Testing a Model for Supernormal Auditory Localization

by

John J. Park

Submitted to the Department of Electrical Engineering and Computer Science

in Partial Fulfillment of the requirements for the degrees of

Bachelor of Science in Electrical Engineering

and Master of Engineering in Electrical Engineering and Computer Science

at the Massachusetts Institute of Technology

May 28, 1996

Author____
Department of Electrical Engineering and Computer Science
May 28, 1996

Certified by____
Nathaniel I. Durlach
Thesis Supervisor

Accepted by____
F.R. Morgenthaler
Chairman, Department Committee on Graduate Theses

Testing a Model for Supernormal Auditory Localization

by

John J. Park

Submitted to the

Department of Electrical Engineering and Computer Science

May 28, 1996

In Partial Fulfillment of the Requirements for the Degree of
Bachelor of Science in Electrical Engineering
and Master of Engineering in Electrical Engineering and Computer Science

ABSTRACT

Previous experiments have concluded that better-than-normal performance in localization could be achieved with "supernormal" auditory localization cues; however, different exposure conditions during training (e.g., the presence or absence of visual cues) affected the subjects' ability to adapt. This study was undertaken to determine which intersensory effects are necessary and sufficient to achieve adaptation to the altered supernormal cues.

Previous experiments have concluded that adaptation (decrease in bias with extended exposure to transformed auditory localization cues) is achieved when explicit visual cues are presented during training. In those experiments, correct answer feedback was provided through the lights in the visual display. When all visual information (explicit visual cues and the visual field) is eliminated using blindfolds, subjects showed no adaptation.

The current study concludes that the visual field can be used to reference head position during training. This information is sufficient to achieve adaptation. Therefore, the visual field is important and useful in achieving adaptation to transformed auditory localization cues.

Thesis Supervisor: Nathaniel I. Durlach
Title: Senior Research Scientist

Acknowledgments

I would like to thank Barbara Shinn-Cunningham and Nat Durlach of the Research Laboratory of Electronics for her patience and giving me the opportunity to perform research that was interesting and worthwhile. I would also like to thank my research partner Greg Lin for his support and help with the project.

I would like to thank Professor Christopher Sawyer-Lauçanno for his understanding and friendship over the last five years at MIT. Finally, I would like to thank my family and friends for all their support.

Table of Contents

I Introduction

Many different studies have been performed in the area of human auditory perception. It is known that the normal human auditory localization system has good resolution in azimuth for sources straight ahead; however, resolution is relatively poor at angles to either side. The development of methods to provide listeners with better-than-normal or "supernormal" auditory localization and the effects of changes in cue presentation will be the focus of the research.

This study is one of many experiments undertaken to help develop understanding of the ability for human subjects to adapt to changes in the way auditory localization cues are presented. This experiment is an extension to previous studies performed by Barbara Shinn-Cunningham of the M.I.T. Research Laboratory of Electronics. This study differs from previous experiments in the type of visual cues presented to subjects in order to train them to adapt to supernormal auditory cues.

This study was motivated by the increasing use of interactive spatial displays (virtual environments) which simulate normal visual and auditory spatial cues. These displays contain various imperfections such as distortion, limited resolution, time delays, and noise. These imperfections can limit performance and must be investigated in order to develop an understanding of how they affect human performance in a virtual environment. Although all forms of sensory input affect perception in a virtual environment, this study will only examine the area of auditory localization and perception. The goal of these studies is to improve on localization ability to allow the operator to adapt and achieve better-than-normal performance.

This experiment will examine the ability to achieve better-than-normal spatial resolution when presented with distortions or transformations of normal auditory spatial cues. Two aspects of performance will be examined in this study: bias and resolution. Bias is defined as an error in mean perceived position induced by a change of cues. Resolution refers to the variability of responses and to the extent to which nearby positions can be discriminated. Supernormal cues can be generated by creating larger-than-normal interaural time differences and/or interaural intensity differences for a given source location. These larger than normal physical cues allow source position to be resolved with more accuracy than source positions using normal auditory localization cues.

There are many cognitive factors and intersensory effects that have a significant influence on auditory localization and adaptation. Studies have been performed to examine the relationships between human sensory perception and the effects of changes in the sensory environment. This study will be concerned with adaptation to auditory rearrangement with changes in visual information provided during training.

In this experiment, we will investigate the relationship between visual information and adaptation. By varying experimental procedures during testing, it is possible to explore the effects of visual information on adaptation.

II Auditory Virtual Environments

The increasing use of virtual displays has led to a large amount of research with the goal of generating an accurate and realistic auditory environment. Despite all the research that has been performed on auditory localization, there is much debate as to which cues are most important in simulating auditory signals.

II.1 Spatial Auditory Localization

Auditory localization cues arise from many different effects, including interaural intensity differences (IIDs), interaural time differences (ITDs), spectral cues, and dynamic cues. The importance of each of these factors in localization and simulating acoustic environments is left to debate; however, the technologies used in creating virtual environments provide the ability to create and manipulate these cues in an auditory spatial display.

II.1.1 Static Cues

II.1.1.1 Head Related Transfer Functions

When both listener and source are stationary, a uniformly-radiating point source in free space can be represented as:

$$R_L(\omega) = H_L(\omega,\theta,\phi,r)S(\omega)$$

$$R_R(\omega) = H_R(\omega,\theta,\phi,r)S(\omega)$$

where $S(\omega)$, $R_L(\omega)$, and $R_R(\omega)$ are the Fourier transforms of the source and the signals at the left and right ears, respectively (where ω is the frequency). $H_L(\omega,\theta,\phi,r)$ and $H_R(\omega,\theta,\phi,r)$ are the complex transfer functions for the left and right ears, respectively. The transfer functions depend on θ, the azimuth to the source relative to the head; ϕ, the elevation of the source relative to the head; and r, the distance from the source to the head. We can simplify these equations for sources more than a few meters away as follows:

$$R_L(\omega) = r^{-1}H_L(\omega,\theta,\phi)S(\omega)$$

$$R_R(\omega) = r^{-1}H_R(\omega,\theta,\phi)S(\omega)$$

In this form, the transfer functions are referred to as Head Related Transfer Functions (HRTFs). HRTFs are filters associated with each ear that capture the direction dependent effects of the head on the signals received at the two ears. HRTFs are convolved with an input acoustic signal, generating a stereo signal with localization cues that are associated with a source from a specific azimuth and elevation relative to the listener.

II.1.1.2 Interaural Intensity and Time Differences

The human ability to localize stationary sounds is dependent upon the ability to separate $H_L(\omega,\theta,\phi)$ and $H_R(\omega,\theta,\phi)$ from $S(\omega)$. This separation is accomplished by using interaural differences in time and intensity.

Interaural intensity differences (IIDs) are the differences in intensity of the signals received at each ear. A signal with higher intensity at the left ear may be perceived as a sound source located to the left of the listener. IIDs only occur for frequencies which have wavelengths that are small relative to head size. Therefore, IIDs provide only weak cues at low frequencies.

Interaural time differences (ITDs) refer to the different arrival times of signals at each ear due to the spatial separation of the two ears. A signal that reaches the left ear earlier than the right ear will be perceived as a sound source located to the left of the listener. For sinusoids, interaural timing information is translated into phase differences between the ears. For high frequencies, phase information becomes ambiguous since the wavelengths are small relative to the distance between the ears. Therefore, ITD cues are weak for narrowband high frequency sources. However, interaural envelope delays do convey information for high frequency sources.

