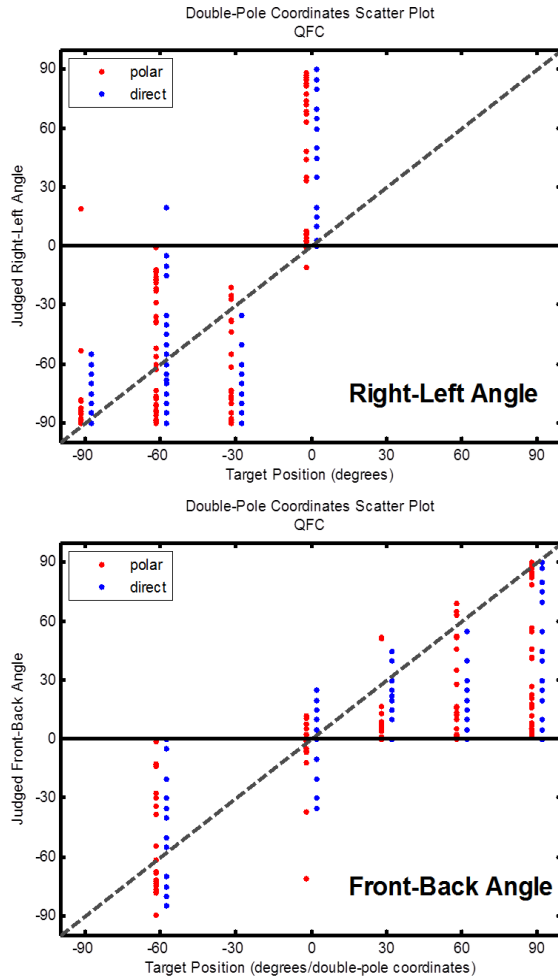
Figure 2.6: Data in the form of Double-Pole coordinates from a single representative listener (Subject ID: QFC) for right-left Angle (top) and front-back angle (bottom) from both GUI response conditions. Red dots show raw azimuth judgments transformed to double-pole coordinates from the polar plot GUI: 10 replications/target azimuth. Blue dots show raw judgments transformed to double-pole coordinates from the direct estimate GUI: 10 replications/target azimuth. Perfectly accurate performance is indicated by the dotted line in each panel. The mean angular error for listener QFC was 27.54° in the polar plot condition and 27.12° in the direct estimation condition. The mean angular error across all subjects was 27.42° in the polar plot condition and 32.85° in the direct estimate condition.

28

Angular error was collapsed across target azimuth and compared between conditions using paired *t*-tests. The polar plot response method ($M = 27.422$, $SD = 4.105$) had significantly less angular error than the direct estimate method ($M = 32.851$, $SD = 7.795$), $t(15) = -3.834$, $p = 0.002$. Azimuth judgments were more accurate in the polar response method than in the direct estimate method.

Figure 2.8a-b displays distributions of azimuth judgments pooled across listeners from each response condition. Each figure includes the mean and standard deviation of the distribution. The mean and standard deviation are similar in the two conditions. Importantly, as mentioned above about the distance response histograms, the shape of the distributions are very similar between the two conditions. Through visually inspecting the distributions there were almost 300 more responses around -90° in the direct estimate GUI condition than in the polar plot GUI condition possibly pointing to a response bias in the direct estimate GUI.
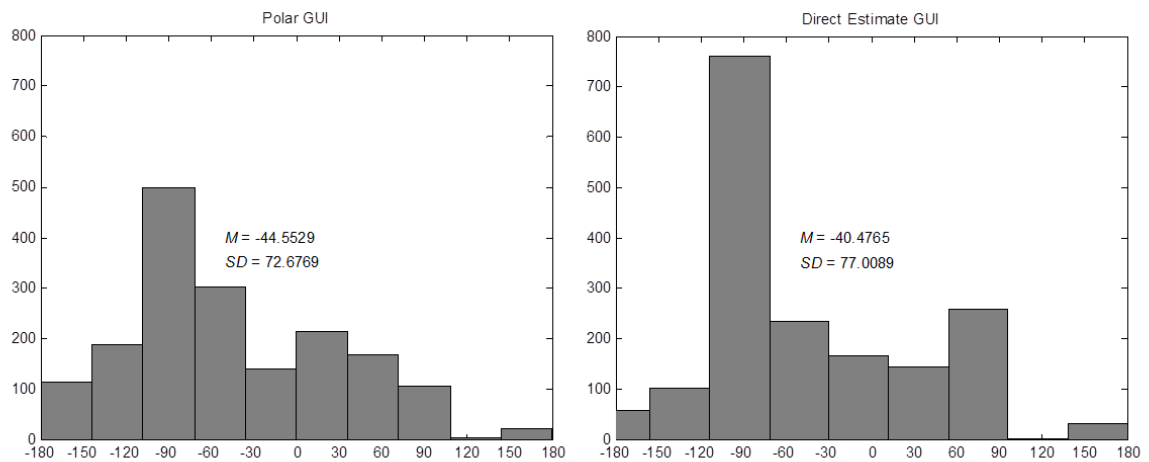
Figure 2.8: Distribution of raw azimuth judgments collapsed across target azimuth from the polar plot GUI (left) and direct estimate GUI (right). Each panel includes the mean and standard deviation of the distribution.

3. Elapsed Time Analysis

The amount of time to complete each condition was measured, in hours, between the first and last trial of each condition. The polar plot method ($M = 0.177$, $SD = 0.048$) was completed in significantly less time than the direct estimate method ($M = 0.399$, $SD = 0.102$) based on a paired samples t-test, $t(15) = -10.086$, $= 0.001$. The polar plot method of response was completed faster than the direct estimate condition.

4. Condition Order Effects

Condition order effects for each response method were analyzed by comparing fit parameters, R^2 values, and angular error between subjects using independent t-tests based on whether listeners responded using the polar plot GUI first ($n = 9$) or second ($n = 7$). The Polar Plot GUI response had no order effects for constants (polar first: $M = 1.587$, SD = 0.5541; polar second: $M = 1.867$, SD = 0.381; $t(14) = -1.090$, $p = 0.294$), exponents (polar first: $M = 0.519$, SD = 0.143; polar second: $M = 0.402$, SD = 0.161; $t(14) = 1.496$, $p = 0.157$), or R^2 values (polar first: $M = 0.498$, SD = 0.103; polar second: $M = 0.5386$, SD = 0.085; $t(14) = -0.803$, $p = 0.435$). The polar response GUI did have an order effect for unsigned azimuth error. When listeners responded using the polar plot method first ($M = 29.300$, $SD = 3.292$) there was significantly more angular error than when listeners responded using the polar plot method second ($M = 25.010$, $SD = 3.950$), $t(14) = 2.372$, $= 0.033$).

Order effects were also analyzed for the direct estimation response method to determine whether listeners responded differently when the direct estimate response method was presented first or second. The direct estimate GUI had no order effects for constants (direct estimate second: $M = 3.377$, SD = 3.614; direct estimate first: $M =$

2.297, SD = 3.171; $t(14) = .601$, $p = 0.557$), exponents (direct estimate second: $M = 0.463$, SD = 0.152; direct estimate first: $M = 0.402$, SD = 0.161; $t(14) = -.119$, $p = 0.907$), or unsigned azimuth error (direct estimate second: $M = 34.896$, SD = 5.951; direct estimate first: $M = 30.222$, SD = 9.505; $t(14) = 1.208$, $p > 0.247$). R^2 values were significantly smaller when listeners responded using the direct estimate method first ($M = 0.484$, $SD = 0.306$) than when listeners responded using the direct estimate second ($M = 0.735$, $SD = 0.123$; $t(14) = 2.225$, $= 0.043$).

Condition order effects were also analyzed for front/back reversals. Table 1.1 displays the number of reversals for all listeners and condition order is noted by the superscript next to the subject IDs. Listeners with a superscript next to their subject ID responded using the polar plot GUI first. Of the nine listeners who responded using the polar plot GUI first, four of the listeners had significantly fewer reversals in the direct estimate condition. Collapsed across listeners who responded using the polar plot GUI first there were significantly more reversals for the polar plot GUI response method ($N = 361$) than in the direct estimation ($N = 217$) condition, $\chi^2(1) = 35.875$, $= 0.001$. Of the seven listeners who responded using the direct estimate method first, three of the listeners had significantly more reversals in the polar plot GUI condition. Collapsed across listeners who responded using the direct estimation method first, there were significantly more reversals in the polar plot condition ($N = 356$) than in the direct estimation ($N = 214$) condition, $\chi^2(1) = 35.375$, $= 0.001$. For both condition orders there were more reversals when listeners responded using the polar plot GUI method. This suggests that practice effects are unlikely to influence the number of front-back reversals for each response method.

31

C. Discussion

Overall, the results from this experiment indicate that both response methods provide similar results for directional localization judgments based on *t*-statistics alone; however, distance judgments should not be collected using the polar plot GUI because of range effects that were not detected by the paired *t*-tests. More reversals in the direct estimate condition does not lead to the conclusion that the polar plot GUI is inappropriate for collecting responses. Non-individualized HRTFs are known to result in front-back reversal rates of about 31.5%, which is numerically greater than the average rates in both the polar plot and the direct response condition (Wenzel, Arruda, Kistler, & Wightman, 1993). As long as reversals are recorded and corrected for in subsequent analyses (Wightman & Kistler, 1989b) the polar plot response GUI would be an appropriate directional judgment response method.

Power functions have been shown to be good fits to distance judgments in the auditory distance studies (Anderson & Zahorik, 2014; Zahorik, Brungart, & Bronkhorst, 2005). Based on R^2 values, power functions were good fits to the distance judgments in both conditions. The R^2 values observed in both the direct estimate and polar plot GUI conditions are in close agreement with past auditory distance perception results (Anderson & Zahorik, 2014; Zahorik, Brungart, & Bronkhorst, 2005). Additionally the fit parameters (*a* and *k*) from both conditions fell within one standard deviation of those reported by Anderson & Zahorik (2014). Fit parameters were also similar between response conditions. From this it can be concluded that there were similar linear and non-linear compression for both response methods.

Azimuth judgments were compared between conditions using a measure of angular error which has been used in past sound directional localization studies (Wightman & Kistler, 1989b; Eberle et al., 2000). Wightman and Kistler's (1989b) listeners were very accurate localizing sounds over headphones with a mean angular error of 19°, however their stimuli were generated using individualized HRTFs. Eberle et al. (2000) used a similar analysis using angular error of judgments collapsed across azimuth to measure directional localization and found a mean angular error of approximately 32° using the a direct estimate response method. A previous study that used a similar polar plot GUI response method measured a mean unsigned error of approximately 27° (Begault, Wenzel, and Anderson, 2001). Ihlefeld & Shinn-Cunningham (2011) also used a similar polar response method; however, the GUI only displayed the front hemi-field. Their listeners' judgments were less accurate at more lateral targets, but they concluded that range effects did not cause the results. Given results of their study, and that the polar plot GUI employed here was a full 360°, there were no concerns that range effects impacted the present results for the azimuth responses. A within-subjects comparison of responses from these two types of response methods has not previously been performed. Based on the present results the polar plot GUI response is suitable to collect directional localization judgments since there were smaller angular errors in the polar plot GUI condition than in the direct estimate GUI condition.

Another component, beyond the accuracy achieved using the two response methods, is the amount of time each method requires to complete data collection. As shown above the polar plot GUI required less time for data collection than the direct estimate GUI while achieving the same amount of accuracy in the distance judgments

33

and more accuracy in the azimuth judgments. Nielsen (1991) reported data collection time using a polar plot GUI response method that was also faster than Anderson & Zahorik (2014) which used a direct estimate method. This provides further support for the use of the polar plot GUI.

Order effects were more complicated to evaluate. While order effects were analyzed across all measures of directional localization and distance judgment accuracy used in this study, only two significant differences were found. When the polar plot GUI was presented second there were smaller angular errors in that condition than when it was presented first. When the Direct estimate GUI was presented second the power function fits explained more variability in that condition than when the Direct estimate GUI was presented first. Both of these results point toward practice or familiarization effects. This effect could be curtailed by continuing to counterbalance the order of conditions or allowing listeners to practice using the response method before performing the task. There were more front-back reversals when listeners responded using the polar plot GUI whether it was presented first or second, so familiarization does not help reduce reversals for either response method.

While these results point to the polar plot response being suitable for subsequent data collection for both distance and azimuth judgments, range effects in the distance domain are a serious concern when using the polar plot response GUI. It became apparent that any distance percepts that may lie beyond the radius of the polar plot figure would be artificially compressed to fit inside the response area. The results show that the polar plot GUI effectively restricts the possible response options, and therefore undermines the validity of task (Perrett and Noble, 1995). Given the inherent variability of distance

34

judgments (Anderson & Zahorik, 2014), the larger variability recorded in the direct estimate GUI condition than in the polar plot condition is important to consider when measuring distance judgments. The shape of the two distributions in figure 2.5 were also important when considering which GUI is suitable for collecting distance responses. Anderson & Zahorik (2014) found that distance responses are normally distributed around the target distances, which is in agreement with the distribution of direct estimate GUI responses but not the distribution of responses from the polar plot GUI condition. Based on these observations the direct estimate GUI is more suitable for collecting distance responses.

As mentioned above, Shinn-Cunningham, Santarelli, and Kopco (2000) used a similar polar plot GUI to collect distance judgments, but they do not mention the effect the GUI may have had on their results. In their study listeners performed near field distance judgments in four conditions: monaural medial sound sources, monaural lateral sound sources, binaural medial sound sources, and binaural lateral sound sources. Distance judgments were plotted as a function of target distance for each condition. Their judgments show most variability both within and between conditions at close target distances. Judgments in all of the conditions show decreasing variability within and between conditions as target distance increases. Their result fits with the present observation that the restricted range of polar plot GUIs in the distance domain created range effects at the far end of the response space.

Based on range effects in distance, the polar plot GUI is inherently flawed for collecting distance judgments. The possibility of increasing the radius of the circle does not ameliorate the range effect problem because the response range will always be

constrained. When listeners have a restricted response range listeners may feel obligated to fit their responses along the entire response space. The direct estimate GUI prevents range effects by allowing listeners to create their own response space. Additionally, if it were possible to make a polar plot GUI with an infinite range, new problems would arise like the response range where the physical sources were located would be dwarfed compared to the rest of the response space.

Range effects were also analyzed in the azimuth domain. Similar means and standard deviations of the distribution of responses in the two conditions are good indications that the two GUIs yield similar response patterns. One difference between the two conditions was that there were almost 300 more responses around -90° in the direct estimate GUI condition than in the polar plot GUI condition. This may be an indication of a response bias which can be avoided by using the polar plot GUI. Based on these results drawn from the response distribution in figures 2.5a-b and 2.8a-b respectively, subsequent distance judgments will be made using the Direct Estimate GUI while subsequent azimuth judgments will be made using the polar plot GUI.

D. Conclusions: Response methods choices for subsequent experiments.

Based on the results of Experiment I, it seems clear that the polar plot GUI is potentially problematic for collecting of distance perception judgements, given the demonstrated range effects. For this reason, traditional direct estimate response methods will be used for subsequent auditory distance perception testing (Experiment II). Because range effect concerns were not evident for direction components of the responses, the polar plot GUI will be used for subsequent directional localization testing (Experiment III), given its advantages in data collection speed and accuracy in terms of angular error.

These advantages outweighed the increase in reversals observed with the polar plot GUI. The explanation for this increase is unknown, but since subsequent testing will all be conducted within listeners using the same response technique, this should have minimal impact on the comparisons.

CHAPTER III

EXPERIMENT II

Experiment II utilized a perceptual weighting paradigm similar to Zahorik

(2002a) to measure the relative perceptual weights assigned to intensity and modulation

depth during distance judgments. In auditory distance perception, multiple cues are likely

combined and weighted to form a single distance percept. The amount of perceptual

weight placed on individual distance cues by the auditory system can be estimated by

correlating physical stimulus parameters with distance judgments. This paradigm uses

multiple regression analysis to estimate the perceptual weights of cues by perturbing the

stimulus parameters of interest and using distance judgments as the response variable.

This is done by independently placing a small random rove, or slight variation, on both

the intensity and modulation depth of the stimulus on each trial. To calculate the

perceptual weights, the intensity, modulation depth, and physical distance are placed into

a multiple regression model with the judged distance as the predictor variable. The

standardized coefficients for each parameter in the model can then be used as a measure

of the perceptual weighting of each parameter. This same paradigm was used by Zahorik

(2002a) to measure the perceptual weights of intensity and the D/R in distance judgments

and found that weights shifted based on stimulus type. The physical distance parameter in

the model is expected to be weighted most strongly since it reflects the listener's usage of

all distance cues that are correlated with physical distance. Modulation depth is expected

to be significantly weighted based on results from behavioral distance judgment tasks

38

where listeners benefit from modulated stimuli (Zahorik & Anderson 2015; Kim et al., 2015), but less weighted than intensity since it is a primary distance cue (Mershon & King, 1975). Unlike stimulus generation in Experiment I where BRIRs were measured in Bigelow Hall, virtual room simulation methods were used to generate BRIRs because the MTFs of the measured BRIRs did not attenuate modulation significantly as a function of distance.

A. Methods

1. Participants

Nine listeners (8 female) ranging in age from 22 to 29 years old participated in this experiment. None of the listeners participated in Experiment I. Listeners were recruited through flyers, email advertisements, and personal contacts. All listeners had normal hearing as verified by audiograms (Maico MA41 audiometer; TDH-39 Headphones) in a sound attenuating booth with less than 25 dB HL at octave frequencies between 250 and 8000 Hz. Listeners also had normal central auditory processing as verified by dichotic digits (Musiek, Gollegly, Kibbe, & Verkest-Lenz, 1991) and masking level difference testing (Wilson, Zizz, & Sperry, 1994) which were performed in the sound booth after audiograms were measured. Listener compensation was in the form of cash payments. All testing was approved by the University of Louisville Internal Review Board.

2. Materials

BRIR Generation. Simulated BRIRs were used to ensure MTFs changed predictably as a function of distance. BRIRs were generated using virtual auditory space techniques as described in Zahorik (2009). This room modeling software simulated early

reflections using an image-model (Allen and Berkley, 1979), while the late reverberant energy was simulated using a statistical model. The direct path and early reflections are filtered using the measured HRTF of an individual listener (ID: SLO) to spatially render convolved stimuli. The non-individualized HRTFs were measured at a fixed distance of 1.4 m from the source in anechoic space, so near-field measurements were unavailable; however, near-field distance cues are useable only within 1 m of the source and are not considered a primary distance cue (Mershon & King, 1975). This simulation technique produces BRIRs that describe transformations of sound between the source and the listeners' ear in a simulated room that are reasonable physical and perceptual approximations to those measured in a real environment (Zahorik, 2009).

A simulated room was used to generate BRIRs because results from pilot data collected using BRIRs measured in Bigelow Hall (described above in Experiment I) indicated listeners had severe trouble performing the distance judgment task. Upon visual inspection of modulation gain as a function of distance for the BRIRs measured in Bigelow Hall, it was clear that modulation gain varied only slightly as a function of distance for the measured distance range, therefore listeners did not have access to the modulation attenuation cue for distance judgments. The simulated room used to generate BRIRs was based on a room used in Brandewie (2012) because of its known acoustic characteristics. The simulated room measured approximately 500 m^3 (10 x 10 x 5 m) with the simulated omni-directional measurement loudspeaker located 2 m from both the front and left wall (2 x 2 x 1.28 m). The simulated microphone was located directly to the left of the simulated measurement loudspeaker (2 m from the left wall) 1.28 m above the ground at 9 distances ranging logarithmically 0.35 to 5.6 m from the loudspeaker. The

broadband energy absorption coefficient, or *early α* parameter, which controls the absorptive properties of the room's surfaces was 0.06. The *late α* parameters, are used for the late statistical portion of the room model at octave bandwidths from 125 to 8000 Hz (125 Hz: 0.06; 250 Hz: 0.06; 1000 Hz: 0.06; 2000 Hz: 0.06; 4000 Hz: 0.06; 8000 Hz: 0.20). The broadband T_{60} (ISO-3382, 1997) was approximately 3 s.

