The Effect of Yaw Based Head Movement on the Perception of Source Elevation

By

Tommy Ashby

Submitted for the Degree of Doctor of Philosophy

Institute of Sound Recording
Faculty of Arts & Social Sciences
University of Surrey

2015

© Tommy Ashby 2015
Declaration of Originality

This thesis and the work to which it refers are the results of my own efforts. Any ideas, data, images or text resulting from the work of others (whether published or unpublished) are fully identified as such within the work and attributed to their originator in the text, bibliography or in footnotes. This thesis has not been submitted in whole or in part for any other academic degree or professional qualification. I agree that the University has the right to submit my work to the plagiarism detection service TurnitinUK for originality checks. Whether or not drafts have been so-assessed, the University reserves the right to require an electronic version of the final document (as submitted) for assessment as above.

Tommy Ashby 19/8/15
Abstract

Yaw movements are the most frequently occurring and largest head movements a listener makes when localising; however, previous research has not resolved whether yaw-based head movements are used in elevation localisation. A series of experiments was devised to investigate the impact of head movement on the elevation localisation response accuracy (LRA) of human listeners.

The experiments were conducted using a laser-guided pointing response method, as this was found to allow listeners to more accurately and consistently report a perceived source location than either verbal or graphical methods. 2 kHz low-pass filtered noise with and without a 6 kHz half-octave bandpass component were both shown to suppress pinna cues, and were therefore used to more clearly separate the effect of head movements.

Head movements were found to improve azimuth and elevation LRA for noise sources. Depending on stimulus and situation, head movements were shown to make an improvement of up to 8.5° in elevation LRA. Head movement improved LRA more for the 2kHz/6kHz filtered noise than it did for broadband noise; when pinna cues are impaired the significance of head movement cues increases. Both forced yaw movements and free movements significantly improved the elevation LRA.

Further experimentation was undertaken to determine whether the improvement in elevation LRA with head movement was caused by greater accuracy when a source is positioned in the listener’s median plane (a static cue), or by the act of moving the head (a dynamic cue). It was found that the static cue did not provide greater accuracy for sources close to the median plane, and hence it was concluded that dynamic cues increased the elevation LRA for yaw head rotations. For octave and half-octave bandwidth sources, static elevation LRA is lower when the listener has turned to face the source.

Yaw head movement improved elevation LRA for high frequency continuous signals, which suggests that dynamic interaural level differences are utilised. Head movements do not improve elevation LRA for programme items with less than an octave bandwidth. For octave programme items, head movements significantly improve elevation LRA, while static LRA shows no improvement; head movement cues are effective at narrower bandwidths than pinna cues. By detailing the nature of head movement cues, one can better inform the localisation model, creating a more accurate representation of the human localisation system.
Acknowledgements

First and foremost I would like to thank my tutors, Tim Brookes and Russell Mason, for their guidance, encouragement and patience. I would also like to thank my colleagues at the Institute of Sound Recording for their companionship, advice and tea break distractions; my listening test subjects who selflessly gave me their time and energy without which this research would not have been possible; my parents and family for their endless support in my efforts to 'push back the frontiers of science'; and Hannah Cummings for getting me out of bed in the morning and listening to my inane PhD related ramblings. The research documented in this thesis was supported by funding from the Engineering & Physical Sciences Research Council (EPSRC).
# Table of Contents

1. Introduction .......................................................................................................................... 1
   1.1 Elevation Localisation Models ...................................................................................... 2
   1.2 Head Movement in Localisation Models ..................................................................... 2
   1.3 Thesis Aim ..................................................................................................................... 3
   1.4 Research Questions and Report Structure .................................................................. 4
   1.5 Coordinates System ..................................................................................................... 7
   1.6 Localisation Response Error (LRE) ............................................................................. 9

2. Head Movements in Localisation .......................................................................................... 10
   2.1 Definition of Rotation Planes ....................................................................................... 11
   2.2 Listener Idiosyncrasy .................................................................................................. 12
   2.3 Rotation Angles ........................................................................................................... 12
   2.4 Rotation Direction ........................................................................................................ 14
   2.5 Movement Reversals ..................................................................................................... 15
   2.6 Programme Items ........................................................................................................ 19
   2.7 Loudspeaker Location .................................................................................................. 19
   2.8 Body Rotations ............................................................................................................ 20
   2.9 Visual Cues ................................................................................................................... 20
   2.10 Summary and Conclusions ......................................................................................... 21

3. The Effect of Experimental Method on LRA ........................................................................ 23
   3.1 Azimuth LRA .............................................................................................................. 23
   3.2 Elevation LRA .............................................................................................................. 25
6.5 Pilot Experiment 1.................................................................82
6.6 Pilot Experiment 2.................................................................91
6.7 Overall Summary and Conclusions .....................................104
7 Localisation Response Methods........................................106
  7.1 Response Methods ..........................................................107
  7.2 Localisation Response Methods Experiment ..................122
  7.3 Experiment Results.......................................................130
  7.4 Conclusions.................................................................154
8 The Effect of Head Movements on LRA..............................159
  8.1 Experimental Setup and Procedure..................................160
  8.2 Results ...........................................................................167
  8.3 Conclusions.................................................................192
9 Head Movement Localisation Without the Presence of Low Frequency ITDs 197
  9.1 Experiment Method.......................................................198
  9.2 Results ...........................................................................201
  9.3 Conclusions.................................................................226
10 The effect of programme item bandwidth on the localisation with head movement ..................................................229
  10.1 Experiment Setup.........................................................229
  10.2 Results...........................................................................233
  10.3 Conclusions.................................................................252
11 Do head movements improve elevation LRA by way of static or dynamic cues? .........................................................255
11.1 Experiment Setup ................................................................. 256
11.2 Results ................................................................................. 260
11.3 Discussion and Conclusions ......................................................... 275
12 Static and Dynamic Cues with Head Movement for Band-Limited Noise Sources ................................................................. 277
12.1 Experiment Setup ..................................................................... 278
12.2 Results ................................................................................... 281
12.3 Conclusions and Further Discussion .............................................. 294
13 Conclusions ............................................................................... 296
13.1 Research Questions .................................................................. 296
13.2 Original Contributions to Knowledge ........................................... 301
13.3 Applications .......................................................................... 303
13.4 Further Work ......................................................................... 305
Appendix A - Reported Location Calculation Method ............................ 308
A.1 Head Tracking System and Experiment Setup ................................ 308
A.2 Coordinate Systems .................................................................. 310
A.3 Calculation Overview ................................................................. 310
A.4 Actual loudspeaker Angle ............................................................ 312
A.5 Coordinates of the Head Centre .................................................. 312
A.6 Calibration .............................................................................. 313
A.7 Reported Location .................................................................... 315
A.8 Reported Loudspeaker Angle ..................................................... 319
A.9 Angle Error..............................320
A.10 Further Calibration..........................320
A.11 Verbal and Graphical Location Calculation..........................320
Appendix B – Loudspeaker Locations for Experiments 7 - 10........321
Appendix C – List of Publications..........................322
References..............................324
List of Figures

Figure 1 - Spherical coordinates system for describing source location [Bloom 1977] .......... 8

Figure 2 - Rotational head movements made when localising and their descriptive terms
......................................................................................................................................................... 11

Figure 3 - A plot from the study of Morikawa et al. [2013]. The red line shows the trajectory of the listener's head while auditioning the sound source; multiple reversals in yaw head movements are made. The green dots show the physical source elevation, while the blue dots show the source elevation reported by the listener. .... 16

Figure 4 - The head movements of a listener in Wightman and Kistler's localisation study
[1999] ................................................................................................................................................. 18

Figure 5 - Path of localisation as a function of centre frequency of narrow band noise from
anywhere on the median plane [Blauert 1997] .................................................................................. 28

Figure 6 - Interaural Level Difference as a function of frequency and azimuth [Feddersen
et al. 1957] ........................................................................................................................................ 35

Figure 7 - Graph shows how Interaural Time Differences varies with source location
[Feddersen et al. 1957] .......................................................................................................................... 37

Figure 8 - LRE with frequency (Hz) [Stevens and Newman 1936] ............................................... 38

Figure 9 – Example cones of confusion given for ITDs (μs) [Wightman and Kistler 1999]. 39

Figure 10 - LRE using LF, HF and Broadband Noise programme items [Ihlefeld and Shinn-
Cunningham 2011] .............................................................................................................................. 42

Figure 11 – Left: Results of listener controlled notch filter frequency against elevation,
Right: Listener HRTF Minima frequency against elevation [Bloom 1977] ................................. 45

Figure 12 - Mean azimuth LRE plotted against programme item duration and separated
into listeners (S:G and S:K) and movement condition [Pollack and Rose 1967] ............. 56

Figure 13 - Mean azimuth LRE separated by listening offset ......................................................... 57

Figure 14 - Proportion of trials where front-back errors occurred plotted against
programme item [Perrett and Noble 1997] ....................................................................................... 65

Figure 15 – Perceived source location against actual location showing up-down errors
[Perrett and Noble 1997] .................................................................................................................... 66

Figure 16 – The acoustic axis definition in the Genelec 8020 Operating Manual ..................... 77

Figure 17 - Laser pointing system .................................................................................................. 79

Figure 18 - User Interface ................................................................................................................. 82
Figure 19 - Absolute elevation LRE separated by loudspeaker number and programme item ..............................................................88
Figure 20 - Absolute elevation LRE plotted against loudspeaker number and programme item ..............................................................96
Figure 21 - Loudspeaker Number, loudspeaker Location and Tukey Grouping (denoted by G) ..........................................................................................................................99
Figure 22 - Plot of reported elevation against loudspeaker number for programme item WN .................................................................................................................................102
Figure 23 - Plot of reported elevation against loudspeaker number for programme item L2_B4 ..........................................................................................................................102
Figure 24 - 16-section system for stage 1 of an iterative localisation response method [Evans 1998] ..................................................................................................................110
Figure 25 - Graphical response method user interface used by Simon et al. [2010]........111
Figure 26 - A Truncated Icosahedron, the shape used to position loudspeakers for the test ........................................................................................................................................124
Figure 27 - Graphical Method User Interface .................................................................................................................................127
Figure 28 - Histogram of ANOVA Standardised Residuals ......................................................................................................................132
Figure 29 - Absolute elevation LRE separated by listener number and response method .................................................................................................................................133
Figure 30 - Absolute elevation LRE plotted against listener number ......................134
Figure 31 - Absolute elevation LRE plotted against loudspeaker and response method.. 137
Figure 32 - Absolute elevation LRE plotted against loudspeaker number and separated by response method. ..................................................................................................................138
Figure 33 - Mean absolute elevation LRE and 95% CI for each response method........141
Figure 34 - Absolute Elevation LRE plotted against loudspeaker number ..................142
Figure 35 - Proportion of front-back errors for each response method .........................144
Figure 36 - Histogram of standardised residuals given by ANOVA ...............................147
Figure 37 - Absolute azimuth LRE separated by loudspeaker number and response method ..................................................................................................................................148
Figure 38 - Absolute elevation LRE plotted against loudspeaker number for each response method group ........................................................................................................................149
Figure 39 - Absolute azimuth LRE separated by listener number and response method. 151
Figure 40 - Mean absolute azimuth LRE and 95% CI for each response method............. 153

Figure 41 - A Truncated Icosahedron; the vertices of this shape were used to position loudspeakers for the test.............................................................................................................. 162

Figure 42 - Absolute azimuth LRE separated by listener number........................................ 168

Figure 43 - Scatter graph of perceived against actual location for all listeners (r = 0.973) ............................................................................................................................................. 169

Figure 44 - Scatter graph of perceived against actual location with Listener 6 removed (r = 0.997)............................................................................................................................................. 169

Figure 45 - Absolute elevation LRE separated by listener .......................................................... 171

Figure 46 - Histogram of ANOVA standardised residuals ......................................................... 174

Figure 47 - Absolute elevation LRE separated by programme item ........................................... 175

Figure 48 - Absolute elevation LRE plotted against movement condition ............................... 177

Figure 49 - Absolute elevation LRE separated by loudspeaker, programme item and movement condition ............................................................................................................................................. 178

Figure 50 - Absolute elevation LRE separated by loudspeaker groups and movement condition. Loudspeaker group 1 - Sources within the field of fixation for all movement conditions; Loudspeaker group 2 - Sources within the field of fixation for the forced and free movement conditions only; Loudspeaker group 3 - Sources within the field for fixation of the free movement condition only ............................................................................. 180

Figure 51 - Absolute elevation LRE separated by programme item and movement condition ............................................................................................................................................. 182

Figure 52 - Absolute elevation LRE plotted against actual elevation angle .............................. 184

Figure 53 - Scatter graph of elevation LRE with actual loudspeaker location ......................... 185

Figure 54 - Signed elevation LRE separated by loudspeaker actual elevation angle................. 186

Figure 55 - Pie chart of front-back errors.................................................................................... 188

Figure 56 - Absolute azimuth LRE separated by programme item and movement condition without front-back errors ............................................................................................................................................. 191

Figure 57 - Frequency spectrum of the HF programme item.................................................... 200

Figure 58 - Histogram of elevation LRE ANOVA standardised residuals ................................. 204

Figure 59 - Elevation LRE separated by loudspeaker and movement condition ..................... 205

Figure 60 - Absolute elevation LRE separated by actual loudspeaker elevation and movement condition ............................................................................................................................................. 207
Figure 61 - Absolute elevation LRE separated by actual loudspeaker elevation and movement condition for the low frequency programme item ................................. 208

Figure 62 - Absolute elevation LRE separated by movement condition ....................... 209

Figure 63 - Elevation LRE separated by programme item and loudspeaker .................. 211

Figure 64 - Elevation LRE separated by loudspeaker elevation and programme item ..... 212

Figure 65 - Elevation LRE separated by programme item .............................................. 213

Figure 66 - Absolute elevation LRE separated by loudspeaker number ....................... 214

Figure 67 - Elevation LRE separated by actual loudspeaker elevation ......................... 215

Figure 68 - Proportion of front back error separated by movement condition and programme item ........................................................................................................... 217

Figure 69 - Histogram of absolute azimuth LRE ANOVA standardised residuals ......... 220

Figure 70 - Absolute azimuth LRE separated by movement condition .......................... 221

Figure 71 - Absolute azimuth LRE separated by movement condition and loudspeaker 222

Figure 72 - Absolute azimuth LRE separated by loudspeaker, movement condition and programme item ........................................................................................................... 224

Figure 73 - Absolute azimuth LRE separated by programme item ............................... 225

Figure 74 - Absolute azimuth LRE separated by programme item with loudspeaker 16 excluded ......................................................................................................................... 226

Figure 75 - Frequency spectrum of 'Octave' programme item ....................................... 231

Figure 76 - Frequency spectrum of 'Half-Octave' programme item ................................ 232

Figure 77 - Histogram of standardised residuals from ANOVA .................................... 236

Figure 78 - Absolute elevation LRE separated by loudspeaker number ....................... 237

Figure 79 - Absolute elevation LRE separated by actual loudspeaker elevation .......... 238

Figure 80 - Absolute elevation LRE separated by programme item and movement condition ..................................................................................................................... 240

Figure 81 - Front-back errors separated by movement condition and programme item. 243

Figure 82 - Absolute azimuth LRE separated by movement condition ......................... 247

Figure 83 - Absolute azimuth LRE separated by programme item .............................. 248

Figure 84 - Absolute azimuth LRE plotted against programme item and movement condition .................................................................................................................... 249
Figure 85 - Absolute azimuth LRE separated by movement condition, loudspeaker number and programme item ................................................................. 250

Figure 86 - Absolute azimuth LRE plotted against actual loudspeaker azimuth and movement condition ................................................................. 251

Figure 87 - Loudspeaker setup in azimuth shown from above ................................................................. 258

Figure 88 - Histogram of ANOVA Standardised Residuals ................................................................. 262

Figure 89 - Mean absolute elevation LRE separated by listener number ........................................ 263

Figure 90 - Absolute elevation LRE separated by loudspeaker number ........................................ 264

Figure 91 - Box plot of absolute elevation LRE plotted against loudspeaker number ...... 265

Figure 92 - Mean absolute elevation LRE with 95% confidence intervals plotted against actual loudspeaker elevation separated by plane number ........................................ 267

Figure 93 - Mean absolute elevation LRE with 95% confidence intervals separated by plane number and loudspeaker elevation number ................................................................. 269

Figure 94 - Mean absolute elevation LRE with 95% confidence intervals separated by plane number and loudspeaker elevation number (excluding speaker elevation numbers 1 and 8) ................................................................................................................................. 270

Figure 95 - Mean absolute elevation LRE with 95% confidence intervals separated by plane number ................................................................................................................................. 271

Figure 96 - Scatter plot of signed elevation LRE against loudspeaker number ......................... 274

Figure 97 - Histogram of standardised residuals given by the ANOVA ................................. 282

Figure 98 - Absolute Elevation LRE separated by loudspeaker number ................................. 283

Figure 99 - Histogram of the Standardised residuals given by the ANOVA ................................. 285

Figure 100 - Absolute elevation LRE plotted against listener number and separated by loudspeaker elevation number. Please refer to Table 1 for the elevation locations of the loudspeaker elevation numbers ................................................................................................................................. 286

Figure 101 - Absolute elevation LRE separated by plane number and loudspeaker elevation number. Please refer to Table 1 for the elevation locations of the loudspeaker elevation numbers ................................................................................................................................. 287

Figure 102 - Absolute Elevation LRE plotted against plane azimuth measured from the median plane ................................................................................................................................. 290

Figure 103 - Absolute elevation LRE separated by loudspeaker elevation number and programme item ................................................................................................................................. 291

Figure 104 - Absolute Elevation LRE separated by plane azimuth and programme item 293
Figure 105 - Setup of tracker for the experiment in Chapter 7 ............................................. 309
Figure 106 - Main locations used to calculate LRA (elevation view) .................................. 311
Figure 107 - Main locations used to calculate LRA (azimuth view) ................................... 311
Figure 108 - Location of head centre with respect to tracker ............................................. 313
Figure 109 - Angles used in calibration point calculation .................................................... 314
Figure 110 - Two-dimensional depiction of piercing point .................................................... 316
List of Tables

Table 1 - Mean azimuth LRE with azimuth angle [Stevens and Newman 1936] .................. 23
Table 2 - Summary of previous studies investigating azimuth LRA [Blauert 1997] ............. 25
Table 3 - Summary of elevation LRA [Blauert 1997] .......................................................... 26
Table 4 - Perceived location of tonal programme items [Pratt 1930] ................................. 27
Table 5 - Wallach’s localisation experiment - angles given in degrees [Wallach 1938] ...... 51
Table 6 - Major Differences in the experiments of Wightman and Kistler [1999] and Perrett and Noble [1997] .................................................................................................................. 69
Table 7 - Loudspeaker number and elevation angle for Pilot Experiment 1 .................... 81
Table 8 - ANOVA of absolute elevation LRE with factors programme item, loudspeaker and listener ................................................................................................................................. 87
Table 9 - Tukey B post-hoc test on the programme item factor ....................................... 89
Table 10 - Loudspeaker number and elevation angle for Pilot Experiment 2 ................. 93
Table 11 - ANOVA of absolute elevation LRE with factors programme item, loudspeaker and listeners ......................................................................................................................... 95
Table 12 - Tukey-B post hoc test groupings ...................................................................... 97
Table 13 - Tukey-B Grouping of loudspeaker number with elevation LRE ..................... 98
Table 14 - Programme items ranked by correlation of actual and reported loudspeaker location .................................................................................................................................... 101
Table 15 - A comparison of head and manual pointing methods in both an unsighted (dark) and virtualised environment [Majdak et al. 2008] ........................................ 117
Table 16 - ANOVA Tables for Absolute Elevation LRE .................................................... 131
Table 17 - Kruskal-Wallis Test of response method on absolute elevation LRE ........... 136
Table 18 - A Tukey-B post-hoc test conducted on the response method factor .......... 140
Table 19 - ANOVA conducted on the absolute azimuth LRE ......................................... 146
Table 20 - Results of Kruskal Wallis Test with test variable Absolute Azimuth LRE and fixed factor Response Method ......................................................................................... 152
Table 21 - Tukey-B post-hoc test on absolute azimuth LRE with response method ....... 153
Table 22 - Comparison of Azimuth and Elevation LRE for each response method ....... 155
Table 23 - Comparison of Azimuth and Elevation LRE for movement and no movement conditions ........................................................................................................................ 156

xv
Table 25 - ANOVA table of all variables and 1st-3rd level interactions for absolute elevation LRE

Table 26 - ANOVA table of all variables and 1st-3rd level interactions for absolute azimuth LRE with front-back errors removed

Table 27 - ANOVA of absolute elevation LRE data

Table 28 - ANOVA of absolute azimuth LRE data

Table 29 - ANOVA of absolute elevation LRE

Table 30 - ANOVA of absolute azimuth LRE

Table 31 - Loudspeaker elevation number and loudspeaker elevations in degrees for each vertical plane

Table 32 - ANOVA of absolute elevation LRE with factors speaker, repeat and listener

Table 33 - ANOVA of absolute elevation LRE with factors plane, loudspeaker elevation, repeat and listener

Table 34 - Pearson's correlation test results for each plane

Table 35 - Pearson's correlation test results for each plane excluding loudspeakers 1 and 8

Table 36 - Loudspeaker elevation number and loudspeaker elevations in degrees for each vertical plane

Table 37 - ANOVA of absolute elevation LRE separated by programme item, loudspeaker number and listener number

Table 38 - ANOVA of absolute elevation LRE separated by programme item, plane number, loudspeaker elevation number and listener number

Table 39 - Results of Mann-Whitney non-parametric test

Table 40 - Loudspeaker numbers, coordinates of loudspeaker locations (cm), and loudspeaker angles from CLP
# 1 Introduction

Identifying the location of a sound can be fundamental to an animal’s survival, and so it is no surprise that humans have evolved very accurate localisation capabilities [Perisa *et al.* 2004]. Localisation is the process by which the perceived location of a sound source is determined [Supper 2005]. On an everyday practical level, the ability to localise allows listeners to orient themselves and interact effectively with the physical world. Human listeners unconsciously localise sounds many times a day, often in challenging acoustical conditions.

The auditory system uses cues, provided by the sound signal arriving at the ears, to localise a source. In this instance a localisation cue is a stimulant (a physical attribute of the sound source signal) that suggests an auditory location. Models can be developed that make use of these localisation cues and these models can be used to predict the response of the listener’s auditory system. Developing auditory models allows one to evaluate perceived sound attributes without the need for time consuming listening tests [Rumsey *et al.* 2008].

Auditory models are developed because audio equipment manufacturers require a measure of perceived sound quality in order to give listeners the highest quality listening experience, which in turn allows them to create a more desirable product. Rumsey *et al.* [2005] found that spatial quality accounts for approximately 30% of the overall perceived audio quality. As such it is imperative that measures of spatial audio quality be included in an overall qualitative analysis of a listening experience. Attributes such as location, width, depth and envelopment have been defined to allow listeners to fully describe their perceived spatial experience. Supper [2005] states that perceived source location is the primary spatial attribute and Dewhirst [2008] justifies this by highlighting the overwhelming dominance of localisation literature in the spatial audio field.
1.1 Elevation Localisation Models

The number of channels used in surround sound systems has seen a rapid increase over recent years, with systems such as 10.2 and 22.2 surround being created in an attempt to improve spatial fidelity. These systems now include elevated loudspeakers that allow listeners to localise in elevation as well as azimuth. Computational auditory models that predict the perceived location of sound sources in terms of azimuth are already available in established research [Pocock 1982; MacPherson 1991; Supper et al. 2004], yet little has been done to predict the perceived elevation of a sound source. A thorough understanding of elevation localisation cues will allow more effective and immersive sound systems to be developed.

The third attribute of location, distance, will not be considered further in this report because there is a large volume of distance location research already available (Blauert [1997], Begault [2000] and Middlebrooks and Green [1991]) and these have shown that it is difficult for listeners to resolve. In this report, unless otherwise stated, the term source location refers to the direction of the sound source and localisation is the process the listener undergoes to determine the direction of the source.