Some previous studies conclude that resolution of distance for a stationary source in free-field is poor (e.g. Coleman, 1968; Molino, 1973; Gardner, 1969; Gardner, 1968a). This is consistent with the fact that the only cue for distance for far-field sources in an anechoic environment is overall level (an ambiguous cue). Subjects also show front-back or up-down confusions (e.g. Wallach, 1939; Wightman & Kistler, 1989b; Young, 1931) in which the IIDs and ITDs corresponding to the apparent source location are roughly equal to the ITDs and IIDs corresponding to the actual source location. Subjects have also shown that resolution in the horizontal plane (azimuth) is far more accurate than in the vertical plane (elevation) (Perrott & Saberi, 1990; Searle et al., 1976a; Oldfield and Parker, 1984a, 1984b). Finally, since gross ITDs and IIDs vary with azimuth, not with elevation, all of these results indicate that the most salient localization cues are the gross ITD and IID.

II.1.2 Binaural and Monaural Spectral Cues

Physical structures of the listener (such as the pinnae and shoulders) also produce localization information. These effects are important at higher frequencies for which the size of the structure is large relative to the wavelengths of the signals. The pinnae act as acoustic filters where the frequency response of the filters depend on the angle of incidence (Blauert, 1969; Butler, 1987; Watkins, 1978). This filtering creates both binaural and monaural spectral cues.

Differences in the physical structures among subjects (or even between individual ears) produce individual HRTFs. It is possible that subjects will achieve the best performance using their own individualized HRTFs, but most studies of auditory localization use HRTFs taken from a single subject. The process used to measure individual HRTFs is difficult and not necessary for this experiment. While the basic characteristics of the HRTFs (gross IID and ITD behavior) are similar between listeners, spectral cues may not be exactly like those the subject normally hears.

Current spatialization systems use measured HRTFs to filter input signals which are presented to the subject over headphones. The result is that all binaural and monaural static cues are provided to the subject.

II.1.3 Dynamic Cues

When the listener moves his or her head, additional localization information can help the subject to distinguish source position. With motion of the head, a different pattern of ITDs and IIDs reach the listener. Head motion is particularly important for resolving positions in front versus positions behind the listener. "Front-back" confusions are common when dynamic cues are not given to subjects. When dynamic cues are included, front-back confusions are much less common. In certain studies (Pollack & Rose, 1967; Thurlow & Runge, 1967), head motion increased localization accuracy. However, some studies have also indicated that head motion does not necessarily improve localization for all conditions (Thurlow & Runge, 1967; Shinn-Cunningham, 1994).

Incorporating dynamic cues in an auditory virtual environments involves the use of tracking systems. Movement of the head is tracked and HRTFs are updated so that the filters applied to the input signal change over time. However, the change in filters is not instantaneous; measurement and communication delays result in some nominal lag between the time a subject moves his/her head and the time the acoustic filters change.

III Adaptation to Supernormal Rearrangement

The rearrangement of acoustic spatial cues may allow better-than-normal or supernormal auditory spatial resolution. Supernormal cues can be generated by creating larger-than-normal differences in localization cues for source locations. The rearrangement of the supernormal cues remap the ITD and IID values to different source locations. This can change the error in mean perceived position (bias), and the accuracy in which nearby positions can be discriminated. By measuring performance over time, these immediate effects and any adaptation over time (any changes in performance due to exposure to the rearrangement) can be observed.

III.1 Supernormal Cues

θ is the "correct" azimuth of a sound source while θ' represents the position normally associated with that sound. Subjects that have not adapted should show a response bias in perceived source azimuth when the correct source location is θ, since they should perceive the source to be at θ'. If the subject has adapted completely after training, he/she should show no bias in response, and θ' should equal θ.

In the current study, we examine localization in azimuth when HRTF cues are transformed such that the HRTFs used to simulate a source at θ are exactly the HRTFs that normally correspond to azimuths $\theta' = f_n(\theta)$, where $f_n(\theta)$ is defined by

$$\theta' = f_n(\theta) = \frac{1}{2}\tan^{-1}\left[\frac{2n\sin(2\theta)}{1-n^2 + (1+n^2)\cos(2\theta)}\right]$$

The parameter n is the slope of the transformation at $\theta = 0$. When $n = 1$, the "normal" cues are presented. Azimuths 0, +90, and -90 degrees map to themselves for all values of n. In the current study, values of $n = 1$ and $n = 3$ were presented. The two mapping functions are shown in Figure 1.

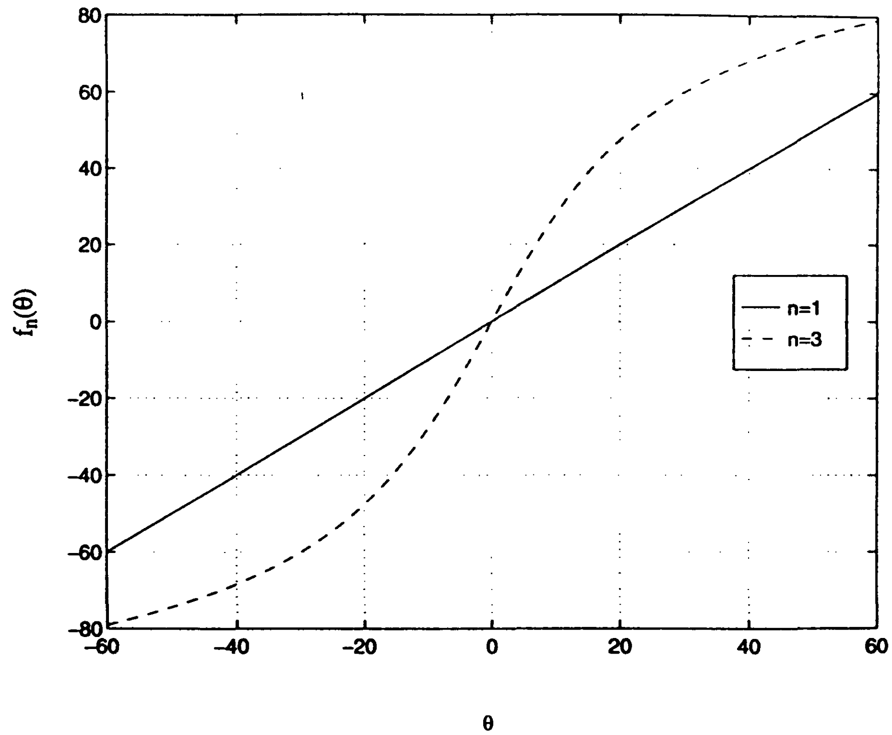


Figure 1. Transformation of $f_n(\theta)$.

Using the transformation $f_3(\theta)$ should affect both bias and resolution. In particular, it is expected that subjects will perceive sources further off center than their actual locations with these cues, causing localization bias. For instance, when n = 3 and a source is to be presented at 10 degrees (position 8), the HRTFs used are those normally corresponding to a source azimuth of 27.88 degrees (position 9.8). The transformations $f_1(\theta)$, $f_3(\theta)$ and their relative position numbers are summarized in Table 1.

The slope of the rearrangement predicts how resolution should be affected. Resolution should improve for sources near center (0°) when the slope is larger than 1 since the difference in physical cues for nearby positions is larger than normal. Figure 1 predicts that for a transformation of n = 3, subjects are expected to show better-than-normal resolution in the front, and reduced localization resolution at the sides.