Stimulus Parameter Determination. The MTFs of the room informed what stimulus parameters to use. MTFs were computed from each of the nine BRIRs based on a technique described by Schroeder (1981). The BRIR is first band-pass filtered using a filterbank of 1/3 octave-wide filters spanning the frequency range of the BRIR. The MTF is then computed as the amplitude at each frequency of the squared band-pass filtered impulse response normalized by the total energy of the impulse response. Figure 3.1 displays how the MTFs of the room change as a function of distance for a source at 90° azimuth for both ears. Each panel displays the modulation gain provided by the room at 0.35 m (top row) and 5.6 m (bottom row). Cooler colors denote decreasing modulation gain. Zero gain indicates no change in modulation depth. When the simulated measurement loudspeaker is 90° to the right there is more modulation depth attenuation at the left ear than at the right ear across distance. At both distances displayed, there is noticeably less modulation attenuation at the lower modulation frequency, which is expected based on how modulation transfer functions are affected by rooms (Houtgast & Steeneken, 1985).
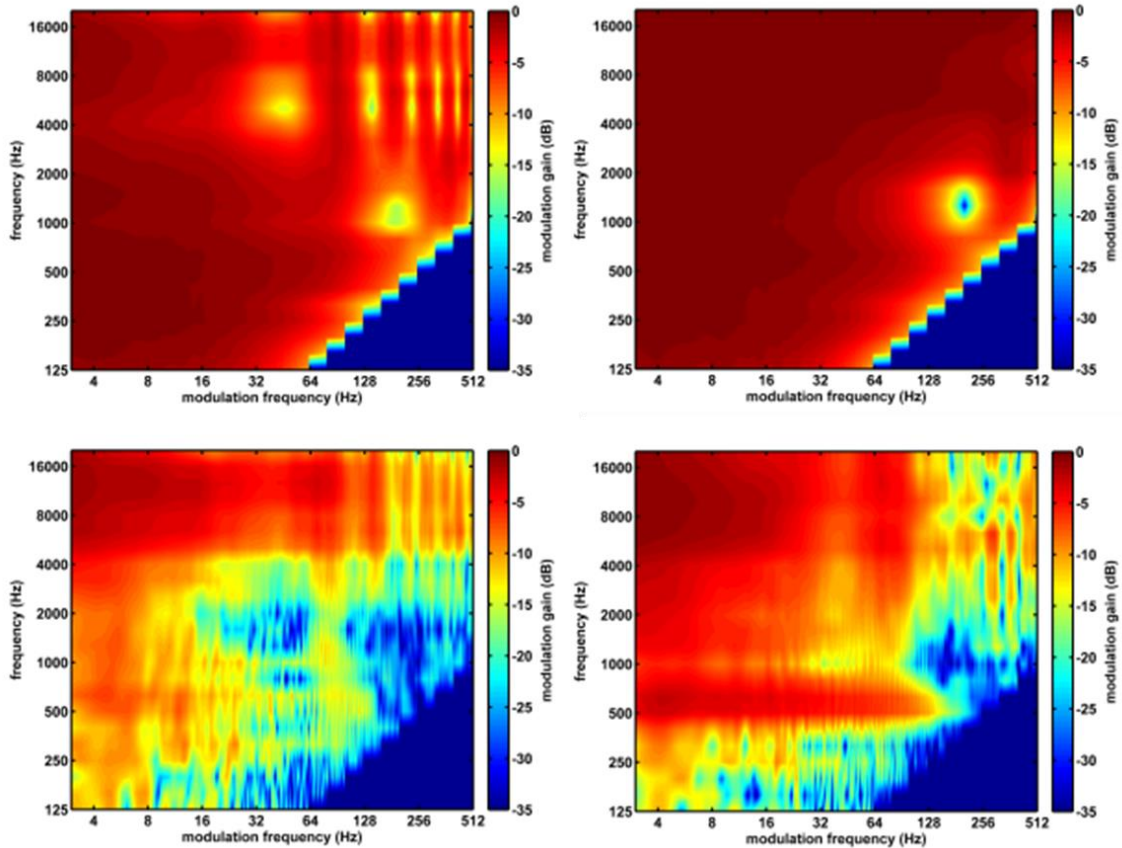
Figure 3.1: Modulation Transfer Functions for the left (left column; contralateral) and right (right column; ipsilateral) ears increasing with distance of 0.35 m in the top row and 5.6 m on the bottom row for a source located 90° to the left. Warmer colors indicate more modulation gain (less attenuation) and cooler colors indicate less modulation gain (greater attenuation). The x-axis denotes modulation frequency (Hz) and the y-axis denotes carrier frequency (Hz).

While the amount of modulation attenuation changes most drastically as a function of distance at the contralateral ear there were concerns about whether D/R cues (and amplitude modulation cues) would be less robust at the contralateral ear since the direct energy will be attenuated by the head shadow. A pilot experiment was performed

where stimuli were presented monaurally to the listeners' contralateral ear, and the stimulus level was equalized for distance by adjusting the gain of the source signal to compensate for the propagation loss of 6 dB per doubling of source distance in anechoic space. The listener's task was to estimate the egocentric distance to the sound source. The results from the pilot experiment indicated that listeners had trouble with the task and seemed to scale reverberation level instead of distance. In this situation the proportion of reverberant energy increases with increasing distance but the level of the direct portion of the source signal remains constant. As reverberation level increases with target distance listeners scale the level of the reverberation and respond to increasing target distances with closer distance judgments. Consequently the decision was made to present stimuli monaurally to the ipsilateral ear and to not compensate for propagation loss. With ipsilateral presentation and not equalizing level the energy in the direct portion of the waveform will remain unattenuated by head shadow while the energy proportion of energy in the direct and reverberant portions of the waveform will vary naturally.

Figures 3.2a-b, derived from the MTFs, plotted modulation gain as a function of target distance for the ipsilateral ear, were used to make further decisions about stimulus parameters. The signal was a 1-octave wide band of noise at different center frequencies. The center frequency increases from 2000 Hz in the top panel to 4000 Hz in the bottom panel. Figures 3.3a-b display the same information for center frequencies of 6000 Hz (top) and 8000 Hz (bottom). Each panel of figures 3.2 and 3.3 displays the modulation gain for the signal at modulation frequencies of 4 (red), 8 (green), 16 (blue), 32 (black), and 64 Hz (cyan). Based on these figures, the 4000 Hz center frequency was chosen because the modulation gain monotonically decreased as a function of distance at

all modulation frequencies and there were larger differences in the amount of gain between modulation frequencies, especially at the farthest target distance. A low and high modulation rate were chosen so there could be one condition where modulation was attenuated less by the room (low modulation rate). This would hopefully allow predictions to be made about how the MTF affects distance judgments. The low modulation frequency was chosen to be 4 Hz, and the high modulation frequency was chosen to be 32 Hz because it was attenuated most at the furthest target location. Since rooms act as low-pass filters in the modulation domain (Houtgast & Steeneken, 1985), it would be expected that the 64 Hz modulation frequency would be attenuated more than the 32 Hz modulation frequency; however, the MTFs in figure 3.1 demonstrate that modulation attenuation is frequency dependent. Frequency dependencies in MTFs can result from early reflections caused by differently sized surfaces in the room that act as frequency dependent reflectors. As a result all rooms have unique physical properties that will result in different in different acoustic characteristics that approximate a low-pass filter in the modulation domain. The signal was 1-octave band of noise because the way the MTF would filter a 1-octave band of noise was easily predictable.
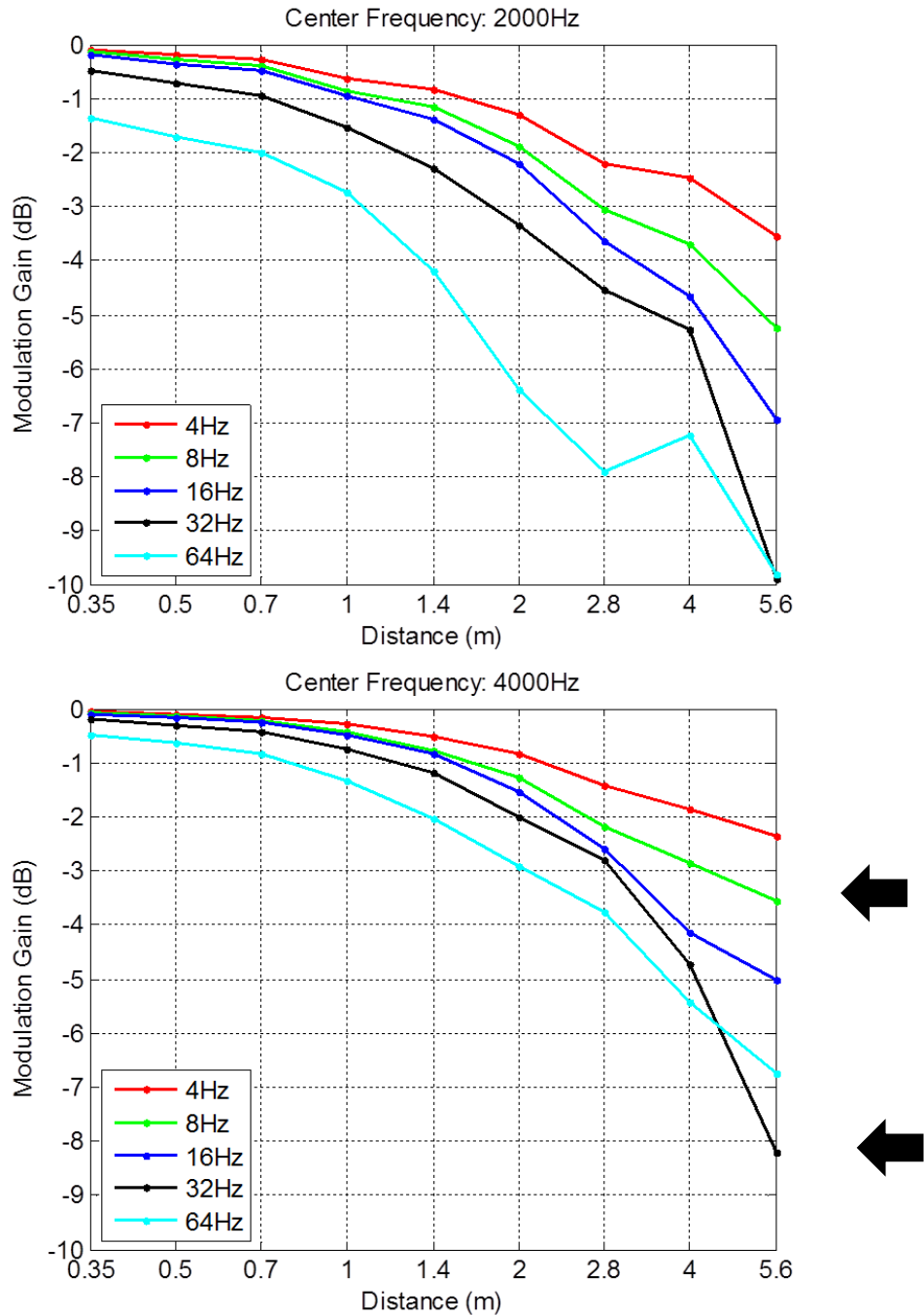
Figure 3.2: Modulation gain as a function of target distance for the ipsilateral (right) ear for center frequencies of 2000 Hz (top) and 4000 Hz (bottom). The carrier signal is 1-octave band noise. Each panel displays the modulation gain for the signal at modulation frequencies of 4 (red), 8 (green), 16 (blue), 32 (black), and 64 Hz (cyan). Arrows indicate the functions that were tested.
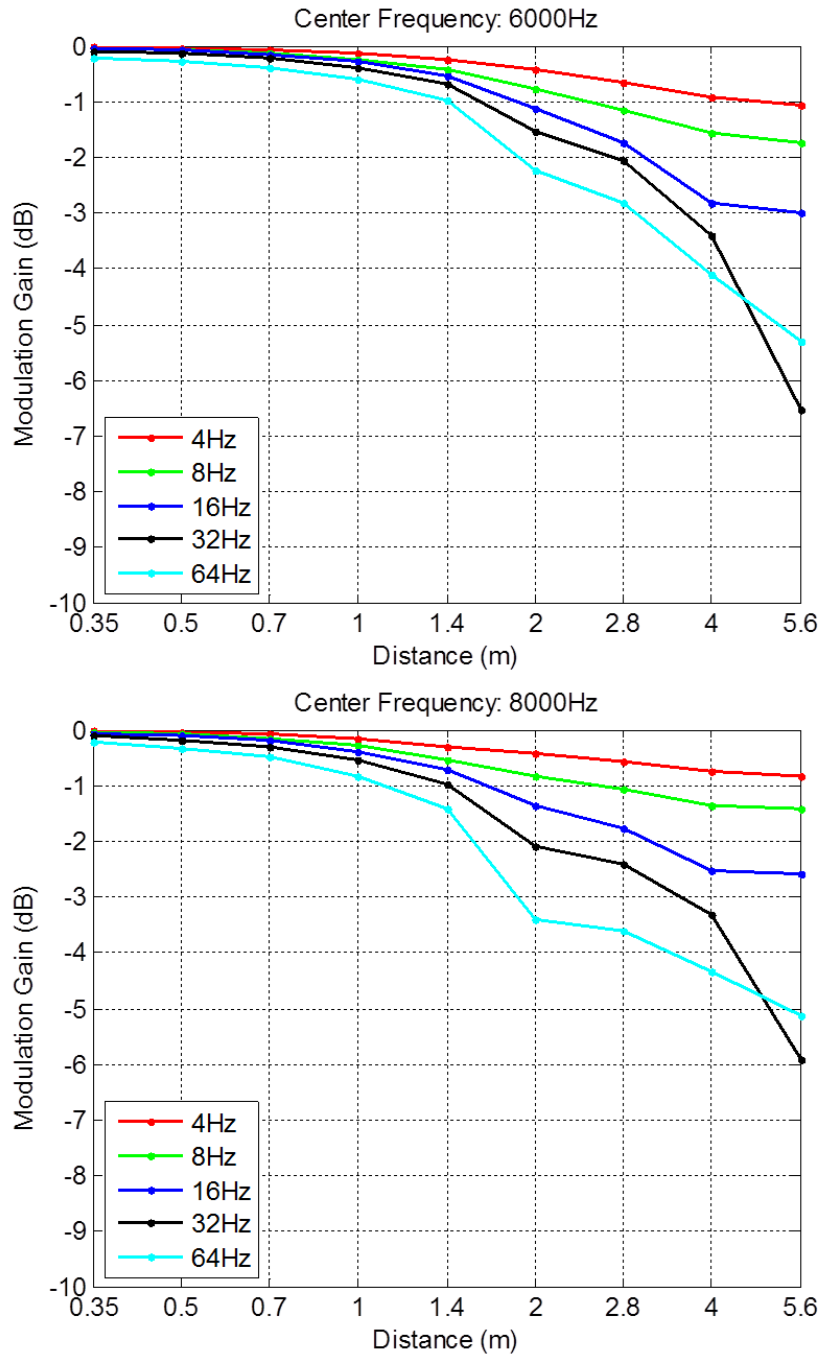
Figure 3.3: Same as figure 3.2, but for center frequencies 6000 Hz (top) and 8000 Hz (bottom).

Source Signal. The source signal was composed of a 1-octave Gaussian noise sample (pseudo-randomly generated before each trial) centered at 4000 Hz with a 2000 ms duration. The signal duration includes a Hanning window of 500 ms rise/fall time.

The noise sample was then sinusoidally amplitude modulated at either 32 or 4 Hz depending on the condition. The source signal was generated and convolved with the BRIR before each trial. Modulation depth attenuation was roved before BRIR convolution using attenuation values randomly sampled from a Gaussian distribution with a mean of 6 dB and a standard deviation of 2 dB. A standard deviation of 2dB, calculated as the modulation depth just noticeable difference (Wakefield & Viemeister, 1990), was used to ensure listeners could extract modulation depth variations placed on the stimuli between trials without drawing listeners' attention overtly to modulation depth changes. The noise signal was then multiplied by the amplitude modulator and convolved with the BRIR. The intensity of the convolved signal was then roved +/- 6 dB based on a random sample from a Gaussian shaped distribution with a mean of -6 dB and a standard deviation of 2 dB. After modulation rove, convolution, and intensity rove, the signal was presented to the listener. Stimuli were presented monaurally to the ipsilateral ear.

3. Design

Listeners were tested using a within-subjects design where egocentric distance judgments to the sound source were made in two conditions: 32 Hz and 4 Hz modulation. The entire experiment took place inside a double-walled sound proof booth (Acoustic Systems, Austin, TX). Listeners responded using the direct estimate GUI as described in Experiment I, except the text box for azimuth judgments was removed from the GUI for this experiment, so only distance judgments were recorded. Each condition was tested within its own block of trials, which included 40 judgments for each of the 9 target distances (360 trials total). The order of blocks was counterbalanced and the order of trials within each block was randomized. Feedback was not provided to the listeners.

4. Procedure

Auditory central processing disorder (ACPD) screening was performed before listeners participated. ACPD was screened because listeners with ACPD have difficulty combining information from both ears, and for subsequent tasks in Experiment III normal binaural integration may be important for exploiting modulation depth cues for directional localization. The screening included dichotic digits (Musiek, Gollegly, Kibbe, & Verkest-Lenz, 1991) and masking level difference testing (Wilson, Zizz, & Sperry, 1994). Before beginning the task, listeners chose between responding in feet or meters. Listeners could listen to the stimulus as many times as they wished before entering their response. Listeners were required to be precise to two decimal places and were instructed to reserve a response of zero for a percept of inside the head locatedness (Blauert, 1997, p. 132). Custom MATLAB software was used to stimulus presentation and data collection.

Data Analysis. Data from both conditions were fit independently using two models: power function and multiple linear regression. Following methods used in previous auditory distance perception studies (Zahorik, et al., 2005; Anderson & Zahorik, 2014), power functions of the following form were fit were fit (least-squares criterion) to the geometric means in each condition: $\hat{y}_r = k\Phi_r^a$ ($\hat{y}_r$ = perceived distance, k = constant, a = power-law exponent, Φ_r = target source distance). Fit parameters, a and k, were used as a measures of accuracy. The exponent indicates non-linear compression ($a < 1$) and expansion ($a > 1$) in the function while the constant indicates linear compression ($k < 1$) and expansion ($k > 1$) in the function. Power functions fit to individual listener's data were used to gauge individual performance on the judgment task.

48

Multiple linear regression was performed on standardized scores fit to individual data to measure perceptual weights placed on distance cues as performed in Zahorik (2002a). This fit is similar to the power function described above except expressed in linear terms ($response' = B_rX_r + C + e$; $response'$ = predicted log-transformed responses, B_r = beta coefficient for log-transformed physical distances, X_r = log-transformed physical distances, C = logarithmic constant, e = error term) with additional parameters added to the model to explain more response variability from the independent perturbations applied to intensity (I) and modulation depth (m). Predicted responses may then be characterized by the following multiple regression equation: $response' = B_rX_r + B_IX_I + B_mX_m + C + e$. Interaction terms were not included in this weighting equation because these terms did not produce statistically significant weights. In the model form above it is difficult to interpret the perceptual weightings since parameters are in different units and to not have the same variance. Therefore all parameters were standardized prior to multiple regression. Standardized predicted responses are given by the model $Zresponse' = B_rZ_r + B_IZ_I + B_mZ_m + e$. In this weighting equation the B_i weights are all partial correlations between the given term (either log-physical distance or cue perturbation value) and the log-response. Weights, B_i, were calculated for each listener individually. R^2 is a measure of the proportion of total variance accounted for by the model as well as how well the weighting model fits the data. The number of trials per condition (360 trials = 9 stimuli x 10 reps x 4 blocks) was based on a previous perceptual weighting study (Zahorik, 2002a) where there were five predictor variables and 480 trials (12 stimuli x 10 reps x 4 blocks).

B. Results

Listeners' central auditory processing abilities were assessed using dichotic digits (DD; Musiek, Gollegly, Kibbe, & Verkest-Lenz, 1991) and masking level difference (MLD; Wilson, Zizz, & Sperry, 1994) disability screening tests. Auditory processing disorder screening was not performed in Experiment I because the experimental AM stimuli were not used. Listeners' performances on these tests were reported in table 2.1. For the DD test a listener would hear 20 trials of 4 digits presented in pairs to each ear simultaneously. The task for the listener was to repeat back all four numbers correctly. Each correct digit is worth 2.5 points where perfect performance was a score of 50 points for each ear. The criterion for normal performance on the DD test is a score of 90% (45 of 50 points) or more of the digits correctly identified for each ear. This is equivalent to missing two digits or fewer at each ear. All listeners in Experiment II had normal results for the DD test for both the left ($M = 95.6\%$, $SD = 4.37\%$) and right ear ($M = 95.6\%$, $SD = 3.69\%$). In the MLD test words were presented to listeners in noise and the listeners had to respond by saying what word they heard. In the first block of words the masking noise was presented in phase to each ear (S_0N_0) and in the second block the noise was 180° out of phase (S_0N_Π). Sets of four words were presented at a given SNR, and the SNR was decreased by 2 dB every time a listener correctly reported all four words. A threshold for each block was considered the lowest SNR at which the listener correctly reported three of the four words. The masking level difference was then measured as the threshold difference between the two conditions. Normal performance was a masking level difference of 6 dB or greater. All listeners in Experiment II had normal results for the MLD test ($M = 6.89$; $SD = 1.45$). Based on these tests all listeners

50

in Experiment II had normal central auditory processing. Because of the small variability of responses and the small number of observations for each screening measure it was difficult to meaningfully relate these results to perceptual weighting measures in Experiment II.

Table 2.1. Measurements from the dichotic digits and masking level difference screening tests for each listener in Experiment II. For the dichotic digits test, the number of items missed for each ear is reported. Normal performance in the dichotic digits task is defined as missing 2 or fewer items at each ear. For the masking level difference test, the masking level difference (in dB) is reported. Normal performance in the masking level difference test is a 6 dB or greater masking level difference. Means and standard deviations for each measure are at the bottom of the table.

	Dichotic Digits		
Subject ID	Left Ear	Right Ear	MLD (dB)
ZFS	1	2	8
ZFU	1	2	6
ZFV	2	1	6
ZFW	0	1	6
ZFY	0	0	6
ZGB	2	1	10
ZGC	0	0	6
ZGD	2	0	6
ZFP	0	1	8
Mean	0.9	0.9	6.9
Std	0.9	0.8	1.5

2. Multiple Regression Analysis (all target distances)

Power functions and distance judgments for all participants are shown in Figure 3.4 for the 32 Hz modulation condition plotted in log space. Dots indicate the raw distance judgments provided by the participant (y), while the open circles represent the geometric mean ($\overline{y}$) for each distance. The power function fits for each condition are plotted as a solid black line ($\hat{y}$), and the diagonal dotted black line represents a perfectly accurate relationship between target distance and estimated distance (i.e., $a = 1$, $k = 1$).