1.2 Head Movement in Localisation Models

A listener will often move their head when asked to locate a sound source [Thurlow et al. 1967]. Yaw movements (rotations on the equatorial plane) are the largest and most frequent head movement listeners will make during localisation [Thurlow et al. 1967]. Wallach [1938] developed a theory based on these yaw head movements; by observing how the localisation cues change as the listener moves their head, localisation in both azimuth and elevation should be possible. However, most localisation models are based on static cues and do not take head movements into account. Results of previous experiments have suggested that head movements may improve static localisation cues by moving the source into a higher area of localisation acuity [Fisher and Freedman 1968] or may create the changing dynamic ‘head movement’ localisation cue [Wallach]
1938]. However, the head movement studies have either been primarily theoretical [Wallach 1938] or have been limited by an aspect of the localisation experiment methodology, such as focusing on a single location plane or using a flawed response method [Perrett and Noble 1997; Wightman and Kistler 1999].

Ashby et al. [2011] showed that a simple head sized sphere model with multiple microphone pairs could be used to simulate the dynamic localisation cues created through head movement and that three-dimensional localisation was possible for any location of azimuth and elevation using this sphere model. However, no study has conclusively shown whether listeners actually use the static or dynamic localisation cues created through head movement, how listeners combine cues from a range of head positions, nor how head movement cues are affected by listening conditions.

The localisation cues created though head movement may be important to elevation localisation and require investigation if an optimum spatial sound system is to be created. Furthermore, an elevation aware spatial auditory model could evaluate and analyse these systems quickly and accurately, giving an indication of the spatial impression listeners would perceive. In order to develop such a model, listening tests must be conducted to discover the cues that facilitate localisation and the acuity of that localisation capability.

### 1.3 Thesis Aim

The majority of studies into localisation cues and localisation models focus on static localisation. Listeners move their heads when localising and there is evidence that these movements might create localisation cues that improve localisation response accuracy. Furthermore most localisation studies focus on azimuth localisation response accuracy on the equatorial plane, ignoring elevation responses altogether. Studies that do focus on elevation localisation predominantly investigate pinna and spectral cues, with head movements often left as an undefined ‘interesting’ footnote. However, Wallach and others have indicated that head movements could provide an integral role in elevation localisation response.
The primary research question this thesis seeks to answer is:

How do yaw based head movements affect a listener’s elevation localisation response accuracy?

A spatial model has already been developed that derives cues from listener head movement [Ashby 2010]. It is necessary to verify whether these cues are actually used by real listeners and how these cues function. The secondary research aims are to optimise localisation experimental methods for elevation and head movement studies; to find out what cues, created by head movement, cause this improved localisation response accuracy; and to find the listening conditions that cause these head movement cues to manifest themselves. These findings will allow a more perceptually accurate head-movement-aware spatial localisation model to be developed.

In this thesis, the localisation response accuracy (LRA) of the listeners is a function of both their accuracy in perceiving the location of a source (localisation accuracy) and their ability to report that location. Many studies report the ‘localisation accuracy’ of a subject when they are describing the LRA; one must consider the effect that visual cues and localisation response methods have on the reported source location.

1.4 Research Questions and Report Structure

In order to answer the primary research question and thesis aims given in Section 1.3, a number of smaller research questions have been formulated.

1. How do listeners move their heads when localising a sound source?

Chapter 2 describes how a listener moves their head when localising. By showing how listeners move their head it is possible to suggest the localisation cues these movements will stimulate. Yaw rotations are show to be the primary head movement when localising. This section also shows the listening conditions in which head movement cues are most likely and how listener head movements change depending on the stimulus presented.
2. How do the experiment conditions affect the LRA of the listeners?

Chapter 3 of this report shows how human LRA changes due to listening conditions. This section will highlight how the localisation experimental methods adopted by the researcher affect the experimental results. This becomes significant in later chapters when trying to analyse and understand the differences between the head-movement localisation studies. It is also important to determine general human LRA if models are to be developed that mimic human localisation response. The more informed a model is by listening test data, the more accurately it can mimic a listener’s subjective response and so the more useful its output.

3. What are the static localisation cues?

Chapter 4 details the major static cues that enable listeners to localise. Interaural time differences (ITDs) and interaural level differences (ILDs) and spectral modification from the pinna are investigated. The listening conditions that cause these cues to fail are also highlighted.

4. What is the current evidence that head movements are an elevation localisation cue?

Chapter 5 introduces head movement as an alternative source of elevation localisation cues and investigates its importance. There have been a number of studies into head movement LRA since it was first proposed by Young [1931] and detailed by Wallach [1938]. Studies using modern localisation test methods have given contrasting results as to the significance of head movement localisation cues. Chapter 5 will describe and discuss the major head movement experiments in detail so that these differences can be explained. This section will also discuss the experimental methods of the previous studies so that an optimised version can be found. The experimental features discussed will be localisation response methods, movement conditions, programme items and loudspeaker placement.
5. How can the effect of pinna cues be suppressed so that head movement cues can be studied more effectively?

Localisation experiments often require the effect of certain cues to be suppressed so that the effects of other cues can be seen more clearly. Chapter 6 explains the methods previous researchers have used to supress pinna cues. Two pilot experiments conducted by the author are included in this chapter. These experiments seek to find programme items that can supress pinna cues and hence more clearly separate the effects of head movement.

6. Which localisation response method allows listeners to most accurately report a source location?

The goal of Chapter 7 is to establish the most effective method for eliciting a location response from listeners in localisation tests. This aspect is often ignored and makes comparing localisation studies difficult. This experiment allows comparisons to be drawn between the studies described in the literature that have used different response methods. Pointing methods are also studied with and without head motion to see how the ability to move their heads affects the LRA of the listeners.

7. Do yaw head movements improve elevation LRA?

It is unclear from previous literature whether yaw head movements improve elevation LRA. Therefore, an experiment investigating head movement as a possible localisation cue was conducted. The experiment design was based on the findings of the literature review and the preceding two experiments. Chapter 8 shows the design and results of this localisation experiment.

8. What cue, created by head movement, causes an improvement in elevation LRA?

Previous literature shows that ITDs, ILDs and pinna cues all operate in different frequency ranges. Chapter 9 describes an experiment investigating head movement localisation in different frequency regions, which will show whether head movement cues are dependent on ITD, ILD or spectral cues. Chapter 10
describes an experiment investigating the bandwidth at which head movement cues fail, which will show whether head movement cues are independent of pinna cues.

9. *Is the improvement in elevation LRA with yaw head movements due to static or dynamic cues?*

The findings of the head movement experiments described in Chapters 8 to 10 do not show whether elevation LRA is higher with head movement due to static or dynamic localisation cues. Static cues are improved by rotating the head until the source is in an area of higher LRA. Dynamic cues are created by observing the change in localisation cue for a given head movement. By restricting dynamic cues it is possible for static cues to be studied alone. Chapter 11 describes an experiment studying static elevation LRA on vertical planes of differing azimuth. When combined with the statement that listeners commonly orient towards the source location, found in Chapter 2, it is possible to conclude whether listener head movements improve their static elevation localisation cues.

10. *Are head movement cues for narrowband programme items dependent on static or dynamic cues?*

Chapter 10 shows that elevation LRA is increased with head movement for some reduced bandwidth programme items. Chapter 12 shows whether this increased elevation LRA is due to static or dynamic cues. The setup of the experiment is the same as the full bandwidth experiment (Chapter 11), except that narrowbandwidth noise programme items are used.

Chapter 13 states the main conclusions of the thesis and answers each of the research questions. This section will also described areas in which the results of this thesis can be applied and findings, tangential to the narrative of this thesis, that could be investigated further.

### 1.5 Coordinates System

In this paper a spherical coordinates system will be used to describe the location of a sound source, with the listener’s head at its centre. Three numbers are used
to represent the source location: azimuth (\(\varphi\)), elevation (\(\theta\)) and radial distance (\(r\)). An example sound source location, A, is given in Figure 1. For the purposes of this paper source elevation will span from \(-90^\circ\) (directly below the listener) to \(+90^\circ\) (directly above the listener).

![Figure 1 – Spherical coordinates system for describing source location [Bloom 1977]](image)

In this report, the equatorial plane is the horizontal plane that dissect the two ears. All points on the equatorial plane have the coordinates \((x, y, 0)\) from Figure 1, where \(x\) and \(y\) can be any value.

The median vertical plane is the vertical plane that dissect the listener’s nose, with each point on the plane being equidistant to the two ears. All points on the median vertical plane have the coordinates \((x, 0, z)\) from Figure 1, where \(x\) and \(z\) can be any value.
1.6 Localisation Response Error (LRE)

In this thesis, the localisation response error (LRE) was the main metric used to measure a listener's LRA. To calculate the LRE in each case, the actual angle to loudspeaker was subtracted from the reported angle to loudspeaker.

\[
\theta_{\text{error}} = (\theta_{\text{reported\_angle}} - \theta_{\text{actual\_angle}})
\]

Both the signed and absolute elevation LRE were calculated. The signed elevation LRE gave an indication of the bias of the listener's response while the absolute elevation LRE indicated the LRA.
2 Head Movements in Localisation

This chapter describes how listeners move their heads when attempting to locate a source. Various characteristics are discussed, including the rotation plane, rotation direction, rotation amplitude and repetition of movement, as well as changes in head movement type due to the influence of experimental factors such as programme item, source position, visual cues and experimenter instruction.

Static localisation cues can be ambiguous when considering sources in three-dimensions. Wallach [1938] showed that head movement could resolve this ambiguity by dynamically varying the localisation cues of the listener. By observing how the localisation cue changes for a given head movement, listeners could resolve both source azimuth and elevation. A multiple microphone sphere model was developed based on Wallach’s dynamic localisation cue [Ashby 2010] and was shown to work effectively, however whether human listeners actually use these cues has not been shown. Other sources suggest that head movements only improve a listener’s LRA by moving the source into a more accurate area of localisation, thereby increasing the listener’s static localisation cues [Pollack and Rose 1967].

To discover whether listeners use dynamic localisation cues created through head movement or static cues gain from a new listening position, first it must be shown how listeners move their heads when localising. The goal of this chapter is to gather information on the common head movements in order to suggest the available localisation cues created through movement. It will also allow the research to focus on the major head movements and highlight effective localisation experimental methods.

It is important to note that studying how listeners move their heads while localising does not indicate whether these head movements actually improve their LRA. Listeners may not be aware of their most effective head movement technique to optimise localisation.
2.1 Definition of Rotation Planes

Thurlow et al. [1967] divide head movements into three types based on their axis of rotation, namely rotation, tip and pivot. Other sources, Kim et al. [2013], Morikawa et al. [2013] and Wightman and Kistler [1999], choose and discuss alternative names for these movements. These movements are:

- **Azimuth, Rotation, Yaw**: movement about the vertical axis that extends from the top of the listener's head through the neck and torso.
- **Elevation, Tip, Pitch, Nodding**: movement about the horizontal axis that joins the ears.
- **Tilt, Pivot, Roll**: movement about the horizontal axis that extends from the listener's nose through the back of their head.

All subsequent discussion of head movements in this thesis will use the terms yaw, pitch and roll to describe their axis of rotation. This will allow head movements to be distinct from descriptions of the source location, which will use the Euler angles, azimuth and elevation. The term ‘rotation’, used by Thurlow et al. to describe the yaw motion, will be used to describe any circular movement about an axis. Figure 2 shows the three rotational head movements that are made when localising a source.

![Figure 2 - Rotational head movements made when localising and their descriptive terms](image)
2.2 Listener Idiosyncrasy

When attempting to localise, listeners' head movements are highly idiosyncratic; listeners move their heads to a lesser or greater degree depending on their own individual tendency and the localisation stimulus presented [Kim et al. 2010; Wightman and Kistler 1999]. Kim et al. [2010] showed that, when localising with head movement, the most significant factor that affects the various modes of head movement is the subject number, i.e. when localising, listeners vary in their use of head movement.

Thurlow et al. [1967] removed a ‘number’ of subjects from their study because they did not move their head at all when attempting to localise. They suggested that these listeners struggled to localise in three dimensions and that they were unaware of their best localisation method. They did not show any LRA data to back this assertion. An alternative hypothesis, based on the listeners’ no head movement, is that these subjects were satisfied using static binaural and spectral cues and had decided they did not need additional head movement cues. Wightman and Kistler [1999] state that listeners who are less able to resolve front-back confusions when listening statically are more likely to employ head movements when listening dynamically. This shows that if a listener has a localisation deficiency using one cue, they can use an alternative cue to compensate. Consequently, if listeners are to be removed from the study then they should be screened based on their LRA, not their method of localisation.

2.3 Rotation Angles

Yaw is the most common head movement direction used when localising. An experiment by Thurlow et al. [1967] compared the mean number of listeners (averaged across the loudspeaker factor) that used each head movement pattern (Table 1). There were a total of 23 listeners in the test.
<table>
<thead>
<tr>
<th>Movement Pattern</th>
<th>0.5 - 1 kHz Noise</th>
<th>7.5 – 8 kHz Noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yaw</td>
<td>9.4</td>
<td>11.2</td>
</tr>
<tr>
<td>Pitch</td>
<td>3.4</td>
<td>3.0</td>
</tr>
<tr>
<td>Roll</td>
<td>1.2</td>
<td>0.6</td>
</tr>
</tbody>
</table>

*Table 1 - Mean number of subjects using each movement plane*

Thurlow *et al.* considered rotations of larger than 3° as a ‘head movement’, which is approximately three times the minimum audible angle (MAA) in azimuth for frontal sources [Mills 1958]. The MAA is the smallest change in physical source location angle that can be perceived by the listener as a change in auditory location. Therefore, head movements that could have given the listener changing localisation cues (between the MAA of approximately 1° and the rotation angle set by Thurlow *et al.* of 3°) were not categorised as ‘head movements’. Basing the movement categories on the MAA would have significantly increased the mean number of subjects using each head movement and given a better indication of the number of subject who had dynamic localisation cues available to them. Chapter 3 will define the MAA in detail and discuss the main factors that affect it.

Yaw movements occur approximately three times more frequently when localising than pitch movements, which in turn occur 3 times more frequently than roll movements. Thurlow *et al.* concluded that yaw movements ‘clearly play an important role’ in attempts to localise sound.

‘Large’ yaw head movements (>10°) were made by 7.4 listeners, much more listeners than either pitch (0.6) or roll (0.4) movements [Thurlow *et al.* 1967]. Nojima *et al.* [2013] showed that listeners are also more likely to make larger head movements (> MAA ± 1°) than smaller ones (< MAA ± 1°), which indicates
that the listeners are significantly varying their localisation cues using head movements.

Combinational head movements that include yaw were also made frequently [Thurlow et al. 1967]: yaw and pitch occurred for 14.4 of the 23 subjects and yaw and roll for 4.4 subjects, while a pitch and roll combination only occurred for 1.4 listeners. Thurlow et al. also state that listeners are more idiosyncratic in their use of pitch and roll movements than yaw movements. This finding suggests that pitch and roll are possibly used to augment the fundamental localisation cues given by a yaw movement.

Thurlow et al. showed that the mean maximum yaw rotation angle was significantly larger than the mean maximum pitch and roll angles for a given source location. Kim et al. [2010] also showed that yaw was the largest head motion in both a localisation style test and a less experimentally controlled computer-game-playing scenario.

These findings show that a yaw rotation is the fundamental movement in localisation with head movement. Yaw movement is central to Wallach's dynamic localisation cue, described in Chapter 5.

### 2.4 Rotation Direction

A universal finding amongst head movement studies is that if the listener chooses to move their head, they are most likely to move it towards the source location. Wightman and Kistler [1999] stated that their compulsory movement condition, which was towards the source location, and their free movement condition, in which the listener could move their head in any way they wished, resulted in little difference in the listener's actual head movement. It could be argued that the compulsory condition may have biased the listener's response to the free movement condition. However, movement towards the source location was also shown by Morikawa et al. [2013] and Thurlow et al. [1967]. Thurlow et al. showed that the maximum movement angle in yaw and pitch is towards the source location.
Morikawa et al. [2013] showed that movements towards the source location are not large enough for the source to be directly in front of the listener, even when head movement was capable of it. This suggests that the listener is not purely trying to improve their static localisation cues. Morikawa et al. define the maximum possible head movement angle as 70° in yaw and state that when judging source azimuth listeners rarely use their full range of motion. Listeners also move their head more when judging source width and envelopment than source location [Kim et al. 2010]. This suggests that listeners are more confident judging source location than the other spatial attributes. It also shows that listeners are physically capable of making larger head movements when judging source location.

Thurlow et al. found that the mean maximum pitch angle for their experiment was approximately 14°, less than half of the mean loudspeaker displacement in elevation from the equatorial plane (30.5°). This means that the listener covered less the half of the elevated displacement of the source with head movement and so, in most cases, did not finish facing the source. Wightman and Kistler [1999] found that although listeners oriented towards the source location, they did not turn to face the source when localising unless specifically told to. The pitch movement produces no change in binaural cues for a given movement for source in elevation on the median plane. Listeners using this movement will not gain Wallach’s dynamic localisation cue.

### 2.5 Movement Reversals

A movement reversal is a change in the direction of head rotation made while the listener is auditioning the stimulus. Thurlow et al. state that, while auditioning a 5 second stimulus, some listeners could use three reversals of magnitude ≥10° while other listeners could use none. Nojima et al. show that listeners are more likely to make multiple head movements from side to side than orient ‘monotonically’ toward the source location when localising in azimuth.

The study of Morikawa et al. showed that listeners use different reversal movements when attempting to localise in azimuth and elevation. In azimuth,
listeners were found to most commonly make a single movement towards the source azimuth, possibly with small adjusting reversals. When localising in elevation on the median plane, listeners use a yaw movement sweeping past the source location employing a number of reversals (Figure 3). In elevation larger head movements appear to be used, between ±60° in azimuth. This movement appears to be used regardless of the source location. They do not discuss any other types of head movements (e.g. tilt or pivot) made during the localisation procedure.

Some head movement studies limited the stimulus so that no reversals were possible [McAnally and Martin 2014]. When studying listener head movement a listener must be allowed to make as many reversals in direction as they wish unless the study is specifically researching the effect of reversals. One of the

Figure 3 – A plot from the study of Morikawa et al. [2013]. The red line shows the trajectory of the listener’s head while auditioning the sound source; multiple reversals in yaw head movements are made. The green dots show the physical source elevation, while the blue dots show the source elevation reported by the listener. The left-most and right-most arrows indicate the scale which should be read for each plot: head rotation angle read from the left-hand scale and presented/perceived location read from the right-hand scale.
main focuses when attempting a head movement study is to allow the listener's movement to be as natural as possible.

Wightman and Kistler [1999] showed that when attempting to resolve front-back confusions, which are a manifestation of the cone of confusion on the equatorial plane, some listeners make frequent movement reversals. These movements, shown in Figure 4, are consistent with Wallach's hypothesised localisation cue. Plot 2 in the figure shows a listener making yaw movement reversals while converging on the source elevation with a smaller pitch movement.
one listener in the virtual source conditions. This listener doesn't help when I do.''

comments such as, ''I don't need to move my head, and it

of listeners who did not make large head movements elicited

ambiguity that led to those confusions. Informal questioning

the fact that there was no feedback regarding the actual source

case that most listeners in the freestyle condition adopted an

ments in the freestyle condition. This is perhaps the most

number of front–back confusions made large head move-

in style. Third, only listeners who demonstrated a substantial

are localizing real or virtual sources. It appeared to be the

interesting result from our visual analysis of head movement

3 degrees or greater in azimuth, elevation, and tilt). Table 1

TABLE I. Percent of trials in the real source restricted condition on which

![Figure 4](image_url) - The head movements of a listener in Wightman and Kistler's localisation study

[1999]
The reversals shown in the plots give an indication of the cues being used by the listeners: a single monotonic movement to the source location suggests that the listener is trying to optimise their static cues; frequent reversals suggest that the listener is trying to create dynamic cues.

2.6 Programme Items

Thurlow et al. used two programme items in their study: ‘low’, 500 – 1000 Hz bandpass filtered noise; and ‘high’, 7500 – 8000 Hz bandpass filtered noise. It was shown that larger yaw movements were made when attempting to locate the low programme item. Furthermore, the low programme item showed a higher number of ‘reversals’ than the high programme item. Pinna cues, described in detail in Chapter 4, are only effective when frequencies above approximately 4 kHz are present in the source signal. The ‘low’ programme item did not allow the use of pinna cues; pinna cues would have enabled listeners to localise more accurately. It is suggested that more extensive head movements are required when listeners are less certain of the source location. When pinna cues or other alternative cues are degraded, head movement cues are relied on more heavily to improve localisation. Morikawa et al. [2013] also showed that a larger yaw movement was used to localise high pass and low pass filtered noise programme items than a full bandwidth white noise programme item. Toyida et al. [2011] state that when only ITD or ILD localisation cues are present, larger head movements are required to resolve the source location. These results show that when the localisation task is more difficult, head movements play a larger role in localisation.

2.7 Loudspeaker Location

Thurlow et al. [1967] positioned loudspeakers at various elevations and azimuths, but no loudspeakers were located on or near the equatorial plane. It appears that Thurlow et al. assumed that sources further in elevation location from the listener would result in larger head movements. Morikawa et al. [2013] showed that, when localising in elevation, similar head movements are made regardless of the source elevation. In contrast to elevation, Morikawa et al.
showed that, when localising in azimuth, the degree to which a listener moves their head is increased as the source location azimuth angle is increased. This rotation will move the source into an area of increased azimuth LRA. These findings indicate that in azimuth listeners used head movements to optimise static cues, while in elevation they are used to create additional dynamic cues.

2.8 Body Rotations

Thurlow et al. [1967] prohibited the listener from using body movements while auditioning the stimulus. Wallach [1940] showed that body movements, or any other movements that stimulate proprioceptive, visual and vestibular senses, are as effective as head movements when attempting to localise a source. In this case proprioceptive senses are stimulated by ‘active bodily movement’ and proprioception is the ability to monitor these movements through muscle engagement and joint positioning [Wallach 1940]. By not allowing the listener to move their head they created unnatural listening conditions and so may have biased the listener’s response. Again the focus when conducting a head movement study must be on the naturalness of experience for the listener.

2.9 Visual Cues

Nojima et al. [2013] showed that visual cues have a profound impact upon head movement rate: when listeners were unable to see the source location (blindfolded) during the trial, they moved their heads for only 30% of the trials, while, when the loudspeakers were visible, head movements were used for 70% of the trials. They define a listener not moving their head as keeping it within the MAA (which they define as ±1˚ in azimuth for median plane sources and ±2˚ in azimuth at 45˚) for the duration of the stimulus. The main conclusion to be drawn from this study is that visual cues profoundly affect the way in which listeners move their head.

An unsighted movement frequency of only 30% is surprisingly low because the programme item used for the experiment was a 500 Hz low-pass filtered noise, which they state gave ‘no front-back localisation cues’ because the filtering suppressed the listener’s pinna cues. Therefore, one would anticipate greater use
of head rotation. However, listeners were only required to localise in azimuth and so interaural difference cues would have resolved the majority of localisation confusion. This may have caused listeners to assume that head movements were unnecessary.

Although the listener was unable to see the loudspeaker locations they may have realised loudspeakers were confined to the equatorial plane. There was not a significant difference in LRA and head movement use between the listeners that had seen that loudspeaker locations before the trial and were subsequently blindfolded, and those that were blindfolded before they entered the room. This indicates that the listeners were aware that sources were confined to the equatorial plane in all visual cases; it is also likely that the experimenters instructed the listeners that they were localising in azimuth only. ITD cues would have subsequently resolved the source location to only two locations on the equatorial plane rather than the whole locus of elevated source positions. Instructing the listeners that they were only localising in azimuth may have caused them to deem head movement unnecessary. However, there were no localisation cues to resolve front-back confusions if head movements were not employed and pinna cues were removed. It was shown that by moving their heads, listeners significantly improved their LRA. These findings can be seen as further evidence that listeners may be unaware of their most effective localisation method.