θ (degrees)	Position Number	$f_1(\theta)$ (degrees)	Relative Position (n=1)	$f_3(\theta)$ (degrees)	Relative Position (n=3)
-60	1	-60	1	-79.10	-0.9
-50	2	-50	2	-74.37	-0.4
-40	3	-40	3	-68.33	0.2
-30	4	-30	4	-60.00	1.0
-20	5	-20	5	-47.52	2.2
-10	6	-10	6	-27.88	4.2
0	7	0	7	0	0
10	8	10	8	27.88	9.8
20	9	20	9	47.52	11.8
30	10	30	10	60.00	13
40	11	40	11	68.33	13.8
50	12	50	12	74.37	14.4
60	13	60	13	79.10	14.9

Table 1. Transformations and Corresponding Position Number.

III.2 Equipment

III.2.1 Visual Display

The visual display used in our experiments consists of 13, 1/2 inch light bulbs, one at each of the positions which were valid responses for the sound source location. Each position is labeled left to right from 1 to 13, each ten degrees apart, spanning -60° to +60°. The lights are positioned approximately at ear level with the subject facing center (0°) at position 7. The

13 positions mark the 13 possible locations of the simulated auditory cues for each experiment. In some experiments, the lights are used to cue the correct response, while in other experiments, the lights and labels simply mark the spatial locations of the possible responses. A diagram of the arc is shown in Figure 2.

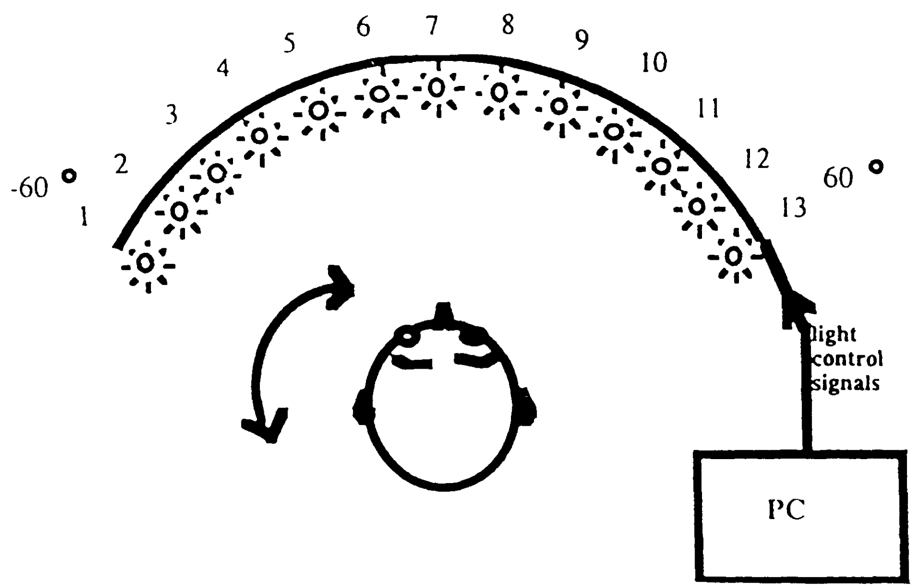


Figure 2: Diagram of visual display from Shinn-Cunningham, 1994.

III.2.2 Auditory Virtual Environment

The auditory virtual environment system that was used for our study contains a head tracker, controlling PC, signal processing hardware, and headphones, all of which were set up in a sound-proof room to eliminate background noise and disturbances. The head tracker reports the head orientation of the subject in absolute coordinates to the PC. The PC then calculates the relative direction from source to subject head position and relays the relative position information to the special hardware. This hardware filters input source waveforms to simulate left- and right-ear signals with the appropriate localization cues (using HRTFs which thus depend on head position as well as source position). The binaural signal is then played to the subject over headphones (Sennheiser 545). A block diagram is shown below in Figure 3.

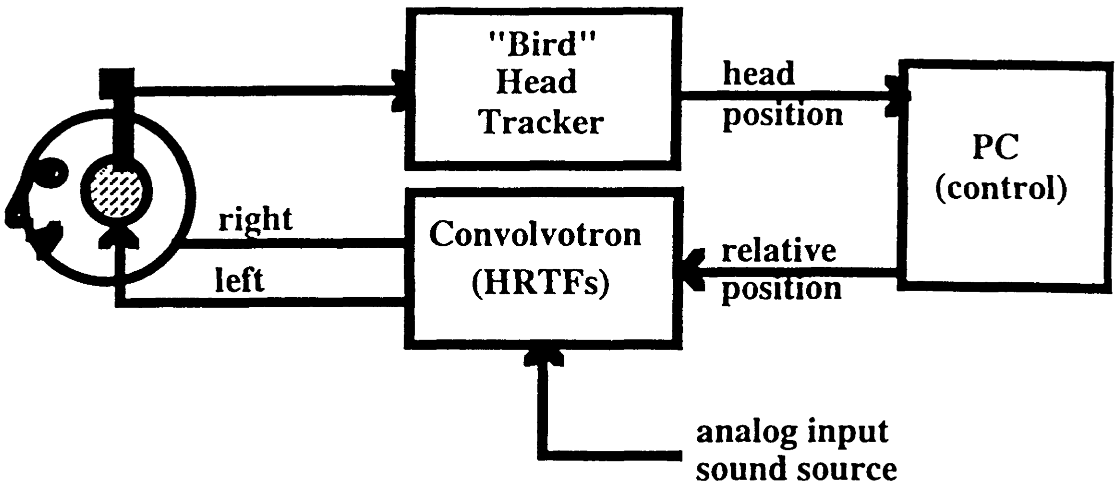


Figure 3: Block diagram of the auditory virtual environment used to simulate acoustic sources from Shinn-Cunningham, 1994.

III.2.2.1 Head Tracker

The head tracker used in this study is the Ascension Technologies Bird. The Bird uses electromagnetic signals to determine relative position and orientation between the transmitter and the receiver, worn on the subjects head. The Bird tracker update rate is 100 Hz, which is dominated by communication delays between the tracking device, the PC, and signal processor. The delay from head movement to when the new head position is reported to the signal processor is approximately 100ms.

III.2.2.2 Controlling PC

Prior to each session, an input file is created which determines the order of experiments to each session, along with the parameters for each run. This script controls the experiment throughout the session, storing data automatically for later analysis. The PC used in the current experiment contains a 486 processor with a 33 MHz clock speed.

III.2.2.3 Signal Processing Hardware

Signal processing is performed by the Convolvotron, a set of computer boards residing in the controlling PC. The Convolvotron can generate a binaural output signal by the simple superposition of the signals from each simulated source. HRTFs are stored in memory within the Convolvotron. The impulse responses of the HRTFs are used to convolve the input signal separately for each ear.

The inputs to the Convolvotron are amplified by a Crystal River Engineering anti-aliasing filter/amplifier before being sampled at 50 KHz and converted to a digital signal. The output signals are passed through an amplifier and sent to the headphones.

III.2.2.4 Headphones

The headphones used in the current experiment are the Sennheiser 545. Early experiments used Ear in-ear headphones along with Bilsom Viking industrial earmuffs which provided significant attenuation of background noise. However, with the introduction of the sound-proof room, the current study was able to utilize a more comfortable pair of Sennheiser 545 headphones that is a standard model used in many applications of auditory research.

III.3 Experimental Procedure

All localization experiments were based on the Head Related Transfer Functions (HRTF) taken from the HRTFs of subject SDO, a petite female, measured by Wightman (Wightman & Kistler, 1989a). Each subject performed eight sessions of multiple runs of testing and training, each one of which lasted approximately two hours. Each session began by testing and training auditory localization with the normal HRTFs followed by multiple test and training runs with the transformed HRTFs. The session concluded by retesting with the normal HRTFs. Each test consisted of 26 trials and each training session lasted 10 minutes. A session is outlined below.