52

Each panel includes the subject I.D. of the listener, the fit parameters (a and k) and the proportion of variability accounted for by the fit (R^2). Figure 3.5 displays the same plots as figure 3.4 but for the 4 Hz modulation condition. Consistent with previous auditory distance estimation studies (Zahorik et al., 2005; Anderson & Zahorik, 2014), power functions in both conditions appear to be good fits to the data evidenced by the R^2 values while the exponent and constant values reflect expected individual differences typical of auditory distance perception studies. Table 2.2 and Table 2.3 display the unstandardized coefficients and R^2 values, from models in the 32 Hz and 4 Hz conditions respectively, with only r' included in the model, only r' and I included in the model, and the full multiple regression model. The R^2 values were compared across the three models to demonstrate that more variability was explained as terms were added to the model. When all of the parameters were included in the model, 18-19% more variability was explained by the model, in the 32 Hz and 4 Hz conditions respectively, compared to when the only parameter in the model was r'. If the predictor values were centered at a mean of zero, the unstandardized coefficients did not change as terms were added to the model.
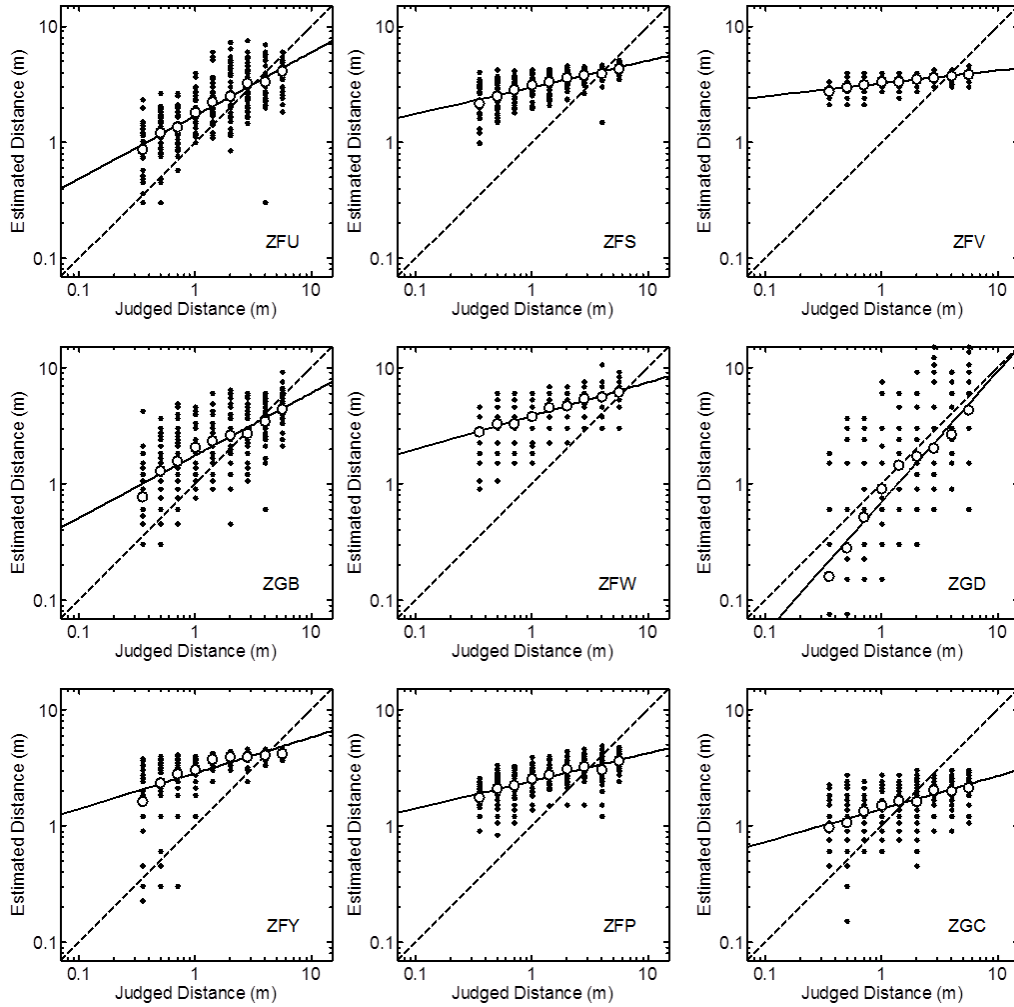
Figure 3.4: Power functions in the 32 Hz modulation condition. Each panel displays data from each listener. Dots show raw distance judgments (y): 40 repetitions/target distance. Open circles indicate geometric means ($\bar{y}$) for each target distance. Data from each listener were fit with a power function ($\hat{y}$; black solid line) of the form $\hat{y}_r = k\Phi_r^a$ ($\hat{y}_r$ = perceived distance, k = constant, a = power-law exponent, Φ_r = target source distance). Each panel includes the subject ID of the listener in the bottom right-hand corner. Perfectly accurate performance is indicated by the dotted line in each panel.

54
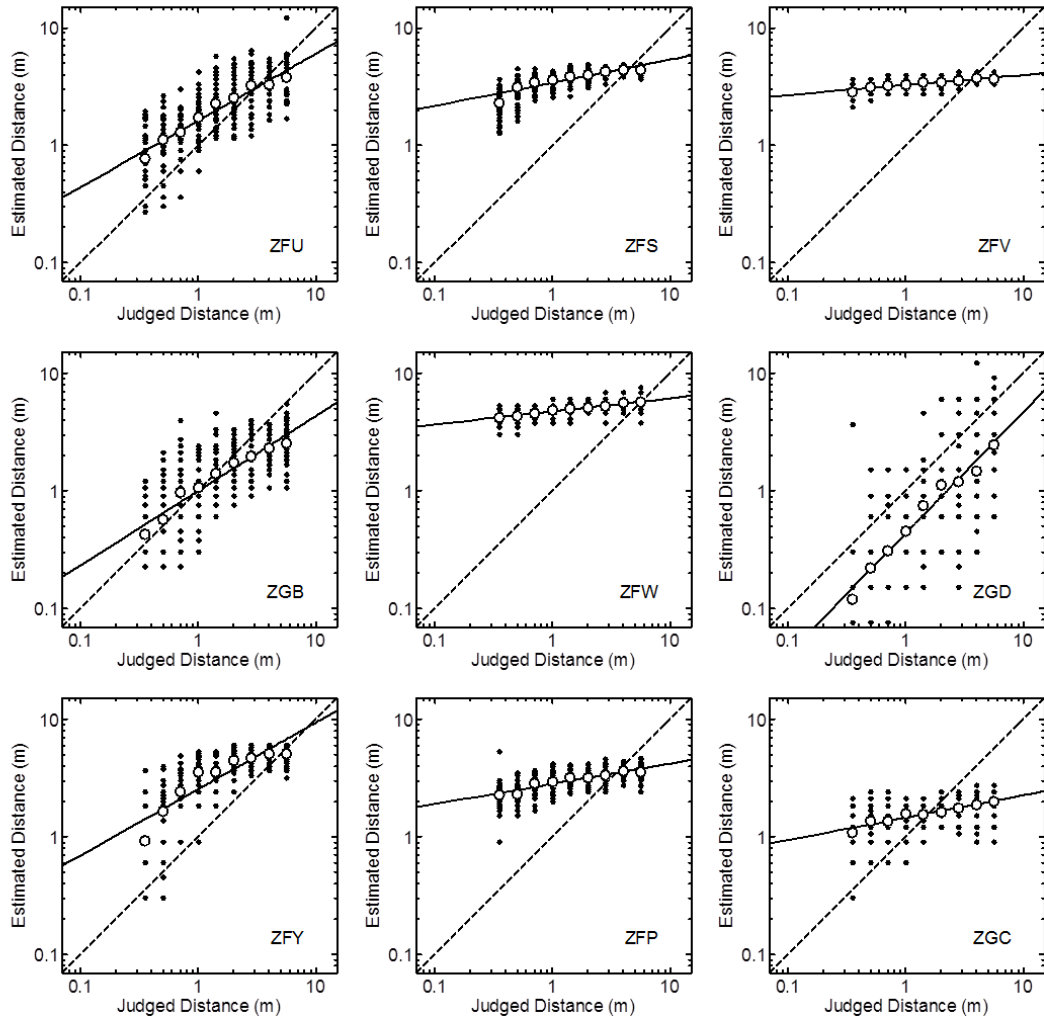
Figure 3.5: Same as figure 3.4 for the 4 Hz modulation condition.

Table 2.2. Unstandardized coefficients from the model when only r' is included as a predictor variable and R^2 values from all three models in the 32 Hz condition for all listeners (From left to right: Only r' in the model which is equivalent to the power function; R^2 for model with intensity rove; R^2 for model with intensity rove and modulation rove). Below the parameters and R^2 values from the listeners are means and standard deviations for each column. At the bottom of the table paired t-tests were performed to compare the R^2 values to the model with only r' included in the model. Associated p-values appear below each t-statistic.

	32Hz Condition (Unstandardized Coefficients)				
	Only r' in model			r' and I in model	r', I and m in model
	constant	r'	R^2	R^2	R^2
ZFU	1.70	0.55	0.55	0.71	0.76
ZFS	2.98	0.23	0.49	0.59	0.63
ZFV	3.22	0.11	0.48	0.79	0.80
ZGB	1.76	0.54	0.45	0.66	0.70
ZFW	3.84	0.29	0.39	0.54	0.58
ZGD	0.70	1.12	0.56	0.65	0.69
ZFY	2.83	0.31	0.30	0.39	0.40
ZFP	2.43	0.24	0.42	0.57	0.64
ZGC	1.40	0.28	0.31	0.37	0.42
Mean	2.32	0.41	0.44	0.59	0.62
Std	1.00	0.30	0.09	0.14	0.14
t-stat(against only r' in model)				-5.79	-7.82
p-value (uncorrected)				0.00	0.00
p-value (Bonferroni corrected)				0.00	0.00

Table 2.3. Same as Table 2.2 for the 4 Hz condition.

	r' in model			*r'* and *I* in model	*r'*, *I* and *m* in model
	constant	*r'*	R^2	R^2	R^2
ZFU	1.62	0.57	0.58	0.71	0.74
ZFS	3.41	0.20	0.39	0.50	0.54
ZFV	3.26	0.09	0.41	0.71	0.74
ZGB	1.00	0.64	0.58	0.70	0.73
ZFW	4.72	0.11	0.40	0.65	0.66
ZGD	0.44	1.03	0.61	0.71	0.75
ZFY	2.54	0.56	0.55	0.63	0.67
ZFP	2.82	0.17	0.39	0.55	0.63
ZGC	1.45	0.19	0.25	0.30	0.37
Mean	2.36	0.39	0.46	0.61	0.65
std	1.35	0.32	0.12	0.14	0.12
t-stat(against only *r'* in model)				-5.54	-7.61
p-value (uncorrected)				0.00	0.00
p-value (Bonferroni corrected)				0.00	0.00

The R^2 values from the multiple regression model above were compared to a
regression model that excluded the intensity and modulation parameters to measure the
amount of variability in the distance judgments that was accounted for by the intensity
and modulation parameters. When intensity was added to the model that only included *r'*
as a predictor, the R^2 value increased in both the 32 Hz ($M = 0.585$, $SD = 0.138$; $t(8) = -$
5.788, $= 0.001$) and 4 Hz ($M = 0.607$, $SD = 0.136$; $t(8) = -5.545$, $= 0.001$) conditions.
When *r'*, intensity, and modulation were included in the model the R^2 value increased in
both the 32 Hz ($M = 0.623$, $SD = 0.139$; $t(8) = -7.825$, $= 0.001$) and 4 Hz ($M = 0.646$, SD
$= 0.125$; $t(8) = -7.614$, $= 0.001$) conditions compared to the models with only *r'* as a
predictor.

Standardized coefficients from the full multiple regression model for all
participants from the 32 Hz and 4 Hz conditions are displayed in figures 3.6a-b
respectively. In these figures white bars indicate significant non-zero weightings. Each

57

bar within a grouping represents the weighting from the model fit to an individual listener. R^2 values from the model fit are displayed to the right of each figure. A few observations can be made by initially looking at the two figures. For all listeners in both conditions, weights placed on physical distance (r') were positive and significant (4 Hz: $M = 0.672$, $SD = 0.092$; 32 Hz: $M = 0.6527$, $SD = 0.083$. The r' parameter was expected to be positive because it reflects the target location of the stimulus and distance judgments increase with physical distance. Intensity rove standardized beta coefficients were negative and significant for all listeners in both the 4 Hz ($M = -0.372$, $SD = 0.105$) and 32 Hz ($M = -0.374$, $SD = 0.091$) conditions. Modulation depth rove standardized beta coefficients were negative and significant for all listeners (except SID: ZFY in the 32 Hz condition) in both the 4 Hz ($M = -0.189$, $SD = 0.059$) and 32 Hz ($M = -0.185$, $SD = 0.066$) conditions. Negative weighting for the intensity and modulation depth parameters were expected since they both have an inverse relationship with physical distance. In both figures it appears that modulation depth weightings were about half as strong as the intensity weighting. This is interesting since modulation depth has not been previously described as an auditory distance cue, yet it is weighted about half that of intensity which is a primary distance cue. R^2 values were also high and significant across participants in both the 4 Hz ($M = -0.646$, $SD = 0.125$) and 32 Hz ($M = -0.623$, $SD = 0.139$) conditions indicating the multiple regression model was a good fit to the data.
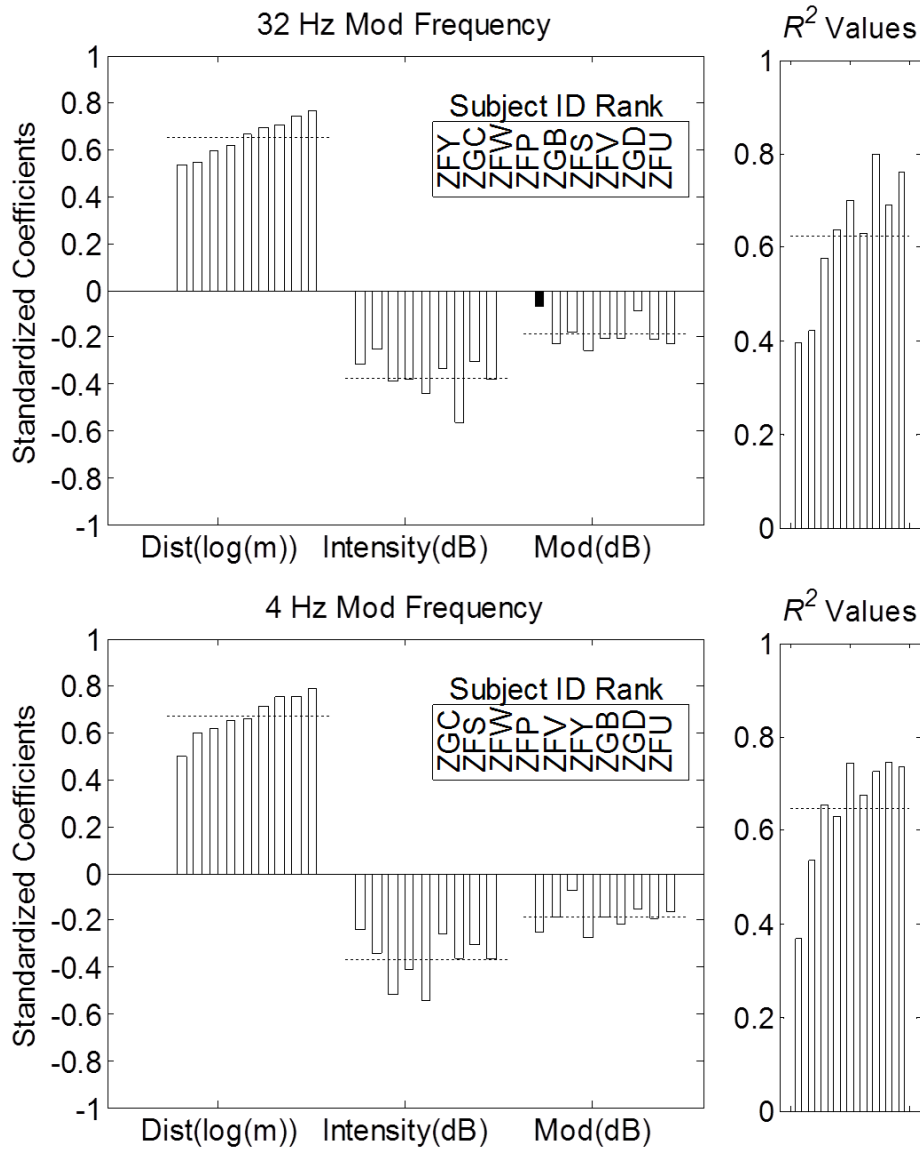
Figure 3.6: Individual listener weights across all target distances with the 32 Hz modulation condition in the top panel (panel A), and the 4 Hz modulation result in the bottom panel (panel b). The order of individual weights is displayed at the top of each panel. The rank from left to right was determined by the r' weight from low to high. Black bars represent weights that are not significantly different from zero (p > 0.05). The proportion of variance explained by the weighting model for each listener, R^2, is displayed on the right side of both panels. Dotted lines indicate mean weights and R^2 values.

Standardized beta coefficients were compared between modulation rate conditions. Using a model with judgments from all target source distances, paired *t*-tests revealed no significant differences between conditions for *r'* ($t(8) = -1.9566, = 0.086$), intensity rove ($t(8) = -0.062, = 0.952$), and modulation depth rove ($t(8) = .159, = 0.878$). This means weighting strategies did not change between the two conditions when including data from all target distances in the model. There was also no significant change of R^2 values between the 32 Hz and 4 Hz conditions ($t(8) = -1.162, = 0.279$).

3. Multiple Regression Analysis (target distances > 2 m)

Figures 3.7a-b show perceptual weights using the same multiple regression analysis as above except only data from target sources greater than 2 m were included in the model. This range was chosen based on figure 3.2 which shows modulation gain for a center frequency of 4000 Hz as a function of distance is greatest beyond 2 m. Figures 3.7a-b have the same labeling scheme as the weighting figures 3.6a-b. High R^2 values indicate good fits to the data (4Hz: $M = 0.468, SD = 0.123$; 32 Hz: $M = 0.406, SD = 0.162$) in both conditions. For all listeners (except listener ZGC) weights placed on *r'* in both conditions were positive and significant (4 Hz: $M = 0.230, SD = .067$; 32 Hz: $M = 0.267, SD = 0.091$). Intensity rove standardized beta coefficients were negative and significant for all listeners in both the 4 Hz ($M = -0.520, SD = 0.147$) and 32 Hz ($M = -0.489, SD = 0.173$) conditions. Modulation depth rove standardized beta coefficients were negative and significant (except listener ZFY 32 Hz condition and listener ZFW in the 4 Hz condition) in both the 4 Hz ($M = -0.304, SD = 0.135$) and 32 Hz ($M = -0.221, SD = 0.108$) conditions.

60
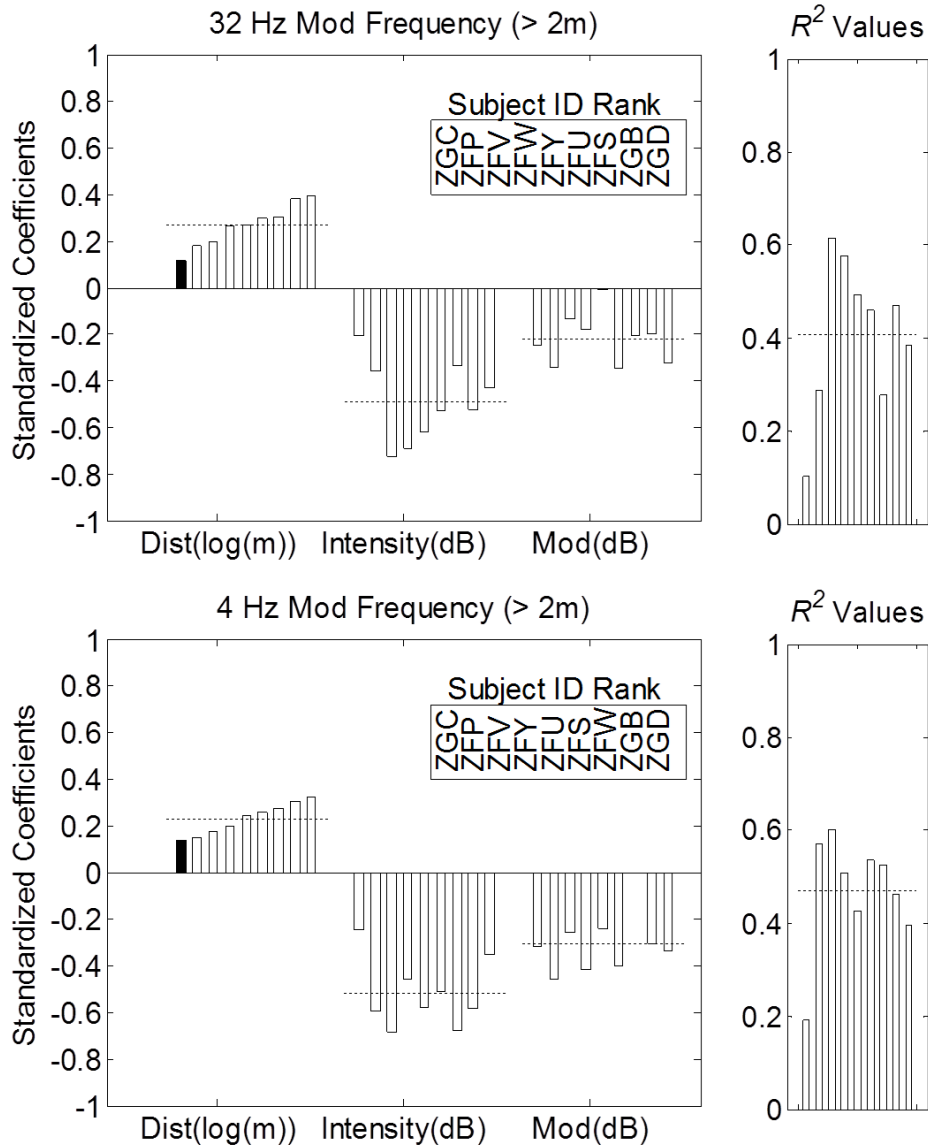
Figure 3.7: Same as figure 3.6 for the 4 Hz condition except only data from targets 2 m or greater were included in the weighting procedures.

Standardized beta coefficients were also compared between conditions for only the far target locations. Paired *t*-tests revealed there was a significant difference between *r'* standardized beta coefficients in the 32 Hz and 4 Hz conditions ($t(8) = 3.615$, $= 0.007$) for far target distances. There were no significant differences between the 4Hz and 32 Hz conditions for the intensity ($t(8) = 0.541$, $= 0.603$) and modulation depth ($t(8)$

= 1.911, = 0.092) standardized beta coefficients for far target distances. This result means that for far source distances the standardized beta coefficients for r' were smaller in the 4 Hz condition than in the 32 Hz condition. There was no significant change of R^2 values in the 4Hz and 32 Hz conditions ($t(8) = -1.453$, = 0.184) for far target distances.

4. Results Summary

In summary, the perceptual weights assigned to modulation depth were significant and approximately half of the perceptual weights assigned to intensity, a primary distance cue, when looking across either all target distances or only targets greater than 2 m from the listener. When analyzing only far target distances, where the attenuation of modulation depth is greatest from the room, r' weights were significantly smaller in the 4 Hz condition than in the 32 Hz condition. This relationship does not hold when analyzing across all source distances.