2.10 Summary and Conclusions

A yaw motion is the most frequently occurring and largest head movement used when localising. Pitch is the second largest head movement, followed by the relatively uncommon roll movement. Yaw appears to be the fundamental to the localisation process, while some listeners gain additional information using the other two rotational movements. The yaw movement is important to Wallach’s dynamic localisation cue hypothesis as it produces the changing interaural differences, which is discussed further in Chapter 5. Pitch movements on the median plane will not produce any changing interaural cues so are unlikely to furnish the listener with any additional dynamic cues. Guided by these findings,
the subsequent localisation studies described in this thesis will focus on the effect of yaw based head movements on the perception of elevation.

Listeners are most likely to orient towards the direction of the sound source location. In azimuth, this allows them to reduce their MAA, and so improves their azimuth LRA. It is unclear from previous research whether orientation toward the source location produces an increase in elevation LRA. This question will be studied further in Chapters 11 and 12.

When localising, a listener’s head movements are idiosyncratic; they depend on individual listener tendency and on the source stimuli. Listeners do not move their heads more than the MAA for every localisation trial. When combined with the studies showing the advantage of head movement, it is further evidence that listeners may not be aware of their most accurate localisation method. Listeners should not be removed from a head movement localisation experiment based on their use of head movement.

Larger head movements are made when listeners are asked to localise a low-pass-filtered noise when compared to a high-pass-filtered noise. The high-pass-filtered noise will allow the listener to make use of pinna cues and ILDs, which will allow a higher LRA in both azimuth and elevation. Head movements are more likely when the sound source is more challenging to localise, and when static localisation cues are compromised or inconclusive. Therefore, in a localisation study investigating head movement cues, a condition in which the effects of pinna cues are reduced is required. Chapter 6 describes experiments investigating pinna cue reduction methods.

When localising, repeated reversals of head movement are more likely than a single movement towards the source location. Head movement reversals result in a more accurate azimuth LRA when compared to a single movement. Reversals are also more likely when localising in elevation than in azimuth. These reversals are consistent with Wallach’s dynamic localisation cue, which is discussed in Chapter 5. Guided by this finding, movement reversals will not be restricted in the subsequent head movement studies.
3 The Effect of Experimental Method on LRA

The goal of this chapter is to show how the listening conditions of the localisation experiment affect the LRA of the listener. This will allow the optimisation of the experimental design used for the studies in this thesis. The LRA data given in this chapter can be compared with those shown in the thesis experiments and can also be used to inform an auditory localisation model.

Azimuth and Elevation LRA will be investigated separately in Sections 3.1 and 3.2 respectively, followed by a review of overall three-dimensional localisation studies (Section 3.3). Finally the concept of minimum audible angle (MAA) will be introduced and the MAA in azimuth and elevation will be discussed.

3.1 Azimuth LRA

This section will discuss the effect of the factors source location and programme item on the azimuth LRA.

3.1.1 Source Location

Azimuth LRA is not constant with sound direction; some sound directions are easier to localise than others. Stevens and Newman [1936] found that azimuth LRA was highest directly in front of or behind the listener (i.e. close to the median plane). Their findings are summarised in the Table 1.

<table>
<thead>
<tr>
<th>Distance from the median plane (°)</th>
<th>0</th>
<th>15</th>
<th>30</th>
<th>45</th>
<th>60</th>
<th>75</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean error (°)</td>
<td>4.6</td>
<td>13.0</td>
<td>15.6</td>
<td>16.3</td>
<td>16.2</td>
<td>15.6</td>
<td>16.0</td>
</tr>
</tbody>
</table>

*Table 1 - Mean azimuth LRE with azimuth angle [Stevens and Newman 1936]*

As the source is moved further in azimuth from the median plane, the azimuth LRE increases. In a similar study conducted by Gardner [1968] it was found that the average azimuth LRE for a source close to the median plane was smaller than Stevens and Newman [1936], at 1.5°. However, there was no indication of the
response method used in this study. It is unclear whether numbered loudspeakers (as used in other papers by Gardner), verbal elicitation or a pointing method was used. They do explain that the 7 speakers used were all visible during the study; therefore any acuity must be a function of the listener's visual acuity and the spacing of the loudspeakers as well as their auditory localisation ability. This is an early example of the effect of localisation response methods on the LRA. Response methods are discussed in detail in Chapter 7.

Other researchers have found a similar trend to Stevens and Newmans in azimuth LRA with source azimuth location. Makous and Middlebrooks [1990] found that ‘along any particular horizontal plane (i.e., at constant elevation), errors tended to increase with increasing azimuth.’ They also found that LREs were smaller for sources in front of the listener than those behind. There is also further experimental evidence in Chapter 8 of this thesis to support these findings.

3.1.2 Programme Item

Azimuth LRA varies depending on the source programme item used in the test. Stevens and Newman [1936] found that hiss and click programme items resulted in smaller mean LREs than a pure tone programme item (Hiss (LRE = 5.6°), click (LRE = 8.0°) and tone (LRE = 13.9°)). The short temporal nature of the click programme item will have excluded any head movement localisation cues, which may explain why the azimuth LRA for this programme item was lower than the hiss programme item. No mention of head movement was made during the study; it can be assumed that the listener was free move their head during the experiment.

Snow [1955] stated that for complex sounds such as speech and clicks the azimuth LRA could be 1° to 2°. The findings of a number of azimuth LRA experiments are shown in Table 2.
Table 2: Summary of previous studies investigating azimuth LRA [Blauert 1997]

<table>
<thead>
<tr>
<th>Author</th>
<th>Absolute Azimuth LRE</th>
<th>Programme item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stevens and Newman [1936]</td>
<td>13.9°</td>
<td>Pure Tone</td>
</tr>
<tr>
<td>Stevens and Newman [1936]</td>
<td>8.0°</td>
<td>Click</td>
</tr>
<tr>
<td>Stevens and Newman [1936]</td>
<td>5.6°</td>
<td>Hiss</td>
</tr>
<tr>
<td>Ford [1942]</td>
<td>6°</td>
<td>Pure Tone</td>
</tr>
<tr>
<td>Ford [1942]</td>
<td>2°</td>
<td>‘Impure Tone’</td>
</tr>
<tr>
<td>Snow [1955]</td>
<td>1° to 2°</td>
<td>Speech and Clicks</td>
</tr>
<tr>
<td>Gardner [1968]</td>
<td>1.5°</td>
<td>Speech</td>
</tr>
<tr>
<td>Makous &amp; Middlebrooks [1990]</td>
<td>2°</td>
<td>150ms noise pulse train (Bandwidth 1.8 – 16 kHz)</td>
</tr>
</tbody>
</table>

Blauert [1997] summarised the findings of a number of studies into forward LRA and states that on the equatorial plane the minimum absolute azimuth LRE is ‘about 1°’.

### 3.2 Elevation LRA

This section will discuss the effect of the factors programme item and source location on the elevation LRA. Studies of elevation LRA on the median plane using different programme items, summarised in Blauert [1997], are shown in Table 3.
### Table 3 - Summary of elevation LRA [Blauert 1997]

<table>
<thead>
<tr>
<th>Author</th>
<th>Programme Item</th>
<th>Absolute Elevation LRE (˚)</th>
<th>Number of Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blauert [1970]</td>
<td>Continuous speech by unfamiliar person</td>
<td>17˚</td>
<td>20</td>
</tr>
<tr>
<td>Damaske and Wagener [1969]</td>
<td>Continuous speech by unfamiliar person</td>
<td>9˚</td>
<td>7</td>
</tr>
<tr>
<td>Wetschurek [1971]</td>
<td>White Noise</td>
<td>4˚</td>
<td>2</td>
</tr>
</tbody>
</table>

It can be seen that, in general, elevation LRA on the median plane is significantly worse than azimuth LRA on the equatorial plane.

The results also suggest that the programme item has a significant impact upon the LRA. Plenge and Brunschen's findings, summarised by Rakerd \textit{et al.} [1999], showed that speech by a familiar person was resulted in a higher elevation LRA than unfamiliar speech. A 90\% correct response rate for familiar speech dropped to 50\% when replaced with unfamiliar speech. These findings are explained by the listeners dependence on spectral cues, which are described in Chapter 4.

For tonal programme items, reported source location is more dependent on the frequency of the tone than of its physical location in space. This was initially highlighted by Pratt [1930], in which he questions the reasons for pitch to be described as 'high' or 'low'. He described the attempts of previous authors to explain the metaphorical grounding of these terms before stating that his experiment shows that:

‘prior to any associative addition there exists in every tone an intrinsic spatial character which leads directly to the recognition of differences in height and depth along the pitch-continuum’.
In his experiment, listeners were instructed to locate the position of the source on a vertical scale numbered from 1 to 15. The vertical scale was 2.5 m in length, extending from the floor to the ceiling of the listening room, and ‘divided into 14 equal parts’. The loudspeaker location was hidden and randomly varied between five different locations on the vertical scale. Six listeners (A-F) participated in the experiment, with each listener making ten judgments. The mean results of these judgments are shown in Table 4.

<table>
<thead>
<tr>
<th>Pitch (Hz)</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>4096</td>
<td>12.4</td>
<td>10.4</td>
<td>13.6</td>
<td>13.4</td>
<td>10.0</td>
<td>14.4</td>
</tr>
<tr>
<td>2048</td>
<td>9.4</td>
<td>9.4</td>
<td>10.7</td>
<td>11.0</td>
<td>9.1</td>
<td>11.8</td>
</tr>
<tr>
<td>1024</td>
<td>7.8</td>
<td>8.3</td>
<td>8.8</td>
<td>8.0</td>
<td>8.4</td>
<td>9.7</td>
</tr>
<tr>
<td>512</td>
<td>6.4</td>
<td>7.1</td>
<td>7.2</td>
<td>7.4</td>
<td>6.8</td>
<td>6.9</td>
</tr>
<tr>
<td>256</td>
<td>4.6</td>
<td>6.2</td>
<td>6.4</td>
<td>5.4</td>
<td>5.8</td>
<td>1.9</td>
</tr>
</tbody>
</table>

Table 4 - Perceived location of tonal programme items on vertical scale numbered from 1-15 and extending from the floor (1) to the ceiling (15) of the listening room [Pratt 1930]

There is a distinct trend in perceived source elevation with frequency that is independent of the (randomly varied) actual source location. Blauert extends this theory stating that for source signals with a bandwidth less than 2/3 of an octave, location on the median plane is dependent only on signal frequency (independent of direction of sound source). The perceived path of this narrow band noise signal on the median plane is shown in Figure 5.
Gardner and Gardner [1972] compared elevation LRA on various vertical planes. From their results it is possible to see a trend in elevation LRA, from a fairly high LRA on the lateral vertical plane to poor LRA on the median vertical plane. Butler and Humanski [1992] also investigated LRA on the lateral vertical plane (LVP) and median vertical planes (MVP). They studied monaural and binaural localisation; both with high and low pass filtered noise. They noted that monaural elevation LRA was significantly lower than binaural elevation LRA on the LVP and suggested that the presence of interaural differences improved LRA.

Both studies used a numbered loudspeaker system as response method, a method whose limitations are discussed in Chapter 7. Perrett and Noble's [1995] paper highlighted flaws in the studies, showing that by using this method certain cues were resolved and so the listening task was simplified. Perrett and Noble concluded that the results of both studies were unreliable and so no previous study has conclusively shown how elevation LRA varies as a function of source azimuth. A pair of experiments investigating elevation LRA on vertical planes of differing azimuth are described in Chapters 11 and 12.
3.3 Three-Dimensional LRA

Oldfield and Parker [1984] attempted to create a map of three-dimensional LRA. The previous studies described in Sections 3.1 and 3.2 confined locations to a particular plane (e.g., the MVP or equatorial plane) to study either purely azimuth or elevation localisation cues. However, by informing listeners that sources were located only on a single plane, spatial ambiguities (such as the cone of confusion described in Chapter 4) were resolved, thus the tester biased the listener’s responses. By not restricting measurements to a single plane, it was hoped that clues to the interaction of pinna cues and interaural difference cues could be observed. The range of elevations was slightly limited, only stretching from -40° to +40° in elevation, but 360° in azimuth was possible.

They found that in normal listening conditions the degree of overall elevation LRE was 8.2°, while the overall azimuth LRE was 9.1°. Azimuth LRA was significantly worse behind the head, from 110° to 170°; while elevation LRE was more consistent throughout azimuth. At elevated positions behind the head, LRE in both azimuth and elevation was increased. They suggested this increase occurred because the structural intricacies of the pinna are not as significant for sources elevated and behind the listener (i.e., it blocks the sound no matter what angle of incidence).

Makous and Middlebrooks [1990] studied both elevation and azimuth LRA in a single experiment, stating that confining a response to a single plane ‘can fail to test the ability of the auditory system to resolve spatial ambiguity.’ They asked each listener to point their head at the perceived location of sound sources. The programme item had an amplitude spectrum flat between 1.8 and 16 kHz. The low frequency cut-off was chosen to create quasi-anechoic conditions [Kim 2009], however, it may have removed certain cues used at lower frequencies such as ITDs.

They found that the smallest LREs were 2° and 3.5° for azimuth and elevation respectively. Both these errors were for sources located at 0° azimuth and -5° elevation. They also found that although azimuth LREs were smaller than
elevation LREs for sources directly in front of the listener, for sources closer to the lateral vertical plane azimuth LREs were, in some instances, actually larger. The largest mean LREs were 16.3° and 19.1° for azimuth and elevation respectively. In both instances the source was at 160° in azimuth and elevated above the equatorial plane.

Both studies of three-dimensional localisation showed a smaller difference between elevation and azimuth LRA than those suggested by the single plane studies. These findings verify Makous and Middlebrooks [1990] assertion that localisation studies should not be confined to a single plane if overall LRA is to be verified. It is the interaction of cues in natural listening conditions that are to be studied if LRA is to be found.

### 3.4 Minimum Audible Angle (MAA)

As defined in the introduction, source localisation is the process by which the perceived location of a sound source is determined. Blauert [1997] defines localisation as 'the mathematical function relating the points of a physical space (sound-source) and those of the auditory space.' By using this definition one is able to separate the physical space in which the sound source is located and auditory space in which the listener perceives it. There may be a significant change in physical space location and no perceived change in auditory space. Alternatively, a change of location can sometimes be perceived in the auditory space when there has been no change in the source's physical location.

The ‘minimum audible angle’ (MAA) or ‘localisation blur’ is the smallest change in physical source location angle that can be perceived as a change in auditory location angle by the listener [Blauert 1997]. The MAA is affected by the location of the source and the type of sound signal generated. Mills [1958] highlighted the MAA as an alternative metric that can be used to show the precision or resolution of auditory localisation. Mills measured the MAA on the equatorial plane using a series of forced choice tests; an initial sound source was played at the test location and then a second source was played at a given test angle either to the left or right of the first one. The listener was asked to state the direction in which the source had moved. This procedure was repeated for a variety of test
angles and tonal programme items. By observing the proportion of correct responses for each test angle, Mills able to conclude the MAA for each test angle and programme item. The MAA is the threshold at which the listener is able to resolve the directional change in location of the source.

Mills concluded that a minimum MAA of $1^\circ$ occurred for 500-700 Hz tone sources directly in front of on the equatorial plane. The MAA increased as the source location was moved towards $90^\circ$ azimuth. At $90^\circ$ azimuth Mills reported a MAA of ‘always more than $40^\circ$’ for tonal sources, which was probably due to the occurrence of front-back confusions. Using a broadband source would have resolved a great number of these errors and reduced the MAA. Furthermore, the listener’s head was fixed in a brace during the trial; even small head movements would have improved the listener’s front-back performance [Wightman and Kistler 1999].

These findings led Mills to state that ‘the resolving power of localisation is at its greatest’ on the median plane (i.e. directly in front of a listener). Whether this is the case for elevation localisation is not shown in the study. A comparison of vertical plane LRA at different azimuths is described in Chapters 11 and 12.

A summary paper by Grantham et al. [2003] states that under optimum conditions of a broadband source presented directly in front of the subject previous studied have found that the MAA is approximately $1^\circ$. Using a KEMAR model to record the stimulus, they found that the mean MAA for broadband white noise on the equatorial plane, $45^\circ$ oblique plane, and median vertical plane were $1.6^\circ$, $2.8^\circ$ and $6.5^\circ$ respectively. However, they found that only 6 of the 20 subjects had a vertical plane resolution of $10^\circ$ or less, which they had set as their initial criterion. This indicates that a number of listeners struggled to gain vertical localisation cues from the KEMAR dummy.

Perrott and Saberi [1990], studied the MAA using a click train programme item, on the equatorial, median and oblique planes. They found that MAAs for sources near the equatorial plane had a mean MAA of approximately $1^\circ$, while sources on the median plane had a MAA of $3.65^\circ$. The minimum MAA found in the study was
0.78°. It should be noted that Perrott and Saberi used only 4 subjects in their trial and so whether these findings are generalizable to a larger population is unclear.

### 3.5 Summary and Conclusions

Azimuth and elevation LRAs vary depending on the spectral and temporal content of the source. In both azimuth and elevation, source programme items with a broader spectrum result in a higher LRA. Tonal programme items tend to have the lowest LRA. In azimuth, a click programme resulted in lower LRA than a continuous white noise programme. The short duration of the click programme may not have allowed the listener to utilise head movement cues effectively.

In azimuth, sources located in front of the listener (on the median plane) result in a higher LRA than those to the side (towards the lateral planes). It is unclear from previous research how elevation LRA changes depending on the azimuth location of the stimulus. The previous studies of Gardner and Gardner [1972] and Butler and Humanski [1992] have compromised their findings by using a flawed response method. Studies investigating elevation LRA on vertical planes of differing azimuth are described in Chapters 11 and 12.

In general, the most up-to-date localisation studies state that mean absolute azimuth LRE is approximately 2° for sound sources with precise cues on the equatorial plane. Studies in elevation are less common and their results are wider ranging, with studies stating elevation LRAs varying from a minimum of 4° to a maximum of 17° depending on the source attributes.

One main conclusion of this chapter is that LRA should be studied in both azimuth and elevation at once if meaningful results are to be obtained. Single plane studies suggest that azimuth LRA is significantly higher than elevation LRA. In the three-dimensional localisation studies there is only a small difference between elevation and azimuth LRA. One such study, by Makous and Middlebrooks [1990], showed errors in azimuth and elevation ranging from 2°-16.3° and 3.5°-19.1° respectively. These results show the large variance in error caused by the location of the source relative to the listener but little difference in LRA between azimuth and elevation.
MAA is the smallest physical change in the location angle of a source that the listener can perceive. The MAA is higher for sources on the median plane than on the equatorial plane. The most commonly reported MAA for ideal listening conditions on the equatorial plane was 1° [Gratham et al. 2003] and the minimum MAA found was 0.78° [Perrott and Saberi 1990].
4 Static Localisation Cues

The goal of this chapter is to define the static localisation cues currently thought to be most important to human localisation. Static cues are those derived from a stationary listening position, i.e. without movement or rotation of the listener's head. The process of localisation is undertaken by the combining of various cues derived from the sound waves as they reach the ears. The auditory system receives a whole series of cues and analyses the plausibility of each, in an effort to localise a source [Hartmann 1990]. Deriving the cues used in localisation is a continuing cumulative process and no complete auditory localisation model has been developed. Defining the prominent static cues allows one to predict how these cues will vary when localising with head movement (i.e. in dynamic listening conditions).

Auditory localisation cues are often split into two categories: binaural and monaural. Binaural cues involve a comparison of the signals arriving at each ear, while monaural cues are derived from the signal at a single ear. Interaural differences, described in Section 4.1, are a binaural cue while pinna-filtering effects, Section 4.2, are a monaural cue.

4.1 Interaural Difference Cues

In 1907, Lord Rayleigh described the interaction between the auditory system's two primary azimuthal localisation cues, Interaural Level Differences (ILDs) and Interaural Time Differences (ITDs). These cues utilise the differing characteristics of the sound signal as it arrives at each ear. These differences occur because the ears are spatially separated on either side of the head.

4.1.1 Interaural Level Difference

Interaural Level Differences (ILD) are calculated by comparing the relative levels of sound at each ear. The listener's head casts a sound shadow that causes the sound pressure level at the ear furthest from the source to be lower in amplitude. The plot in Figure 6 shows how ILD varies with frequency and source location.
As shown, ILDs are larger when the frequency of the source signal is higher. A 3 kHz sine wave at 90° results in 10dB of attenuation while 6 kHz results in 20dB [Begault 2000]. At low frequencies, ILDs become negligible as the incident sound wave can diffract around the listener's head. Moore states that ILDs become ‘negligible’ at below 500Hz, while Begault suggests this value is nearer to 1kHz. There is no absolute cut-off as the reduction in ILDs is gradual and varies depending on the source signal, listener and listening conditions.

**4.1.2 Interaural Time Difference**

Interaural Time Difference (ITD) is the difference in time of arrival of the signal at the ears. If the signal reaches one ear before the other then it is likely that the source is closer to the preceding ear [Cheng and Wakefield 2001]. With a simple sinusoidal signal, the brain compares the phase of the two signals to work out the time difference. Above 750 Hz these phase comparisons become inconclusive.
as it is impossible for the brain to discern which waveform is leading [Moore 2004: 237]. Furthermore, above approximately 4-5 kHz frequency the ear’s inner hair cells no longer fire at regular intervals during the wave cycle [Moore 2004: 45]. This is known as a ‘breakdown in phase locking’ and means that, for sinusoidal signals, ITDs are no longer indicative of source location.

Woodworth [1938] simplified the head’s physiology down to a spherical model and deduced Equation 2 to describe the ITD. This assumes the source is distant enough for the sound waves to be considered plane [Kim 2009].

\[
ITD = \frac{r(\theta + \sin \theta)}{v}
\]  

\( \theta \) is the lateral angle of the sound source (the azimuth displacement from the median plane, as described in Section 1.5), \( r \) is the radius of the head, \( v \) is the speed of sound and \( ITD \) is the interaural time difference. Using this equation, the maximum interaural time delay is approximately 661µs for a head diameter of 17.5cm (diameter from Blauert [1997]). Feddersen et al. [1957] plotted interaural time difference over azimuth (Figure 7). Moore [2004] states that in practice this time delay varies with frequency and so this graph is only an approximation. The plot does show that ITDs vary consistently over azimuth and so can be used as an azimuth cue.
As the source’s angle of azimuth increases towards 90°, the ITD increases, in a near linear trend. From 90° to 180° the inverse trend is noted; as source azimuth is increased, the ITD decreases.

4.1.3 The Duplex Theory

The ‘Duplex Theory’, formulated by Rayleigh [1907], suggests that sound localisation in azimuth is dependent on ITDs at low frequencies and ILDs at high frequencies. As previously stated, ITDs are limited by the breakdown in phase locking and phase confusions at high frequencies, while ILDs are limited by diffraction at low frequencies. Therefore, it is likely that for a broadband programme item the cues derived from ITDs and ILDs are combined using this duplex theory.

Mills [1958] found that for frequencies between 1.5 and 6 kHz, dichotic presentation of pure ILDs via headphones gave an azimuth location response that very closely mimicked the response curves for actual physical localisations. This finding indicated that ILDs were the main cues used at those frequencies.
The duplex theory indicates that there is a middle zone in which neither cue is effective. Stevens and Newman [1936] found that error in localisation was highest at 3 kHz, the crossover between ITD and ILD cues (Figure 8). Mills [1958] found that the MAA was greatest at approximately 2 kHz, again showing that in the crossover between ITDs and ILDs azimuth LRA is poor.