Test 1 using normal cues (26 trials)
Training with normal cues (10 minutes)
Test 2 using normal cues (26 trials)
Break
Test 3 using *altered* cues (26 trials)
Training with *altered* cues (10 minutes)
Test 4 using *altered* cues (26 trials)
Training with *altered* cues (10 minutes)
Test 5 using *altered* cues (26 trials)
Training with *altered* cues (10 minutes)
Test 6 using *altered* cues (26 trials)
Break
Test 7 using *altered* cues (26 trials)
Training with *altered* cues (10 minutes)
Test 8 using normal cues (26 trials)
Training with normal cues (10 minutes)
Test 9 using normal cues (26 trials)
Training with normal cues (10 minutes)
Test 10 using normal cues (26 trials)

For each trial of a test, the subject was presented with a series of clicks and then asked to respond as to the perceived source location of the source. In most experiments, the subject responded by entering the number corresponding to the perceived source location. In an earlier experiment (described in section V.2), the subjects were blindfolded, and unable to respond accurately by typing. In this experiment, subjects responded by turning to face the apparent source location of the source.

During the training sessions, the subject was asked to track the source by turning his/her head to the correct location. The series of clicks continued until the subject turned his/her head to within 5 degrees of the location of the source.

IV Data Processing and Analysis

Bias and resolution will be the two aspects of performance that will be examined. By creating larger-than-normal interaural time differences and/or interaural intensity differences for a given source location, it is possible to generate supernormal cues that can increase localization sensitivity and enhance accuracy, so that localization error is reduced.

IV.1 Bias

Bias is defined as the error in mean perceived position induced by a change of cues. Bias is traditionally used to measure adaptation and is estimated as the difference between mean response and correct response, normalized by the standard deviation for the position. Bias β is

$$\beta_i = \frac{i - m_i}{\sigma_i},$$

where i is the cue location, m_i is the mean subject response for cue location i, and σ_i is the standard deviation of the subject response to cue location i. With the introduction of the transformation in Test Run 3, an immediate bias should be induced. Adaptation is measured by any decrease in bias over extended exposure to the transformed cues.

Changes in performance are investigated over the course of specific sessions. Test 1 (first normal cues), provided a control test to compare other results to. Test 3 (first altered cues) provided a measure of immediate effect of the transformed cues. Comparing test 3 and 7 (last altered cues), provided a measure of any decrease in the effect of the transformations. Finally, test 8 (return to normal cues) showed any negative after-effect from the exposure to the altered cues. The last set of training and test runs were performed to help the subjects re-adapt to normal cues.

IV.2 Resolution

Resolution is defined as the accuracy with which nearby positions can be discriminated. Resolution between adjacent positions is estimated as the difference in mean responses normalized by the average of the standard deviations of the position. Resolution d' between locations i and $i+1$ is

$$d'_{i+1,i} = \frac{m_{i+1} - m_i}{\sqrt{\sigma_{i+1}\sigma_i}}$$

where m is the mean subject responses, and σ is the standard deviation. Resolution measures the ability to discriminate between source locations. If the perceived source locations are perceptually close, resolution decreases as they become more difficult to discriminate from one another. The transformation remaps source locations such that acoustic location cues for positions in the front are spread over a wider physical range and acoustic cues for positions at the edges of the range are remapped closer together. Therefore, with the introduction of the transformed cues, it is predicted that resolution will be enhanced at the center and decreased at the edges.

V Previous Studies

There have been many studies of adaptation with variations in cues during training. The current study will be compared to two specific experiments performed by Barbara Shinn-Cunningham of the Research Laboratory of Electronics.

The same general procedure described in Section III.3 was used in previous experiments (Experiments A and C). However, differences in the procedure of earlier experiments resulted in very different results. These earlier experiments and results are reviewed in the following sections.

V.1 Experiment A

V.1.1 Experiment A Procedure

Previous studies of adaptation to rearranged auditory cues (Shinn-Cunningham, 1994c) show that subjects adapt to transformed auditory cues when they are given visual-spatial information about how auditory cues were rearranged. In Experiment A, the lights were used to show correct auditory position during training. During training, the light at the "correct" location of the auditory source was turned on at the same time as the auditory cues. When subjects turned their heads within 5 degrees of the source position, the light and sound were turned off. Thus, in this experiment, the subject received explicit visual information about the auditory transformation. During testing, responses in Experiment A were given by typing in the number corresponding to the heard source location (lights were not used).

V.1.2 Results of Experiment A

Previous results for bias and resolution for Experiment A (Shinn-Cunningham, 1994c) are shown in Figure 4 and Figure 5. Normal-cue runs (Test 1 and Test 8) are plotted with a solid line; altered-cue runs (Test 3 and Test 7) with dashed lines. The x's represent runs prior to altered-cue training exposure (Test 1 and Test 3), while the o's represent runs after exposure to altered-cue training (Test 7 and Test 8).

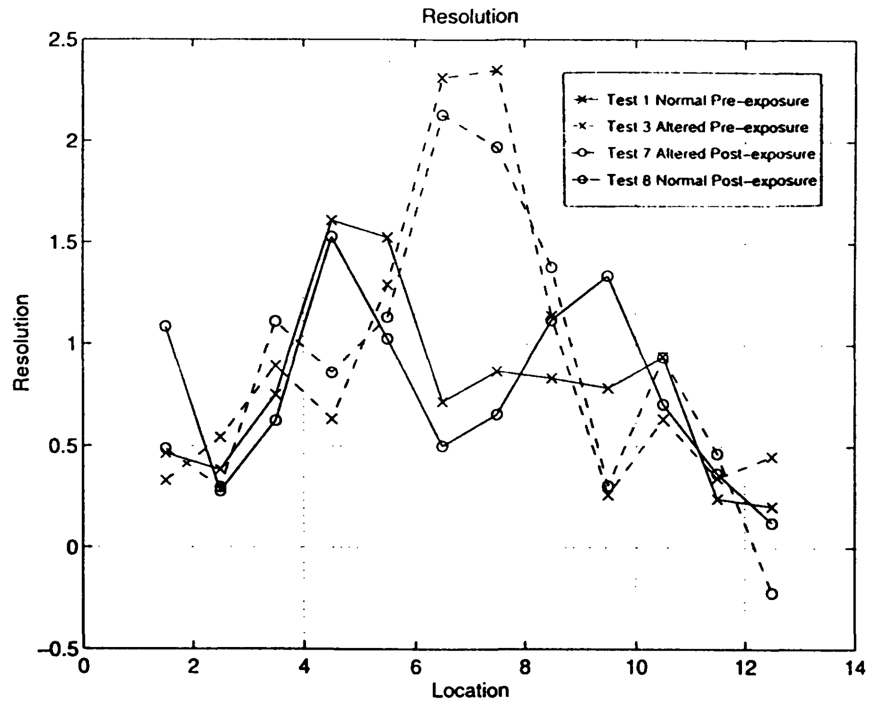


Figure 4. Resolution results as a function of cue location for Experiment A.