C. Discussion

Overall, the results from this experiment indicated that for monaural stimuli, the modulation depth attenuation cue improves the accuracy of distance judgments for sources beyond 2 m. Past auditory distance studies (Zahorik, 2001; Zahorik, 2002a; Zahorik et al., 2005; Anderson & Zahorik, 2014; Zahorik & Anderson 2015; Kim et al., 2015) used power function fit parameters, a and k, as measures of distance judgment accuracy. Fit parameters from the present investigation and previous studies that used monaural modulated stimuli (Zahorik & Anderson 2015; Kim et al., 2015) were compared to baseline auditory distance judgment investigations with large sample sizes (Zahorik et al., 2005; Anderson & Zahorik, 2014). Zahorik et al. (2005), collected mean fit parameters and R^2 values from across 21 auditory distance studies and reported mean

fit parameters of $a = 0.54$ and $k = 1.32$. Anderson and Zahorik (2014) collected auditory distance judgments using binaural unmodulated stimuli from a large sample of listeners ($N = 62$) and reported mean fit parameters of $a = 0.61$ and $k = 2.22$. Auditory distance studies using monaural AM stimuli reported constant values between 2.3 and 3.3(Zahorik & Anderson, 2015; Kim et al., 2015), which is similar to the reported constant values in baseline studies using binaural unmodulated stimuli (Zahorik, et al., 2005; Anderson & Zahorik, 2014). Exponents in studies using monaural AM stimuli (Zahorik & Anderson, 2015; Kim et al., 2015) are much lower (a = approximately 0.2) than studies using binaural unmodulated stimuli (Zahorik, et al., 2005; Anderson & Zahorik, 2014). The constants in the present study were in agreement with both sets of auditory distance studies; however, the exponents in the present experiment were in closer agreement with those reported in auditory distance studies using binaural unmodulated stimuli. The smaller exponents in the previous studies using monaural AM stimuli may have resulted from the use of level normalization that compensated for the 6 dB propagation loss for each doubling of source distance (Zahorik and Anderson, 2015; Kim et al., 2015). Even in the monaural unmodulated conditions of Zahorik and Anderson (2015) and Kim et al. (2015) exponents were below 0.10. The general agreement of the present fit parameters with past auditory distance studies collected using large sample sizes and binaural unmodulated stimuli (Zahorik et al., 2005; Anderson & Zahorik, 2014) indicates that the listeners in Experiment II responded almost as accurately as listeners in auditory distance baseline studies.

The R^2 values reported here were low compared to other auditory distance investigations. In auditory distance perception studies using binaural unmodulated

63

stimuli, R^2 values were measured to be 0.91 (Zahorik, et al, 2005) and 0.64 (Anderson & Zahorik, 2014). Zahorik & Anderson (2015) reported a high mean R^2 value of 0.90 and used modulated stimuli presented monaurally and level was normalized. The lower R^2 values reported here may have been due to the stimuli varying on three independent dimensions (distance, modulation depth, and intensity) while in Zahorik and Anderson (2015) level was normalized and no rove was placed on modulation depth, so stimuli only varied along two independent dimensions: physical distance and intensity rove.

Results from the weighting analysis suggest that modulation depth is a perceptually weighted, salient auditory distance cue. Large, significant R^2 values from all listeners in both conditions indicate that the multiple regression model was a good fit to the data in both conditions. The model was very similar to Zahorik (2002a) where multiple regression analysis was used to determine the perceptual weightings of intensity and D/R for auditory distance judgments. Zahorik's model included interaction terms for both intensity x r' and D/R x r', however interaction terms were excluded from the present model because they were not significant. The model used here included prediction terms for physical distance, intensity, and modulation depth. Like Zahorik's model fits all of the R^2 values from the present model fits were significant and relatively large indicating that the models were good fits to the data for both the 4 Hz and 32 Hz conditions.

Perceptual weightings for the two conditions, when data from all target distances were included in the model, have several similarities such as all terms in the model were significantly weighted by all listeners (except for the modulation depth weighting of one listener in the 32 Hz condition). In both conditions physical distance was positively weighted while intensity and modulation depth were negatively weighted.

64

The direction of these weightings were all expected based on what is known about auditory distance cues. Physical distance is expected to be positively weighted because judged distances increases with physical distance. The physical distance model term includes all distance cues in the stimulus not accounted for in the model by intensity and modulation depth, like D/R. Because D/R is a highly salient auditory distance cue (Mershon & King, 1975), the physical distance term was expected to be strongly weighted. The negative weighting intensity rove was also predicted because increasing physical distance results in decreasing intensity, as evidenced by the inverse square law where intensity decreases by 6 dB for every doubling of physical distance. Modulation depth was also expected to be negatively weighted because with increasing distance modulation depth decreases in a room based on the filter characteristic of rooms in the modulation domain (Houtgast & Steeneken, 1985). An unexpected observation from both conditions is that modulation depth is weighted approximately half of intensity. Intensity is a very salient distance cue (Mershon & King, 1975), so for this new correlate of D/R to be weighted almost half that of intensity is surprising. Previous investigations have found other possible correlates to D/R that the auditory system may use for distance judgments including temporal onset/offset duration cues (Zahorik, 2002b), spectral variance (Jetzt, 1979; Larsen, 2008), and interaural correlation (Bronkhorst & Houtgast, 1999). The high perceptual weighting of modulation depth relative to intensity provides strong evidence that it can be added to the list of correlates to D/R. However the results only apply to monaural stimuli. It remains uncertain how binaural presentation would impact the results.

The amount of benefit from the modulation depth attenuation cue may be predicted by the MTF of the room. It was determined that distance judgments were more accurate when modulation was more attenuated by the room. Based on the modulation gain analysis in figure 3.2b (4000 Hz center frequency) there was relatively little difference in modulation attenuation between the 4Hz and 32 Hz modulation conditions for near distances – at most 1 dB difference between 32 Hz and 4 Hz modulation. There were no significant perceptual weight changes between the 4 Hz and 32 Hz conditions when distance judgments from all target source distances were included in the regression model. The modulation gain analysis indicates that larger modulation gain differences between the two modulation rates exist at farther distances. At the farthest target distance (5.6 m) the modulation depth in the 32 Hz condition is about 6 dB more attenuated than the 4 Hz condition. When the multiple regression analysis was performed using only judgments from physical distances 2 m or more away, an interesting weight change emerged between conditions: r' weights were significantly larger in the 32 Hz condition than in the 4 Hz condition. The unstandardized coefficient for the r' parameter is mathematically identical to the alpha parameter, or slope value, from the power function fits. This indicates that distance judgments were more accurate and less compressed when modulation was more attenuated by the room (at far distances in the 32 Hz condition). The acoustic parameter space where modulation was attenuated most appears to provide the most benefit to listeners making distance judgments. This offers evidence that the room's MTF can be used to predict the modulation rate that will most benefit auditory distance judgments based on the amount of modulation depth attenuation at that modulation rate, for the stimuli and parameters used here. Since the intensity and

66

modulation depth weights remain unchanged between conditions, even at far distances, this means that the weights placed on intensity and modulation depth are independent of the amount of modulation depth attenuated by the room. This is important because it indicates that listeners did not optimally use the modulation depth cue since the perceptual weighting of modulation depth would be expected to increase when it is more correlated with physical distance (as in the 32 Hz condition). While listeners use modulation depth as a distance cue they may not understand how modulation depth is associated with physical distance.

In Kim et al, (2015) and Zahorik and Anderson (2015) listeners performed distance judgment tasks were stimuli were either modulated or unmodulated. Both studies observed that judgments were more accurate in the modulated condition than in the unmodulated condition. The results from the present experiment expanded on these previous studies in several ways: 1) The stimuli used here were not normalized to compensate for the for the propagation loss of 6dB per source distance doubling. Therefore intensity cues were preserved by allowing source level to vary with distance; 2) the relative contribution of modulation depth was measured using the perceptual weighting paradigm; and 3) the acoustic parameters that facilitate the use of AM as a distance cue were determined.

The use of monaural stimuli qualifies the results from Experiment II. Under naturalistic listening conditions, input is present at both ears. It is unknown whether modulation depth will remain an effective cue under binaural presentation. More distance information is available under binaural presentation in the near-field where ITD and ILD cues are present as a function of distance. Zahorik and Anderson (2015) included a

condition in which listeners performed the distance judgment task using modulated and unmodulated binaural stimuli. Their results indicated there was no difference between the binaural modulated and unmodulated conditions, indicating the modulation depth cue was not effective when stimuli were presented binaurally. However, their stimuli were normalized for level, so it is uncertain how the results would change using the stimuli in the present experiment with binaural presentation where more distance information is available to the listener.

D. Conclusions

A few main conclusions can be drawn from this experiment. First, power functions were good fits to listener's data in both the 4 Hz and 32 Hz conditions. Second, in both the 4 Hz and 32 Hz conditions modulation depth is weighted approximately half that of intensity. Third, the MTF of a room can be used to predict the modulation rate that will benefit auditory distance judgments because the r' weight changes between the 4 Hz and 32 Hz conditions for sources more than 2 m from the listener, which coincides with a region of greater modulation attenuation. These conclusions together provide strong evidence that modulation depth cues are used by listeners when estimating sound source distance monaurally; however, the modulation depth cues are not optimally used by listeners.

CHAPTER IV

EXPERIMENT III

Unlike auditory distance perception, directional auditory localization in the

horizontal plane (azimuth angle) is highly accurate with a minimum audible angle as

small as 1° at midline (Mills, 1958). In general, reverberant environments have been

found to reduce the accuracy of directional auditory localization judgments for

unmodulated signals (Hartmann, 1983; Giguere & Abel, 1993; Ihlefeld & Shinn-

Cunningham, 2011). However, the manner in which MTFs attenuate modulation depth as

a function of azimuth may provide additional acoustic information related to direction at

the two ears. This in turn could potentially increase directional localization accuracy in

reverberant environments for modulated signals.

To understand the additional acoustic information present for modulated

signals in rooms, consider a source positioned 90° to a listener's side where interaural

modulation depth difference is greatest. The modulation depth is more attenuated at the

contralateral ear due to that ear having more exposure to the modulation depth

attenuating reverberation of the room. As a sound source moves toward midline, the

interaural modulation depth difference of the signal reaching the ears is reduced because

at midline the ears have the same amount of exposure to the room. The extent to which

this additional acoustic information can lead to changes in directional localization

performance in reverberation is unknown. Experiment III tests the hypothesis that

modulated source signals in a reverberant environment can improve directional

localization accuracy.

A. Methods

1. Participants

For Experiment III, 12 listeners (10 female) ranging in age from 20 to 32

years old participated. Nine of the listeners also participated in Experiment II. Listeners

were recruited through flyers, email advertisements, and personal contacts. All listeners

had normal hearing as verified by audiograms with less than 25 dB HL at octave

frequencies between 250 and 8000 Hz. Listeners also had normal central auditory

processing as verified by dichotic digits (Musiek, Gollegly, Kibbe, & Verkest-Lenz,

1991) and masking level difference tests (Wilson, Zizz, & Sperry, 1994). Five of the 12

listeners were left out of analysis because of task compliance issues. The criterion for

exclusion from analysis was that the more than 10% of the listener's judgments for

targets more than 15° from midline were placed in the wrong left-right hemisphere. This

was a good indicator of task compliance especially since all target sources were in the left

hemisphere. Listeners were compensated in the form of cash payments. All testing was

approved by the University of Louisville Internal Review Board.

2. Stimuli

BRIR Generation. BRIRs were recorded in Bigelow Hall using the same

specifications described in Experiment I. Measurements were recorded 4.88 m from the

measurement microphone in -15° steps from 0° to 180° with recordings to the left of

KEMAR. For the anechoic condition BRIRs were windowed so only the direct portion of

the BRIR remained. The start of the direct energy of the BRIR was determined by finding

the first time the energy exceeded 20 dB above the noise floor, and the duration of the direct portion was defined as 2.5 ms from the start of the direct energy (Zahorik, 2002b). ITDs were maintained in the BRIR during extraction of the direct portion.

Stimulus Parameters. The MTF of the room was measured using the same techniques described in Experiment II. Figure 4.1 displays the MTF of the BRIRs described above. The left column displays MTFs from the left (ipsilateral ear) and the right column displays MTFs from the right (contralateral ear). The top panels display MTFs measured at 0°, and the bottom panels display MTFs measured at -90° to the listener's left. Each panel displays the modulation gain of the room at varying center frequencies as a function of modulation frequency. Cooler colors denote greater modulation removed by the room. Moving from midline to -90° interaural modulation depth difference increases. At 0° only very small interaural modulation depth differences were noticeable with careful observation. However, at -90° interaural modulation depth differences are more visible especially around 3000 Hz. In the right ear there are some blues present (greater modulation depth attenuation) at 3000 Hz for modulation frequencies above 16 Hz, while in the left ear the same region remains red (very little modulation depth attenuation).
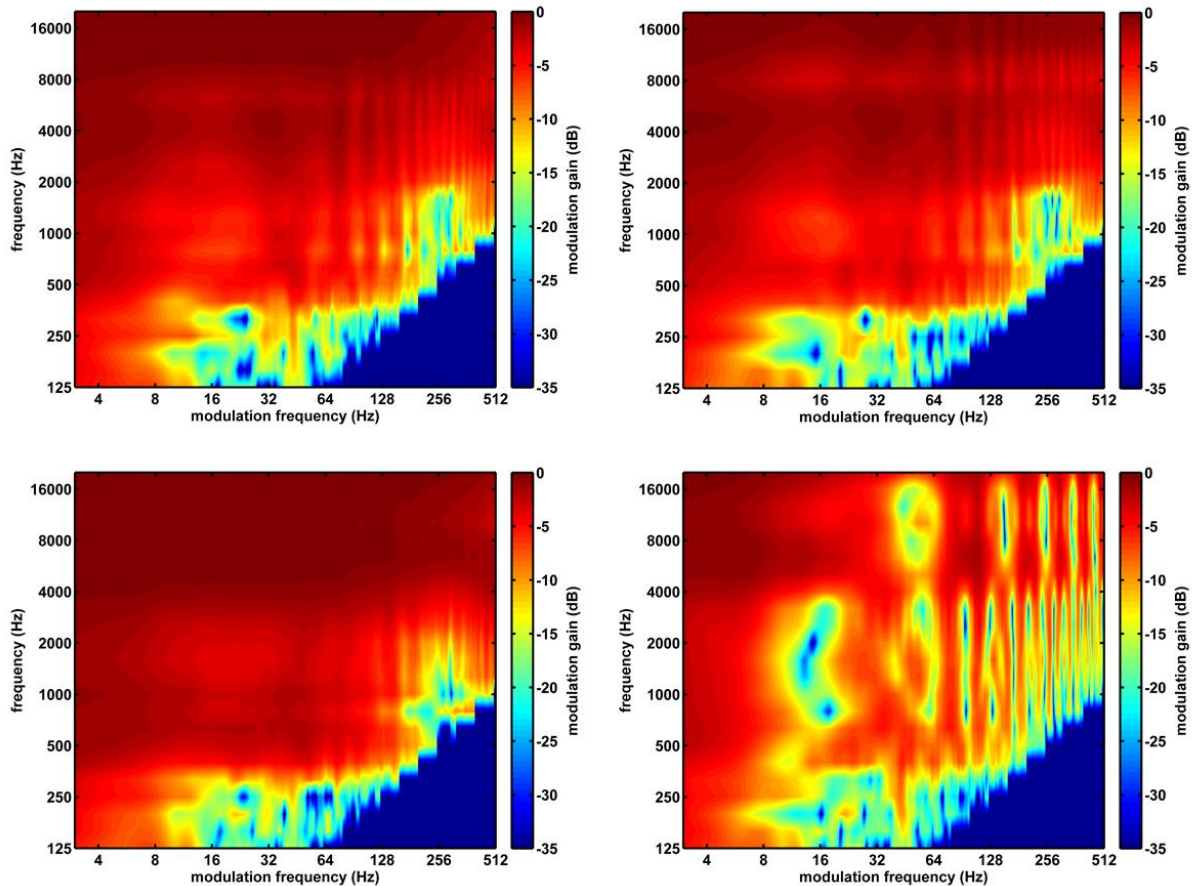
Figure 4.1: Same as figure 3.1. MTFs for sources as a function of azimuth for the ipsilateral (left) and contralateral (right) ears. The source location is 4.88 m. Source azimuth is 0° in the top row and -90° on the bottom row.

Figures 4.2a-b and 4.3a-b, derived from the MTFs, display modulation gain as a function of azimuth at the ipsilateral (left) and contralateral (right) ears. These plots were used to make further decisions about stimulus parameters. The center frequencies in 4.2a-b are 2000 Hz (top) and 4000 Hz (bottom). The center frequencies in 4.3a-b are 6000Hz (top) and 8000 Hz (bottom). Each panel in figures 4.2 and 4.3 displays the modulation gain for the signal at modulation frequencies of 4 (red), 8 (green), 16 (blue), 32 (black), and 64 Hz (cyan). The rebound in the in the amount of modulation gain for

highly lateralized source locations may be due to early reflections. ITDs are known to be a primary directional localization cue in the horizontal plane (Wightman and Kistler, 1992). Listeners are less sensitive to high frequency ITD cues (Moore, 2008, p. 236), so higher center frequencies, than those used in Experiment II, were preferred in for this experiment to limit directional localization information from ITDs. The ILDs, however, were not limited.

By limiting the ITD cue, other cues, like interaural modulation depth differences, may be used instead. A higher modulation rate is also necessary because of the low-pass characteristic of rooms in the modulation domain. With these two stimulus requirements in mind, an 8000 Hz center frequency was chosen for all azimuth stimuli and a 64 Hz modulation frequency was chosen for the modulated condition. These same stimulus parameters were used to generate stimuli in the no modulation condition.
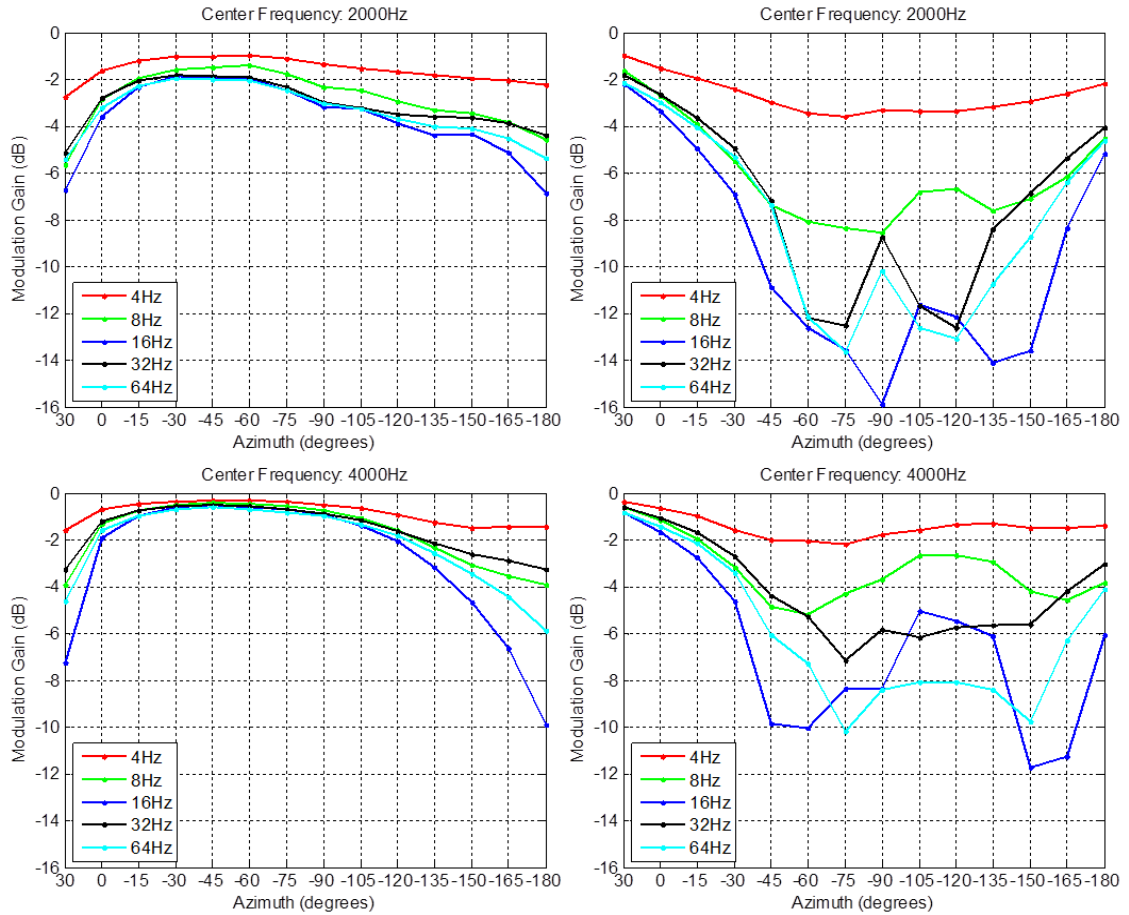
73

Figure 4.2: Same as figure 3.2 for locations as a function of target azimuth for center frequencies 2000 Hz (top) and 4000 Hz (bottom) for the ipsilateral (left) and contralateral (right) ears.
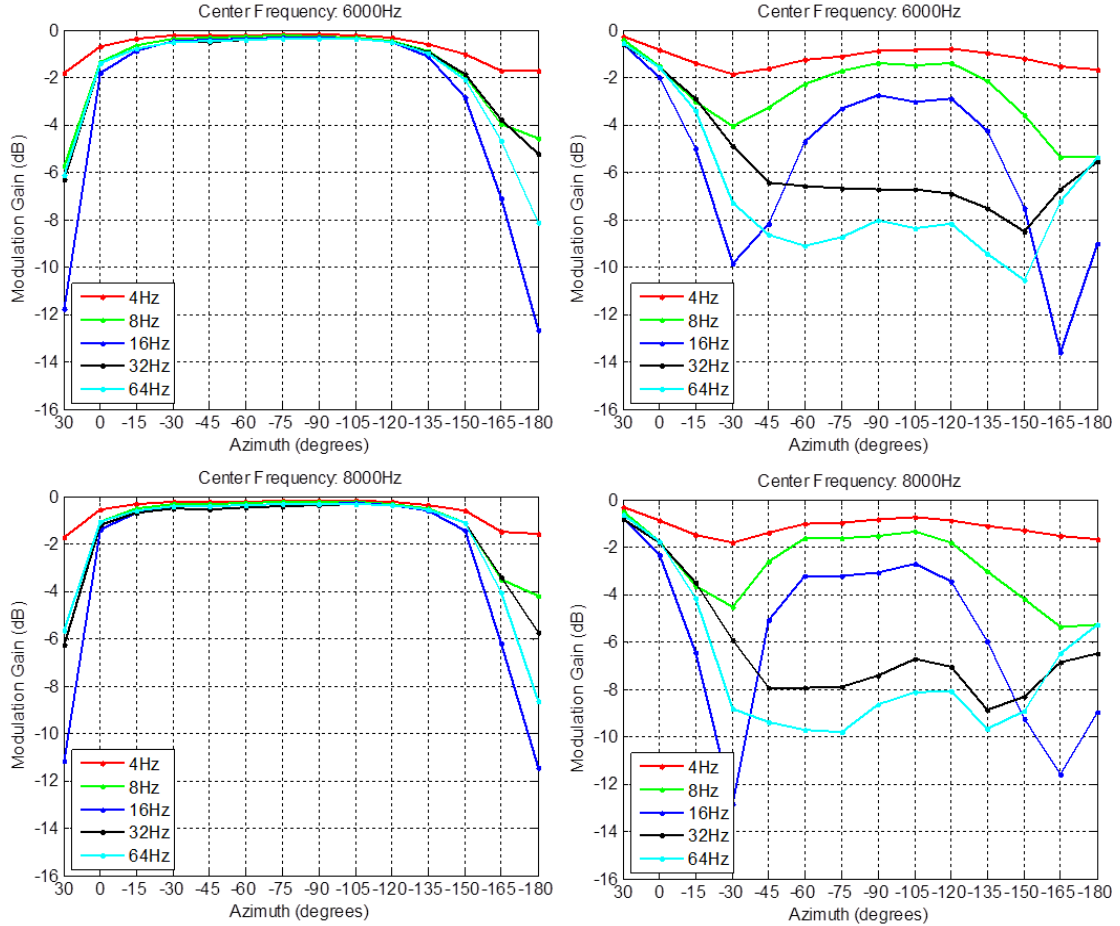
Figure 4.3: Same as figure 4.2 for center frequencies 6000 Hz (top) and 8000 Hz (bottom).