4.1.4 Amplitude Envelope ITD

Henning [1974] showed that ITDs could be detected for high frequency signals if their amplitude was modulated by a low frequency tone. It was shown that with 300 Hz modulation of a 3.9 kHz carrier, the location could be was determined as accurately as when localising a pure 300Hz tone. It has also been shown that time delay information can be derived from high frequency transients or the onset and offset of high frequency sounds [Nuetzel 1976]. Therefore, when creating programme items to be used in listening tests, care must be taken to avoid transient onset and offsets as they may create addition undesired localisation cues.

4.1.5 Cone of Confusion

When considering sources placed in three-dimensions, interaural differences suggest a number of possible locations resulting in a 'Cone of Confusion'. The most obvious example of the cone of confusion ambiguity occurs on the median vertical plane that separates the two ears. There are no interaural differences at any elevation on this plane so it is impossible for the listener to use these cues to
differentiate between source locations. For every interaural difference there is a corn of confusion, which means that three-dimensional localisation using static interaural differences alone is impossible. Figure 9 shows cones of confusion as lines of equal ITD [Wightman and Kistler 1999].

Figure 9 – Example cones of confusion given for ITDs (µs) [Wightman and Kistler 1999]

4.1.6 Front – Back Confusion

The most common example of the cone of confusion occurs on the horizontal plane, where listeners often cannot discern whether a source is in front of or behind them.
‘Actually, interaural differences, be they time or intensity, do not aid us in making a front-rear discrimination if the sounds are tonal stimuli’ [Musicant and Butler 1983]

Stevens and Newman [1936] found that for tonal sources below 3 kHz discrimination of front-back location was little better than chance, while for sources above 3 kHz the frequency of front-back confusions fell, occurring in approximately \( \frac{1}{6} \) of trials. They suggested that there was a cue at higher frequencies that enabled listeners to discriminate front and back.

To test this further they randomly varied the intensity of the high frequency source programme item (3-7 kHz) to compare it with the constant intensity source and found that the front-back confusion rate climbed from 18.6% with the constant intensity source to 47% with random intensity source. Stevens and Newman suggest that in the constant intensity test the listener had gained reference loudness for sources in front of them and had compared each subsequent stimulus against it. They suggest that the “sound-shadows from the pinna” cause this difference of intensity and when the stimuli were not replayed at a constant level this reference could no longer be used. This was one of the earliest suggestions that the shape of the ear and pinna could be a cue in localisation, a theory discussed in detail in Section 4.2.

### 4.1.7 Interaural Differences in Reverberant Conditions

Most studies that measure LRA conduct their tests under anechoic conditions. Using anechoic conditions allows the tester more control of variables and makes computer modelling of the test less complex. However, anechoic conditions do not simulate a realistic environment for most natural listening situations and so removing reverberant information may reduce the relevance of the test results.

Ihlefeld and Shinn-Cunningham [2011] tested how a listener’s azimuth LRA changed in the presence of reverberant energy. They suggested that listeners should be able to adapt to reverberant conditions and find the optimum combination of interaural cues.
They used three programme items in their experiment: low-pass-filtered noise, high-pass-filtered noise and broadband noise. The low-pass noise was created to give the listener mainly ITD cues, while the high-pass noise was created to give mainly ILD cues.

They found that for sources more than 45° from the median plane, the perceived location of the source was biased towards the median plane. This effect was found in all localisation instances but was more pronounced when the direct to reverberant ratio was decreased. This means that when reverberation was higher, the LRE was also higher. This is easily interpreted in the ILD case, as reverberation is common to both ears, so reduces the difference in level ratio between the ears. It is more difficult in the ITD case, where time differences remain constant. However, the highly reflective listening environment may make it more difficult for the listener's auditory system to calculate the ITD, thus resulting in a higher response error.

When the direct to reverberant ratio was low, high frequency programme items were found to result in a higher LRA than low frequency. This suggests ILDs are more reliable for localisation than ITDs when in reverberant conditions. They then tested the broadband programme item against the HF and LF noise. They found that the location of a programme item containing only high frequency content was more accurately reported than a broadband programme item (Figure 10). This suggests that weighting of low and high frequency information for localisation is not always optimised for a reverberant space.
The results of Ihlefeld and Shinn-Cunningham’s tests show that if listeners were to weight cues differently depending on their reverberant conditions, a more accurate LRA would be possible. The imprecise weighting shows that a listener’s localisation system is not optimised to most effectively use the localisation cues available.
4.2 Pinna Cues

The pinna is often cited as the hearing system's primary vertical and front-back localisation mechanism [Perisa 2004; Begault 2000]. The complex structure of the pinna causes time delays and comb filtering effects that change the spectrum of the sound wave before it arrives at the tympanic membrane [Rogers 1981]. This results in different transfer functions for each vertical location.

Although early researchers such as Rayleigh [1907] and Wallach [1938] suggested that the pinna has an impact on sound localisation, most research focused on interaural differences. Wallach [1938] described how front-back and elevation localisation was effective without head movements, which he suggested was made possible by the ‘selective sound shadow of the pinnae’. He described the elevation localisation resolution provided by the pinna as ‘crude’, but provided no results to explain why he thought this was the case. He did state that front-back discrimination was reliable using just the pinna.

A demonstration performed by W. B. McLean was one of the first instances in which the importance of the pinna cues was highlighted [Batteau 1967]. He showed that by contorting the shape of the pinna it was possible to confuse the subject's localisation capability. This demonstration suggested that the pinna performs an acoustical transform upon the incident sound wave that is unique for each sound source location. To allow localisation, the brain must learn the spectral patterns resultant from these transforms for every angle of incidence. Thus, the listener’s response to pinna cues is cumulative and may change or improve over time.

Batteau [1967] states that localisation by persons totally deaf in one ear is commonplace which shows that localisation is possible using monaural pinna cues alone. He provided a mathematical model of the delays given by the pinna in the time domain. The shape of the pinna causes delays of a few microseconds that produce comb filtering in the source’s spectrum at the eardrum [Middlebrooks and Green 1991].
However, static spectral cues do not completely resolve source elevation. To discover the effects of the pinna on the perceived signal it is necessary to know the spectrum of the original source signal [Hofman and Opstal 1998]. The original source material may already contain spectral notches that suggest an elevation angle. This means a model based on spectral effects is hard to develop.

In Chapter 3 it was shown that speech by a familiar person was localised more accurately than unfamiliar speech [Rakerd et al. 1999]. This is because *a priori* knowledge of the frequency content of the source programme item allows any spectral modifications due to the source’s location to be highlighted. Blauert [1997] stated that the familiarity in Plenge and Brunchen’s study was established only a ‘short time’ before the actual experiment, suggesting that it did not take long for a listener to become familiar with a source’s spectrum.

Based on the theory that every source elevation has a differing spectral characteristic at the listener’s ears, Bloom [1977] suggested that by varying the spectrum of the sound source signal replayed over a loudspeaker, different illusory source elevations could be created. One of the main premises of Bloom’s experiment was that if the loudspeaker was at or above 60˚ elevation and the subject was listening monaurally then “one may neglect the pinna’s effect or at least consider it constant for frequencies below 10 kHz.” So they state that by using octave wide white noise centred around 8 kHz and a high source elevation, one could produce a spectral ‘blank canvas’.

Bloom’s pilot experiment tested whether a single notch filter moved smoothly across frequency, used to replicate the shifting sensitivity minima of the ear’s Head Related Transfer Function (HRTF), could produce a smooth elevation shift in the listener’s perception. In an informal test, conducted during an Audio Engineering Society Convention presentation, Bloom found that over 75% of the listeners did indeed perceive a smooth elevation change. He was able to shift the source from +40˚ to -30˚ in elevation, as the notch filter was moved from 10 kHz to 6.3 kHz respectively.

Bloom then went on to probe the relationship between the phantom source location and the notch frequency value that suggests it. A loudspeaker was
located at +60° elevation and 90° azimuth, and octave wide white noise centred on 8 kHz was produced. The notch filter frequency of this loudspeaker could be varied by the subject and was used as a “moveable pointer”. A second loudspeaker, located at various elevations from -60° to +45° was added. This produced unmodified octave wide noise and was used as the ‘target’. The subject was asked to change the spectrum of the first loudspeaker until it appeared to be at the same elevation as the second one. The objective was to see how the frequency of notch selected by the listener corresponded to the minima noted in the listener’s HRTF response. The results, shown in Figure 11, compare the notch filter centre frequency selected by the listener when trying to replicate the location of the target loudspeaker using the moveable pointer (left), with the minimum frequency found in the listener’s HRTF response curves (right). Both are plotted against source elevation.

![Graph](image1.png)

**Figure 11 –** Left: Results of listener controlled notch filter frequency against elevation, Right: Listener HRTF Minima frequency against elevation [Bloom 1977]

It can be seen that reported elevation increases linearly as notch centre frequency is increased. This trend is also apparent in the graph plotting HRTF minima against elevation. These results not only imply that elevation can be suggested by changing the spectrum of the signal (i.e. each location has its own unique notch filter frequency), but also that the frequency of the notch filter is the same as the most prominent trough in the listener’s HRTF for each elevation.
Langendijk and Bronkhorst [2002] outlined three methods for analysing spectral elevation cues. Firstly, they suggested measuring the HRTFs for a number of source locations, plotting the subsequent spectral graphs and looking for significant trends in the spectra of the HRTFs as the source location changes. A number of cues can be taken from the HRTF analysis and it is difficult to tell which ones actually make a difference to elevation localisation. Secondly, localisation experiments involving band-limited programme items can be used to isolate certain frequency areas for study. However, they themselves suggest:

“It is uncertain if cues derived from band-limited signals can explain localization of broadband sounds.” [Langendijk and Bronkhorst 2002]

To take the extreme example, can cues and subject localisations based on sine tones be extrapolated to gain insight into full bandwidth white noise? Thirdly, the pinna can be occluded in some way to suppress spectral cues. This has been proven to have a significant effect upon LRA [Gardner and Gardner 1972; Perret and Noble 1997]. However, it is difficult to suppress the spectral cues in a systematic way that may give some clue as to how localisation is taking place. Furthermore occluding the pinna often risks damaging the listener's hearing. Chapter 6 discusses pinna cue suppression methods and attempts finds an optimised method for head movement localisation studies.

4.3 Summary and Conclusions

This chapter has shown that there are two main static localisation cues, namely interaural differences (ITDs and ILDs) and pinna/spectral cues. Static interaural differences are binaural, as they depend on the signal to both ears, while static pinna cues are monaural and rely on the spectra at each ear individually.

Interaural differences occur because the ears are located on opposite sides of the head. This spacing causes timing differences (ITDs) in the signals reaching each ear, while level differences (ILDs) are caused by the acoustic shadowing of the head. For constant amplitude signals, ITDs are not used as a cue above approximately 4 kHz due to a breakdown in phase locking of the hair cells and phase confusions for high frequency sine wave sources. In contrast, ILDs are
negligibly small at frequencies below 500Hz and become larger at higher frequencies. ITDs can be used at higher frequencies if the signal is amplitude modulated by a signal below 4 kHz. Interaural differences were shown to follow Rayleigh’s duplex theory, which states that the auditory system predominantly uses ITDs at low frequencies and ILDs at high frequencies.

The pinna produces comb-filtering effects that change the spectrum of the incoming signal. Listeners learn the spectral signature of each location and so can use the change in spectrum to localise the source. In static listening conditions, for complex spectra, *a priori* knowledge of the source spectrum is required so that any change in spectrum is noted. Further facets of pinna cues, such as the bandwidth required for pinna localisation, are discussed in Chapter 6.

In conclusion, static localisation cues are limited; static interaural cues result in the cone of confusion, while static pinna cues require *a priori* knowledge of the original source spectrum. Head movements offer a solution to the uncertainties of both interaural cues and pinna cues. By noting how interaural differences vary as the head is moved a single point upon the cone of confusion can be resolved [Wallach 1938]. Head movements might also allow the underlying spectrum of the source to be revealed so that pinna cues can be used effectively regardless of the original source spectrum. Head movement cues are discussed in detail in Chapter 5.
5 Dynamic Localisation Cues

Chapter 2 showed how a listener moves their head while attempting to localise a sound source. The goal of this chapter is to find out whether these head movements can be used to facilitate improved LRA. In general, head movements could cause improved LRA by allowing the listener to dynamically process changing cues or by moving the source into an area of higher static LRA.

The main studies into localisation using head movement are individually discussed in the following sections. The results and the experimental methods of the studies will be compared. It will be shown that the experimental methods had a profound effect on the experimental outcomes. The results will show whether listeners used dynamic cues to increase their LRA. The findings will also inform the experimental methods adopted in this thesis.

5.1 Wallach [1938]

Wallach was one of the first to suggest that the listener's head movements could provide a vital cue in azimuth and elevation localisation [Wallach 1938]. He stated that head movements offer dynamically varying interaural differences that allow a discrete point upon the cone of confusion to be resolved.

Wallach suggested that when the head is moved it is the change in interaural differences that the listener uses to ascertain the vertical location of a source. If the source is directly overhead then yaw head movements will not affect the interaural differences. If the source is level with the ears then the same yaw head movements will produce a significant variation in interaural cues.

For a given head movement, the listener must observe the change in lateral angle. Wallach expresses this mathematically as:

\[
\gamma = \frac{\partial \psi}{\partial \beta} \tag{3}
\]
In this equation, $\psi$ is the lateral angle given by the interaural differences (the angle from the axis of the ears to the source), $\beta$ is the angle of head movement in azimuth and $\gamma$ is an index describing the rate of change of lateral angle for a given head movement angle. For a source directly above the listener, yaw head movements will not affect the interaural differences (the angle from the axis of the ears to the source stays at 90° so $\partial\psi = 0$ and so the rate of change $\gamma = 0$). If the source is level with the ears then the same head movements will produce an equivalent change in lateral angle, so $\partial\psi = \partial\beta$ and $\gamma = 1$.

This only gives the source’s elevation and not its hemisphere, so whether a source is above or below the listener will still be unresolved. Wallach proposed that,

“For discrimination between these two equivalent directions another head movement around a different axis is required.”

Although it was mathematically proven that this cue could allow source elevation to be perceived, whether humans actually used it was unverified. To test this Wallach synthesised head-movement-related cues in an effort to alter the listener’s perception of source elevation. Wallach’s experiment used an array of loudspeakers surrounding the listener on the equatorial plane, and a mechanical head-tracking device. When the listener moved their head, the tracking device would change which individual loudspeaker from the array was used. Thus, the sound source could be moved varying amounts when the listener moved their head. According to the hypothesis, this should alter the listener’s perceived source elevation. It also tested another of Wallach’s hypotheses:

“it should be possible to present during a head movement a sequence of lateral angles representing a certain sound direction without presenting the sound direction itself”

Wallach felt if this were found to be the case it would give further proof that head movements were the auditory systems main elevation cue.
In the first experiment, head movements were exactly mirrored by loudspeaker movement so that the source appeared with no interaural differences at any head angles. Using Wallach’s hypothesised metric, this should suggest a source directly above. Of the listeners who were able to localise elevation (10 out of 17), all of them heard the sound originate from directly above. The source stayed directly above even after head movement stopped. However, Wallach’s suggestion that 7 of the 17 listeners were unable to localise elevation at all is questionable.

In the second experiment, intermediate elevation angles between 0° and 90° were synthesised. To do this, the distance between adjacent speakers was varied, which changed the degree to which the loudspeaker source moved for a given head movement. Wallach calculated that \( \alpha = \beta(1 - \cos \nu) \), where \( \nu \) is the elevation angle of the synthesised source, \( \beta \) is the head movement angle and \( \alpha \) is the angular distance between two loudspeakers in the array. He used this formula to synthesise various elevation angles and then asked listeners to localise them. The results show a very close match between the projected angle and the listener’s perceived angle (Table 5).
Wallach’s experimental findings agreed with the theory. However, the analysis was not exhaustive, testing only a few angles of elevation. Wallach states:

“It was not the aim of this experiment to gather data on the accuracy with which sound directions can be synthetically produced. We wished merely to obtain confirmation of the theory.”

Furthermore, the loudspeaker switch system meant each loudspeaker was turned on in succession. Therefore, no intermediate points were available and so the natural movement of the source was only approximated.

One of the main contentions with Wallach’s experiment is that it does not replicate natural listening conditions [Perrett and Noble 1997]. The unnatural events that occur within the experiment create conflicting auditory cues. The pinna cues suggest that the source is level with the ears, while dynamic

<table>
<thead>
<tr>
<th>Projected Angle</th>
<th>60</th>
<th>78</th>
<th>120</th>
</tr>
</thead>
<tbody>
<tr>
<td>Listener 1</td>
<td>64</td>
<td>80</td>
<td>130</td>
</tr>
<tr>
<td>Listener 2</td>
<td>61</td>
<td>75</td>
<td>122</td>
</tr>
<tr>
<td>Listener 3</td>
<td>60</td>
<td>75</td>
<td>122</td>
</tr>
<tr>
<td>Listener 4</td>
<td>55</td>
<td>75</td>
<td>119</td>
</tr>
<tr>
<td>Listener 5</td>
<td>50</td>
<td>70</td>
<td>118</td>
</tr>
<tr>
<td>Listener 6</td>
<td>48</td>
<td></td>
<td>115</td>
</tr>
<tr>
<td>Listener 7</td>
<td>43</td>
<td></td>
<td>114</td>
</tr>
</tbody>
</table>

Table 5 - Wallach's localisation experiment - angles given in degrees [Wallach 1938]
interaural differences suggest an elevation angle given by the source movement. It could be argued that Wallach actually found that, when localising a source's elevation, the presence of head movement cues cause pinna cues to be ignored. Furthermore, the experiment tested whether head movements could be used to give the impression of an elevation angle; not whether, when a source is elevated, head movements improve elevation LRA. Does the former assertion inexorably prove the latter?

5.2 Wallach [1940]

Wallach [1938] stated that a dynamic binaural localisation cue requires the listener to monitor two information streams simultaneously; they must know their changing binaural cues, and they must know how their body is moving in order to create these changing cues. In Wallach [1940], the second information stream is studied in detail, to answer their primary research question: are head movements required to allow localisation using dynamic cues or is any kind of perceivable listener movement adequate?

Wallach lists three senses with which the listener can monitor their own movement with respect to a sound source: vestibular stimulation, visual stimulation from the eyes, and proprioceptive stimulation from the muscles that cause the movement of the head and/or torso. Do each of these sensory streams allow dynamic binaural localisation or is it dependent on only proprioceptive stimulation created through the movement of the head? Wallach conducted a series of experiments that separated out the three senses and studied them individually.

Wallach's experiment used the same method described in Wallach [1938], equatorial plane loudspeakers that were dynamically varied to create an elevated perceived source location. It was shown that using vestibular stimulation only (by blindfolding the listener and moving the chair for them) 6 of the 15 listeners were still able to localise with a ‘satisfactory’ response, which appears from Wallach’s results to be a response within 5° to 10° of the +60° desired location. Wallach [1940] states that vestibular stimulation only monitors
a change in acceleration, not velocity; therefore calculations of overall head displacement with head movement may be inaccurate. Furthermore, this test presented confounding cues that indicated different locations; if a listener relies predominantly on spectral cues to localise in elevation then they would perceive the source on the equatorial plane, whereas if head movement cues were dominant then the desired elevated location would be +60°. Once the blindfolds were removed from the listeners, 13 of the 15 subjects localised with a satisfactory accuracy. When the dynamic cues are solidified by both visual and vestibular movement cues, they appear to take precedence over pinna cues.

Wallach concluded that passive movement of the listener, i.e. when the experimenter moves the listener, results in localisation equally as accurate as localisation when the listener moves their own head. Furthermore Wallach states that visual cues allow listeners to localise accurately with dynamic cues while vestibular stimulation give a ‘fairly accurate representation of listener movement’ and in some cases can be used to localise accurately.

By using a rotating screen, Wallach managed to use perceived listener movement (the screen was moving, not the listener) to trigger a dynamic localisation cue and cause sources on the equatorial plane to be perceived as elevated above the listener. Wallach used this to show that visual cues alone can allow dynamic localisation using listener movement.

Wallach also discusses the ‘selective principle’, in which listeners choose the static or slowest moving possible source location given by the dynamic localisation cues presented. Listeners assume that the simplest source position is the most likely. This principle is the underlying theory behind Wallach’s experimental method and is shown from the results to be effective.

Interestingly Wallach allowed the listeners to choose between two localisation response methods, a verbal method and a pointing method, and showed both sets of the results where necessary. Localisation response methods are discussed further in Chapter 7.
5.3 Pollack and Rose [1967]

Pollack and Rose [1967] studied the effect of head movement on azimuth LRA on the equatorial plane. They stated that head movements contribute to the cocktail party effect by allowing the listener to focus on a specific sound source within a mixture of multiple sources. The goal of their experiment was to discover whether head movements improve the azimuth LRA of a single source against a quiet background. Their experiment had 5 stages, studying how the type and length of programme item and the type of head movement affected the listeners' azimuth LRA.

5.3.1 Results

Stages 1, 2 and 3 of their experiment used a click of 15ms duration. This click was either initiated either by a button push, in the static case, or by the movement of the listeners' heads, in the dynamic case. Since the click was such short duration, the listeners were unable to dynamically monitor continuously changing cues for a given head movement. Thus, the cue described by Wallach [1938] could not be used.

There were a number of differences between the experiments in stages 1 and 2:

- *stage 1 used a larger loudspeaker separation (9.6°) than stage 2 (3.2°).*
- *stage 2 used a constrained head movement condition in which the listener was instructed to follow a relay-light timing system; stage 1 appeared to used no head movement control method, only instructions that the listener should move their head in a single direction.*
- *stage 1 used three different initial static body positions with the head pointing straight ahead.*
- *stage 1 used anechoic chamber while stage 2 used a smaller non-anechoic room.*

The results from stages 1 and 2 both showed no improvement in azimuth LRA due to head movement. For stage 1 they state that the azimuth LRA was 'largely unaffected' by the initial body positions, although an initial orientation to the
right did appear to reduce the azimuth LRA for all listeners. This stage also showed that, for static listening conditions, the loudspeaker position variable was significant. The location of loudspeakers further from the median plane were less accurately reported. This result agrees with the findings of Stevens and Newman [1936] described in Chapter 3.

For stage 2, the mean azimuth LRE for the non-moving head was approximately 2°, while the mean azimuth LRE for a moving head was approximately 11°. This would strongly suggest that head movement actually degrades azimuth LRA on the equatorial plane.

For stage 2, listeners were instructed to follow a ‘relay-light-timing-system’, used to control the head movement. Pollack and Rose suggest that the attention of the listener may have been focused on following the relay-light, not localising the sound source. They state that this, coupled with the forced head movement being unnaturally confined to a single plane, may have caused the reduction in azimuth LRA. However, other experiments have used both light-guided head movement and single-plane head movement, and have still shown an improvement in LRA with head movement [Perrett and Noble 1997].

The main conclusion that can be drawn from these experiment stages is that a single click programme item does not allow dynamic localisation through head movement to be used. Pollack and Rose [1967] themselves state that the goal of stages 1 and 2 was to show that the use of extremely short programme items does not allow time for reorientation of the head in relation to source and so in this case head movement does not improve azimuth LRA. It is suggested that forcing the listener to move their head for such short programme items may confuse the listener more than stationary listening.

For stage 4 they used white noise programmes of varying lengths and compared the movement and no-movement accuracies for both cases. They used programme items of up to 3 second in duration, which allowed ‘sufficiently long to permit head reorientation to the source’. The results of the study are shown in Figure 12.
The plot and a subsequent ANOVA showed that, in general, the no head movement condition gave a higher azimuth LRA than the free head movement. An interaction of movement condition*programme item duration was also shown to be significant. From 0.1s to 3s signal duration, as the programme item duration is increased, the difference in error between the no movement and free movement conditions is reduced. With signal duration of 3s the azimuth LRE of the movement condition is actually less than the no movement condition for listener S:G. However, whether this finding was significant is unclear from the plot. They state that there is no ‘clear cut’ improvement in azimuth LRA given by head movement but they offered no statistical analysis, such as a planned contrast for the 3s duration programme item, to verify this statement.