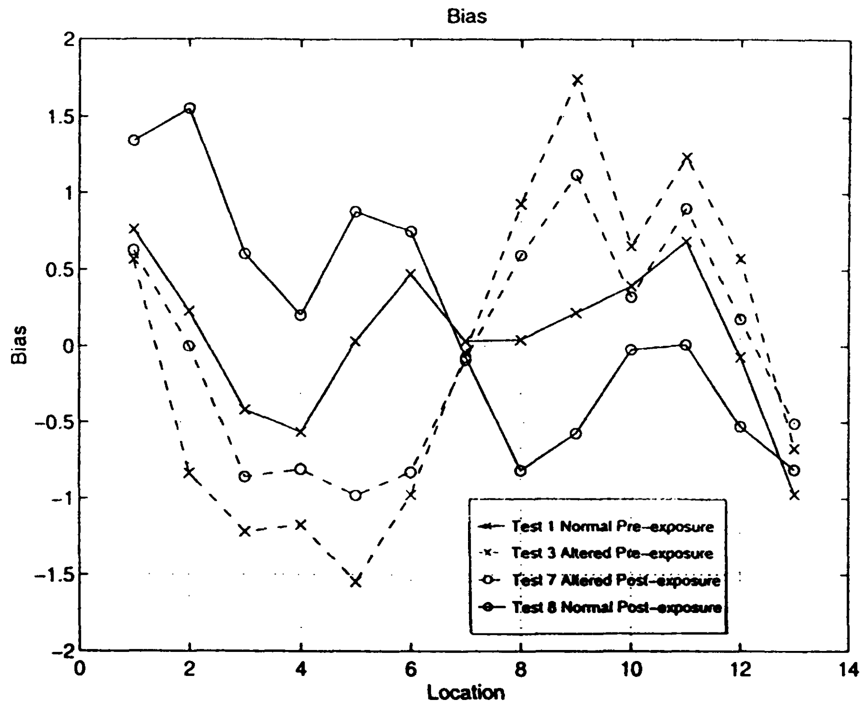


Figure 5. Bias results as a function of correct cue location for Experiment A.

V.1.3 Discussion for Experiment A

Results show that resolution increased at the center (straight ahead) and decreased at the sides, as expected, when exposed to the transformation remapping. A decrease in resolution was observed when comparing results prior to exposure (Test 3) to those following exposure (Test 7) with the supernormal transformation. A small decrease was also seen in the resolution results for Test 8 (first test with normal cues after exposure to supernormal cues) and Test 1 (first normal cue test).

Bias results show that a strong bias occurred in Test 3 (first altered test) in the direction predicted by the transformation. Results from Test 7 (last altered test) show a clear reduction in bias over all positions tested. However, this adaptation was not complete. Bias was reduced by roughly 30 percent. The first normal-cue test after exposure showed a strong after-effect (bias in the opposite direction) after training with supernormal cues.

V.2 Experiment C

V.2.1 Experiment C Procedure

A similar study (Experiment C) showed no adaptation when subjects were blindfolded throughout the experiment. The blindfolds eliminated both the visual field as well as explicit visual information about the auditory cue rearrangement throughout the experiment. Thus in Experiment C, the only cues available to the subject during training were the felt position of the head and how these cues corresponded to the auditory stimuli. Subjects had to turn their heads until they faced the auditory source, and the source was then turned off. During testing, subjects could not type accurately due to their blindfolds. Responses, in this case, were measured by having subjects turn to face the apparent source location after the stimulus ended.

V.2.2 Results of Experiment C

Bias and resolution results for Experiment C are shown below in Figures 6 and 7.

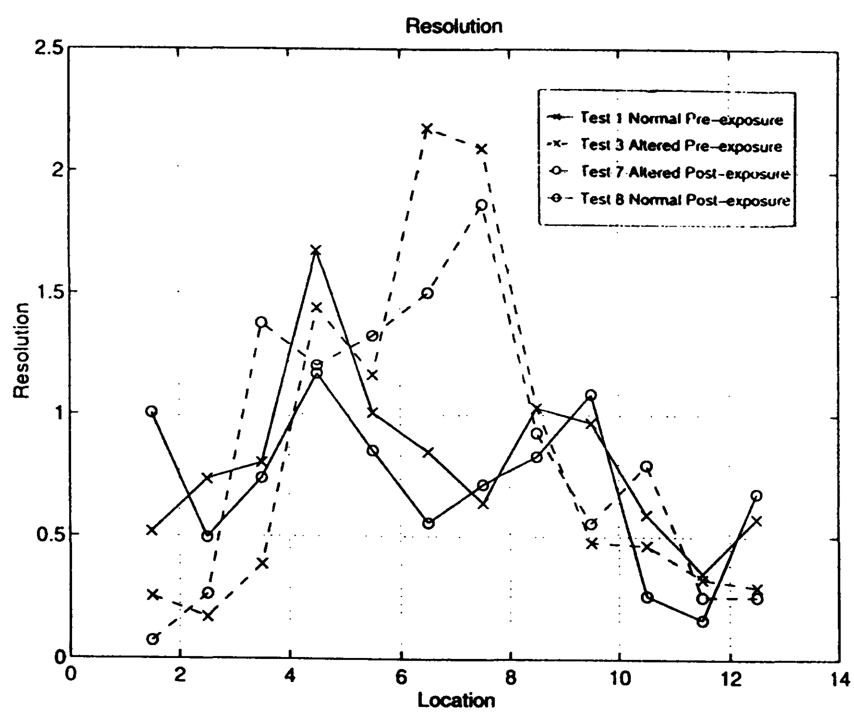


Figure 6. Resolution results as a function of cue location for Experiment C.

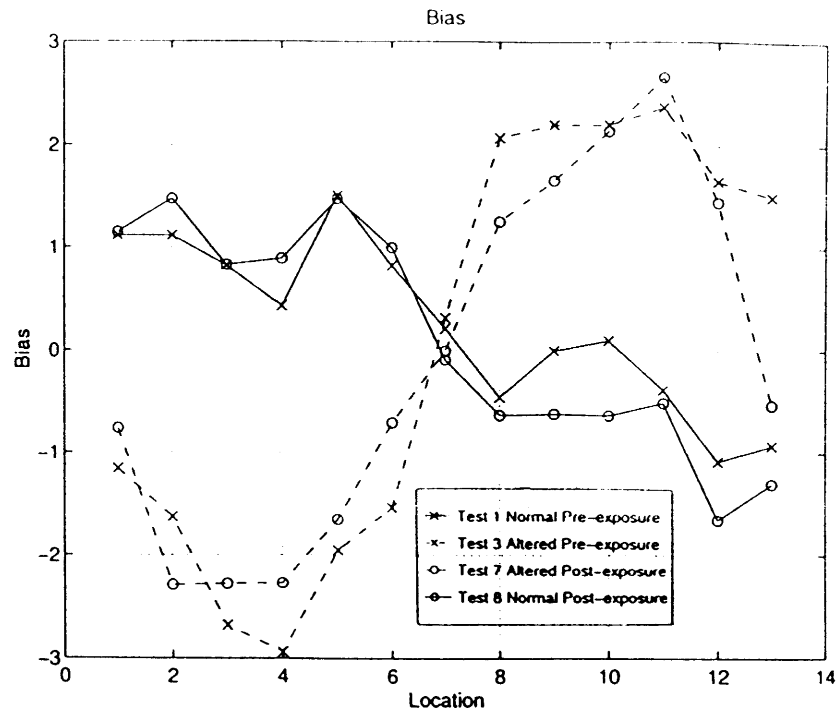


Figure 7. Bias results as a function of cue location for Experiment C.

V.2.3 Discussion for Experiment C

Resolution results for Experiment C show some enhancement in the center region (both before and after exposure) with altered cues. A slight decrease in resolution was also seen with exposure to the altered cues.

Bias results in Figure 7 show no reduction in bias with exposure to altered cue training. Therefore there was no adaptation.