The source signal was composed of a 1-octave Gaussian noise sample centered at 8000 Hz with a 2000 ms duration. The signal duration includes a Hanning window of 500 ms rise/fall time. The source was either unmodulated or modulated at 64 Hz before convolution with the BRIR. All stimuli were presented binaurally over headphones.

3. Design

Listeners were tested using a 2 x 2 within-subjects design where directional localization judgments of a target source were made in anechoic and reverberant conditions with modulated and unmodulated signals. The entire experiment took place inside a double-walled sound proof booth (Acoustic Systems, Austin, TX). Listeners

75

responded using the polar plot GUI as described in Experiment I. Listeners were told to judge both the azimuth and distance to the target, however only azimuth judgments were analyzed. Each level of each condition was tested within its own block (modulated and unmodulated anechoic; modulated and unmodulated reverberant) which included 10 judgments for each of the 13 target locations (130 trials x 4 blocks). A limitation of the design was that the anechoic condition was run before the reverberant condition with the intention of collecting baseline data from listeners before reverberant conditions. Anechoic responses were collected at least 5 weeks before responses in the reverberation condition. The ramifications of this will be discussed in more detail below. The order of modulation presentation within each environment was counterbalanced and the order of trials within each block was randomized. No feedback was provided to the listeners.

4. Procedure

Auditory central processing disorder screening was performed before listeners participated. The screening included dichotic digits (Musiek, Gollegly, Kibbe, & Verkest-Lenz, 1991) and masking level difference testing (Wilson, Zizz, & Sperry, 1994). Like in Experiment I, at the beginning of each block the listener chose to respond using either feet or meters and the GUI would update to reflect their choice. Listeners could listen to the stimulus as many times as they wished before making a response. Custom MATLAB software was used for stimulus presentation and data collection.

Data Analysis. Data from directional conditions were analyzed similarly to the azimuth data in Experiment I using unsigned angular errors with front/back errors resolved and recorded. Unlike Experiment I, angular error was measured for each target azimuth. Subsequent analyses on angular errors were divided into three regions based on

amount of mean angular error across listeners for each target azimuth (Wightman &
Kistler, 1989b): front (0-45°), side (60-120°) and back (135-180°). Angular error was
analyzed between conditions using a 2 x 2 within-subjects ANOVA for each of the three
regions. The distribution of front/back errors between conditions was analyzed between
conditions using Chi-square tests of independence.

B. Results

1. Auditory Processing Disorder Analyses

As in Experiment 2, listeners' central auditory processing was tested using
dichotic digits (DD; Musiek, Gollegly, Kibbe, & Verkest-Lenz, 1991) and masking level
difference (MLD; Wilson, Zizz, & Sperry, 1994) screening tests. Listeners' performances
on these tests are reported in table 3.1. All listeners in Experiment III had normal results
(score >90% i.e., missing 2 or fewer digits) for the DD test for both the left ($M = 95.0\%$,
$SD = 3.78\%$) and right ear ($M = 95.0\%$, $SD = 3.78\%$). All listeners in Experiment III had
normal results (MLD $\geq$ 6 dB) for the MLD test ($M = 7.43$; $SD = 1.90$). Based on these
tests all listeners in Experiment III had normal central auditory processing. Like
Experiment II, the small variability of responses and the small number of observations for
each measure made it difficult to meaningfully relate these results to further localization
results from this experiment.

77

Table 3.1. Measurements from the dichotic digits and masking level difference screening tests for each listener in Experiment III. For the dichotic digits test, the number of items missed for each ear is reported. Normal performance in the dichotic digits task is defined as missing 2 or fewer items at each ear. For the masking level difference test, the masking level difference (in dB) is reported. Normal performance in the masking level difference test is a 6 dB or greater masking level difference. Means and standard deviations for each measure are at the bottom of the table.

	Dichotic Digits		
Subject ID	Left Ear	Right Ear	MLD (dB)
ZFS	1	2	8
ZFW	0	1	6
ZFU	1	2	6
ZGB	2	1	10
LHI	1	0	10
ZGA	0	1	6
ZGD	2	0	6
Mean	**1.0**	**1.0**	**7.4**
Std	**0.8**	**0.8**	**1.8**

2. Double Pole Coordinate Plots

Double pole scatter plots of azimuth judgments for left-right and front-back angles in both reverberant and anechoic conditions, similar to figure 2.6 in Experiment I, were plotted for all participants in figures 4.4 through 4.10. Each plot includes subject I.D., and data from both the modulated (red) and unmodulated (blue) conditions. The diagonal dashed line represents a perfectly accurate relationship between double pole coordinates of the target azimuths and judged azimuths. Small random jitter was added to the target azimuths on the x-axis for visualization purposes. Assessing these figures it is apparent that there are a few different response patterns between listeners in all four

78

conditions. For example listener ZFU, in figure 4.6 (top), appears to have an effect of modulation in reverberation where front-back judgments in the unmodulated condition were concentrated in the front hemisphere, while their judgments in the modulation condition were mostly in the back hemisphere. And Listener ZGB, in figure 4.7, appears to have an effect of reverberation where their left-right responses in anechoic were less variable than in reverberation.
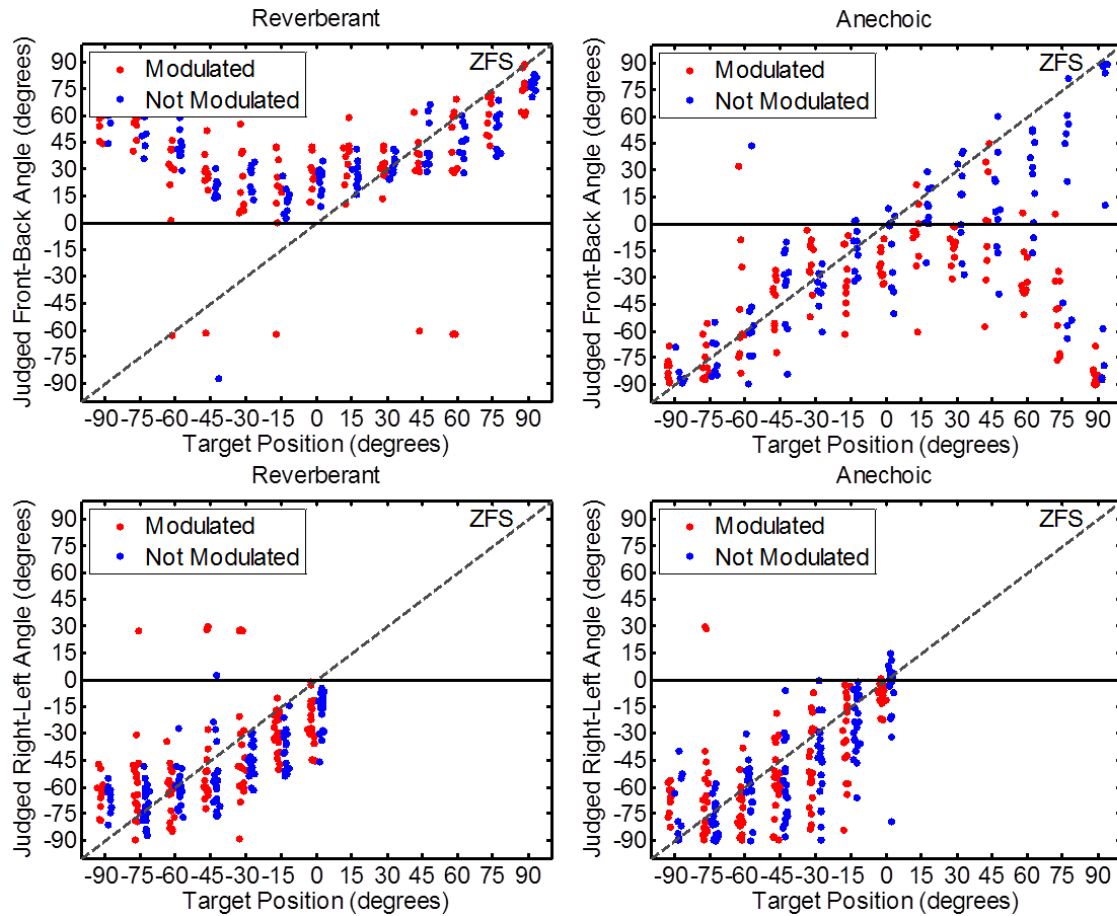
Figure 4.4: Azimuth data in the form of Double-Pole coordinates from listener ZFS for front-back angle (top) right-left angles (bottom). Data from the reverberant environment are plotted on the right, and data from the anechoic environment are plotted on the left. Red dots indicate raw azimuth judgments transformed to double-pole coordinates from the modulated condition: 10 replications/target azimuth. Blue dots show raw judgments transformed to double-pole coordinates from the no modulation condition: 10 replications/target azimuth. Perfectly accurate performance is indicated by the dotted line in each panel

Figure 4.5: Same as Figure 4.4 but for listener ZFW.
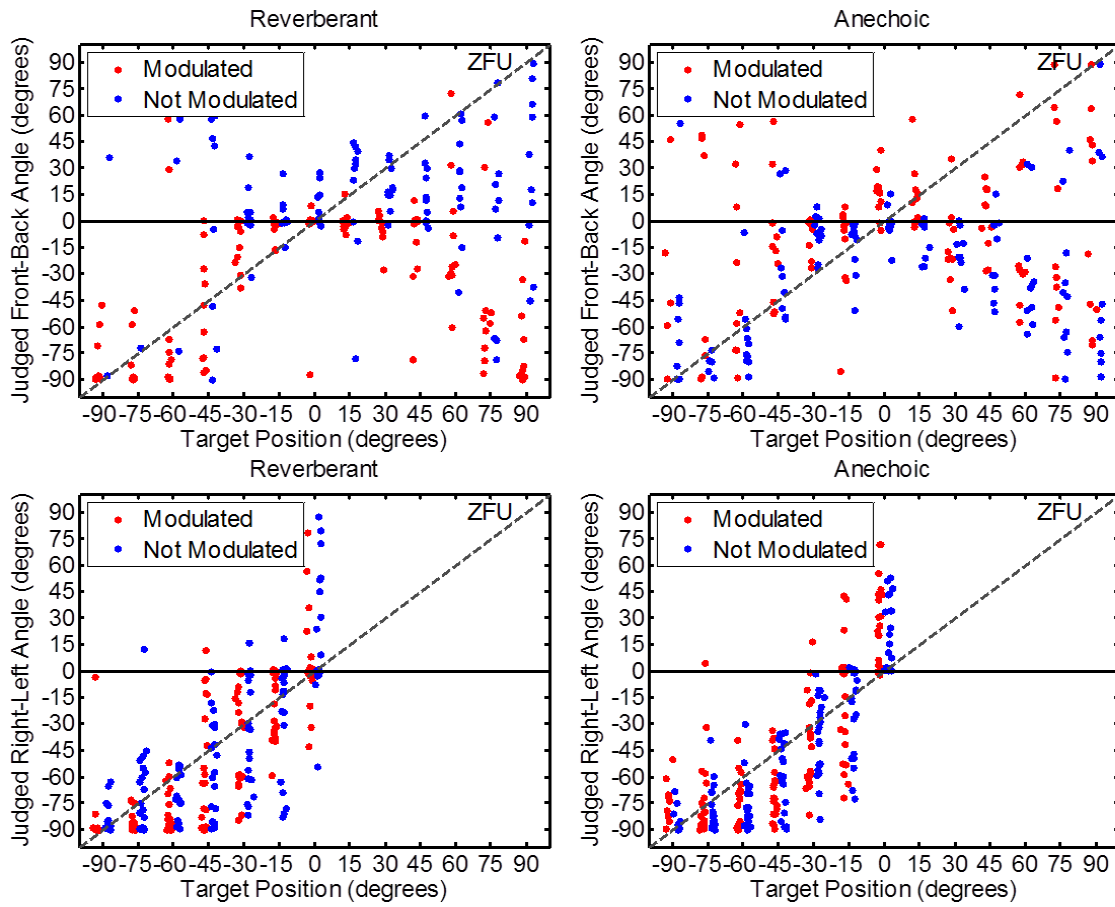
Figure 4.6:Same as Figure 4.4, but for listener ZFU
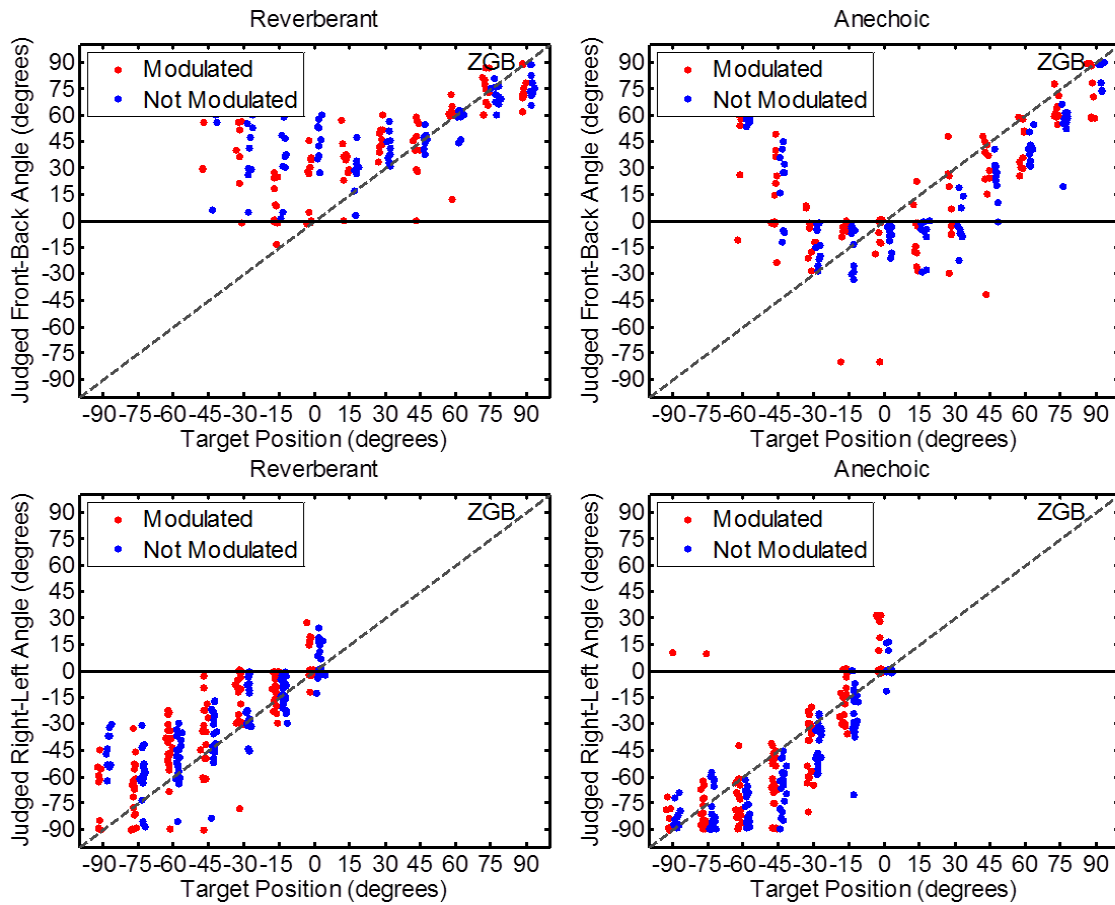
Figure 4.7: Same as Figure 4.4, but for listener ZGB.
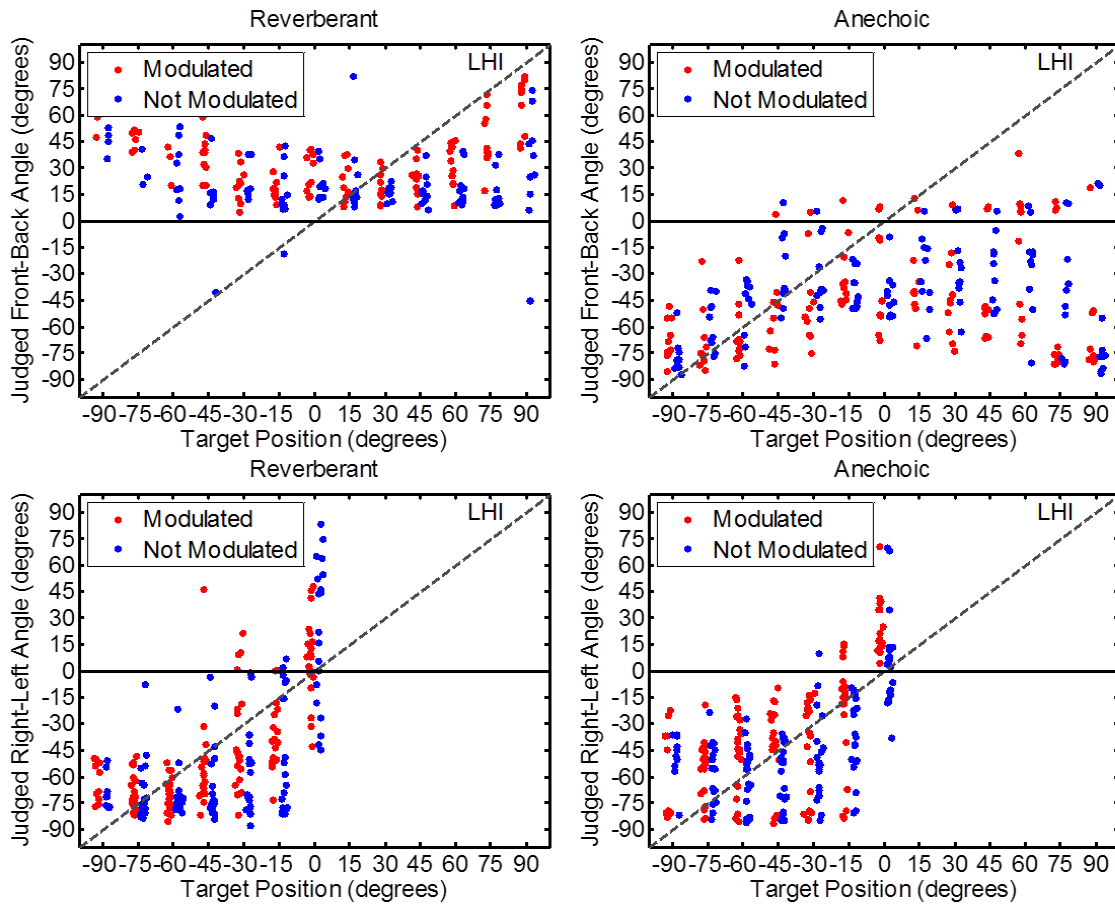
Figure 4.8: Same as Figure 4.4, but for listener LHI.
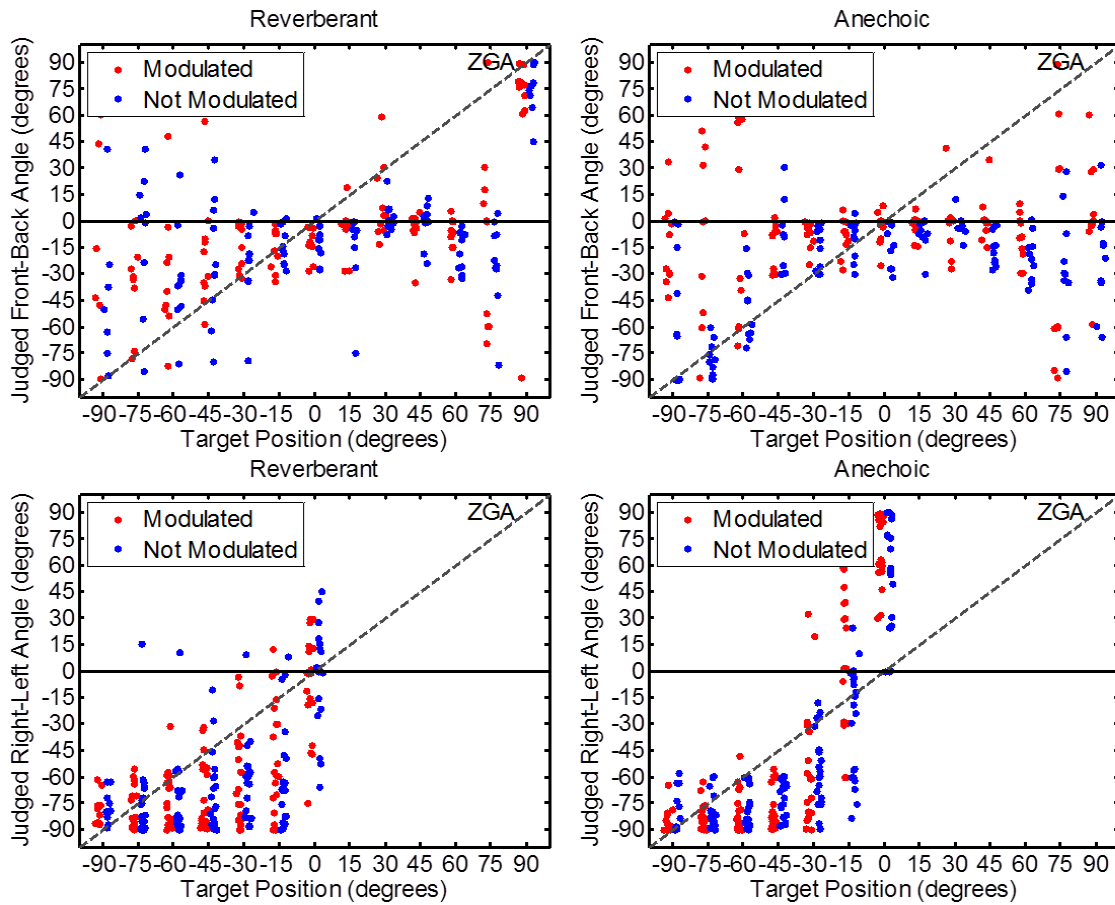
Figure 4.9: Same as Figure 4.4, but for listener ZGA.