Stage 5 studied how azimuth LRA varied as the sound was moved further in azimuth from the listener’s median plane. A plot of the mean azimuth LRE
against listening offset for two separated listeners and both movement conditions is shown in Figure 13.

The plot shows that when the listening offset is 0° (the source directly in front of the listener) the movement condition results in a lower mean azimuth LRE than the no-movement condition. However, whether this difference is significant is unclear as this finding was unreported by Pollack and Rose. The main trend they note is that, as the listening offset is increased, the difference in azimuth LRA between the free movement and no movement conditions is also increased. They state that only when they allowed time for the listener to orient towards the source location, thereby improving the listener's static localisation cues, did localisation with free head movement show an improved azimuth LRA. They state that only when they have 'attempted to stack conditions in favour of obtaining a positive effect of head movement upon localization’ did any improvement due to head movement occur.

Figure 13 - Mean azimuth LRE separated by listening offset
5. Dynamic Localisation Cues

5.3.2 Experimental Issues

Pollack and Rose asked the listener to report the loudspeaker from which the sound source appeared to originate; the listeners chose from a series of numbered loudspeakers located on the equatorial plane. Wallach hypothesized that head movements might improve LRA because they allows resolution of the cone of confusion, which is caused by ambiguous interaural localisation cues. Showing the listeners the loudspeaker positions, located on the equatorial plane, caused the cone of confusion to be resolved and this visual cue rendered any improvement due Wallach’s head movement cue to be ineffective. There is no mention of front-back confusions at any point in the paper. It seems unlikely that a static azimuth localisation study would make no mention of these confusions unless the listener’s response had been biased by knowing that there were no loudspeakers behind them. Since the listener’s response was biased enough to remove all front-back confusions, it is likely that their azimuth LRA, with front-back confusions removed, was also biased. The issue of localisation response method bias is discussed further in Chapter 7.

They also started with a loudspeaker separation of 9.6°, which was reduced down to 3° by their final experiment stage. This large separation angle limited the listeners’ responses to significantly greater than the MAA (approx. 1° in azimuth). Furthermore, they then reported a LRE of lower than the loudspeaker separation, which is questionable. The response method used in this experiment biased the responses of the listeners.

Pollack and Rose used few listeners for their test, between two and three listeners for each experiment stage. As shown in studies such as Wallach [1938], there is large variation in listeners responses and so making generalisations based on only two listeners is likely to be inaccurate. They also failed to fully randomize the trial presentations, in one case presenting all free movement conditions after all no movement conditions and thus possibly biasing the listeners’ responses. Furthermore their results analysis was mainly confined to plotting and comparing means; more powerful statistical analysis tools such as
ANOVA were occasionally referred to but not clearly discussed and confidence intervals were not used at all.

### 5.3.3 Summary

Pollack and Rose showed that when the listener is given a short programme item, such as a click programme item, head movement does not improve azimuth LRA. Further conclusions they made about the effectiveness of Wallach's head movement cue are compromised by their choice of programme item, response method, control of movement condition and statistical analysis. Their experiment did not actually test whether Wallach's cue was effective for a number of reasons: they did not use long enough duration programme item for Wallach's cue to be used effectively; the effects of Wallach’s cue were hidden by static cues for long duration programme items; the effects of Wallach’s cue, which were implied by the results plots, were not investigated or discussed.

Pollack and Rose were the first study to suggest that the improvements in LRA with head movement could be due to improved static cues. Pollack and Rose state that azimuth LRA is higher for sources in front of the listener than for sources at the side of the head. Therefore head movement improves LRA by ‘reorienting the ears into more acute listening positions.’ Whether this finding is also true for elevation LRA is shown in Chapter 11.

Their results also suggest that as the programme item duration is increased the effectiveness of head movement cues increases and that as the source is moved further from the median plane the effectiveness of head movements cues increases. However, as discussed earlier, their experimental methodology has left it unclear whether these conclusions are valid.

### 5.4 Thurlow and Runge [1967]

Following an extensive study of listener head movement, Thurlow and Runge went on to investigate whether these head movements improve LRA. They used two programme items, low pass filtered noise (cut-off 7.5-8 kHz) and high pass filtered noise (cut-off 0.5-1 kHz) and a pointing response method. They used a
number of head movement conditions: no movement, controlled head movements and free head movement and studied both azimuth and elevation LRA. The controlled head movements consisted of yaw (45˚), roll (15˚), pitch (15˚) and a combination of yaw and roll.

5.4.1 Results

For all programme items, a forced yaw movement completely removed all instances of front-back error (0% error). Front-back errors were present to some degree for all programme items in the no movement condition and at its highest this error occurred for 90% of stimuli (low-band noise). Furthermore, the combined yaw and roll movement removed all confusions for noise programme items and the majority of confusions for click programme items. Results suggest that both ITDs and ILDs contribute to reducing front-back errors.

Yaw, roll and combinational yaw and roll movements were all shown to significantly improve elevation LRA for low-band noise. This finding indicates that head movement cues are dependent upon dynamic ITD cues. Thurlow and Runge stated that the main error for the no movement condition was an ‘underestimation of the departure of the source direction from the horizontal.’ This underestimation was also shown in Wallach [1940] and appears to be a feature of static elevation localisation.

There were no significant differences in elevation LRA between movement conditions for the high-band noise. Thurlow and Runge suggest that the location of high-band noise was accurately reported without the need for improvement offered by head movement. The static spectral cues given by this source resulted in significantly higher LRA when compared to the low-band noise programme. For the high-band programme item there was a significant interaction between loudspeaker number and movement condition for the yaw and no movement conditions; yaw movements gave a consistent response across loudspeaker location while no movement varied significantly across loudspeaker. For example, loudspeaker 1 resulted in an error of approximately 0˚ for the yaw movement and +20˚ for the no movement condition. This shows that yaw head movements increase the consistency of elevation LRA across source location.
For repeated click programme items, none of the forced movement conditions significantly improved the elevation LRA. It is suggested that the click programme item did not produce continually varying cues, therefore Wallach’s dynamic localisation cue could not be used.

A forced pitch movement was shown to offer no improvement in either elevation or azimuth LRA. Pitch movements offer no change in interaural differences for a given head movement and so the dynamic localisation cues described by Wallach [1938] cannot not be used.

Free head movements were shown to significantly improve elevation LRA for the low-band noise. Further analysis showed that there were no significant differences between the yaw movement and the free movement for either elevation or azimuth LRA. This suggests that yaw motion offers all of the dynamic cues necessary to improve LRA.

Yaw head movements were shown to improve azimuth LRA for both high-pass and low-pass filtered noise. However their filter characteristics were not based on the interaural cues; the low pass filtered programme item contained frequencies up to 8 kHz, while the high pass filtered programme contained energy down to 0.5 kHz. Therefore, both programmes items contained some ITDs, ILDs and pinna cues. The presence of pinna cues in the programme item would have limited the improvement offered by head movement, especially for elevation localisation. The experiment described in Chapter 9 studies the effect of head movement on LRA with the presence of ITD and ILD cues individually.

5.4.2 Experiment Limitations

The conclusions given by Thurlow and Runge are confounded by various experiment limitations. In the controlled movement condition listeners were permitted to have only one direction of movement; no reversals were allowed. As shown in Chapter 2, listeners make frequent reversals when attempting to localise a sound and their choice of head movement changes depending on the stimulus presented. Not allowing movement reversals may have impeded significant improvements in LRA due to head movement.
They controlled head movement by mounting the listener’s head in a clamp and using a motorised frame to move the head. Although they installed failsafe measures such as emergency stop buttons for both listener and experimenter, this method poses a significant risk to the listeners’ safety.

When choosing the controlled head movements to be auditioned by the subject, Thurlow and Runge chose yaw and roll movements and a combined yaw and roll movement for one study and a single pitch movement in a separate study. It is unclear why they chose to combine yaw and roll movements. In an earlier study they showed that a combination of yaw and pitch movements was significantly more common. One possible justification could be that pitch movements on the median plane produce no changing interaural cues and so if they were looking to study Wallach’s localisation cue then studying pitch movement may have initially been thought to be unnecessary. They did go onto show that pitch movements produced no improvement in either elevation or azimuth LRA.

Thurlow and Runge used only 7 loudspeakers for each programme item. Furthermore the loudspeakers were grouped in only a few distinct elevation and azimuth locations. Thurlow and Runge state that they did not repeat any trial for fear that the listener may have remembered the loudspeaker location. However, the loudspeakers to right and left of the listener were either at between –21° to –26° elevation or between +30° to +35° elevation; the two frontal sources were both at +41° elevation. Since there were only three possible elevation areas from which a source could originate, generalisations made to overall elevation LRE are questionable. The small number of loudspeaker locations and small overall data set mean that the results of Thurlow and Runge might not be generalisable to the population. Loudspeaker locations should be spread evenly in elevation and azimuth so that the listeners are not biased in their response and so that conclusions can be generalisable to all possible loudspeaker locations.

Wallach [1940] showed that any kind of listener movement, not only proprioceptive movement, can be used to create dynamic cues. When experiments, such as Thurlow and Runge [1967], do not allow rotation of the body/chair, or any alternative motion to simple head rotation, then they are
impeding the natural response of the listener. In a dynamic localisation study comparing movement conditions, it is important to have a free movement condition and to allow body movement as well as head movement.

Thurlow and Runge used an impressive pointing response method considering the experiment was conducted before head tracking and the computerisation of audio experiments. The listener would point a metre stick at the source location with a straight arm, and the experimenter would align a semi-circular arc with the pointer and read both the azimuth and elevation angle from a scale. The problem with this method is that it is extremely time intensive and, as a result, Thurlow and Runge were only able to test a limited number of stimuli, 10 listeners for each movement condition and only 14 trials for each listener. This resulted in a small statistical power, making significant differences harder to find and conclusions less robust.

Interestingly Thurlow and Runge calculated the signed error for each loudspeaker location and then combined them using an absolute sum to calculate the overall error. This method would find the bias for each loudspeaker location, not the overall error which is based on both bias and localisation uncertainty. This method removes the variance in response around the source location; if listeners are less sure of a sources location then they are likely to produce a larger variance. Therefore, it is not a complete measure of listener LRA. If the absolute error for each listener response had been calculated then the conclusions may have been different.

The findings of Thurlow and Runge were confounded by their experiment method, namely: the calculation of LRA; the limited number of loudspeaker locations; the time-consuming response method and the forced movement condition that offered no movement reversals.

### 5.5 Perrett and Noble [1997]

The issues and ambiguities of Wallach’s head movement localisation studies prompted Perrett and Noble [1997] to create an experiment that attempted to test Wallach’s hypothesis from an alternative perspective using modern
experimental methods. They positioned sources at a number of elevation angles and asked listeners to identify the source’s location in a variety of different listening conditions. The experiment tested four listening conditions:

- Motionless
- Motionless and Pinna Occluded
- Rotational Movement
- Rotational Movement and Pinna Occluded

In motionless conditions, the listener was instructed to keep a laser, which was attached to their head, in a fixed position on the screen in front of them at 0° azimuth and elevation. In rotational movement conditions, the listener was instructed to move the laser between points at ±30° azimuth, with at least two complete oscillations per three-second stimulus signal. The pinnae were occluded by inserting a tapered plastic tube into the listener’s ears.

The programme items used in the experiment were:

- White noise low-pass-filtered at 1, 2 and 4 kHz
- White noise high-pass-filtered at 1, 2, and 4 kHz
- Broadband white noise

This series of programme items allowed the dynamic cue’s most significant frequency ranges to be highlighted. By referring to the duplex theory, this may also reveal whether ITDs or ILDs are the most important contributor to the head movement cue. Sound energy above 3 kHz is required to allow significant spectral pinna cues, so by using a low-pass programme item, pinna cues can be negated [Algazi et al. 2004].

Their first experiment showed that head movements substantially reduced the front-back errors regardless of the source signal (Figure 14). On average, 27% of all trials without head movements resulted in front-back errors. This value increased to 35% when the pinna was occluded. With head movement, the occluded pinna resulted in 0.6% front-back error, while unoccluded listeners
had no errors (0%). These results suggest that head movements provide the auditory system’s primary cue for front-back error, while the pinna also provides significant cues.

![Diagram](image.png)

**Figure 14 - Proportion of trials where front-back errors occurred plotted against programme item [Perrett and Noble 1997]**

They found that when judging the source elevation, the frequency range of the source signal was important. Localisation of low pass noise of 4, 2, and 1 kHz was heavily dependent on head movement; while high pass filtered noise was more dependent on pinna cues. For broadband programme items, the listener’s LRA was depended on both the pinna and head movement cues. They concluded that,

“when signals do not provide energy below 2 kHz, the rotation cue fails.”

This statement, when combined with the duplex theory of localisation, places more weight on ITDs than ILDs when using head movements to localise.

Their first experiment was confined to the upper hemisphere of the Median Vertical Plane (MVP) with loudspeakers positioned at 30° intervals. In their second experiment they studied the left vertical plane (LVP) and the MVP extending both above and below the listener using only the 2 kHz low-pass-filtered programme item. It was found that even with head movements, up-down confusions were prominent. This error was only found for sources on the median
to the rear or directly below the listener (Figure 15). The auditory system seems to assume sources are in the upper hemisphere unless it receives cues to the contrary.

![Figure 15 - Perceived source location against actual location showing up-down errors](Perrett and Noble 1997)

It may have been interesting to include broadband programme items in the test to see whether the pinna cue's interaction with dynamic cues could resolve up-down confusions.

Perrett and Noble only studied four elevation angles (0°, 30°, 60° and 90°), so only a general trend in elevation LRA could be observed. In addition, it would be interesting to investigate how the elevation LRA changes with the displacement of the source in azimuth.

During the experiment all sources above the equatorial plane were localised below their actual location. This was also found in Thurlow and Runge [1967] and Wallach [1938]. This may be a response that the listener has learned through experience; listeners assume a source is near the equatorial plane unless they receive localisation cues to the contrary.
The experiment was confined to the MVP and LVP. This significantly reduced the difficulty of the localisation task, which means its relevance in natural listening conditions may be limited [Makousa and Middlebrook 1990]. Their statement that ‘the greatest benefit for localization from rotation of the head appears to be gained for sources positioned on the MVP’, suggests that when all angles of elevation and azimuth are considered, the effects of head movements will have a less significant impact upon the elevation LRA.

5.6 Wightman and Kistler [1999]

Wightman and Kistler [1999] did not limit their experiment to a single plane, using a variety of source locations in both azimuth and elevation. Their experiment had three listening conditions:

- Motionless
- Freestyle movement
- Compulsory movement

In freestyle movement conditions, listeners were encouraged to use head movement to localise the source. In compulsory movement, the listener was instructed to orient their head towards the localised position of the sound source.

The listener verbally described the perceived sound source location using a “standard spherical or world coordinate system”. There is no indication that any visual references were used to anchor these coordinates. It is suggested that it is difficult to give a precise angle without a reference. This means that the perceived location given in this experiment may have been less accurate than the head tracker method used by Perrett and Noble. Evans [1998] states that verbal responses are unintuitive and so likely to result in inaccurate responses from the listeners. Localisation response methods are discussed in detail in Chapter 7.

Unlike Perrett and Noble, their experiment did not include a condition in which the pinna was occluded. Therefore, the degree to which head movement improved LRA was only tested in the presence of pinna cues. It is likely that
without the pinna spectral cue, more weight is placed on the head movement cue.

The experiment was also carried out in an anechoic chamber. Reverberant conditions, common to natural listening, produce confusing cues that make localisation more difficult. In these conditions head movement cues may be relied upon more heavily to resolve the localisation ambiguity.

Their experiment used only one programme item, broadband white noise. Therefore, pinna spectral cues were always present, making the study of dynamic cues more difficult.

It was shown that head movements significantly reduce the occurrence of front-back confusions. However, they concluded that there was no indication that head movements significantly increase a listener's elevation LRA.

Based on this experiment it is possible that when pinna cues are present, the additional cues provided by head movement do not significantly improve elevation LRA. It would be interesting to test whether, when head movement cues are present, the additional cues provided by the pinna significantly improve elevation LRA. Can the auditory system accurately resolve source elevation using either cue, so that the addition of the other causes no notable improvement?

The overall similarity of Perrett and Noble's and Wightman and Kistler's experiments is obvious, however, the reason they came to differing conclusions about the importance of head movement as a cue is unclear. They both positioned sources at a number of locations in elevation and azimuth and asked the listeners to localise them. The major notable differences are summarised in Table 6.
### Table 6 - Major Differences in the experiments of Wightman and Kistler [1999] and Perrett and Noble [1997]

<table>
<thead>
<tr>
<th>Perrett and Noble</th>
<th>Wightman and Kistler</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restricted locations to Median Vertical Plane and Left Lateral Plane</td>
<td>Unrestricted – any location in azimuth and elevation</td>
</tr>
<tr>
<td>Real sources only</td>
<td>Both real and virtual sources</td>
</tr>
<tr>
<td>Broadband, high-pass and low-pass filtered white noise</td>
<td>Broadband white noise</td>
</tr>
<tr>
<td>Pinna both occluded and unoccluded</td>
<td>Pinna unoccluded</td>
</tr>
<tr>
<td>Semi-anechoic listening room</td>
<td>Anechoic Room</td>
</tr>
<tr>
<td>• Set head rotation (±30°)</td>
<td>• Freestyle with head movements encouraged</td>
</tr>
<tr>
<td>• No movement</td>
<td>• No movement</td>
</tr>
<tr>
<td>• Orient towards sound source</td>
<td>• Orient towards sound source</td>
</tr>
</tbody>
</table>

**5.7 Nojima et al. [2013]**

Although the goal of Nojima *et al.* was to find how listeners moved their head when localising, they also analysed how these natural head movements affected the azimuth LRA of the listeners. They did not use separate movement and non-movement conditions; instead they observed how the listener moved their head when listening naturally and correlated the movements to their LRA. The advantage of this method is that the listener’s responses are natural, not biased
by an experimenter’s instruction. However they do not guarantee a complete data set; there is a chance a single listener may always or may never move their head, this makes comparisons and statistical analysis difficult.

Nojima et al. found that the rate of correct responses in azimuth was significantly higher when the listeners moved their heads. They showed that larger head movements (>MAA±1°) are more effective than smaller ones (<MAA±1°) at improving listeners’ azimuth LRA. Furthermore head movements that make use of reversals also improve the azimuth LRA. This is further evidence of the Wallach dynamic localisation. However they use the lack of head movement overall to argue against Wallach’s localisation cue. When the listeners were unable to see the source loudspeakers they moved their heads for 30% of the trials; when they could see the loudspeakers head movements occurred for 70% of the trials. This is further evidence that the listener may not be aware of their most accurate localisation method. It also shows that a listener’s visual cues are an overriding factor when attempting to localise and so should be controlled carefully.

5.8 Further Conflicting Views on Dynamic Cues

Middlebrooks and Green [1991] state that:

“there is essentially no evidence to suggest that the information gained from two head locations (and the information gained from two cones of confusion) is substantially better than the information gained from a single head position.”

A number of sources state that dynamic head movements cues do not improve azimuth LRA [Fisher and Freedman 1968; Pollack and Rose 1967]. However, the ability to move one’s head would allow the listener to move the source into the localisation system’s highest area of acuity, thereby improving their static localisation cues.

Furthermore, as highlighted by Makous and Middlebrook [1989], and Oldfield and Parker [1984] any study that reduces possible source locations to one plane effectively simplifies the localisation task. To study the importance of dynamic
cues in localisation one must position the loudspeakers in a variety of both azimuth and elevation locations.

5.9 Conclusions

The goal of this chapter was to investigate whether dynamic cues, created through the movement of the head, are used to improve a listener’s elevation LRA. After analyzing the previous literature the answer is still unclear. Wallach [1938] showed how dynamic interaural differences could be used to calculate both the elevation and azimuth of a source. Empirical evidence that listeners use these dynamic cues to localise in elevation was given. Further studies have indicated that head movements improve elevation LRA for certain experiment conditions [Perrett and Noble 1997; Thurlow and Runge 1967]. Others state that there is no improvement in elevation LRA with head movement [Wightman and Kistler 1999; Middlebrooks and Green 1991].

Pollack and Rose [1967] state that improvements in azimuth LRA with head movements are only due to improved static cues created by turning to face the source. They state that they could find no evidence for Wallach’s dynamic cue. However, numerous studies have shown that yaw head movements significantly reduce front-back confusions for both noise and click programmes [Thurlow and Runge 1967; Perrett and Noble 1997; Wightman and Kistler 1999; Pollack and Rose 1967].

In conclusion, whether dynamic cues given through head movement improve elevation LRA is still unverified. Chapter 8 will investigate whether head movements increase LRA and whether these improvements are due to dynamic cues, improved static cues or a facet of response method errors.

The experiments of Perrett and Noble [1997] and Thurlow and Runge [1967] suggested that head movement cues are more prevalent when pinna cues are limited. Both studies stated that head movements improve elevation LRA using low-pass filtered noise programme items. This indicates that dynamic ITD cues improve LRA with head movement. Neither study investigated whether dynamic ILD cues improve elevation localisation LRA. To study dynamic ITD and ILD cues
alone, the effect of pinna cues must be reduced. Chapter 6 will investigate how to reduce pinna cues so that head movement cues can be studied more effectively. Chapter 9 will investigate whether ILD cues can be used in conjunction with head movements to improve elevation LRA.

Differences in experimental methodology are likely to account for the differing outcomes of the head movement studies described in this chapter. It is important to note these differences so that an optimised experimental method can be used in this thesis. The head movement localisation studies described in this chapter used differing localisation response methods to ascertain the perceived source location. These studies came to differing conclusions and the response method was highlighted as a significant factor affecting these conclusions. Chapter 7 shows how the localisation response method can significantly change the LRA of listeners and highlights the most effective localisation response method.

In summary, the literature described in this chapter has left some questions unanswered:

- Do yaw head movements improve a listener’s elevation LRA?
- What extra cues do head movements create that allow improved elevation LRA?

An in-depth experiment investigating the importance of head movement cues in localisation is proposed. However, before any test is conducted an accurate method must be found to elicit the perceived loudspeaker location (Chapter 7), and method for reducing pinna cues, which hide the effects of head movements, must be found (Chapter 6).
6 Pinna Cue Suppression Methods

The goal of this chapter is to find a way to suppress pinna cues and hence more clearly separate the effects of head movement. This pinna cue suppression method will then be applied to subsequent elevation localisation studies in this thesis.

As stated in Chapter 4, the shape of the listener's pinna causes variations in the spectrum of an approaching sound signal. By noting these spectral changes, the listener can resolve the location of a source, as each location has its own spectral signature. Many experiments have tried to suppress or remove pinna cues in order to observe the resultant degradation in LRA [Gardner and Gardner 1973; Perrett and Noble 1997]. This allows the importance of pinna cues to be highlighted and allows other cues to be studied that may otherwise be masked.

This chapter will discuss methods of pinna cue suppression used by previous studies and will describe two experiments conducted by the author to find an alternative pinna cue suppression method. In Sections 6.1, 6.2 and 6.3, pinna cue suppression methods are split into three broad categories: physical occlusion, virtualization, and source signal manipulation, and the merits and weaknesses of each category are discussed. Sections 6.4, 6.5 and 6.6 describe two experiments conducted by the author to find a programme item that will suppress pinna cues without any need for physical occlusion.