V.3 Vision and Adaptation

There are three possible hypothesis to explain the previous experimental results. These different hypothesis lead to different predictions for the current experiment.

A) Adaptation did not occur in Experiment C because proprioceptive information about felt head position is inaccurate in the absence of visual field cues. However, when the visual field is available to subjects, head orientation is registered accurately. With this hypothesis, subjects will show full adaptation as long as the visual field is available during training.

B) Both explicit visual cues and the visual field are important and useful during training. In this case, we expect subjects to adapt in the current experiment, but not as completely as in Experiment A.

C) Subjects need explicit visual-spatial information about the auditory rearrangement or no adaptation will occur.

If hypothesis A is true, we expect to see adaptation when the visual field provides information to relate a reference in the visual field to head position during training. If hypothesis B is true then we predict an intermediate amount of adaptation to occur if only the visual field is provided to the subjects. Finally, if hypothesis C is true, then we expect no adaptation if the lights are not ever turned on as no explicit visual information about correct auditory location is introduced.

VI Current Study

In the current experiment, data was collected for 5 subjects. Each subject performed 8 sessions described in the experimental procedure of Section III.3. The data were averaged across all subjects and all sessions and compiled into 10 composite matrices (1 for each Test).

VI.1 Current Experiment Procedure

The experimental procedure used for the current study did not use the lights during training. Otherwise, the procedure was identical to the one used for Experiment A. The testing procedure was identical to the testing for Experiment A. This experiment explicitly tested the three hypotheses described in section V.3.

VI.2 Results of Current Experiment

Using the processing methods described in section IV, we observed the results shown in Figures 8, 9, and 10. Normal-cue runs (Test 1 and Test 8) are plotted with a solid line; altered-cue runs (Test 3 and Test 7) with dashed lines. The x's represent runs prior to altered-cue training exposure (Test 1 and Test 3), while the o's represent runs after exposure to altered-cue training (Test 7 and Test 8). Figure 8 shows the subjects' response and the relationship to the correct cue location. The responses were averaged across all subjects and across all sessions. Figure 9 shows the results on resolution and Figure 10 on bias.

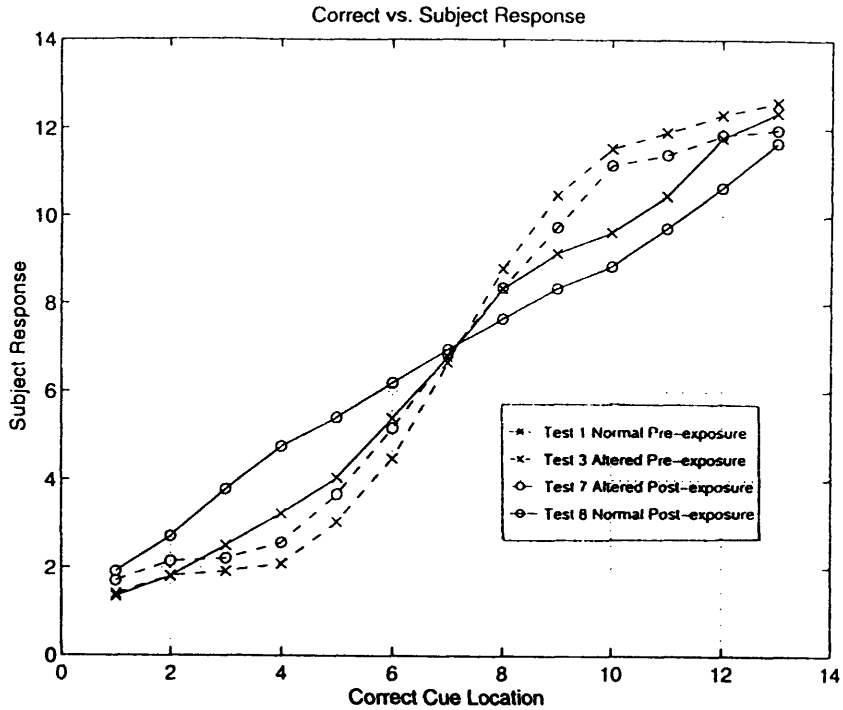


Figure 8. Subjects' responses as a function of correct cue location for current experiment.

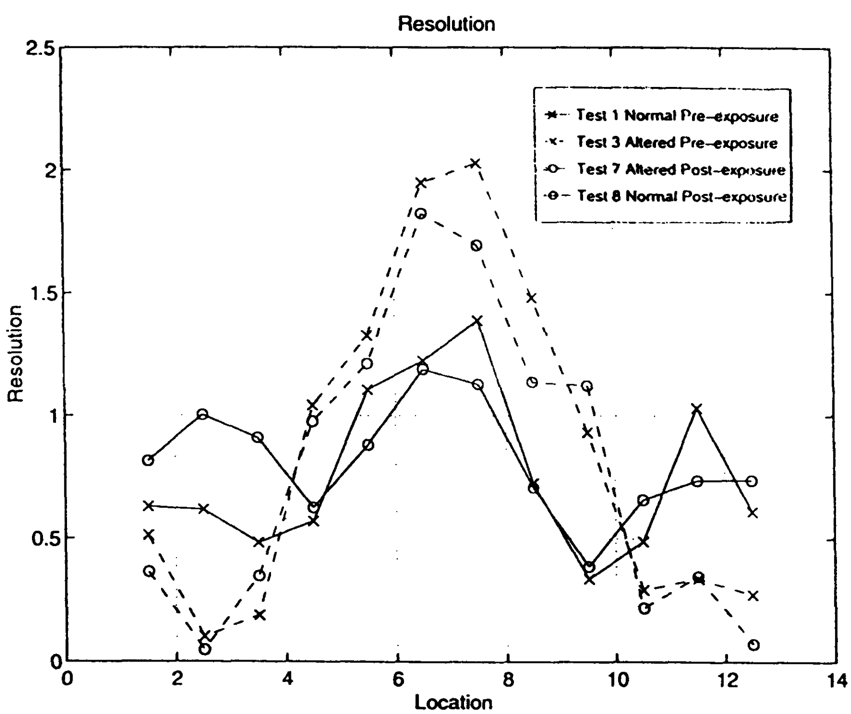


Figure 9. Resolution results as a function of cue location for current experiment.

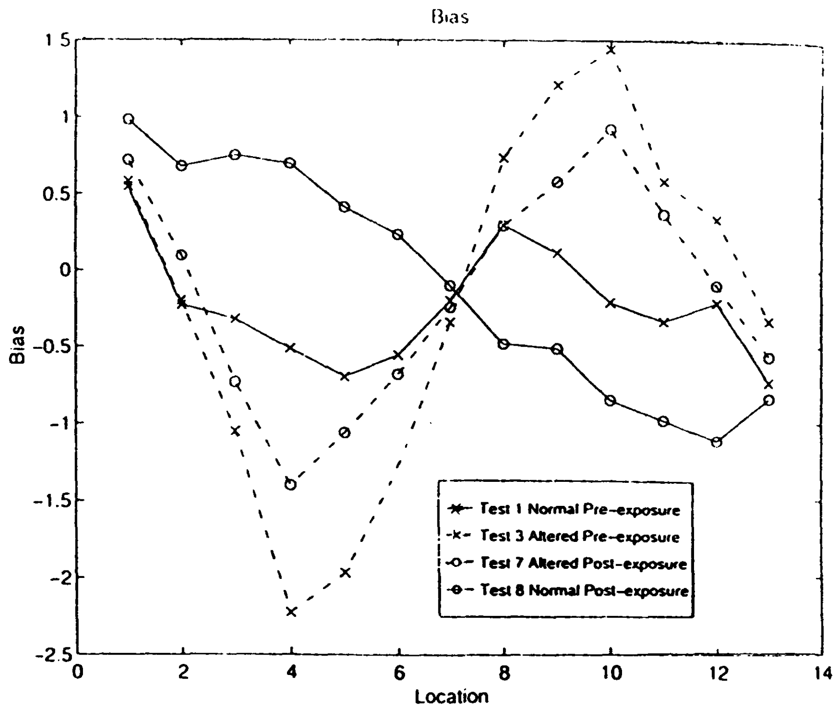


Figure 10. Bias results as a function of cue location for current experiment.