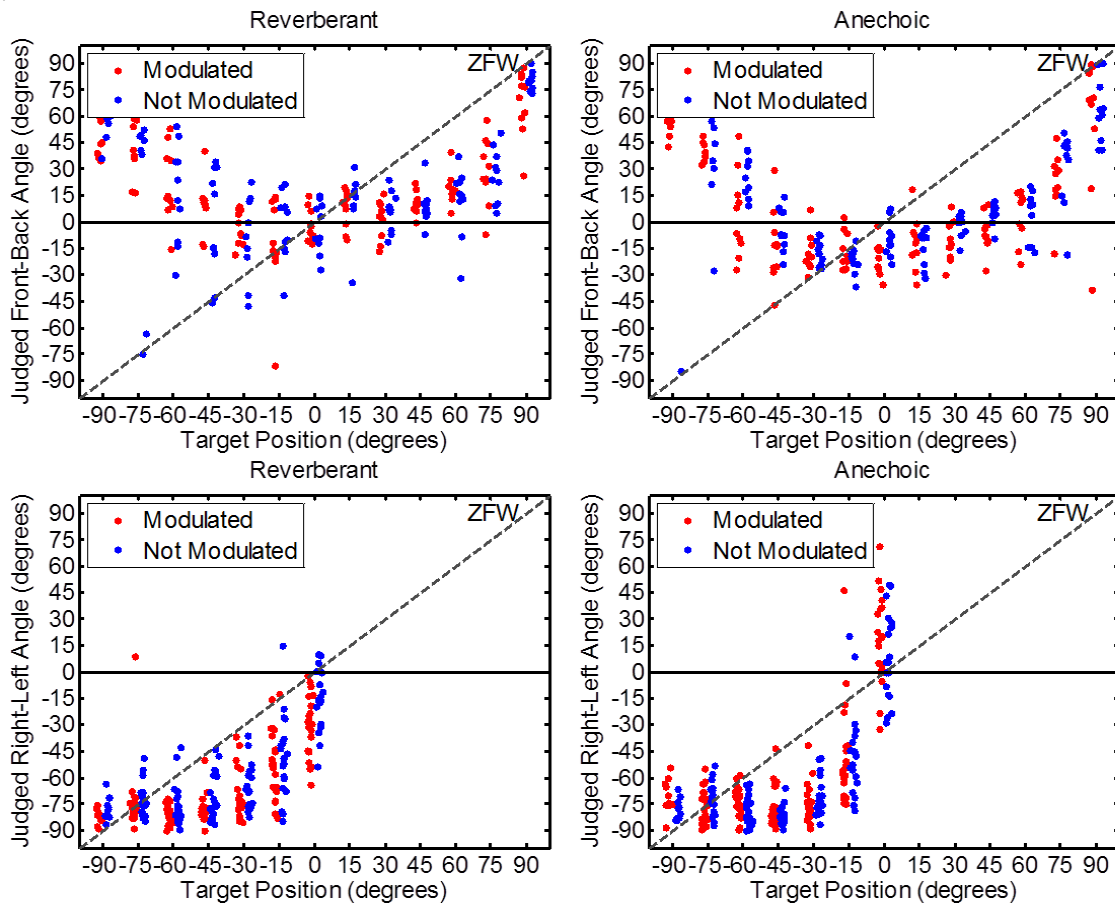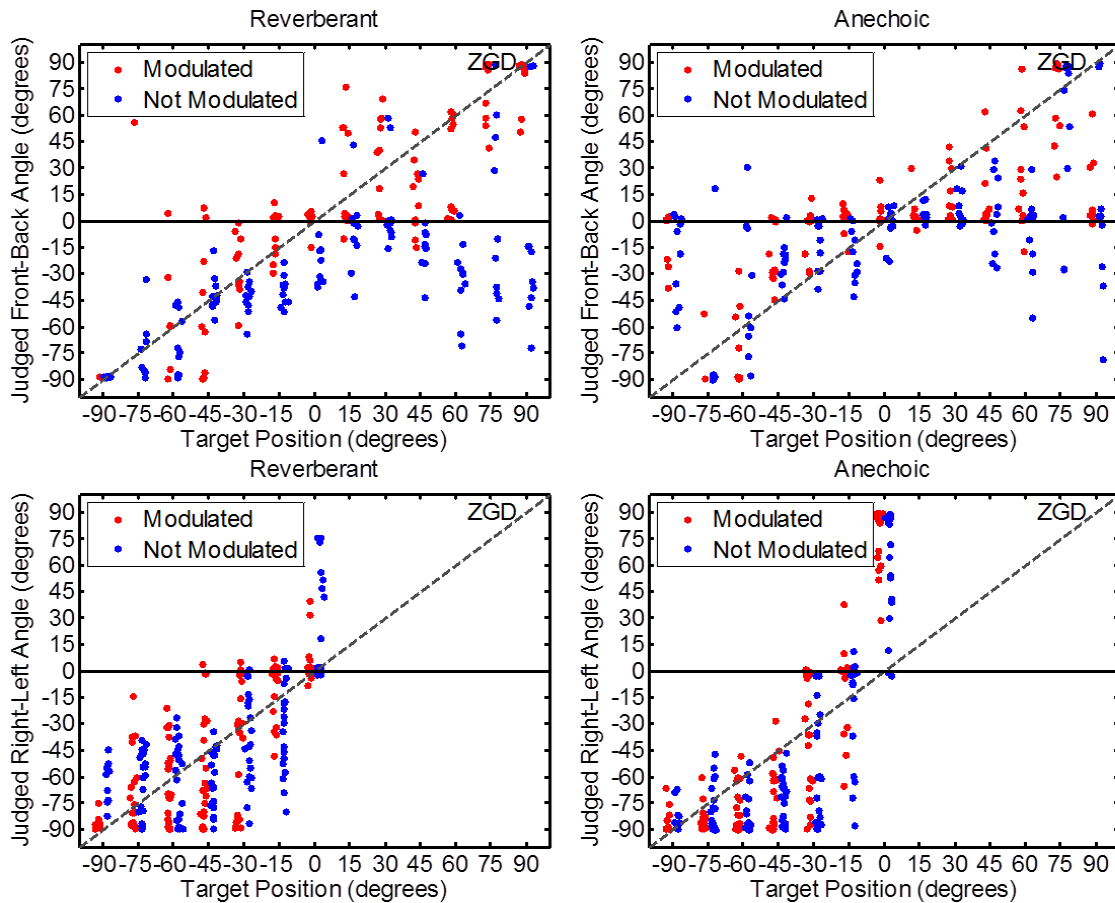Figure 4.10: Same as Figure 4.4, but for listener ZGD.

Across listeners, there are some indicators that listeners are accurately performing the task. Based on responses in the left-right angle plots almost all responses are correctly placed in the left hemisphere. The exception is near midline where some listeners appear to have difficulty discriminating between hemispheres. In the front-back plots there are a few consistent response patterns that emerge between participants. Listeners who show a 'U' or inverted 'U' pattern, like ZGB, place all of their responses in one hemisphere but are more accurate near 0°. Front/back reversals are common in azimuth studies especially when stimuli are generated using non-individualized HRTFs (Wenzel, Arruda, Kistler, & Wightman, 1993). Their study found that listeners using

nonindividualized HRTFs had an average increase in front-back reversals of 12.2% when compared to free-field localization.

3. *Angle of Error Analyses*

Figure 4.11 shows the mean average angle of error for all participants at each target azimuth for each condition. Dots represent the individual mean for each participant while the circles represent the mean azimuth error pooled across all participants, with standard error bars included. Across the four panels it is evident that mean angle of error is dependent on the region in which the target lies. These three regions (front, side, and back) are divided by vertical lines in each panel. Most notable when comparing across region is that angular error decreases in the side locations (-60 to -120°). For this reason, subsequent analyses and figures for angle of error were pooled across region. Similar procedures have been used previously in the directional localization literature (e.g. Wightman & Kistler, 1989b).
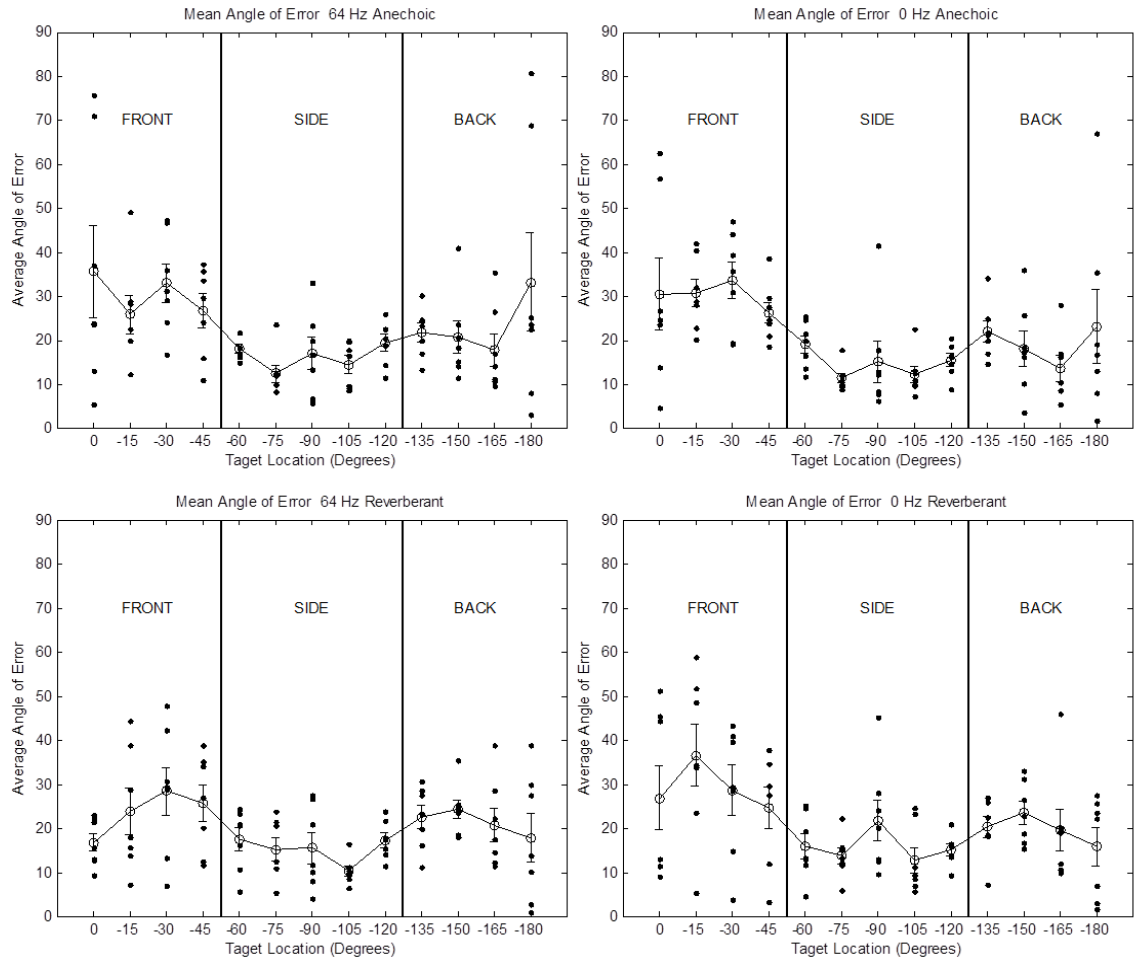
Figure 4.11: Individual (n = 7) average angle of error (computed with reversals resolved, see text for details) as a function of azimuth for the no modulation in anechoic (top right), 64 Hz modulation in anechoic (top left), no modulation in reverberation (bottom right), 64 Hz modulation (bottom left). Dots indicate individual means. Circles indicate group means with standard error bars. Vertical lines in each plot separate the front (0° - -45°), side (-60° - -120°), and back (-135° - -180°) regions.

Figures 4.12, 4.13, and 4.15 display grouped bar charts of mean angle of error for all conditions for the front, side, and back region respectively. Each grouping represents a performance of an individual listener with the right-most grouping displaying the mean (with error bars representing standard deviation) across all listeners. Conditions are represented by differently shaded bars (black = modulated anechoic; dark gray = unmodulated anechoic; light gray = modulated reverberation; white = unmodulated reverberation). These plots are used to call attention to response strategies between listeners.
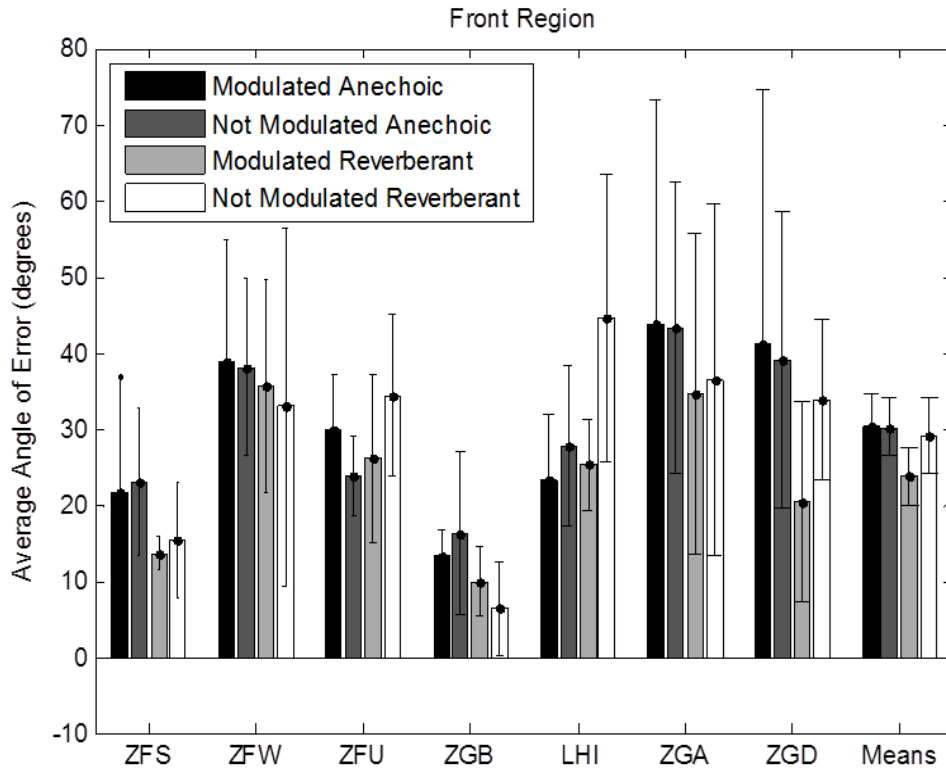
Figure 4.12: Average angle of error (degrees) for each listener and across all listeners in each condition for the front region. Each group of bars represents a listener and includes error bars representing 95% confidence intervals. The right most group displays mean angle of error across listeners with standard error bars. Conditions are displayed by shaded colors (Black = modulated anechoic; Dark Gray = not modulated anechoic; Light gray = modulated reverberation; White = not modulated reverberation) within each group.
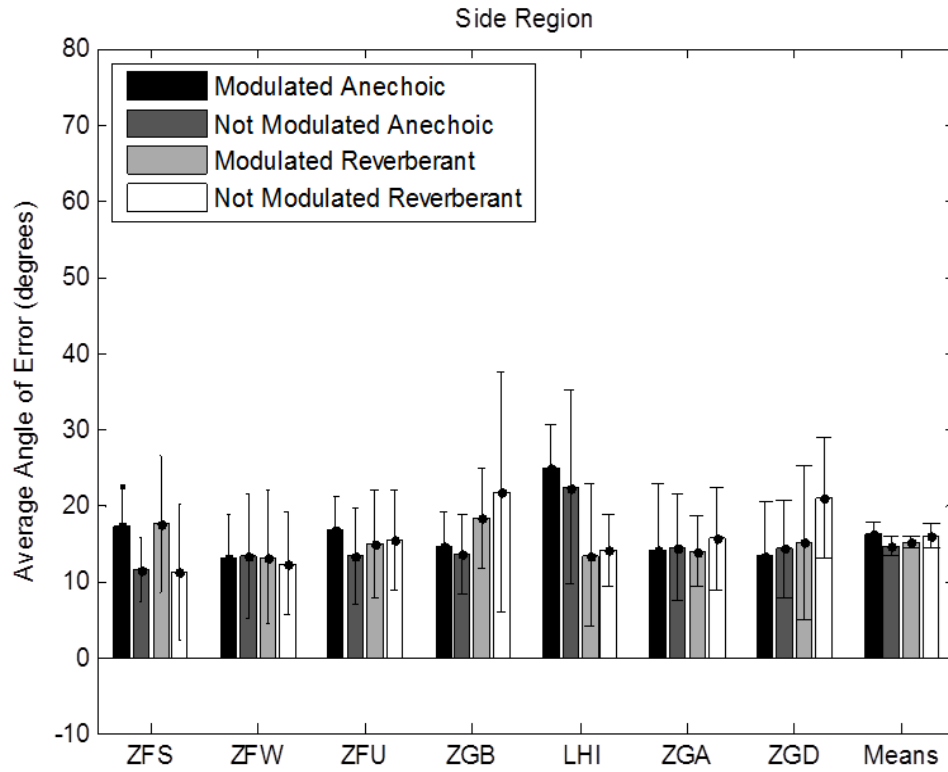
Figure 4.13: Same as figure 4.12, but for the side region.
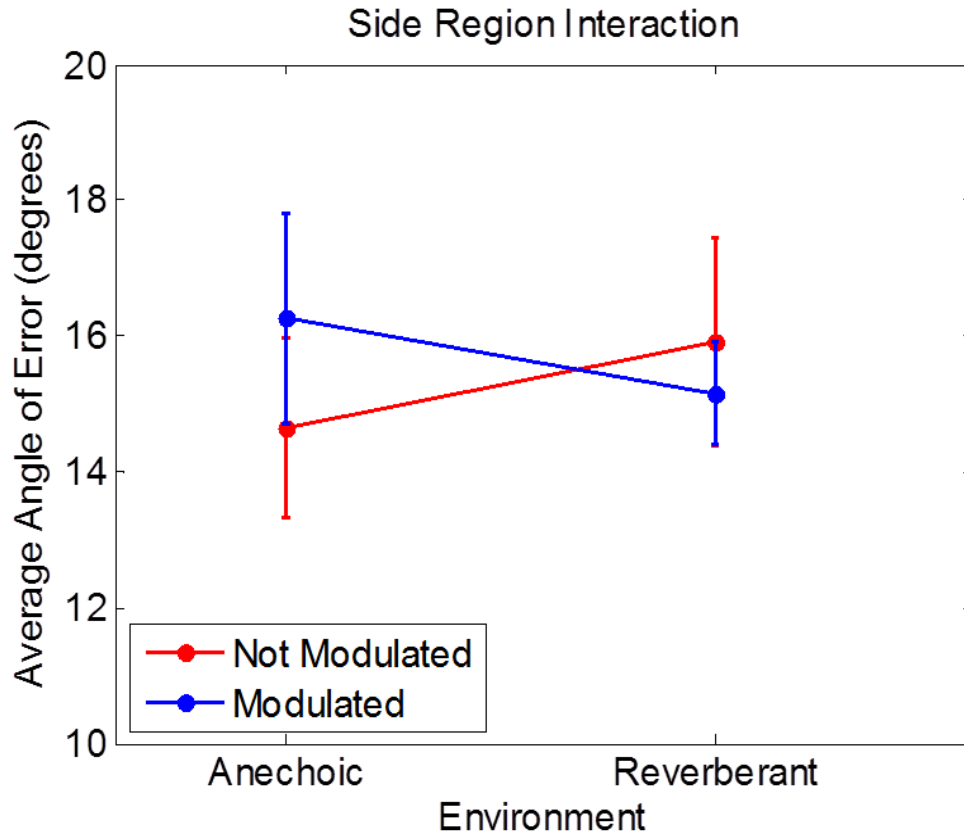
Figure 4.14: Interaction of average angle of error in the side region. Data from the not modulated conditions are plotted in red, and data from the modulated conditions are plotted in blue. Error bars represent standard errors. Each environment is placed on the x-axis.
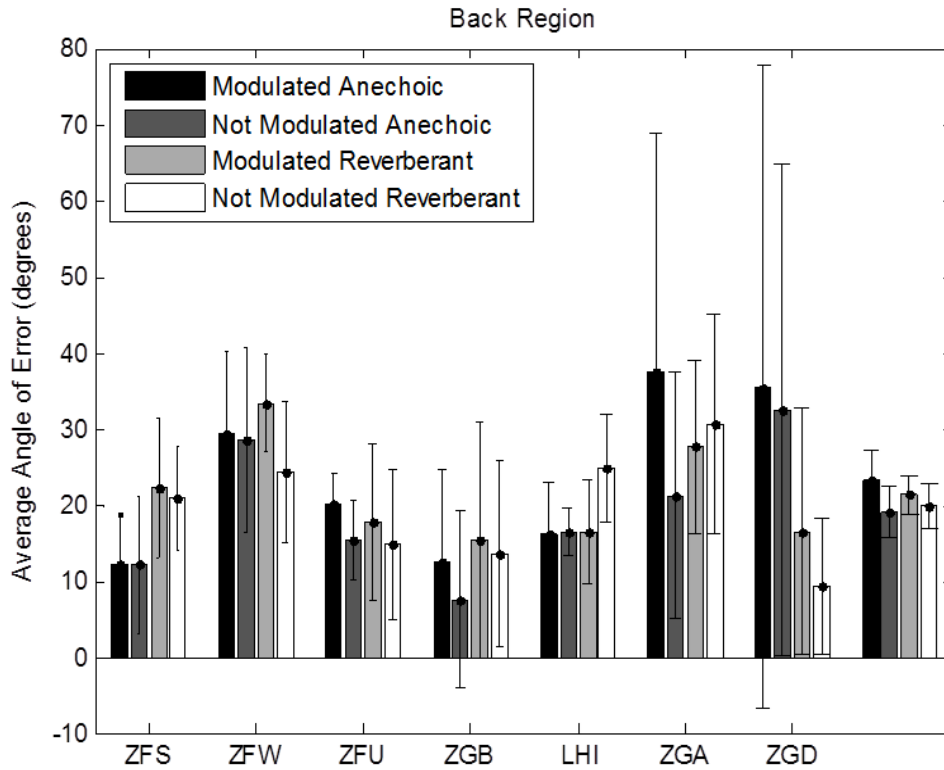
Figure 4.15: Same as figure 4.12, but for the back region.