6.1 Physical Occlusion

Physical occlusion of the pinna involves changing the way in which the incoming sound wave interacts with the pinna. Blocking or reshaping the undulations of the pinna changes the spectral variations added to the sound wave. The listener has learned their own spectral response for each location, a response that is mainly dependent on the shape of their pinna. By changing this shape the listener is unable to localise accurately as their spectral cues will change. Most physical occlusion methods attempt to stop pinna cues altogether by removing all spectral variations created by the pinna.
In Perrett and Noble's [1997] elevation localisation experiment, a tube was inserted into the ear to suppress the spectral cues given by the pinna. This method of pinna occlusion is aggressive and can damage the extremely thin skin within the ear canal, which can lead to infection of the ear. It may also alter the effective radius of the listener's head. This will change the interaural differences of the listener, thus causing any dynamic interaural difference cues to be confused. Furthermore, it may change the resonance of the ear canal; again creating cues that may lead to confusion and localisation inaccuracy.

One of the most extensive pinna occlusion experiments was conducted by Gardner and Gardner [1972]. They used a combination of two mould making rubber formulas to create the occlusion. Time was spent matching the flexibility and softness of the moulds to that of a human pinna. Different sections of pinna were occluded to test which area was most important in median plane localisation. A small hole was left to allow the sound waves to enter the ear canal. This method does not elongate the ear canal, nor does it encroach on the sensitive skin of the ear canal. The interaural differences are much less affected by the mould occlusion than when using the tube method, which is important if head movement cues are to be studied. The moulds will remove the small undulations in the shape of the listeners pinna; these undulations affect the level of the high frequency signals reaching each eardrum. Therefore the moulds will still cause small variations in the high frequency ILDs of the listener.

Each listener had a number of moulds made to fit their individual pinnae making the method both time consuming and expensive. Gardner and Gardner [1972] noted that each mould took four days to cure. Creating the moulds takes an experienced technician and errors may damage the listener's ears.

In their study, Butler and Humanski [1992] used a ready-made ear moulding formula called ‘Audi-sil’ and created a mould for each subject. Audisil claims to be ‘instant’, which suggests a setting time considerably less than four days. However, this does not remove the extensive time taken to give each individual listener a pinna mould. Does the individuality of the each listener's pinna make enough of a difference to warrant individual moulds?
Use of plasticine to fill the pinna is a less time consuming method of occlusion [Keen 1925]. Furthermore, it does not need to be individualized beforehand as it can be quickly moulded into shape. However, care must be taken not to allow any plasticine to enter the ear canal, as this may cause infection and damage.

### 6.2 Virtualisation

It is possible to suppress pinna cues without occluding the pinna. The difficulty is to do so whilst maintaining an otherwise natural experience for the listener.

Creating virtual sources to be played to the listener via headphones is one way to suppress pinna cues. Headphones could be used to present constant interaural level and time differences with no spectral variation. However, much has been written about the difficulty in externalization for a sound source recreated on headphones [Hartmann and Wittenberg 1996]. It is suggested that certain spectral notches are necessary for externalization and these would not be present in the system suggested above. Therefore, they may not give a natural listening experience, making localisation difficult. A further problem using headphones is that the sound source still passes through the pinna and the ear canal, just from a different location (right next to the ears). This may create confusing spectral cues that suggest a different source location and may be a cause of the ‘in the head’ experience.

### 6.3 Source Signal Manipulation

The source signal of a localisation test can be manipulated in various ways to suppress the cues given by the pinna. This method is less time intensive than either physical occlusion or virtualisation, as it can be universally extended with no need to study listeners individually.

Sources suggest various frequency regions required for pinna localisation. Perrett and Noble [1997] suggest that pinna cues are only effective at frequencies above approximately 4 kHz, while Algazi et al. [2001] state that sound energy above 3 kHz is required as this is the point at which the wavelength of the sound wave becomes comparable to size of the pinna. Roffler
and Butler [1968] found that, “sound energy above 7 kHz is needed for MP [Median Plane] localisation” [Hebrank and Wright 1974: 935]. Gardner and Gardner’s study [1972] indicated that mid-frequency band noise around 2kHz and 3kHz resulted in significant errors in localisation even with an unoccluded pinna. Using this evidence, Perrett and Noble used low-pass filtered white noise to suppress pinna cues. It can be seen that if white noise with a low-pass cut-off frequency of 2kHz is used then pinna cues will be suppressed according to all studies.

In an experiment it may be of interest to study elevation localisation cues at high frequencies, such as dynamic ILDs. This will give a better indication of the interaction between head movement cues and pinna cues. If only low-pass filtered noise is used, these interesting characteristics will be impossible to study. Therefore, an alternative method to disable pinna cues whilst maintaining high frequencies is sought.

6.4 Pilot Experiments

Two pilot experiments were devised to look for a programme item that contained high frequency content yet did not stimulate pinna cues. This section describes the setup of these two experiments. The only difference between the experiments was the set of programme items the listener was asked to localise.

Elevation LRA on the median plane will be significantly reduced when the listener has no pinna cues and is unable to move their head [Perrett and Noble 1997; Gardner and Gardener 1972]. By conducting a listening test on elevation localisation with no head movement, one should be able to determine whether the pinna cues have been disabled. If a listener’s elevation LRA is significantly reduced by a certain programme item, then it is suggested that the programme item has degraded the listener’s pinna cues.

6.4.1 Setup

The experiment took place in the live room of Studio 2 in the Institute of Sound Recording at the University of Surrey. This room has a high ceiling, which means
that loudspeakers can be placed at extreme elevations without reflections from
the ceiling and floor becoming problematic to localisation cues. The room is not
anechoic, but was a pop studio by design so had a short reverb time (RT60 of
235 ms [Coleman et al. 2014]). If the room was too reverberant then localisation
cues could be masked or confused. The test required localisation cues to be as
distinct and precise as possible, so that any differences due to the test variables
could be observed. An anechoic chamber was not chosen for the study as the
results from anechoic studies may give unrepresentative results and not be
transferrable to natural listening conditions, as discussed in Chapter 4.

Six Genelec 8020b loudspeakers were positioned on the median vertical plane at
elevation angles from -9° up to +58° elevation. The loudspeakers were arranged
to have approximately equal spacing between them; an exact spacing between
the loudspeakers was avoided to reduce the chances of bias or habituation.
Loudspeaker locations were limited by their proximity to the floor and the
maximum height of the loudspeaker stands. The loudspeakers’ positions for all
experiments described in this thesis were measured from their ‘acoustic centre’
as defined in Genelec’s 8020a ‘Operating Manual’ and shown in Figure 16.

![Figure 1: Location of the acoustic axis](image)

**Figure 16 – The acoustic axis definition in the Genelec 8020 Operating Manual**

During the stimulus playback, listeners were instructed to keep their head
stationary and the laser pointer fixed on a calibration point at 0° elevation. Since
all sources were on the median plane, their interaural differences were
approximately zero. Therefore, interaural differences could not be used as an alternative elevation cue. The loudspeakers were held in place using microphone stands and were concealed behind an acoustically transparent but visually opaque curtain. The curtain was used to reduce any localisation bias given by visual cues.

All loudspeakers were positioned at a radius of 1.5m from the centre of the listener's head and angled to face the listener. This avoided spectral colouration due to the off axis response of the loudspeakers or the proximity of the loudspeakers to the listener.

### 6.4.2 Response Method

A laser pointing localisation response method was used. The listener was instructed to point a laser attached to their head at the perceived sound source. A Polhemus Patriot head tracking system was used to measure the perceived source location. The tracker gave its location with six degrees of freedom: Cartesian coordinates X, Y, Z in inches and orientation angles, azimuth (φ), elevation (θ) and tilt (τ) in degrees.

It was first suggested that attaching the laser to the head tracker would be the most effective way of gaining the perceived source location [Perrett and Noble 1994]. However, the Patriot was a magnetic based tracker system and the laser interfered with its output when they were positioned close together. A hard-hat was used to distance the laser pointer from the head tracker. The tracker was positioned inside the hard-hat while the laser was attached to the top of the hat using a microphone clip (Figure 17).
It was thought that the edges of the hard-hat might interfere with the listener’s spectral cues. Therefore, large areas of either side of the hat were removed.

A calibration point was positioned at 0° elevation in front of the acoustically transparent curtain. A cardboard cross, positioned on a microphone stand, marked this calibration point. Its position was kept constant throughout the experiment. Before each test the listener was instructed to point the laser at this calibration point. This was used to define the relationship between the laser and the tracker for each test.
During the test setup it was necessary to find the actual angle to each loudspeaker location. The author pointed the laser at the calibration point and then at each loudspeaker in turn and logged the head tracker orientation for each position. This was carried out before the curtain was used to conceal the loudspeakers, which meant that the laser point was visually guided to each loudspeaker location. The calibrated ‘actual angle’ was calculated:

\[
\theta_{\text{actual\_angle}} = \theta_{\text{head\_tracker}} - \theta_{\text{calibration\_point}} \quad (4)
\]

\(\theta_{\text{head\_tracker}}\) was the elevation angle output from the Polhemus tracker when the laser was pointed at the loudspeaker, while \(\theta_{\text{calibration\_point}}\) was the elevation angle output from the Polhemus tracker when the laser was pointed at the calibration point.

The same calibration calculation was used to work out the perceived angle during the listening tests:

\[
\theta_{\text{perceived\_angle}} = \theta_{\text{perceived\_head\_tracker}} - \theta_{\text{calibration\_point}} \quad (5)
\]

\(\theta_{\text{perceived\_head\_tracker}}\) was the elevation angle output from the Polhemus tracker when the laser was pointed at the (visually hidden) loudspeaker location as perceived by the listener, while \(\theta_{\text{calibration\_point}}\) was the elevation angle output from the Polhemus tracker when the laser was pointed at the calibration point.

Loudspeaker numbers and their respective elevation angles for pilot experiment 1 are shown in Table 7. The loudspeakers evenly spanned locations at intervals angles of approximately 13°. All loudspeakers were located on the median vertical plane.
### Table 7 - Loudspeaker number and elevation angle for Pilot Experiment 1. All loudspeakers were located on the median vertical plane

<table>
<thead>
<tr>
<th>Loudspeaker Number</th>
<th>Elevation Angle (Degrees)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-9.4</td>
</tr>
<tr>
<td>2</td>
<td>3.5</td>
</tr>
<tr>
<td>3</td>
<td>15.3</td>
</tr>
<tr>
<td>4</td>
<td>28.7</td>
</tr>
<tr>
<td>5</td>
<td>42.4</td>
</tr>
<tr>
<td>6</td>
<td>57.8</td>
</tr>
</tbody>
</table>

The method of calculation outlined above did not take into account the forward and back movement of the listener's head between tests. This movement was minimised by fixing the chair to a specific position before each test. Due to the distance difference between the laser pointer and the listener's eyes, the curtain location could have affected the perceived angle. However, the curtain was close enough to the actual loudspeaker locations to make any angle difference negligible. Furthermore, the test looked for differences in the LRA between the programme items, rather than absolute LRA of the listener. Use of this pointing method was justified as long as it was accurate enough to reveal these differences.

To analyse the results the LRE was calculated. To calculate the LRE in each case, the actual angle to loudspeaker was subtracted from the reported angle to loudspeaker.
\[ \theta_{\text{error}} = (\theta_{\text{reported\_angle}} - \theta_{\text{actual\_angle}}) \]  \hspace{1cm} (6)

Both the signed and absolute elevation LRE were calculated. The signed elevation LRE gave an indication of the bias of the listener's response while the absolute elevation LRE indicated the LRA.

### 6.4.3 User Interface

In both tests the listener was presented with the same user interface using MaxMSP (Figure 18).

![Figure 18 - User Interface](image)

Listeners clicked ‘Start’ to play the stimulus. Once the stimulus had played listeners could either replay the stimulus using the ‘Replay’ button or point their head at the perceived loudspeaker location and click ‘Log’ to save the perceived loudspeaker location orientation. The ‘Log’ button moved the listener onto the next trial stimulus. The trial number was displayed in the bottom right corner of the interface and a message box displayed listener instructions.

### 6.5 Pilot Experiment 1

#### 6.5.1 Movement Condition

Before beginning the test listeners were instructed to keep their heads stationary during the stimulus playback. Only once the stimulus had finished
playing were listeners permitted to move their head to point at the perceived source location. During stimulus playback listeners were instructed to keep the laser point directed at the calibration position.

6.5.2 Programme items

Full bandwidth white noise was used as a baseline against which the elevation LRE of the other programme items was judged. All other programme items were derived from this white noise programme item but processed in some way to suppress localisation cues.

Each programme item was three seconds long, with an onset and offset ramp of 300ms. The onset and offset ramps were included to avoid creating additional sounds such as clicks or distortions. These would be wide bandwidth and so would create extra localisation cues that could compromise the study of the test stimulus. Furthermore, other cues given by reflections from the floor or from the listener's shoulders would be more apparent with a distinct onset or offset.

In an effort to find a programme item that disabled pinna cues, two main groups were tested:

- *Low Pass Filtered Noise*
- *Spectrally Varied Noise*

Low Pass Filtered Noise (LPF Noise)

Although the goal of the experiment was to find a programme item that contained high frequency content and did not stimulate pinna cues, low pass filtered noise was included for two reasons:

- *To verify the findings of other studies, which stated that low pass filtered noise was an effective method of reducing pinna cues.*
- *To use this programme item as a second baseline comparison, against which the effectiveness of the other programme items could be measured.*
The low pass filtered noise was created using white noise filtered at 2 kHz. The filtered noise had a stop band attenuation of 80dB. Each programme item was auditioned to check that any pass band ripple or distortions were undetectable.

Spectrally Varied Noise

The spectrally varied noise was chosen based on research by Bloom [1977], described in Section 3.2. By filtering broadband white noise with a single notch filter Bloom intended to imitate the main notch in spectrum of sources located at different elevations. A single loudspeaker located at 60° elevation was used to replay the programme items for all test cases. It was shown that the listener perceived the elevation at the location indicated by the notch in the programme item, not by the actual location of the loudspeaker. This showed that the listener could by tricked into perceiving elevations other than the source’s actual location by varying its spectrum.

Based on this research it was suggested that by using a source with a constantly varying spectrum, the listener’s spectral cues could be confounded, leaving the listener unable to discern the actual location of the source. The spectrally varied noise was created using two different methods:

- **HRTF Based Random Spectral Swept Noise (SSN)**
- **Dynamically Filtered Noise (DFN)**

HRTF Based Random Spectral Swept Noise (SSN)

This programme item group was created by rapidly crossfading between different convolved noise samples. A white noise programme item was convolved with head related transfer function (HRTF) responses for random locations in elevation. These convolved noises were then combined by crossfading between them at regular intervals. This resulted in a noise signal whose spectrum changed rapidly at higher frequencies while remaining fairly constant at lower frequencies. The HRTF database was taken from Gardner and Martin [1994].
It has been shown through headphone auralisation that elevation can be perceived using convolved HRTFs. The spectrum of these auralised sources gives the correct cues to suggest a certain location. By constantly and quickly varying the HRTFs and by making these HRTFs different from the actual source location, it may be possible to mask the actual source elevation.

An informal test was conducted to choose the optimum rate of change between each convolved HRFT. Two rates of change were chosen, 60ms and 150ms.

**Dynamic Filtered Noise (DFN)**

Lida *et al.* found that three filters, two notch and one peak, could be used to replace a listener's HRTF to provide 'almost the same localisation accuracy' [Blanco-Martin *et al.* 2011]. In this experiment these three filters were replicated using MaxMSP. Each individual filter was randomly assigned a centre frequency between 6 and 12 kHz. The randomly assigned frequency changed periodically. Three period times were used: 50, 100 and 200 ms. These three filters were combined to create an overall dynamically changing filter.

When this filter was applied to the white noise it created a signal whose spectrum constantly changed for high frequencies and a signal that sounded like the HRTF swept noise. The advantage of this method was that at low frequencies the signal remained constant and that the signal was simple to create.

### 6.5.3 Listeners

Five experienced listeners took part in the pilot experiment, all between the ages of 18 and 35 and with no reported hearing problems. A reported location was elicited from each listener for every programme item and loudspeaker location combination, resulting in 42 trials per listener. A familiarity session of 10 trials was included at the start of the test. The test took approximately 20 minutes to complete.

### 6.5.4 Aim

Experiment 1 aimed to answer the following question:
6. Pinna Cue Suppression Methods

- Do the test programme items, Low Pass Filtered Noise or Spectrally Varied Noise, suppress the pinna cues of the listener?

If the Low Pass Filtered Noise or Spectrally Varied Noise resulted in a significantly less accurate elevation LRA when compared to the broadband noise then it could be concluded that pinna cues were suppressed.

6.5.5 Test Output

For every trial the following data were recorded:

- Listener Number
- Trial Number
- Programme item
- Loudspeaker Number
- Tracker Cartesian Coordinates
- Tracker Orientation

A text file was created for each test. These files were imported into Microsoft Excel where the calculations described in Section 6.4.2 were carried out. The reported source elevation was calculated from the tracker data.

6.5.6 Results

The data analysis described in this section was conducted in the statistics software, SPSS. For ease of description, short hand terms will be used to describe the programme items in the results section. Programme items will be labeled as follows:

- HRTF Based Random Spectral Swept Noise - SSN
- Dynamically Filtered Noise - DFN

The time following the acronym denotes the speed at which the filter settings were varied in milliseconds, for example: ‘SSN 60 ms’ used the HRTF Based Random Spectral Swept Noise and randomly varied the HRTF every 60 ms.
6.5.6.1 Absolute Elevation LRE

The absolute difference between the reported and actual loudspeaker locations is a metric used by many localisation studies to represent the LRA [Perrett and Noble 1997; Thurlow and Runge 1967]. An ANOVA was conducted on the absolute elevation LRE with fixed factors ‘programme item’ and ‘loudspeaker’, and the random factor ‘listener number’. Validation tests conducted on the absolute elevation LRE prior to the ANOVA showed 38 of the 42 factor comparisons to be normally distributed. The results of the ANOVA are shown in Table 8.

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>Hypothesis</td>
<td>30415.755</td>
<td>1</td>
<td>30415.755</td>
<td>216</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>564.484</td>
<td>4</td>
<td>141.121</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Programme_Item</td>
<td>Hypothesis</td>
<td>624.128</td>
<td>6</td>
<td>104.021</td>
<td>4.45</td>
<td>.004</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>560.807</td>
<td>24</td>
<td>23.367</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loudspeaker</td>
<td>Hypothesis</td>
<td>5414.209</td>
<td>5</td>
<td>1082.842</td>
<td>5.92</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>3657.477</td>
<td>20</td>
<td>182.874</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Listener</td>
<td>Hypothesis</td>
<td>564.484</td>
<td>4</td>
<td>141.121</td>
<td>.764</td>
<td>.561</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>3710.672</td>
<td>20.1</td>
<td>184.726</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loudspeaker * Listener</td>
<td>Hypothesis</td>
<td>3657.477</td>
<td>20</td>
<td>182.874</td>
<td>8.50</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>2581.809</td>
<td>120</td>
<td>21.515</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Programme_Item * Listener</td>
<td>Hypothesis</td>
<td>560.807</td>
<td>24</td>
<td>23.367</td>
<td>1.09</td>
<td>.370</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>2581.809</td>
<td>120</td>
<td>21.515</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Programme_Item * Loudspeaker</td>
<td>Hypothesis</td>
<td>1285.681</td>
<td>30</td>
<td>42.856</td>
<td>1.99</td>
<td>.005</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>2581.809</td>
<td>120</td>
<td>21.515</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8 - ANOVA of absolute elevation LRE with factors programme item, loudspeaker and listener

The factors programme item (F (6, 210) = 4.45, p = 0.004, $\eta^2 = 0.527$) and loudspeaker (F (5, 210) = 5.92, p = 0.002, $\eta^2 = 0.597$) were both shown to be significant. The secondary interactions programme item * loudspeaker (F (30, 210) = 1.99, p = 0.005, $\eta^2 = 0.332$) and loudspeaker * listener (F (20, 210) = 8.5, p > 0.001, $\eta^2 = 0.586$) were also shown to be significant. A Kolmogorov-Smirnov test conducted on the ANOVA model’s standardised residuals showed a normal distribution, indicating that the ANOVA model provided a good fit to the data.
Programme Item

The ANOVA showed that the programme item factor significantly affected the absolute elevation LRE. However the secondary interaction programme item * loudspeaker was also significant and will be discussed first. Figure 19 shows the absolute elevation LRE separated by loudspeaker number and programme item.

![Figure 19 - Absolute elevation LRE separated by loudspeaker number and programme item](image)

Loudspeaker 6 shows a significantly increased error in localisation when compared to the other loudspeaker locations for the DFN 200 ms, DFN 100 ms LPF, SSN 150 ms noise programme items. Loudspeaker 6 was located at the highest elevation tested; this indicates that loudspeakers located further in elevation from the equatorial plane results in a lower elevation LRA. No particular trend in programme item with loudspeaker location can be seen from this plot. Therefore the primary factor programme item was investigated alone.
Bonferroni and Tukey-B post hoc tests were conducted on the programme item factor of the ANOVA. Both tests showed that the Low Pass filtered noise programme item was significantly different from all other programme items (Table 9).

<table>
<thead>
<tr>
<th>Programme Item</th>
<th>N</th>
<th>Subset</th>
</tr>
</thead>
<tbody>
<tr>
<td>White Noise</td>
<td>30</td>
<td>10.2569</td>
</tr>
<tr>
<td>SSN 60 ms</td>
<td>30</td>
<td>11.0283</td>
</tr>
<tr>
<td>DFN 50 ms</td>
<td>30</td>
<td>11.2384</td>
</tr>
<tr>
<td>SSN 150 ms</td>
<td>30</td>
<td>11.4799</td>
</tr>
<tr>
<td>DFN 100 ms</td>
<td>30</td>
<td>11.5804</td>
</tr>
<tr>
<td>DFN 200 ms</td>
<td>30</td>
<td>12.7572</td>
</tr>
<tr>
<td>LPF Noise</td>
<td>30</td>
<td>15.9026</td>
</tr>
</tbody>
</table>

Table 9 - Tukey B post-hoc test on the programme item factor

The spectrally varied program items, both the DFN and SSN methods, were not significantly different from the white noise programme item; furthermore, when the ANOVA was repeated with LPF noise removed the programme item factor was no longer significant. This shows that the LPF noise programme significantly reduced the listeners static pinna cue LRA when compared to a broadband noise programme item. No other programme item was shown to significantly reduce pinna cues using the ANOVA analysis.

The Tukey B table also shows that Low Pass Filtered Noise had the highest mean elevation LRE and the broadband white noise had the lowest. This shows that using the LPF programme item made localising the source’s elevation most challenging; this finding is in agreement with previous studies described in Chapter 4 [Perrett and Noble 1997]. It can be seen that all the spectrally varied programme items did have a slightly higher mean elevation LRE than the broadband noise programme item, however this was not shown to be significant.
6.5.7 Conclusions

The low pass filtered noise programme item was shown to significantly degrade the elevation LRA of the listeners. This finding corroborates the results of other studies described in Section 6.3.

The main spectral variations due to the filtering of the pinna occur above 4 kHz. A low pass filtered noise programme item does not contain significantly energy at frequencies above 2 kHz and so does not excite the pinna filtering frequencies. This means that the spectral cues required for elevation localisation are not present and so LRA is reduced. Elevation LRA using low pass filtered noise was not reduced to random guessing and so some localisation cues were still present. It is suggested that these might be due to shoulder and floor reflections creating spectral cues at low frequencies.

Results showed that only LPF Noise significantly reduced the elevation LRA of the listeners when compared to broadband white noise. The spectrally varied noise did exhibit a slightly higher mean elevation LRE than the baseline white noise programme item, however this was not found to be significant. These findings show that the spectrally varied noise programme item did not degrade the elevation LRA of the listeners significantly.

The listener had no knowledge \textit{a priori} of the source spectrum, which was constantly changing and never flat, so the instantaneous spectrum could not be used to calculate the source location. To resolve the correct location, the listener must have been able to separate the underlying spectrum given by pinna filtering from the constantly changing source spectrum. If the listener's hearing system had integrated the spectrum of the source over a long time period, then they may have been able to derive the underlying spectrum created by their pinna from the varying spectrum of the programme item. Investigation of this finding is not necessary for the narrative of this thesis but is an interesting area of further study.
6.6 Pilot Experiment 2

In Pilot Experiment 1, Low Pass Filtered Noise was shown to significantly reduce elevation LRA. Spectrally varied noise was ineffective at suppressing pinna cues, as it did not reduce LRA significantly. It would be useful to find a programme item that both contained high frequency content and suppressed pinna cues. Pilot Experiment 2 tested an alternative programme item type, to see whether it effectively suppressed pinna cues.