VI.3 Discussion of Current Experiment

From Figure 9, we see that the introduction of the transformation increased resolution at the center of the field, while decreasing resolution at the edges. We see this when comparing normal- and altered-cue tests (solid lines and dashed lines, respectively). This was expected and predicted, given the transformation employed.

Figure 9 also shows a slight decrease in resolution after exposure to the altered cue training. This was observed in previous studies and explained by the model used by Shinn-Cunningham, 1994. In many experiments, this decrease in resolution enhancement was found with extended exposure to the transformed cues during which there was a decrease in bias (adaptation).

The model predicts adaptation, as exhibited by a reduction of bias, causes a change in resolution. Therefore, the assumption in previous models, (that resolution and bias are independent) is not true. The model (Shinn-Cunningham, 1994) explains this relationship between bias and resolution as deriving from a single, central-processing mechanism.

Figure 10 shows bias results which are traditionally used to measure adaptation. From the figure, we see that a strong bias occurred in the first altered-cue run (Test 3) in the direction predicted by the transformation.

Adaptation can be measured by comparing the first altered-cue run (Test 3, dashed line x's) to the last altered-cue run (Test 7, dashed line o's). Figure 10 shows that subjects adapted to the transformed auditory cues as there is a clear reduction in bias over the range of positions tested. However, this adaptation was not complete.

The bias results also show a negative after-effect. This is seen by the increase in bias in the opposite direction when the subjects were retested with normal cues after extended exposure to the altered cues (comparing Test 1 to Test 8 in Figure 10).

VII Discussion

VII.1 Comparison of Previous and Current Results

By comparing resolution and bias results of Experiments A, C, and the current study, we can observe the effects of the differences in experimental procedure. Figure 11 and Figure 12 show the resolution and bias results for the previous experiments and the current study.

In Experiment A subjects were presented with explicit visual cues (i.e. correct answer feedback from lights) during training. In Experiment C, subjects were blindfolded, eliminating all visual cues during the entire experiment. Subjects had to rely solely on felt head position to register the auditory rearrangement in this case. In the current study, subjects were not blindfolded and the lights were not used. The presence of the visual field allowed them to associate a visual reference or position with head position during training. The subjects received no explicit visual cues during the current experiment. Table 2 summarizes the experimental procedures for these experiments.

Experiment	Testing	Visual Information during Testing	Training	Visual Information during Training
Experiment A	Subjects entered number	III	Subjects turned head	I
Experiment C	Subjects turned head	II	Subjects turned head	II
Current Experiment	Subjects entered number	III	Subjects turned head	III

I Visual field (labels) and explicit visual information (lights)
II None (blindfolded)
III Visual field (labels)

Table 2: Summary of experimental procedures.

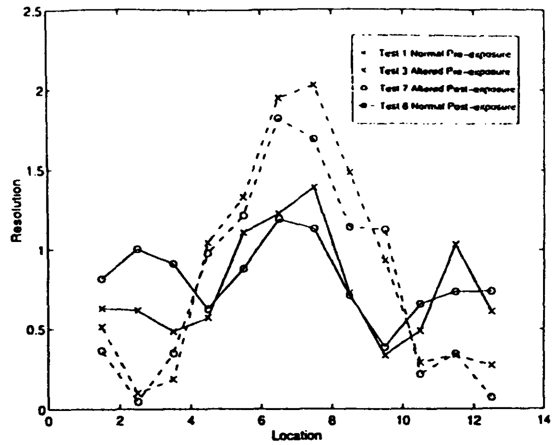
By using transformations in the mapping from auditory localization cues to source position, spatial resolution was enhanced directly in front of the subject and decreased at the edges of the range (sides of the subject). Figure 11 shows this was consistent in all three experiments.

Figure 11 also shows a slight decrease in resolution after exposure to the altered cue training for all experiments. This was observed in previous studies and explained by the model used by Shinn-Cunningham, 1994.

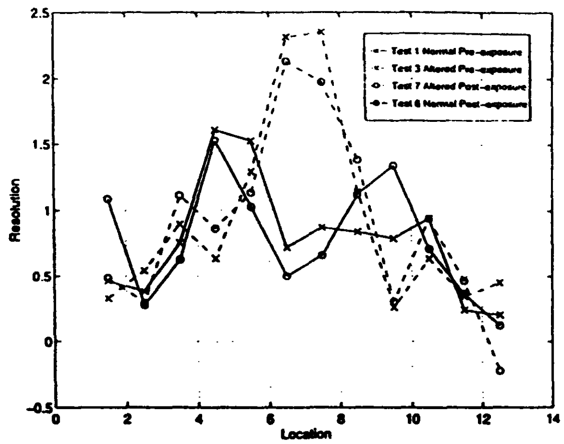
Figure 12 shows a clear reduction in bias after extended exposure to the altered cues in Experiment A and the current experiment. This indicates adaptation. Experiment C yielded no adaptation. When comparing the results on bias (adaptation) of the current experiment to the results of Experiment A (see Figure 12), subjects achieved approximately an equal level of adaptation in the current experiment when presented only with the visual field. Since bias results from Experiment C yielded no adaptation, subjects showed more adaptation when comparing the current experiment and Experiment C.

Given the variations in visual cue presentation and corresponding adaptation results, the experiments conclude that the presence of the visual field is important for achieving adaptation to transformed auditory localization cues. From the bias diagrams in Figure 12, the amount of adaptation appears to be approximately equal when explicit visual-spatial feedback is presented compared to conditions when no such information is available. It is clear that the presence of the visual field provides sufficient information to accurately register head position, and that adaptation to altered auditory localization cues is not dependent on explicit visual-spatial feedback.

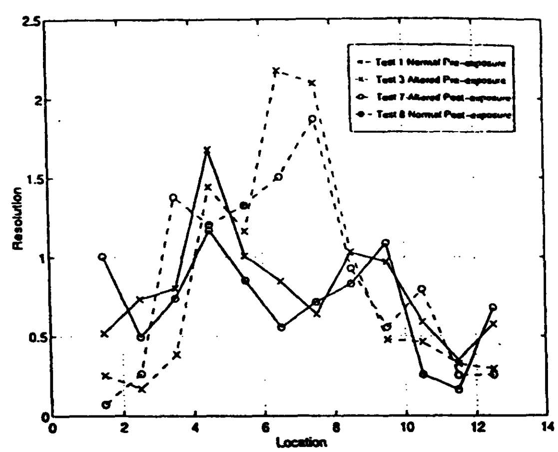
The observed adaptation in the current experiment may be due to the explicitly marked possible locations in the visual field. The lights and labels of the source locations may provide important information and may be another cue which is needed for adaptation to the transformed auditory localization cues. This possibility suggests a future study to examine if removing these visual markers affects adaptation when a visual field is present.



a) Resolution results for current experiment.