In the front region, individual differences between listeners are prominent; however a few patterns can be obtained from figure 4.12. Five listeners (Subject IDs: ZFS, ZFW, ZGB, ZGA, and ZGD) appear to be using a response strategy related to reverberation. A within-subjects 2 x 2 ANOVA was performed on the group means within the front region. There was no significant main effect for reverberation ($F(1,6)$ = 1.720, = 0.238) or modulation ($F(1,6)$ = 2.375, = 0.174) and no significant interaction ($F(1,6)$ = 2.550, = 0.161). Therefore in the front region modulation and reverberation have no effect on angular error.

The side region in figure 4.13 has noticeably less average angle of error and less variability than the front and back regions. Listeners ZFS and ZGA appear to use a response strategy related to modulation while listeners ZGB and LHI appear to use a

response strategy related to reverberation. A within-subjects 2 x 2 ANOVA was performed on the group means within the side region. There were no main effects (reverberation: $F(1,6) = 0.002$, $p= 0.969$; modulation: $F(1,6) = 0.141$, $p= 0.720$), but there was a significant interaction between modulation and reverberation ($F(1,6) = 6.215$, $p= 0.047$). The interaction is displayed in figure 4.14 with environment plotted on the x-axis and modulation condition plotted with difference colored lines (red: not modulated; blue: modulated) In the side region there was more angular error for the unmodulated stimulus in a reverberant environment than in anechoic, and for modulated stimuli there was more error in the anechoic environment than in reverberant environment. Post-hoc simple effects were performed comparing angular errors between the anechoic and reverberant environments for both modulated stimuli and for unmodulated stimuli using paired t-tests. For modulated stimuli there was no significant difference between the anechoic ($M = 16.254$, $SD = 4.117$) and reverberant ($M = 15.050$, $SD = 2.010$) conditions, $t(6) = 0.597$, $p = 0.573$. For unmodulated stimuli, there was also no significant difference between the anechoic ($M = 14.656$, $SD = 3.506$) and reverberant ($M = 15.906$, $SD = 4.064$) conditions, $t(6) = 0.609$, $p = .565$. Simple effects were also analyzed between the reverberant and anechoic environments. There was no significant difference between the modulated ($M = 15.159$, $SD = 2.010$) and not modulated ($M = 15.906$, $SD = 4.064$) conditions in the reverberant environment, $t(6) = 0.515$, $p = 0.625$. In the anechoic environment there was no significant difference between the modulated ($M = 16.254$, $SD = 4.117$) and not modulated ($M = 14.656$, $SD = 3.506$) stimuli, $t(6) = -1.682$, $p = .144$. The lack of significant simple effects may be due to power issues that result from a small sample size.

94

A few response patterns emerged in the back region in figure 4.14, but they are not as well defined as in the front region. Three listeners (Subject ID: ZFS, ZGB, ZGD) appear to be using a response strategy related to reverberation while listeners ZFW and ZFU appear to be using a response strategy related to modulation. A within-subjects 2 x 2 ANOVA was performed on the group means within the back region. There was no main effect for reverberation ($F(1,6) = 0.031$, $p= 0.865$) or modulation ($F(1,6) = 4.313$, $p = 0.083$) and no significant interaction ($F(1,6) = 0.606$, $p= 0.466$).

Given the variability within individual listeners, 95% confidence intervals comparing conditions within each listener were plotted in figures 4.12, 4.13, and 4.15 (front, side, and back regions respectively) to examine whether individual listeners showed effects of reverberation and/or modulation. In each figure anechoic conditions are displayed with square symbols and reverberant conditions are displayed with circles. The symbols are filled for modulated conditions and unfilled for unmodulated conditions. The symbols represent individual means and the error bars are 95% confidence intervals. Overlapping bars within a subject indicate that there was no significant difference between overlapping conditions.

In all three regions confidence intervals overlap for all listeners in all conditions. The only exception is listener LHI in the side region where the modulated anechoic condition does not overlap with the not modulated reverberant condition. For this listener the modulated anechoic condition has more angular error than the unmodulated reverberant condition. The lack of individual effects confirms the results from the ANOVAs performed above that there were no effects of reverberation or modulation for directional localization. The absence of significant simple effects for the

interaction in the side region, and the large variability both between and within listeners may have been the result of low power from the small sample size in this experiment.

4. Front-Back Reversal Analyses

Table 3.2 displays the percentage of reversals for each condition. Chi-square tests of independence were performed for each participant to determine whether the number of front-back reversals differed between conditions. Table 3.2 also shows the results of the chi-square tests (degrees of freedom = 3) for each listener. Only one of the seven listeners showed a significant difference in the number of front-back reversals between conditions. Based on these results it can be said that generally the number of front-back reversals does not differ between conditions. It is interesting that listener 'ZFS' was the only listener with a significant Chi-square test of independence, and s/he also showed possible response strategies in all three directional regions above. This listener appeared to have used more distinct response strategies compared to other listeners. In the front and back regions, listener ZFS used a response strategy related to reverberation and in the side region a strategy related to modulation.

Table 3.2. The percentage of responses that were front-back reversals in each condition for all listeners. The bottom displays the means across all listeners. On the right the χ^2 statistic and associated *p*-values are displayed comparing the number of reversals in each condition for all listeners.

Subject ID	Modulated Anechoic %Reversals	Unmodulated Anechoic %Reversals	Modulated Reverb %Reversals	Unmodulated Reverb %Reversals	χ^2	p
ZFS	38.5	22.3	46.2	45.4	12.5	0.01
ZFW	46.2	40.0	40.8	35.4	1.9	0.60
ZFU	41.5	47.7	38.5	45.4	1.5	0.68
ZGB	40.0	41.5	43.1	46.2	0.6	0.89
LHI	37.7	40.0	46.2	45.4	1.6	0.67
ZGA	36.2	43.1	34.6	40.8	1.6	0.67
ZGD	29.2	26.9	26.9	36.9	2.9	0.40
Mean	**38.5**	**37.4**	**39.5**	**42.2**		

Chi-square analyses were also performed to determine whether the number of front-back reversals differed between conditions within each hemisphere. Table 3.3 displays the percentage of front-back reversals separated by whether the target was in the front or back hemisphere. The Chi-square statistic is displayed in the table for each listener in the table along with the associated *p*-value (degrees of freedom = 3). For targets in the front hemisphere all listeners, except one, showed a significant difference in the number of reversals between conditions. Follow-up Chi-square tests were performed, with Bonferroni correction, comparing the mean number of reversals between the anechoic and reverberant environments and between the modulated and unmodulated conditions in the front hemisphere. There were significantly more reversals in the anechoic environment than in the reverberant environment ($\chi^2(1) = 12.967$, $p = 0.003$), but there was no significant difference between the modulated and unmodulated

97

conditions ($\chi^2(1) = 0.215$, $p = 0.003$). For targets in the back hemisphere all listeners, except one, showed a significant difference between conditions. The same follow-up Chi-square tests as above were performed for the back hemisphere. More reversals were recorded in the reverberant environment than in the anechoic environment ($\chi^2(1) = 15.246$, $p = 0.003$), and there was no significant difference between the modulated and unmodulated conditions ($\chi^2 = 0.042$, $p = 0.837$). Synthesizing these two results, it can be said that in anechoic conditions there are more responses in the back hemisphere while in the reverberant conditions there are more responses in the front hemisphere.

Table 3.3. Percentage of responses that were front-back reversals for each condition divided by whether the targets were in the front or back hemisphere for all listeners. Means are displayed below individual percentages for each hemisphere. On the right the χ^2 statistic and associated *p*-values are displayed comparing the number of reversals in each condition for all listeners.

Hemi-field	Subject ID	Modulated Anechoic %Reversals	Unmodulated Anechoic %Reversals	Modulated Reverb %Reversals	Unmodulated Reverb %Reversals	χ^2	*p*
Front	ZFS	37.7	13.8	2.3	0.0	86.2	0.00
	ZFW	23.8	15.4	6.2	5.4	23.3	0.00
	ZFU	25.4	37.7	33.1	11.5	19.0	0.00
	ZGB	10.8	13.1	0.0	0.0	31.6	0.00
	LHI	35.4	38.5	0.0	0.8	93.4	0.00
	ZGA	24.6	40.8	25.4	28.5	7.3	0.06
	ZGD	3.1	14.6	3.1	34.6	62.3	0.00
	Mean	**23.0**	**24.8**	**10.0**	**11.5**		
Back	ZFS	0.8	8.5	43.8	45.4	86.1	0.00
	ZFW	22.3	24.6	34.6	30.0	4.3	0.23
	ZFU	16.2	10.0	5.4	33.8	37.1	0.00
	ZGB	29.2	28.5	43.1	46.2	9.0	0.03
	LHI	2.3	1.5	46.2	44.6	103.9	0.00
	ZGA	11.5	2.3	9.2	12.3	9.1	0.03
	ZGD	26.2	12.3	23.8	2.3	29.4	0.00
	Mean	**15.5**	**12.5**	**29.5**	**30.7**		

C. Discussion

Double-pole scatter plots were used as a visualization method to assess performance of individual listeners. From these visualizations it is clear that there are individual differences between listeners both in the size of their angular errors and in the strategies used for directional localization. Part of the reason for this could be related to

99

the stimuli used. Narrowband stimuli degrade directional localization and results in systematic errors like front-back reversals (Middlebrooks, 1992). Additionally the use of ITDs would have been limited since listeners are less sensitive to high frequency ITDs (Moore, 2008, p. 236). The use of nonindividualized HRTFs recorded with KEMAR is also known to hurt directional localization performance (See Moller, Hammershoi, Jensen, & Sorensen, 1999 for extensive lit review). Moller et al. found directional localization errors increased in the median plane from 16% in the free field to 38.5% when stimuli were generated from recordings made with a KEMAR mannequin. While reverberation can be detrimental to directional localization (Hartmann, 1983; Giguere and Abel, 1993; Ihlefeld and Shinn-Cunningham, 2011), it can also assist with externalization of stimuli (Plenge, 1974). When interpreting the results it is important to remember that the stimuli were generated to be difficult to localize with narrow bandwidths, high center frequency, (Stevens & Newman, 1936; Yost, 2013; Middlebrooks & Green, 1991) and gradual onsets and offsets (Wagenaars, 1990; Wallach, Newman, and Rosenweig, 1949).

Analysis of angular error in the side region using the 2 x 2 within-subjects ANOVA revealed a significant interaction between modulation and reverberation. For unmodulated stimuli, there were larger angular errors in the reverberant environment than in the anechoic environment. While for modulated stimuli there were larger angular errors in the anechoic environment than in the reverberant environment. This result for the unmodulated stimuli fits other past directional localization experiments showing that directional localization is worse in reverberant environments (Hartmann, 1983; Giguere and Abel, 1993; Ihlefeld and Shinn-Cunningham, 2011). It is interesting that for modulated stimuli, angular errors were larger in the anechoic environment. This does

support the hypothesis that directional localization benefits from modulation in reverberant conditions; however, the benefit was only found in the side region. The side region is also where minimum audible angle (Mills, 1958) is worst. Therefore the benefit from modulation in reverberation in the side region may result from directional localization ability being impaired in the side region. Eberle et al. (2000), which did not separate azimuth judgments by region, found modulation decreased angular error for very high modulation frequencies, but believe it was due to the sidebands increasing the bandwidth of the stimuli. In the present experiment, the modulation rate was not high enough for sidebands to increase the stimulus bandwidth. In Experiment III, a deliberate effort was made to minimize ITD cues to force listeners to use other directional localization strategies, like modulation depth. The side region may have benefitted from AM in reverberation because other directional localization cues were minimized sufficiently to allow listeners to benefit from modulated stimuli.

There are a few limitations evaluating the results of the 2 x 2 ANOVA for angular error in the side region. One limitation is that compared to the front and back regions, angular errors were much smaller across all listeners. The smaller variability in the side region may have been driving the significant interaction. While the ANOVA results support the hypothesis of modulation benefiting directional localization in a reverberant environment, the lack of consistent response patterns between listeners makes it difficult to definitively conclude that either modulation or reverberation affected angular errors. Additionally, a limitation across Experiment III is that the anechoic condition was presented before the reverberant condition, so a benefit in reverberation may be due to practice effects, and not from the modulation itself. However, if that were

101

the case, then performance should be improved in reverberation for both modulated and unmodulated stimuli, which was not observed.

The results from the 2 x 2 within-subjects ANOVA on mean angle of error in the front and back regions showed that there was no effect of modulation or room on angular error. The lack of significant main effects and interactions in these regions maybe due to individual differences which would make effects difficult to detect by the ANOVAs. Additionally the statistical power was relatively low with only seven subjects, and some of the listeners demonstrated conflicting response strategies.

Front-back errors were analyzed using a Chi-square test of independence to compare the number of front-back errors between conditions. In the anechoic environment there were more responses in the back hemisphere, while in the reverberant environment there were more responses in the front hemisphere. There was no effect of modulation on front-back reversals. Past studies acknowledge the prevalence of front-back reversals in directional localization studies (Wenzel, Arruda, Kistler, Wightman, 1993; Wightman & Kistler, 1989b), particularly when spectral cues are limited due to decreased signal bandwidth (Middlebrooks, 1992).

When the number of front-back reversals was separated based on whether the target was in the front or back hemisphere a response bias became evident. Again Chi-square tests of independence were used to determine whether the number of front-back reversals differed between conditions within each hemisphere. This test revealed a tendency under anechoic conditions for listeners to place responses in the back hemisphere, while under reverberant conditions listeners tended to place response in the front hemisphere. Typically there are more front-to-back reversals than back-to-front

102

(Zahorik, 2006; Wightman and Kistler, 1999). There is no precedent in the literature for an interaction between reverberant and anechoic environments and front/back reversals. Reverberation has been shown to help with spatializing sounds (Plenge, 1974) and this may impact front-back response shifts, but it does not explain why there would be an overall tendency for directional localizations to be in the front hemisphere. In more ecologically valid conditions, resolution of front-back errors can be mediated by moving either the head of the listener or the source during stimulus presentation (Wightman & Kistler, 1999), however this was not possible in the current set up.

Two listeners (ZGB and LHI) had larger thresholds (better performance) than the rest of the listeners in the MLD diagnostic test. This task is sensitive to central auditory processing problems in the auditory brain stem (Wilson, Zizz, & Sperry, 1994). The angular errors from these listeners were visually compared to the results of other listeners to determine whether better performance in the MLD task had any corresponding performance changes on the localization task. The only noticeable trend was that both listeners appeared to use a response strategy related to reverberation in the side region; however, there were no significant effects for either listener.

Interaural modulation gain differences in the side region may have facilitated improved localization performance. Comparing the modulation gain between the two ears in Figure 4.2 for the 8000 Hz center frequency shows that modulation depth at contralateral ear is more attenuated than the ipsilateral ear as a function of azimuth. The largest modulation gain difference between ears is in the side region at -90° where modulation is attenuated about 8 dB more at the contralateral ear than at the ipsilateral ear. Near midline there is relatively little interaural modulation gain difference. The large

103

interaural modulation depth gain differences in the side region may create dynamic ILDs that listeners can use as a localization cue. Dynamic ILDs are level differences between the two ears that fluctuate in time. Sinusoidal amplitude modulation (SAM) can be thought of as sinusoidally varying stimulus level. If modulation gain is identical at both ears, then when a SAM signal reaches the ears the level at each ear will be identical across time (assuming no phase shift between ears): That is there would be no dynamic ILDs. If modulation is more attenuated at the contralateral ear than the ipsilateral ear, then when a SAM signal reaches the ears, the level at the contralateral ear will vary in amplitude less than the signal at the ipsilateral ear: That is, there will be larger dynamic ILDs. Dynamic ILDs will be greatest when the interaural modulation depth attenuation is greatest, and that is in the side region under reverberant conditions. In practice, where phase shifts would be present, dynamic ILDs would be more prominent because the difference between the waveforms reaching each ear would be larger across time.

D. Conclusions

Individual differences between listeners were prevalent in the directional localization task and were an impediment for studying the impact of modulation on directional localization. Nevertheless, the significant interaction in the side region indicated that both the acoustic environment and modulation influenced directional localization accuracy.

CHAPTER V

GENERAL DISCUSSION

The results from Experiments II and III collectively provide evidence that AM attenuation can facilitate both distance (Exp. II) and directional (Exp. III) sound localization under specific conditions. For distance (Exp. II), the facilitation from modulation is most evident for targets beyond 2 m where the AM attenuation is physically the greatest. For direction (Exp. III), the facilitation is evident only on the side region where the dynamic fluctuation in ILD caused by AM attenuation is the greatest. Both cases of facilitation require listening in a reverberant soundfield, and listening to signals that are amplitude modulated. Neither reverberation (Hartmann, 1983; Giguere and Abel, 1993; Ihlefeld and Shinn-Cunningham, 2011) nor modulation alone (Eberle et al., 2000) enhance localization performance. It is, therefore, truly the combination of reverberation and amplitude modulation that must be present for enhanced localization. These constraints are interesting because they are representative of the vast majority of our everyday listening experience: most signals have AM, and most listening is done in reverberant soundfields.

A. Effects on Speech

It is as though the auditory system has evolved mechanisms specifically to extract localization and speech information in reverberant environments. Research in speech intelligibility (Brandewie & Zahorik, 2010) has found that intelligibility improved when listeners have prior exposure to a room environment. This enhancement is

eliminated when the prior exposure to speech is in an anechoic environment instead of a reverberant environment. This provides further evidence that the auditory system includes mechanisms that specifically extract information from degraded signals specifically in a reverberant environment.

The ability to discriminate two sound sources in space is important for obtaining spatial release from masking in azimuth (Plomp & Mimpen, 1981) and distance (Brungart & Simpson, 2002). The present findings are applicable to improving spatial release from masking in the distance domain within the modulation range of speech stimuli. By taking advantage of how a room attenuates modulation as a function of distance, spatial release from masking between a signal and a masker may be enhanced. The envelope spectrum of speech shows a maximum around 3 Hz (Houtgast & Steeneken, 1985) which is below the range of modulation rates tested in Experiment II. Results from experiment II showed no benefit from AM stimuli in the 4 Hz conditions; however, the modulation gain as a function of distance did not predict a benefit from modulation in the lower range of modulation frequencies. The envelope spectrum of speech ranges between 2 and 50 Hz, with even higher modulation rates (50 - 500 Hz) associated with periodicity in speech (Rosen, 1992). These higher modulation rates that were more attenuated by the room as a function of distance provided a greater benefit to distance judgments, like in the 32 Hz condition. Since the modulation frequency range over which AM benefits distance judgments overlaps with the modulation spectra of speech, spatial release from masking should be enhanced in the distance domain for speech stimuli. An additional benefit of using speech as a stimulus is that listeners are already very familiar with the stimulus.

106

B. Role of Familiarity

An important consideration for the use of modulation-based localization cues relates to prior knowledge of the source signal modulation characteristics. For modulation depth to be used as a localization cue, it seems likely that listeners must make assumptions about the modulation depth of the source signal. The modulation depth cue exists because the room attenuates the depth of modulation before the signal reaches the ears, but listeners do not know how much modulation depth is removed by the room compared to how much modulation depth was in the original signal. This ambiguity implies that listeners must be making an assumption about the modulation depth of the original signal. In the case of speech, where listeners are familiar with the stimulus, it is conceivable that the modulation depth of the signal is known, so the amount of modulation attenuated by the room would be easier for listeners to discern. Kim et al. (2015) also suggested that the modulation depth of the source must be known by the listener to take advantage of the AM depth cue for distance perception.

Without prior knowledge about the source signal's modulation characteristics, modulation depth can only serve as a relative distance cue. In auditory distance perception D/R is generally considered to be an absolute distance cue because in a reverberant environment. This is in contrast to anechoic environments, where only intensity is available as a distance cue, and distance judgments change with practice, presumably because listeners must hear presentations of stimuli at different relative physical distances to create a stable distance percept (Mershon & King, 1975). Similarly listeners may shift their distance judgments after exposure to successive trials with varying modulation depths. Not only does this provide information about the original

107

source signal, but also how modulation is attenuated by the specific room as a function of distance. To determine whether modulation depth is a relative distance cue, distance judgments from Experiment II were pooled across listeners and split between the first half and second half of trials in the 32 Hz and 4 Hz conditions to test whether listeners' first set of judgements were different from their second set. If there is a significant difference between the first set and the second set then this supports AM being a relative distance cue. In the 32 Hz condition, distance judgments were significantly smaller for the first half of judgments ($M = 3.032$, $SD = 1.877$) than the second half ($M = 3.181$, $SD = 1.223$), $t(1619) = -2.760$, $p = 0.006$, but in the 4 Hz condition there was no significant difference between the first half ($M = 2.865$, $SD = 1.569$)and second half ($M = 2.844$, $SD = 1.816$) of judgments, $t(1619) = 0.487$, $p = 0.627$. This suggests that AM is a relative distance cue, but only when modulation depth changes appreciably as a function of distance (in the 32 Hz condition). The range of modulation frequencies attenuated by the room as a function of distance will change depending on room acoustics.