6.6.1 Programme Items

Langendijk and Bronkhorst [2002] investigated the effects of removing spectral cues on LRA. They measured direction specific HRTFs for each listener and 976 loudspeaker locations, ranging 360° in azimuth and from -50° to +90° in elevation. To create a generalised HRTF that would have no directional cues, which they called the average transfer function (ATF), they averaged the HRTFs across all 976 positions. By subtracting the average transfer function from the direction specific HRTF, a ‘Directional Transfer Function’ (DTF) was created. By combining the DTFs and ATFs within different frequency bands, the importance of the spectral cues within those bands could be found. It was found that removing spectral cues in half-octave bands, by replacing the DTF with the ATF within that band, left the elevation LRA unchanged. This was true for all half-octave bands tested, namely 4-5.7, 5.7-8, 8-11.3, and 11.3-16 kHz. For each programme item there was sufficient spectral information outside of the half-octave band to allow elevation localisation. When one and two octave bands of ATF were used the elevation LRA was significantly reduced.

This experiment was dependent on removing the cues from certain bandwidths of the programme item. The alternative study would test whether, by including only a certain bandwidth of noise, localisation using spectral cues was still possible. Based on the study of Langendijk and Bronkhorst, it is proposed that narrow half-octave bands of noise do not contain enough spectral information to allow localisation using pinna cues. A programme item was devised that combined low-pass-filtered noise with narrow band half-octave high frequency
filtered noise. In Pilot Experiment 1, Low Pass Filtered Noise was shown to significantly reduce elevation LRA. It is suggested that combining the Low-Pass-Filtered Noise with a narrow band of high frequency noise will not improve this LRA.

To create this group of programme items the two components, low-pass-filtered noise and high frequency narrow bandpass filtered noise, were created separately and then combined. Four programme items were used:

- 2 kHz Low Pass Noise with a 4-5.7 kHz narrow bandpass component
- 2 kHz Low Pass Noise with a 5.7-8 kHz narrow bandpass component
- 2 kHz Low Pass Noise with a 8-11.3 kHz narrow bandpass component
- 4 kHz Low Pass Noise with a 5.7-8 kHz narrow bandpass component

The half-octave bands were chosen to be within the frequency range affected by pinna filtering. An informal listening test using the setup of the actual experiment was used to reduce the number of programme items to the four above.

This programme item type was tested against Broadband White Noise (WN) and Low-Pass-Filtered Noise (LN). The Broadband White Noise was used as a baseline against which the elevation LRE of the other programme items was judged. Each programme item was three seconds long, with an onset and offset ramp of 300ms.

In Experiment 1, Low Pass Filtered Noise was shown to weaken pinna cues, indicated by the significant reduction in elevation LRA. It was included in this experiment as a second baseline against which the degradation in LRA performance of the other programme items could be compared.

This experiment used two cut-off frequencies for the Low Pass Filtered Noise – 2kHz and 4kHz. In Experiment 1, a 2kHz cut-off was used to suppress pinna cues. Using the 4kHz cut-off frequency noise will indicate the importance of frequencies between 2 and 4 kHz in elevation LRA. Previous literature gives conflicting evidence as to the importance of this frequency range.
6.6.2 Loudspeaker Locations

The loudspeaker locations for pilot experiment 2 (Table 10) were very similar to those in pilot experiment 1. All loudspeakers were located on the median vertical plane.

<table>
<thead>
<tr>
<th>Loudspeaker Number</th>
<th>Elevation Angle (Degrees)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-8.7</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>15.8</td>
</tr>
<tr>
<td>4</td>
<td>28.0</td>
</tr>
<tr>
<td>5</td>
<td>42.5</td>
</tr>
<tr>
<td>6</td>
<td>58.4</td>
</tr>
</tbody>
</table>

*Table 10 - Loudspeaker number and elevation angle for Pilot Experiment 2. All loudspeakers were located on the median vertical plane.*

6.6.3 Listeners

Six experienced listeners took part in the pilot experiment, all between the ages of 18 and 35 and with no reported hearing problems. A reported location was elicited from each listener for every programme item and loudspeaker location combination, resulting in 42 trials per listener. A familiarity session of 10 trials was included at the start of the test. The test took approximately 20 minutes to complete.
6.6.4 Results

The data analysis described in this section was conducted in the statistics software, SPSS. For ease of description, short hand terms will be used to describe the programme item in the results section. Programme items will be labeled as follows:

Broadband White Noise – WN

Low Pass Filtered Noise Group - LN

2 kHz Low Pass Filtered Noise – L2

4 kHz Low Pass Filtered Noise – L4

Low Pass Filtered Noise with Narrow Bandpass Component Group - BN

2 kHz Low Pass Noise with a 4-5.7 kHz narrow bandpass component – L2_B4

2 kHz Low Pass Noise with a 5.7-8 kHz narrow bandpass component – L2_B5.7

2 kHz Low Pass Noise with a 8-11.3 kHz narrow bandpass component – L2_B8

4 kHz Low Pass Noise with a 5.7-8 kHz narrow bandpass component – L4_B5.7

Overall programme item categories are shown in bold. The number following B gives the lower cut-off frequency of the half-octave bandpass filter.

Absolute Elevation LRE

An ANOVA was conducted on the absolute elevation LRE with fixed factors 'programme item' and 'loudspeaker', and the random factor 'listener number'. Validation tests on the absolute elevation LRE prior to the ANOVA showed 32 of the 42 factor comparisons to be normally distributed. The results of the ANOVA are shown in Table 11.
Table 11 - ANOVA of absolute elevation LRE with factors programme item, loudspeaker and listeners

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>Hypothesis</td>
<td>65427.868</td>
<td>1</td>
<td>65428</td>
<td>307.9</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>1062.631</td>
<td>5</td>
<td>212.5a</td>
<td>.000</td>
<td>.375</td>
</tr>
<tr>
<td>Programme_Item</td>
<td>Hypothesis</td>
<td>1084.974</td>
<td>6</td>
<td>180.83</td>
<td>2.997</td>
<td>.020</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>1810.019</td>
<td>30</td>
<td>60.33b</td>
<td>.000</td>
<td>.375</td>
</tr>
<tr>
<td>Loudspeaker</td>
<td>Hypothesis</td>
<td>14909.679</td>
<td>5</td>
<td>2981.9</td>
<td>6.181</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>12060.863</td>
<td>25</td>
<td>482.4c</td>
<td>.000</td>
<td>.553</td>
</tr>
<tr>
<td>ListenerNo</td>
<td>Hypothesis</td>
<td>1062.631</td>
<td>5</td>
<td>212.53</td>
<td>.416</td>
<td>.834</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>14137.980</td>
<td>27.7</td>
<td>511.6d</td>
<td>.000</td>
<td>.717</td>
</tr>
<tr>
<td>Loudspeaker * ListenerNo</td>
<td>Hypothesis</td>
<td>12060.863</td>
<td>25</td>
<td>482.43</td>
<td>15.19</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>4765.566</td>
<td>150</td>
<td>31.77e</td>
<td>.000</td>
<td>.717</td>
</tr>
<tr>
<td>Programme_Item * ListenerNo</td>
<td>Hypothesis</td>
<td>1810.019</td>
<td>30</td>
<td>60.334</td>
<td>1.899</td>
<td>.007</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>4765.566</td>
<td>150</td>
<td>31.77e</td>
<td>.000</td>
<td>.275</td>
</tr>
<tr>
<td>Programme_Item * Loudspeaker</td>
<td>Hypothesis</td>
<td>3012.262</td>
<td>30</td>
<td>100.41</td>
<td>3.160</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>4765.566</td>
<td>150</td>
<td>31.77e</td>
<td>.000</td>
<td>.387</td>
</tr>
</tbody>
</table>

The factors programme item (F (6, 210) = 2.997, p = 0.02, η² = 0.375) and loudspeaker (F (5, 210) = 6.181, p = 0.001, η² = 0.553) were both shown to be significant. The secondary interactions programme item * loudspeaker (F (30, 210) = 3.16, p > 0.001, η² = 0.387), loudspeaker * listener (F (25, 210) = 15.19, p > 0.001, η² = 0.717) and programme item*listener (F (30, 210) = 1.899, p = 0.007, η² = 0.275) were also shown to be significant. A Kolmogorov-Smirnov test conducted on the ANOVA model’s standardised residuals showed a normal distribution, indicating that the ANOVA model provided a good fit to the data.

Programme Item

Figure 20 shows the absolute elevation LRE plotted against loudspeaker number and programme item.
There is an overall trend of lower LRA for higher and lower loudspeaker locations. Loudspeakers 3 and 4, located at 15.8° and 28.0° respectively, were the most accurately reported. This finding is investigated further in the 'loudspeaker' section below.

Loudspeaker 3 does appear to highlight some differences in elevation LRA between programme items. The L2_B5.7 programme item has a significantly higher elevation LRE than the WN, L4 and L2_B4 programme items. However this finding is only based on the 6 listener responses for each bar. The programme item variable was subsequently investigated alone, so that overall trends with a larger more robust data set could be found.
Bonferroni and Tukey-B post hoc tests conducted after the ANOVA showed that the broadband white noise programme item had a significantly lower absolute elevation LRE than all other programme items (Table 12).

<table>
<thead>
<tr>
<th>Programme_Item</th>
<th>N</th>
<th>Subset</th>
</tr>
</thead>
<tbody>
<tr>
<td>WN</td>
<td>36</td>
<td>1</td>
</tr>
<tr>
<td>L4_B5.7</td>
<td>36</td>
<td>11.4782</td>
</tr>
<tr>
<td>L2</td>
<td>36</td>
<td>15.0296</td>
</tr>
<tr>
<td>L2_B8</td>
<td>36</td>
<td>16.4671</td>
</tr>
<tr>
<td>L4</td>
<td>36</td>
<td>17.2121</td>
</tr>
<tr>
<td>L2_B4</td>
<td>36</td>
<td>17.3024</td>
</tr>
<tr>
<td>L2_B5.7</td>
<td>36</td>
<td>17.6258</td>
</tr>
</tbody>
</table>

Table 12 – Tukey-B post hoc test groupings

The low-pass with an additional high frequency bandpass component programme items were not significantly different from the low-pass programme items alone. This shows that the addition of the high frequency bandpass component did not improve the listeners’ elevation LRA; the low pass filtered noise with an additional bandpass component suppressed pinna cues to a similar level to the low-pass filtered noise. Actually, the mean error of BN is slightly higher than that of LN. The WN programme item has the lowest absolute elevation LRE, indicating that it is the easiest programme item to localise.

Loudspeaker

The loudspeaker number was found to be significant, $F (5, 210) = 6.181$, $p = 0.001$, $\eta^2 = 0.553$. This means that the elevation LRE was significantly different between loudspeakers. Previous elevation studies on the median plane have shown that as the loudspeaker elevation gets closer to ± 90°, the mean perceived loudspeaker location is biased towards the equatorial plane [Perrett and Noble 1997]. This ‘compression’ effect means that listeners perceive the loudspeaker location lower (if above the equatorial plane) than its actual location.
A Post Hoc Tukey B test was conducted on the factor 'Loudspeaker Number' (Table 13).

### Table 13 - Tukey-B Grouping of loudspeaker number with elevation LRE

<table>
<thead>
<tr>
<th>Loudspeaker</th>
<th>N</th>
<th>Subset</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Tukey B(^{a,b})</td>
<td>42</td>
<td>6.7807</td>
</tr>
<tr>
<td>3</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>42</td>
<td>14.6131</td>
</tr>
<tr>
<td>2</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>42</td>
<td></td>
</tr>
</tbody>
</table>

The mean elevation LRE for each loudspeaker is displayed. The Tukey B test assigns statistically significantly different factors to different subsets. The test showed the loudspeakers divided into three non-overlapping subset groups. In this instance, the subset groups corresponded to the distance of the loudspeakers from the median loudspeaker location, i.e. the outermost, middle, and innermost loudspeakers were grouped into pairs (Figure 21).
Speakers 3 and 4 were located at 15.8° and 28° respectively and so did not span the equatorial plane. This finding suggests that it was not the distance from the equatorial plane that caused the compression effect but the distance from the median possible response angle. In the experiment the acoustically transparent sheet covered locations from approximately 5° above the highest loudspeaker location to 5° below the lowest loudspeaker location. During the test listeners were aware that loudspeakers were only placed behind the curtain, therefore their response angles were limited to those locations. It is likely that the compression effect is a function of this limited response angles. At middle loudspeaker elevations it is suggested that listener responses form a normal distribution around the actual loudspeaker location. At the extreme angles, just below the top of the curtain or just above the bottom of the curtain, listener response will not follow a normal distribution around the actual location because listeners are aware that the loudspeaker location cannot be outside of the curtain (this would make it visible to the listener) and so their mean
response angle will be biased towards the centre of the curtain. This suggests that the compression effect is a characteristic that occurs when the listener response is limited by the curtain’s location. It is suggested that if a curtain were used to conceal all possible response angles, therefore enabling the listener to respond with any perceived angle then the effect will be reduced. The effect of loudspeaker position on the elevation LRA is discussed in greater detail in Chapter 8.

**Correlation of Reported and Actual Loudspeaker Elevation**

A further metric that gives an indication of the elevation LRA is the correlation between the reported location of the loudspeaker and the actual loudspeaker location. If these were shown to correlate closely then that would indicate that the listener was gaining a significant impression of the actual loudspeaker elevation.

The overall data file was split by programme item and the correlation between the factors ‘Actual Elevation’ and ‘Reported Elevation’ was checked using a Pearson Correlation. All programme items showed some correlation between the actual and reported elevation. As stated by Field [2009], effect sizes should be interpreted ‘within the context of the research literature’. It was anticipated that high correlations would occur between reported and actual location in this localisation test. The best method of interpretation is a comparison between cases within the test.

L2_B4 was the only programme item to result in a non-significant correlation, $r = 0.326$, $p = 0.053$. Programme item WN had the highest correlation $r = 0.757$, $p < 0.001$. Field states that ±0.5 indicates a large effect, and only WN (0.757) and L4_B5.7 (0.592) were above this value. Table 14 shows all the programme items ranked in order of correlation.
<table>
<thead>
<tr>
<th>Programme item</th>
<th>Pearson Correlation r Value</th>
<th>Significance p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2_B4</td>
<td>0.326</td>
<td>0.053</td>
</tr>
<tr>
<td>L2_B5.7</td>
<td>0.343</td>
<td>0.041</td>
</tr>
<tr>
<td>L4</td>
<td>0.372</td>
<td>0.026</td>
</tr>
<tr>
<td>L2_B8</td>
<td>0.402</td>
<td>0.015</td>
</tr>
<tr>
<td>L2</td>
<td>0.464</td>
<td>0.004</td>
</tr>
<tr>
<td>L4_B5.7</td>
<td>0.592</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>WN</td>
<td>0.757</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

*Table 14 - Programme items ranked by correlation of actual and reported loudspeaker location*

Interestingly, L2 was the most highly correlated of the medium effect programme items (r value between 0.3 and 0.5). However, the differences in this medium group are small and are probably due to the small sample size. It is possible to state from the correlation calculations that the BN programme item group with L2 component confuses elevation localisation cues at least as much as the LN group.

Figure 22 and Figure 23 plot the actual loudspeaker elevation against reported location programme items for the programmes with the highest (WN) and lowest (L2_B4) Pearson correlation value.
Figure 22 - Plot of reported elevation against loudspeaker number for programme item WN

Figure 23 - Plot of reported elevation against loudspeaker number for programme item L2_B4
The trend of increasing reported elevation with increasing loudspeaker elevation can be easily observed for the WN programme item; this finding shows that listeners could perceive the changing elevation of the loudspeakers. For the L2_B4 programme item no significant trend is apparent. This shows that the L2_B4 programme item effectively reduced the listener’s elevation LRA.

### 6.6.5 Conclusions

White noise was shown to be the easiest programme item to localise using a number of statistical methods. The programme item factor was shown to be significant using an ANOVA on the absolute elevation LRE and a Tukey-B post-hoc test highlighted white noise as giving a significantly higher LRA than all the other programme items. The correlation tests showed that the perceived location of the loudspeaker varied in a similar way to the actual location of the loudspeaker ($r = 0.757$, $p < 0.001$) for the white noise programme. White noise was predicted to be the easiest programme item to localise due to its wide bandwidth. Spectral cues from a large range of frequencies could be used to localise the source.

L2_B4 was shown to be the programme item with the least correlation between reported and actual loudspeaker location, while L2_B5.7 had the largest mean absolute elevation LRE. All spectrally altered programme items were shown to be significantly different from the white noise programme item using a Tukey-B post hoc test. The hypothesis, based on the research of Langendijk and Bronkhorst [2002], was that no significant improvement in LRA would be gained by adding a half-octave band of high frequency noise. Results indicate that not only was no improvement found but possibly a slight degradation in LRA took place. It is suggested that this degradation was due to the splitting or smearing of the perceived location of the stimuli, described by the listeners auditioning the BN programme item group.

The elevation LRA of listeners to all of the BN programme items was similar to that of the LN programme items. This shows that no significant improvement in elevation LRA was gained by the addition of the narrow bandwidth high frequency component. Spectral cues require a larger bandwidth than half-octave
to allow accurate localisation. The BN programme item group was found to reduce elevation LRA to the same level as LN while also containing a high frequency component. This means the BN group can be used as a programme item in a localisation experiment in order to suppress pinna cues.

The L4 programme item had a lower mean absolute elevation LRE than the L2 programme item and showed a higher correlation between the reported and actual loudspeaker location. However, these differences were not found to be significant; the main conclusion from the experiment was that both the L2 and L4 programmes were effective at suppressing a listener’s pinna cues.

A compression effect was noted on listener responses, with sources further from the median source location resulting in a larger bias towards this median. It is suggested that this is due to the placement of the curtain and listener expectation. This finding is shown in all the experiments of this thesis and is discussed in greater detail in Chapters 8 and 11.

6.7 Overall Summary and Conclusions

Various methods used to suppress pinna cues have been discussed in this chapter: namely pinna occlusion, virtual sources and low-pass filtered noise. Methods of physical pinna occlusion can cause the spectrum of the sound source to be unnaturally altered and can damage the listener’s ears. Virtual sources often lack externalisation cues and are time intensive to create if individual HRTFs are to be obtained.

Low-pass filtered noise proved an effective method of supressing pinna cues in other localisation tests [Perrett and Noble 1997] and in both pilot studies described in this chapter. The goal of the pilot studies was to find a programme item that supressed pinna cues and also contained high frequency content. The lack of high frequency content in the low-pass filtered programme item is potentially a problem as further studies may require other cues at high frequencies to be studied. A novel programme item, Low-Pass Filtered Noise with Narrow High Frequency Bandpass Component, was shown to suppress pinna cues in the pilot test and was inspired by previous studies conducted by
Langendijk and Bronkhorst [2002]. The importance of this programme item is that its high frequency component can be used to study other cues at high frequencies, such as dynamic ILDs given by head movement. This novel programme item will be included in the head movement localisation study described in Chapter 8. It will test whether high frequency dynamic ILD cues can be used to improve elevation LRA, whilst the cues available from pinna cues are suppressed.
7 Localisation Response Methods

During a localisation test listeners will be asked to judge the apparent location of sound sources. A response method is required to communicate where the listeners perceive the location of the sound sources. There is no standardised method for eliciting the perceived loudspeaker location from a listener during a localisation test [Evans 1998]. A number of different localisation response methods have been used in the localisation studies described in Chapters 2-6. The localisation response method has a significant effect upon the LRA of the listeners and so assessing and comparing the results of localisation studies is difficult. The aim of this chapter is to find the most accurate and effective localisation response method for three-dimensional localisation studies that focus on head movements.

The response method used in the pilot experiments described in the previous chapter was adequate to highlight differences in median plane elevation LRA for the various programme items in the study. However, when the smaller differences between responses are investigated, such as the effect of head movement on the azimuth LRA of a broadband white noise source, a more accurate method is sought.

Previous tests have used various response methods: verbal and non-verbal response; continuous and quantised response locations; graphical and physical pointing methods. Frank et al. [2010] split the most common tests into 4 categories: verbal methods, graphical methods, methods of adjustment and pointing methods.

The method will vary depending on what the tester is trying to measure. If the listener's absolute localisation capability is to be measured then the response method must have a higher resolution angle than the listener's hearing system [Makous and Middlebrooks 1990]. Therefore, an angle resolution of less than the listener's Minimum Audible Angle (MAA) is required and so the focus is on accuracy of measurement. If the experiment is testing whether a listener can discern whether a source is in front or behind, above or below, then a less
accurate measurement system is required and so the focus may shift to the intuitiveness of response or technical simplicity.

Makous and Middlebrooks [1990] highlighted a number of criteria they deemed necessary for a localisation response method:

- *Sources free to vary in azimuth and elevation with no a priori quantization due to having a limited number of response angles*
- *Responses easily learned*
- *Reasonable data collection rate*

Two further criteria are important for a localisation response method:

- *Accurate response*
- *Ease of set up for tester*

Section 7.1 of this chapter reviews the suitability and effectiveness of response methods used in previous localisation studies by referencing these criteria. Response methods will be split into the 4 broad sections outlined above and the attributes of each will be discussed. Section 7.2 of this chapter describes an experiment that compares various response methods in an effort to find the most accurate and effective.

### 7.1 Response Methods

#### 7.1.1 Verbal Methods

In Wightman and Kistler’s [1999] experiment they instructed listeners to verbally report the location of the source using a ‘world coordinate system’. The listeners must visualise a line connecting the centre of their head to that of their perceived location and then use this line as an aid to report the azimuth and elevation of the perceived source location. They used a 30-minute training session to allow the listeners to get familiar with the test. This system is technically simple to implement, requiring no complex computation from the tester.

Evans [1998] states that verbally reporting source coordinates is direct and can be very precise (even without fractions of a degree, responses in units of degree still give an error approximately equal to the MAA), however, it is unintuitive to
the listener and so results are likely to be inaccurate. To become even fairly accurate with a world coordinates system would require much training and result in large inter-listener variation. The majority of Evans’ statements are conjecture with no statistical evidence, but they do highlight the issue of using a response method without justification.

Evans also suggests that using this method, due to the listener’s egocentric view of space, response coordinates are likely to be skewed. A source located on the equatorial plane at 70° azimuth would appear to the listener as further from the median plane (0° azimuth) than a source at 70° azimuth and 70° elevation, resulting in an incorrect azimuth elicitation. It is suggested that this effect could be removed through training but a more intuitive and accurate response method should possibly be sought.

Some form of visual anchor, such as a grid laid out in degree increments in front of the listener, is likely to allow the listener to report the perceived source location more accurately. However, this would require very accurate experimental setup and would require the grid to be the same radial distance from the listener as the loudspeakers. If this were not the case then listener head movement would cause the grid to be wrongly placed and so the results would be skewed.

Numbering loudspeakers and asking the listener to report the loudspeaker number from which the sound source originated is another verbal response method. This removes the difficulty of reporting coordinates with no reference or prior knowledge and is an intuitive, simple task.

Perrett and Noble [1995] investigated the biases introduced during elevation localisation testing when using a set number of visible loudspeakers. They discussed the findings of Butler and Humanski [1992], who concluded that listeners are able to localise the elevation of sources using pure static interaural differences. Butler and Humanski [1992] found that on the lateral vertical plane, listeners were able to localise effectively even with low-pass filtered noise (used to suppress pinna cues). Perrett and Noble suggested that showing the position of the loudspeakers from which the sources originated effectively constrained the listener's response. The cone of confusion was resolved because there was
only one visual loudspeaker location with the correct interaural cues. They hypothesised that if the loudspeakers were obscured or the listeners were shown a variety of cone of confusion locations, localisation would become impossible.