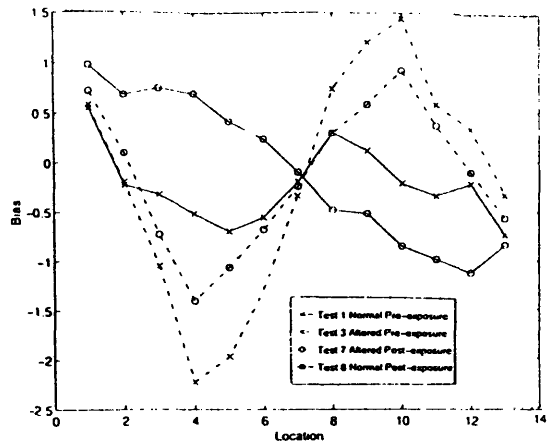


b) Resolution results for Experiment A.

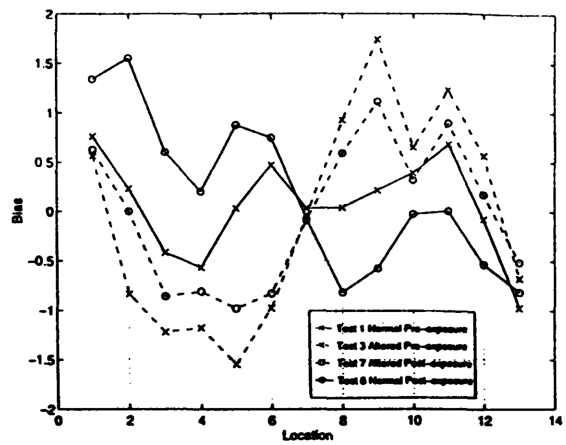


c) Resolution results for Experiment C.

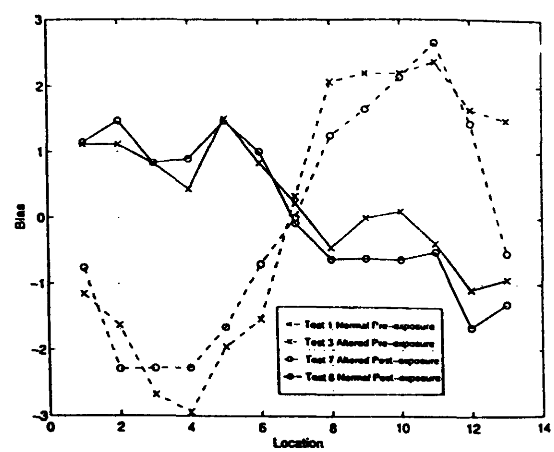
Figure 11: Resolution results for previous experiments and current study.



a) Bias results for current experiment.



b) Bias results for Experiment A.



c) Bias results for Experiment C.

Figure 12: Bias results for previous experiments and current study.

VII.2 Experimental Considerations

There are several considerations of the experiment that may have influenced the results of the study. Different subjects were used in different experiments which may result in inconsistencies. Physical differences, and the use of premeasured HRTFs (measured from a petite female) may have influenced the data. It is conceivable that using personalized HRTFs would produce the most accurate results and these differences may have affected individual performance. Also, physical deficiencies may have contributed to skewed data. For instance, it is possible that a subject may unknowingly have a slight impairment in one ear or a subject may have had an ear infection during a session that affected his/her performance.

Another consideration is the transfer of environments during the experiment. The first few sessions for each subject were performed in a room where background noise was present. With the addition of the sound-proof room, all (except the buzzing of the head tracker receiver) background noise was eliminated. This change in environment may have affected performance however, the averaging of all subjects and across all sessions normalized this inconsistency.

Finally, it is possible that the physical state of the subject may have affected performance. This is an uncontrollable factor that could have a great affect on the results. An example would be if a subject was extremely tired and fails to be attentive during training sessions. This would have a great influence on the bias and adaptation results.

Despite these problems, the results of the current experiment are consistent with previous experiments and show strong effects of training. In particular, results of each subject are similar to the results shown in the summary graphs in Section VII.1. As such, it is probable that these issues contribute to subject variability; however, the effects being measured are sufficiently robust so that such noise does not obscure our results.

VIII Conclusion

From the results of the current experiment, we tentatively conclude that hypothesis A is true. Hypothesis A states that proprioceptive information about felt head position is not salient in the absence of visual field cues. Subjects are unable to learn auditory remapping during training when blindfolded. In the presence of the visual field, subjects use the visual field to accurately register head position. Thus, adaptation is achieved as bias is decreased over extended exposure to the altered auditory localization cues.

The explicit markings of possible source locations (lights/labels), may provide another important cue to the subject during training. Further experiments are necessary to examine if different types of visual fields (i.e. fields not explicitly marked) provide any important cues that are needed for subjects to achieve adaptation.

References

Blauert, J. (1969). Sound localization in the median plane. Acustica, 22, 205-213.

Butler, R. A. (1987). An analysis of the monaural displacement of sound in space. Perception and Psychophysics, 41(1), 1-7.

Coleman, P.D. (1968). Dual role of frequency spectrum in determination of auditory distance. Journal of the Acoustical Society of America, 44(2), 631-632.

Durlach, N.I., Held, R.M., & Shinn-Cunningham, B.G. (1993). *Supernormal Auditory Localization*, Vol. 2 (pp. 89-103). Cambridge, MA.

Durlach, N.I. (1968). *A Decision Model for Psychophysics,* Communication Biophysics Group, Research Laboratory of Electronics, MIT. Cambridge, MA

Gardner, M. B. (1969). Distance estimation of 0-degree or apparent 0-degree-oriented speech signals in anechoic space. Journal of the Acoustical Society of America, 45(1), 47-53.

Gardner, M. B. (1968a). Historical background of the Haas and/or precedence effect. Journal of the Acoustical Society of America, 43(6), 1243-1248.

Molino, J. (1973). Perceiving the range of a sound source when the direction is known. Journal of the Acoustical Society of America, 53, 1301-1304.

Oldfield, S. R., & Parker, S. P. (1984a). Acuity of sound localization: a topography of auditory space. I. Normal hearing conditions. Perception, 13, 581-600.

Oldfield, S. R., & Parker, S. P. (1984b). Acuity of sound localization: a topography of auditory space. II. Pinna cues absent. Perception, 13, 601-617.

Perrott, D.R., & Saberi, K. (1990). Minimum audible angle thresholds for broadband noise as a function of the delay between the onset of the lead and lag signals. Journal of the Acoustical Society of America, 85, 2669-2672.

Pollack, I., & Rose, M. (1967). Effect of head movements on localization of sound in the equatorial plane. Perception and Psychophysics, 2, 591-596.

Searle, C.L., & Braida, L.D., Cuddy, D.R., & Davis, M.F. (1976a). Binaural pinna disparity: Another auditory localization cue. Journal of the Acoustical Society of America, 57, 448-455.

Shinn-Cunningham, B.G. (1994). *Adaptation to Supernormal Auditory Localization Cues in an Auditory Virtual Environment,* Massachusetts Institute of Technology Research Laboratory of Electronics. Cambridge, MA.

Thurlow, W. R., & Runge, P. S. (1967). Effect of head movements on localization of direction sounds. Journal of the Acoustical Society of America, 42, 480-488.

Wallach, H. (1939). On sound localization. Journal of the Acoustical Society of America, 10, 270-274.

Watkins, A. J. (1978). Psychoacoustical aspects of synthesized vertical locale cues. Journal of the Acoustical Society of America, 63, 1152-1165.

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

Young, P.T. (1931). The role of head movements in auditory localization. Journal of Experimental Psychology, 14, 95-124.