C. Room Acoustics

While modulation depth can be used as an auditory distance cue, the conditions that allow listeners to extract distance information from modulation depth are dependent on both stimulus parameters and acoustical properties of the listening environment. In Experiment II, AM provides the most benefit for the listener under acoustic conditions where the modulation depth of the signal is significantly attenuated as a function of distance by the room. Therefore it is important to take room acoustics into account to determine the amount of benefit provided by modulation depth. In the room used in Experiment II the modulation depth cue was most useful for targets greater than

108

approximately 2 m from the listener and for higher modulation rates. While this combination of stimulus parameters was ideal for this room, they may change depending on acoustic characteristics of other environments. Acoustic properties like reverberation (Mershon & King, 1975) and T_{60} (Bronkhorst & Houtgast, 1999) have been shown to influence distance judgments. Unpublished data from our lab show that varying the T_{60} of a room will shift the constant value of listeners' power function fits while leaving exponents relatively unchanged indicating that the amount of reverberation in a room linearly scales distance judgments.

Both the measured and simulated BRIRs used to generate stimuli in Experiments II and III, respectively, were from large halls with long reverberation times (T_{60}). Although it may be difficult to directly generalize the results from these rooms to other rooms, the general techniques for room acoustic analysis may be applied: The MTF of a room can be used to predict whether the room removes a significant amount of modulation depth as a function of distance and azimuth for specific stimulus parameters. Large rooms were used here because they typically have a long T_{60} and because larger rooms allow for a wider range of source positions to be measured. T_{60} can shift depending on room size, absorption coefficients of surfaces in the room, and where the measurements were made in the room. Increasing the T_{60} of a room would generally increase the amount of reverberation for all source locations in the room, which would result in greater modulation depth attenuation for all sources. Houtgast & Steeneken (1985) theorized that increasing T_{60} results in greater modulation attenuation across modulation rates. In the distance domain, as mentioned above, this may push distance judgments farther away because there will be more reverberation and all stimuli will have

109

less modulation depth. For directional judgments the modulation depth of the signal reaching the contralateral ear would be even more attenuated compared to the ipsilateral ear. If dynamic ILD cues are driving the benefit of modulation depth in the side region then increased reverberation would be predicted to make ILD fluctuations between the two ears even larger because the modulation gain difference between the two ears would be increased. So far, these results can only be applied to normal hearing listeners because no hearing impaired listeners participated in experiments II and III.

D. Hearing Impaired Listeners

The impact of hearing impairment on acoustic cue processing related to modulation gain is difficult to predict because of complexities introduced by hearing aids. It is unlikely that hearing impaired listeners with current hearing aid technology, can take advantage of these modulation-based localization cues. Hearing aids can deteriorate intensity cues because of compression algorithms that may degrade both the intensity information of the original signal as well as the modulation depth of the signal. To use the AM distance cue, the listener must know the modulation depth of the original signal, but this is complicated if the AM depth is attenuated by both the room and compression algorithms. Directional localization is also complicated by the compression algorithms because dynamic ILDs may be necessary to take advantage of interaural AM depth as a function of azimuth. Assuming a listener has bilateral hearing aids, the compression algorithms at each ear will impose compression independently at each ear, which will artificially shift dynamic ILDs. If the listener has asymmetrical hearing loss and uses a unilateral hearing aid then predicting whether AM cues can be used becomes even more complicated. In Experiment II the benefit from AM on distance judgments was only

shown for the ipsilateral ear, so the amount of benefit would depend on both the amount of hearing loss and whether or not the ipsilateral ear had the hearing aid. In directional localization a unilateral hearing aid would also complicate the use of dynamic ILDs. If the modulation depth at one ear is attenuated by the compression algorithms, while the modulation depth at the other ear is only attenuated by the room then the dynamic ILDs will be different than if the modulation depth arriving at each ear were left to vary naturally. Based on the complications of hearing aid algorithms and bilateral vs. unilateral fitting, the impact of hearing aids on the use of AM depth localization cues is too complicated to hypothesize whether hearing aid listeners can also benefit from AM depth cues.

REFERENCES

Allen, J. B. And Berkley, D. A. (**1979**). "Image method of efficiently simulating small-room acoustics," *Journal of the Acoustical Society of America, 65*, 943 - 950.

Anderson, P. W., & Zahorik, P. (**2011**). "Auditory and visual distance estimation." In *Proceedings of Meetings on Acoustics*, *12*, 050004.

Anderson, P. W., & Zahorik, P. (**2014**). "Auditory/visual distance estimation: Accuracy and variability," *Frontiers in Psychology, 5*, 1 - 11.

Begault, D. R., and Wenzel, E. M. (**1993**). "Headphone localization of speech," *Human Factors, 35*, 361 - 376.

Begault, D. R., Wenzel, E. M., and Anderson, M. R. (**2001**). "Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source," *Journal of the Audio Engineering Society, 49*, 904 - 916.

Bernstein, L. R., and Trahiotis, C. (**1985b**). "Lateralization of sinusoidally amplitude-modulated tones: Effects of spectral locus and temporal variation," *Journal of the Acoustical Society of America, 77*, 1868 - 1880.

Best, V., Carlile, S., Jin, C., and van Schaik, A. (**2005**). "The role of high frequencies in speech localization," *Journal of the Acoustical Society of America, 118*, 353 - 363.

Blauert, J. (**1997**). *Spatial hearing: The psychophysics of human sound localization*. Cambridge, MA: MIT press.

Brandewie, E. J. (**2012**). "Factors affecting speech intelligibility improvement with exposure to reverberant room listening environments," Ph.D. Thesis, University of Louisville, USA.

Brandewie, E. J., and Zahorik, P. (2010). "Prior listening in rooms improves speech intelligibility," *Journal of the Acoustical Society of America, 128*, 291 - 299.

Bronkhorst, A. W., and Houtgast, T. (**1999**). "Auditory distance perception in rooms," *Nature, 397*, 517 - 520.

Brungart, D.S., and Simpson, B.D. (**2002**). "The effects of spatial separation in distance on the informational and energetic masking of a nearby speech signal." *Journal of the Acoustical Society of America, 112*, 664 - 676.

Eberle, G., McAnally, K. I., Martin, R. L., and Flanagan, P. (**2000**). "Localization of amplitude-modulated high-frequency noise," *Journal of the Acoustical Society of America, 107*, 3568 - 3571.

Giguere, C., and Abel, S. M. (**1993**). "Sound localization: effects of reverberation time, speaker array, stimulus frequency, and stimulus rise/decay," *Journal of the Acoustical Society of America, 94*, 769 - 776.

Grantham, D. W. (**1984**). "Discrimination of dynamic interaural intensity differences," *Journal of the Acoustical Society of America, 76*, 71 - 76.

Hartmann, W. M. (**1983**). "Localization of sound in rooms," *Journal of the Acoustical Society of America, 74*, 1380 - 1391.

Houtgast, T., and Steeneken, H. J. M. (**1985**). "A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria," *Journal of the Acoustical Society of America, 77*, 1059 - 1077.

Ihlefeld, A., and Shinn-Cunningham, B. G. (**2011**). "Effect of source spectrum on sound localization in an everyday reverberant room," *Journal of the Acoustical Society of America, 130*, 324 - 333.

ISO – 3382 (**1997**). "Acoustics – Measurement of the reverberation time of rooms with reference to other acoustical parameters," (International Organization of Standardization, Geneva).

Jetzt, J. J. (**1979**). "Critical distance measurement of rooms from the sound energy spectral response," *The Journal of the Acoustical Society of America, 65*, 1204 - 1211.

Kim, D. O., Zahorik, P., Carney, L. H., Bishop, B. B., and Kuwada, S. (2015). "Auditory distance coding in rabbit midbrain neurons and human perception: monaural amplitude modulation depth as a cue," *The Journal of Neuroscience, 35*, 5360 - 5372.

Kistler, D. J., and Wightman, F. L. (**1992**). "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction," *Journal of the Acoustical Society of America, 91*, 1637 - 1647.

Larsen, E., Iyer, N., Lansing, C. R., and Feng, A. S. "On the minimum audible difference in direct-to-reverberant energy ratio," *Journal of the Acoustical Society of America, 124*, 450 - 461.

Mershon, D. H., and King, L. E. (**1975**). "Intensity and reverberation as factors in the auditory perception of egocentric distance," *Perception & Psychophysics, 18*, 409 - 415.

Mershon, D. H. and Bowers, J. N. (**1979**). "Absolute and relative cues for the auditory perception of egocentric distance," *Perception, 8*, 311 - 322.

Middlebrooks, J. C. (**1992**). "Narrow-band sound localization related to external ear acoustics," *Journal of the Acoustical Society of America, 92*, 2607 - 2624.

Middlebrooks, J. C., and Green, D. M. (**1991**). "Sound localization by human listeners," *Annual Review of Psychology, 42*, 135 - 159.

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

Moller, H., Hammershoi, D., Jensen, C. B., and Sorensen, M. F. (**1999**). "Evaluation of Artificial Heads in Listening Tests," *Journal of the Audio Engineering Society*, *47*, 83 - 100.

Moore, B. C. J. (2008). *An introduction to the psychology of hearing.* Bingley, UK: Emerald Publishing.

Musiek, F. E., Gollegly, K. M., Kibbe, K. S, and Verkest-Lenz, S. B. (**1991**). "Proposed screening test for central auditory disorders: Follow-up on the dichotic digits test," *The American Journal of Otology, 12*, 109 - 113.

Nielsen, S. H. (**1991**). "Distance perception in hearing," Ph.D. Thesis, Aalborg University, Denmark.

Perrett, S., and Noble, W. (**1995**). "Available response choices affect localization of sound," *Perception and Psychophysics, 57*, 150 - 158.

Plenge, G. (**1974**). "On the difference between localization and lateralization," *Journal of the Acoustical Society of America, 56*, 944 - 951.

Plomp, R., and Mimpen, A. M. (**1981**). "Effect of the orientation of the speaker's head and the azimuth of a noise source on the speech-reception threshold for sentences," *Acta Acustica united with Acustica, 48*, 325 - 328.

Rife, D. D., and Vanderkooy, J. (**1989**). "Transfer-function measurement with maximum-length sequences," *Journal of the Audio Engineering Society, 37*, 419 - 444.

Rosen, S. (1992). "Temporal information in speech: acoustic, auditory and linguistic aspects," *Philosophical Transactions: Biological Sciences, 336,* 367 - 373.

Schroeder, M. R. (**1981**). "Modulation transfer functions: definition and measurement." *Acustica, 49*, 179-182.

Shinn-Cunningham, B. (**2001**). "Localizing sound in rooms," In *Proceedings of the ACM, SIGGRAPH and EUROGRAPHICS Campfire*, Snowbird, Utah.

Shinn-Cunningham, B., Santarelli, S., and Kopco, N. (**2000**). *Distance perception of nearby sources in reverberant and anechoic listening conditions: Binaural vs. monaural cues*. Poster presented at the 23rd Annual MidWinter meeting of the Association for Research in Otolaryngology, St. Petersburg, FL.

Stevens, S. S., and Newman, E. B. (**1936**). "The localization of actual sources of sound," *The American Journal of Psychology, 48*, 297 - 306.

Strutt, J. W. (**1907**). "On our perception of sound direction," *Philosophical Magazine, 13*, 214 - 232.

Trahiotis, C., and Bernstein, L. R. (**1986**). "Lateralization of bands of noise and sinusoidally amplitude-modulated tones: Effects of spectral locus and bandwidth," *Journal of the Acoustical Society of America, 79*, 1950 - 1957.

Wallach, H., Newman, E. B., and Rosenweig, M. R. (**1949**). "The precedence effect in sound localization," *The American Journal of Psychology, 62*, 315 - 336.

Wagenaars, W. M. (**1990**). "Localization of sound in a room with reflecting walls." *Journal of the Audio Engineering Society, 38*, 99 - 110.

Wakefield, G. H., and Viemeister, N. F. (**1990**). "Discrimination of modulation depth of sinusoidal amplitude modulation (SAM) noise," *Journal of the Acoustical Society of America, 88*, 1367 - 1373.

Wenzel, E. M., Arruda, M., Kistler, D. J., and Wightman, F. L. (**1993**). "Localization using nonindividualized head-related transfer functions," *Journal of the Acoustical Society of America, 94*, 111 - 123.

Wightman, F. L., and Kistler, D. J. (**1989a**). "Headphone simulation of free-field listening. I: stimulus synthesis," *Journal of the Acoustical Society of America, 85*, 858 - 867.

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

Wightman, F. L., and Kistler, D. J. (**1992**). "The dominant role of low-frequency interaural time differences in sound localization," *Journal of the Acoustical Society of America, 91*, 1648 - 1661.

Wightman, F.L., and Kistler, D. J. (**1999**). "Resolution of front-back ambiguity in spatial hearing by listener and source movement," *Journal of the Acoustical Society of America, 105*, 2841 - 2853.

Wilson, R. H., Zizz, C. A., and Sperry, J. L. (**1994**). "Masking-level difference for spondaic words in 2000-msec bursts of broadband noise," *Journal of the American Academy of Audiology, 5*, 236 - 242.

Yost, W. A. (**2013**). "Sound source localization: Bandwidth and envelope," In *Proceedings of Meetings on Acoustics, 19*.

Zahorik, P. (**2001**). "Estimating sound source distance with and without vision," *Optometry and Vision Science, 78*, 270 - 275.

Zahorik, P. (**2002a**). "Assessing auditory distance perception using virtual acoustics," *The Journal of the Acoustical Society of America, 111*, 1832 - 1846.

Zahorik, P. (**2002b**). "Direct-to-reverberant energy ratio sensitivity," *The Journal of the Acoustical Society of America, 112*, 2110 - 2117.

Zahorik, P. (**2009**). "Perceptually relevant parameters for virtual listening simulation of small room acoustics," *Journal of the Acoustical Society of America, 126*, 776 - 791.

Zahorik, P., and Anderson, P. W. (**2015**). The role of amplitude modulation in auditory distance perception. In *Proceedings of Meetings on Acoustics, 21*, 050006.

Zahorik, P., Brungart, D. S, and Bronkhorst, A. W. (**2005**). "Auditory distance perception in humans: A summary of past and present research," *Acta Acustica United with Acoustica, 91*, 409 - 420.

Zahorik, P., Bangyan, P., Sundareswaran, V., Wang, K., and Tam, C. (**2006**). "Perceptual recalibration in human sound localization: Learning to remediate front-back reversals," *Journal of the Acoustical Society of America, 120*, 343 - 359.

CURRICULUM VITA

Paul W. Anderson, M.S.
Department of Psychological and Brain Sciences
2301 South Third Street
University of Louisville
Louisville, KY 40292
Phone: 404-386-3686 (cell)
Email: paul.anderson@louisville.edu

EDUCATION

Expected August 2015	Ph.D., Experimental Psychology
	University of Louisville, Louisville, KY
2012	M.S., Experimental Psychology
	Concentration: Vision and Hearing Science
	University of Louisville, Louisville, KY
2010	B.S., Psychology
	Georgia College & State University, Milledgeville, GA

RESEARCH EXPERIENCE

July 2010 – June 2015	Graduate Research Assistant
	Zahorik Auditory Perception Lab
	PI: Pavel Zahorik, Ph.D.
	Department of Psychological and Brain Sciences
	University of Louisville, Louisville, KY
July 2013 – June 2014	Graduate Research Assistant
	Auditory Perception and Processing Lab
	PI: Christian E. Stilp, Ph.D.
	Department of Psychological and Brain Sciences
	University of Louisville, Louisville, KY

TEACHING EXPERIENCE

Spring 2013	Graduate Teaching Assistant
	PSYC 301 Quantitative Methods in Psychology
	University of Louisville, Louisville, KY

Fall 2012	Instructor
	PSYC 451 Introduction to Computer Programming for Psychology
	University of Louisville, Louisville, KY
2009 – 2010	Supplemental Instructor
	PSYC 2700 Statistics for Behavioral and Health Sciences
	Georgia College & State University, Milledgeville, GA

GRANTS/AWARDS

February 2015	GSC Travel Funding ($250)
September 2014	GSU Travel Award ($100)
July 2010 – June 2012	Graduate School Fellowship, University of Louisville
August 2006 - May 2010	Georgia HOPE Scholarship

MEMBERSHIPS

January 2014 – Present	Association for Research in Otolaryngology
January 2011 – Present	Acoustical Society of America
August 2014 – June 2015	Experimental Psychology Graduate Student Peer Mentor
July 2012 – June 2013	Psychology Graduate Student Department Representative
September 2011 – June 2012	Graduate Student Council Representative

PUBLICATIONS

Stilp, C. E., **Anderson, P. W.**, & Winn, M. B. (2015). Predicting contrast effects following reliable spectral properties in speech perception. *Journal of the Acoustical Society of America, 137*(6), 3466 - 3476.

Anderson, P.W., & Zahorik, P. (2014). Auditory/visual distance estimation: accuracy and variability. *Frontiers in Psychology, 5*: 1097. doi: 10.3389/fpsyg.2014.01097

Stilp, C. E., & **Anderson, P. W.** (2014). Modest, reliable spectral peaks in preceding sounds influence vowel perception. *Journal of the Acoustical Society of America, 136(5)*, EL383 – EL389.

PROCEEDINGS (PEER REVIEWED)

Zahorik, P., & **Anderson, P. W.** (2015). The role of amplitude modulation in auditory distance perception. *Proceedings of Meetings on Acoustics, 21*, 050006.

Zahorik, P., & **Anderson, P. W.** (2013). Amplitude modulation detection by human listeners in reverberant sound fields: Effects of prior listening exposure. *Proceedings of Meetings on Acoustics, 19*, 050002.

Zahorik, P., Kim, D. O., Kuwada, S., **Anderson, P. W.**, Brandewie, E., Collecchia, R., & Srinivasan, N. (2012). Amplitude modulation detection by human listeners in reverberant sound fields: Carrier bandwidth effects and binaural versus monaural comparison. *Proceedings of Meetings on Acoustics,* **15**, 050002.

Anderson, P. W., & Zahorik, P. (2011). Auditory and visual distance estimation. *Proceedings of Meetings on Acoustics,* **12**, 050004.

Zahorik, P., Kim, D. O., Kuwada, S., **Anderson, P. W.**, Brandewie, E., & Srinivasan, N. K. (2011). Amplitude modulation detection by human listeners in sound fields. *Proceedings of Meetings on Acoustics,* **12**, 050005.

PRESENTATIONS

Stilp, C.E., **Anderson, P. W.**, Assgari, A. A., Ellis, G.M., & Zahorik, P. (2015, May). "Reverberation increases perceptual calibration to reliable spectral peaks in speech." Paper accepted for presentation at the 169th Meeting of the Acoustical Society of America, Pittsburgh, PA.

Anderson, P. W., & Stilp, C. E. (2014, October). *Adaptive compensation for reliable spectral characteristics of a listening context in vowel perception.* Poster presented at the 168th Meeting of the Acoustical Society of America, Indianapolis, Indiana.

Zahorik, P., **Anderson, P. W.**, Brandewie, E., & Srinivasan, N. K. (2014, October). *Improved speech understanding and amplitude modulation sensitivity in rooms: wait a second!* Paper presented at the 168th Meeting of the Acoustical Society of America, Indianapolis, Indiana.

Stilp, C. E., & **Anderson, P. W.** (2014, October). *Modest, reliable spectral peaks in preceding sounds influence vowel perception.* Paper presented at the 168th Meeting of the Acoustical Society of America, Indianapolis, Indiana.

Zahorik, P., & **Anderson, P. W.** (2014, May). *The role of amplitude modulation in auditory distance perception.* Poster presented at the 167th Meeting of the Acoustical Society of America, Providence, RI.

Anderson, P. W., & Stilp, C. (2014, May). *Defining essential characteristics of reliable spectral properties that elicit spectral contrast efforts.* Poster presented at the 167th Meeting of the Acoustical Society of America, Providence, RI.

Zahorik, P., & **Anderson, P.** (2014, February). *Enhanced amplitude modulation sensitivity in reverberant soundfields: Effects of prior listening exposure, soundfield, and modulation frequency.* Poster presented at the 37th Annual MidWinter Meeting of the Association for Research in Otolaryngology, San Diego, CA.

Anderson, P. W., & Stilp, C. (2014, February). *Perceptual calibration to modest, predictable spectral peaks in precursor sounds influences vowel identification.* Poster presented at the 37th Annual MidWinter Meeting of the Association for Research in Otolaryngology, San Diego, CA.

Zahorik, P., & **Anderson, P.W.** (2013, June). *Amplitude modulation detection by human listeners in reverberant sound fields: Effects of prior listening exposure.* Poster presented at the 165th Meeting of the Acoustical Society of America, Montreal, Quebec, Canada.

Zahorik, P., Kim, D. O., Kuwada, S., **Anderson, P. W.**, Brandewie, E., Collecchia, R., & Srinivasan, N. K. (2012, May). *Amplitude modulation detection by human listeners in reverberant sound fields: carrier bandwidth effects and binaural versus monaural comparison.* Poster presented at the 163rd Meeting of the Acoustical Society of America, Hong Kong, People's Republic of China.

Anderson, P.W., & Zahorik, P. (2011, May). *Auditory and visual distance estimation.* Poster presented at the 161st Meeting of the Acoustical Society of America, Seattle, WA.

Zahorik, P., Kim, D. O., Kuwada, S., **Anderson, P. W.**, Brandewie, E., & Srinivasan, N. K. (2011, May). *Amplitude modulation detection by human listeners in sound fields.* Poster presented at the 161st Meeting of the Acoustical Society of America, Seattle, WA.

Professional Service
Ad hoc reviewer: *ACM Transactions on Applied Perception*