Perrett and Noble repeated the experiment of Butler and Humanski but placed further loudspeakers horizontally around the listener to give alternative cone of confusion positions. With low-pass filtered noise, accurate elevation localisation response was no longer possible and errors corresponded to the cone of confusion as hypothesised. This study showed that by using a set number of visible loudspeakers, localisation responses were biased. Localisation tests that involve numbered loudspeakers are all subject to these biases. Therefore, any test should allow the listener to respond with any angle of azimuth and elevation, and not confine the listener’s response to certain locations. Even if these numbered loudspeakers correspond to cone of confusion areas, on the median plane for example, the listeners do not have the whole spectrum of angles to select from and so their responses would have a higher accuracy than during an unsighted test.

The use of numbered loudspeaker positions also means that auditory localisation acuity cannot be tested as listener responses are quantised to loudspeaker positions and so are inherently biased by their visual cues and acuity. A subject’s visual acuity is approximately two orders of magnitude greater than their auditory acuity, thus giving an unrealistic auditory LRA [Blauert 1997].

Evans [1998] suggests that a system based on a clock face is sometimes suitable due to the listener’s familiarity with the layout. Larger ‘gross errors’, such as front-back confusions, can be elicited using this response method. However, an angle resolution of 30° is unsuitable for applications such as testing listener LRA between different conditions. Extending this system into three-dimensions will also be difficult as listeners may be unable to split elevation and azimuth and report them separately.

Evans also describes a more complex and accurate experimental setup in which the listener must make a number of iterative judgements. The listener must first
judge which of a possible 16 sections a sound source is located within (Figure 24).

![Diagram of 16-section system for stage 1 of an iterative localisation response method]

*Figure 24 – 16-section system for stage 1 of an iterative localisation response method*

[Evans 1998]

In the second stage a marker sound will be played from the section chosen in the previous stage and the listener must respond with the source’s position relative to that marker. Evans states that the marker sound must be different from that being localised to avoid confusion. It is also important that the sound must be a different loudness so that the listener cannot use comparative cues to resolve spatial attributes. If a listener is confused about whether a source is in front of or behind them then a marker may be used to compare loudness and so resolve source location or change their response.

If, for example, in stage 1 the source is described as ‘right, forwards, above’ then the listener can choose from four further options once the marker has played:

- *Further right and higher*
• Further forward and higher
• Further right and lower
• Further forward and lower

This allows the listener to choose a direction from a set of 64 possible positions. By adding loudspeakers at these positions a further fourfold increase is possible allowing 256 possible directions. However, using this method a large number of loudspeakers are required to gain an angle accuracy still significantly lower than those reported in elevation localisation acuity studies [Makous and Middlebrooks 1990].

7.1.2 Graphical Methods

Simon et al. [2010] conducted a localisation study in azimuth and allowed the listener to report results graphically rather than aurally. Their GUI was marked at 15° angle increments and the listener could chose locations from a continuous rotational scale (Figure 25). A circular metal structure displayed 5° increments around the listener at a radius of 1m. They used this metal structure to ease the listener's shift from an egocentric perspective to a graphical one. The listener was also represented graphically within the GUI to further ease this shift.

Figure 25 – Graphical response method user interface used by Simon et al. [2010]
It is difficult to extend this GUI into three-dimensions and maintain response accuracy. A GUI could display two graphics, one for elevation and one for azimuth, however, as stated by Frank et al. [2010], mapping three-dimensional space onto two dimensions creates systematic errors.

If a sphere was used to represent the listening space then listeners could map their response onto the outside of it, therefore avoiding any two to three dimension mapping. Creating a virtual environment is one way of doing this (as used in [Majdak et al. 2008]), however it is likely to be more accurate if the sphere is a real object, as this will allow the listener to physically interact with it. A motion tracker’s orientation could be used to calculate the reported location by being held on the surface of the sphere. Alternatively, the sphere could be marked in 10° increments and the subject could read their response from the grid.

It may be difficult for listeners to perceive their own location within the sphere and so mapping a source location could be difficult. To ease this mapping, reference marks could be made on the sphere and these could correspond to reference loudspeakers placed around the listener.

### 7.1.3 Methods of Adjustment

When using methods of adjustment the listener is asked to move a marker until its location matches that of the source. One of the difficulties of this method is how to create a moveable marker that is controllable by the listener. Often the marker is a second acoustic stimulus. This requires a loudspeaker arm that gives no extra cues through mechanical noise and is capable of moving in both azimuth and elevation whilst maintaining the same distance from the listener. Engineering and building a silent moveable arm is extremely difficult and expensive. Methods of adjustment can prove effective in virtualised environments where there are different engineering restrictions.

Pulkki and Karjalainen [2001], in their investigation of stereophonic panning, allowed the listener to adjust the amplitude panning of a stereophonic system so that the phantom image it produced matched the location of a real loudspeaker. This method could not be used in a three-dimensional localisation acuity study.
as stereophony cannot be used to change a source's elevation. Furthermore, stereophonic localisation acuity is different from monophonic localisation acuity. The stereophonic setup is used to give the impression of a source at an azimuth location but not all of the cues are the same as that of monophonic listening.

Pulkki and Karjalainen state that methods of adjustment can be quick and intuitive for the listener. They highlight difficulties with biasing caused by the initial conditions of the adjustable marker. The initial location of the marker may affect the resultant response of the listener.

A further problem with the method of adjustment is that the listener does not have to construct a spatial image, instead they can note when the two sources timbrally and spatially sound the same. This is not the same task as asking listener to find the location of a source. A similar problem is encountered in MAA experiments. The listener will note any small change in timbre between two sources without having to construct two spatial images and compare them [Makous and Middlebrook 1990].

Morimoto and Iida [1995] placed 72 Light Emitting Diodes (LEDs) at 2.5° intervals on the horizontal plane and asked listeners to adjust a dial to light the two that best displayed the perceived apparent source width of the stimuli. A single LED could be used to measure the perceived location in a localisation study. In a localisation experiment the angle interval between the LEDs would have to be smaller than the listener's MAA for no response accuracy to be lost. Morimoto and Iida's setup allowed 180° coverage on the horizontal plane. To use such a setup for three-dimensional localisation studies would require an impractically large number of LEDs.

7.1.4 Physical Pointing Methods

Physical pointing methods come in a number of forms, for example: a finger point, a laser point or a head point. One of the advantages of this method is that the listener is allowed to choose from a spectrum of responses and is not quantised to respond in a certain way. However, as stated by Evans [1998]:
"The usefulness of this style of test is limited by the accuracy with which the user can point in the perceived direction, and the accuracy with which the indicated direction can be read."

Makous and Middlebrooks [1990] concluded that a head pointing method fulfils all the criteria they outlined for an effective localisation response method. An advantage of this technique is the naturalness of response; as shown in Chapter 2, listeners want to orient their head towards the source.

Frank et al. [2010] suggested that pointing is the most intuitive method while maintaining accuracy. However, the listener's pointing accuracy varies dependent on source location. For example, in an experiment conducted by Middlebrook [1992] it was observed that sources behind the listener were less accurately localised than those in front (median RMS error in front = 13.2° and behind = 16.3°). This highlights a problem present in all localisation response methods: There is no way of separating the accuracy of localisation from the accuracy of response. It is likely that sources behind the listener are both harder to localise and harder to point at, but which is causing the largest error is difficult to resolve. In this instance the error did not limit the test because another factor, the centre frequency of programme item, was the cause of much larger LREs. This shows that if the factor under test results in a larger change in angle than the measurement method, then the measurement method is valid.

Thurlow and Runge [1967] instructed listeners to point at the perceived source location using a metre stick. A semi-circular metal arc was built to surround the listener, with its ends fixed at points directly above and below. It was aligned with the listener's metre stick point and angles were read from a scale marked in elevation and azimuth. The subject's head was fixed in place using a clamp that stopped movement of the centre of the listener's head. This removed any inaccuracy due to misalignment of the head centre but was unnatural for the listener and posed a serious risk if the equipment were to have malfunctioned.

It is hard to engineer a metal arc capable of pivoting correctly and measure to within a degree in both azimuth and elevation. This test method requires the tester to align the arc for each trial, which causes the response measurement to be extremely slow and intensive.
Oldfield and Parker [1984] used a pointing method, stating that it is the most common and robust measuring technique. They explained that pointing does not restrict or bias the listener’s response. In their experiment, listeners pointed a ‘gun’ at the source which, when fired, triggered multiple cameras used to triangulate the reported location. This experiment required a camera setup which is now unnecessarily complicated. With current technology such as headtracking or the WII game remote, it should be possible to find an easier and equally accurate alternative.

The WII remote, released with Nintendo’s ‘Wii’ games console, has motion-sensing capability, which makes it a possible pointing method. It is simple to use and is likely to have some listener familiarity. However, its motion accuracy is unverified and it has limitations in its range of movement (it can’t sense all angles of elevation and azimuth).

Frank et al. [2010] described a pointing method that used a toy gun to ‘pierce’ a surrounding surface onto which loudspeakers were mounted. They suggested that it is easier for a listener to point with an object they can see, such as a finger or hand held object, than with a self centred object they can’t, such as an eye, head or torso.

They encountered a number of problems often found when using a motion-tracked system. When the listener extended their arm in either direction, the tracker transmitter moved out of the sensor’s range and so the listener’s pointing accuracy was greatly reduced. If the transmitter were kept within the sensor’s range then it could offer an accuracy of better than 1° (under the human MAA). If the pointing method was conducted by an object more spatially fixed, such as the listener’s head or torso, then it would be easier to control this range.

A further pointing method involves the listener directing a visible laser pointer at the perceived source location on the acoustically transparent sheet. This sheet would be marked with degree angles in elevation and azimuth. The listener would note down the angle of elevation and azimuth at which they perceive the source on a piece of paper. Unfortunately, this method makes data collection slow and inefficient. To create the sheet marked with degree angles will be complex and may lead to inaccuracy in the displayed angle.
Eye tracking or ‘gaze’ methods of localisation pointing provide an accurate response. However, their setup is complex, expensive and may require a trained medical physician. Their response is also limited to the listener’s field of vision. To gain larger response angles the tester must track both the head movement and eye movement of the listener and combine their response.

In Perrett and Noble’s [1997] elevation localisation test, the listener was instructed to orient their head towards the perceived sound source location after the offset of the stimulus. A head tracker logged the coordinates of this location. This method was simple for the listener; had a fast data collection rate; did not quantise the listeners’ responses; and the head-tracker was not limited in azimuth or elevation angle.

Pointing techniques allow a simple, intuitive and repeatable elicitation of a subject’s perceived location. They minimise the biases that are present in other methods and do not quantise the listener’s response.

7.1.5 Response Method Comparisons

It is difficult to make comparisons between response methods that have been used in different experiments. To directly compare the accuracy of response methods, each method must be used in the same experiment. Two studies that have compared response methods are described in this section.

Haber et al. (1993) studied nine response methods to find the optimum method used in localisation tests for blind subjects. The methods were:

**Pointing Methods**

- *Nose*
- *Chest*
- *Finger*
- *Cane*
- *Stick*

**Graphical Methods**

- *Waist centred graphical pointer*
• *Table centred graphical pointer*

• *Draw*

**Verbal Methods**

• *Clock*

They found that pointing methods were the most accurate and least varied response method. Chest point gave the lowest LRE of approximately 6.5˚ while the cane offered a mean error of 6.9˚. Subjects blind from birth would not have calibrated their visuomotor connections and so may be less aware of how a physical movement of their body changes their orientation in space. Therefore, findings of a study conducted on blind people cannot easily be transferred to sighted subjects.

Majdak *et al.* [2008] compared head and manual pointing methods in a virtual environment study. They measured source location using lateral angles ranging from -90˚ to 90˚ and polar angles ranging from -30˚ (front, below eye-level) to 210˚ (rear, below eye-level). Their results are summarised in Table 15.

<table>
<thead>
<tr>
<th></th>
<th>Dark Head</th>
<th>Dark Manual</th>
<th>Virtual Env. Head</th>
<th>Virtual Env. Manual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Lateral Error</td>
<td>15.9˚</td>
<td>16.7˚</td>
<td>13.3˚</td>
<td>13.5˚</td>
</tr>
<tr>
<td>Mean Raw Polar Error</td>
<td>52.5˚</td>
<td>52.5˚</td>
<td>51.7˚</td>
<td>48.2˚</td>
</tr>
<tr>
<td>Mean Corrected Polar Spread</td>
<td>32.3˚</td>
<td>32.3˚</td>
<td>30.5˚</td>
<td>30.5˚</td>
</tr>
</tbody>
</table>

*Table 15 - A comparison of head and manual pointing methods in both an unsighted (dark) and virtualised environment [Majdak et al. 2008]*

They concluded that “in localization tasks where high accuracy in the vertical dimension is required, manual pointing is the better choice.” However, an initial
overview of their results table shows only a slight improvement for manual when compared with head pointing and this is only the case in the virtual environment. Furthermore, both errors are very large (manual 48.1° and head 51.7°) which suggests the HRTF virtualisation may not have been accurately rendered.

Since the system used polar angles ranging from -30° to 210° rather than elevation angles, front-back confusions were wrongly included in the 'Mean Raw Polar (Elevation) Error'. They corrected for this, producing the 'Mean Corrected Polar Spread', in which all front-back confusions were flipped onto the correct hemisphere. For this corrected value both head and manual pointing produced exactly the same results. This finding suggests that there is no difference between head and manual pointing in elevation localisation, therefore rendering their conclusions inaccurate.

There is some improvement when using the virtual environment when compared with unsighted darkness, suggesting the visual reference cues do have an impact upon LRA. Could this test be repeated but in a real rather than virtual environment? It is suggested that in this environment elevation LRE would be significantly reduced.

7.1.6 Summary

Four main localisation response methods have been considered in this chapter: verbal methods, graphical methods, methods of adjustment and pointing methods.

The challenges facing testers using each method have been highlighted. It is difficult to accurately report the source location using verbal methods without some form of visual anchor. It is problematic to position anchors without biasing the listener response. Reporting numbered loudspeakers is undesired as it biases the listener's response. Listeners should be free to respond with any location in azimuth and elevation.

Graphical methods most commonly require the listener to project their three-dimensional percept onto two two-dimensional interfaces. Using graphical
methods, it must be made apparent where the listener is located with reference to the graphical representation.

Methods of adjustment are difficult to set up and are most commonly realised in a virtualised environment. With methods of adjustment it is often the MAA that is measured not the localisation acuity.

Pointing methods are intuitive and simple for the listener to understand. The ideal pointing mechanism is still unverified, though it is suggested that a pointer that the listener can see is better.

An overview of localisation response methods assessed by the criteria outlined by Makous and Middlebrook is given in Table 16.
Table 16 - A comparison of localisation response methods using the criteria of Makous and Middlebrook [1990]

<table>
<thead>
<tr>
<th>Methods of Adjustment</th>
<th>Pointing Methods</th>
<th>Verbal Methods</th>
<th>Graphical Methods</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal</td>
<td>Fast (manual)</td>
<td>Fast</td>
<td>Fast</td>
<td>Response point can be time intensive (response point can only respond within movement of marker)</td>
</tr>
<tr>
<td>Graphical</td>
<td>Fast (though may take one for elevation)</td>
<td>Fast (though other requires two adjustments – azimuth and elevation)</td>
<td>Fast</td>
<td>System is fast</td>
</tr>
<tr>
<td>Methods</td>
<td></td>
<td></td>
<td></td>
<td>World coordinates</td>
</tr>
<tr>
<td>of Adjustment</td>
<td></td>
<td></td>
<td></td>
<td>Depends on method</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data collection rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response easily learned</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sources time to vary in azimuth and elevation with no response quantisation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problems transferring from 3D to 2D perspective</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D perspective</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problems transferring from 3D to 2D perspective</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Listeners have difficulty understanding 3D perspective</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>World coordinates</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System can be difficult</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Listeners have difficulty understanding 3D perspective</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Haber et al. [1993] compared localisation response methods for blind subjects. Whether their results are transferrable to sighted subjects is unclear. It is noted that the azimuth LRE in these studies is higher than would be expected for sighted subjects. It is suggested that blind subjects may not have iteratively calibrated their auditory system using their visual system, causing them to have less accurate localisation ability. This may be especially prevalent in the use of pinna cues, which have been shown to be a learned localisation cue. The study was also confined to a single plane, which will have a significant impact upon the optimal choice of response method. Majdak et al. [2008] conducted their response study in a virtual environment. They found an abnormally high LRE, which suggests that the rendering of their virtual environment was inaccurate and so localisation cues were imprecise. If the localisation cues were imprecise then it is likely that the error of localisation was higher than the error of the response method, so no meaningful optimal method could be found.

It is tentatively suggested that pointing methods are the most accurate of the four response methods, as suggested by the comparison studies. However, flaws in the studies of both Haber et al. [1993] and Majdak et al. [2008] make their conclusions inapplicable to a three-dimensional head movement localisation study. Conclusions can be suggested by comparing separate studies on response methods but methods must be compared in the same conditions for accurate conclusions to be reached. There has not been a conclusive study comparing the LRA of localisation response methods.
7.2 Localisation Response Methods Experiment

Section 7.1 compared previous experiments that have used different localisation response methods. Each method had its advantages and disadvantages and it was concluded that based on previous evidence there was no obvious optimal response method. It was also found that direct comparisons based on separate experiments using different response method experiments are difficult.

The studies of Perrett and Noble [1997] and Wightman and Kistler [1999], described in Chapter 5, came to different overall conclusions on the importance of head movement in elevation localisation. One difference highlighted in the comparison of the studies was the response method. Observing the impact of the response method on the LRA may explain the differences in the outcomes of these studies.

A study was devised to compare the LRA of localisation response methods. This study should allow a better understanding of what comparisons can be made between localisation studies that have used different methods. It will also indicate the optimal response method, which can be used to justify its use in further localisation studies. Since the localisation studies in this thesis focus on head movement in localisation, it is also necessary to verify the effect of head movement on LRA. This study will allow the studies of Perrett and Noble [1997] and Wightman and Kistler [1999] to be better compared.

To avoid the limitations of previous experiments, the proposed study will have real sources located at various points in both azimuth and elevation to be localised by sighted subjects who will respond with various response methods.

The proposed study has two goals:

- **To find the most accurate localisation response method**
- **To measure how the LRA varies when the listener is able move their head.**

Listeners will be asked to give the perceived location of an easily localisable programme item. Loudspeaker sources will replay this programme item at varying locations in azimuth and elevation. The experiment results will also
enable one to map localisation response acuity across azimuth and elevation for each response method.

Section 7.2.1 explains how the experiment was set up, including details on the listeners, the programme item and the source loudspeakers’ locations. Section 7.2.2 describes how the reported source location was calculated for each response method. Section 7.2.3 explains the test procedure, while Sections 7.3.1 and 7.3.2 detail the results of the test and Section 7.4 shows how these results fulfil the goals laid out for the test.

7.2.1 Experimental Setup

7.2.1.1 Programme Item

The programme item was a pulse train of 150ms white noise bursts with 300ms of silence between pulses. Using a wide bandwidth source allowed changes in the source’s spectrum to be easily observable. Pinna cues, an important elevation localisation cue, are dependent on spectral variation of the source and so a wide bandwidth source allows for a higher LRA. Interaural time difference cues were given by the onset and offset of the stimulus. It was hoped that by using an easily localisable programme item, the differences in LRA between response methods would be most apparent.

7.2.1.2 Loudspeaker Locations

Eighteen Genelec 8020b loudspeakers were used to replay the test stimuli. Since it has been shown that localisation is unchanged between hemispheres [Oldfield and Parker 1994], loudspeakers were placed only on the right hemisphere. The loudspeakers were spread in azimuth from 0°, directly in front of the listener, to 180°, directly behind the listener. They were spread in elevation from -35° to +90°, due to the restricted height of the listener above the floor.

The loudspeakers were not confined to a single plane as this is thought to bias the localisation tests [Makousa and Middlebrook 1990]. Instead, they were located at the vertices of a Truncated Icosahedron (Figure 26). Although this shape has 60 vertices, only loudspeaker elevations from -35° to +90° in one
hemisphere were used, which reduced the required number of loudspeakers to eighteen. Using a regular shape allowed sources to fairly evenly span locations in azimuth and elevation. The loudspeaker numbers, coordinates and angles used in the experiment are given in Table 40 in Appendix B.

![Figure 26 – A Truncated Icosahedron, the shape used to position loudspeakers for the test](image)

The loudspeakers were oriented to face the listener so that the frequency response of the loudspeaker was consistent and did not vary with azimuth and elevation. The loudspeakers were concealed behind an acoustically transparent but visually opaque sheet to avoid the biasing provided by visual cues that was described by Perrett and Noble [1995]. The loudspeakers were mounted on microphone stands.

### 7.2.2 Response Methods

As discussed in Chapter 7, localisation response methods are often grouped into four main categories: verbal methods, graphical methods, methods of adjustment, and pointing methods. As concluded in Chapter 7, methods of adjustment are extremely time intensive and complex to set up in a non-
virtualised environment. For these reasons methods of adjustment were not included in the study.

The test investigated three main localisation response methods:

1) Laser Pointing – The listeners were instructed to point a laser, attached to their heads, at the perceived location of the source.

2) Verbal – The listeners were instructed to verbally describe, using a world coordinates system, the perceived location of the source.

3) Graphical – The listeners were instructed to direct two pointers, located on a computer’s graphical user interface, at the perceived location of the source. One pointer was used to identify elevation and one to identify azimuth.

Each method is to be described in detail in the following sections.

7.2.2.1 Laser Pointing

The laser, produced by ‘Digiflex’ and measuring 65x15x15mm, emitted a narrow beam of red light. It was clipped into the headband of a head torch, which was wrapped around the circumference of the listener’s head. The laser was aligned so that it pointed out from between the listener’s eyes.

A Polhemus Patriot system was used to track the motion of the listener’s head, logging the perceived location of each trial. The Patriot is an electro-magnetic based tracker that measures the listener’s position with six-degrees-of-freedom, the X, Y, Z coordinates and the pitch, yaw and roll orientation. It has an update rate of 60Hz and was connected via USB to the Macbook Pro where drivers, created by Kim [2010], allowed it to interface with MaxMSP. The patriot’s sensor was mounted on the listener’s head using a hair band; the band was small and flat in profile and positioned away from the listener’s ears to avoid colouring their spectral cues. The base emitter was positioned on a microphone stand within close proximity of the listener. The emitter position varied for each of the tests but was kept within the sensor limitations defined in the Polhemus user manual.
The laser pointing test was conducted in two movement conditions while the programme item was playing:

1. **No head movement**
2. **Free head movement**

These two conditions were included in the experiment to test how head movements affected LRA. The model described in Ashby [2010] was based on listener head movement and the majority of the experiments detailed in subsequent chapters will use both conditions so it is important to quantify how head movements affect LRA.

The no head movement condition was enforced using the Polhemus tracker and MaxMSP code. The code muted the programme item if the listener moved their head beyond ±2.5°, half the value recommended by Simon (5°) [2009]. In free movement conditions, the listener was able to move their head in any way they wished.

### 7.2.2.2 Verbal

Listeners were instructed that the calibration point directly in front of them represented 0° in azimuth and all sources must be measured as the change in angle in degrees from this point. In elevation, sources below their horizon (level with the floor) were given negative values, while sources above were positive. At the extremes, a source directly overhead was +90° and a source directly below the listener was -90°. The listeners input the reported azimuth and elevation source locations in two number boxes. Listeners were unable to proceed onto the following trial until the values in each number box had been entered.

### 7.2.2.3 Graphical

In the graphical method listeners were presented with the interface shown in Figure 27.
The listener would rotate the pointers on the dials to the reported location of the source in azimuth and elevation. A number box showed this angle in degrees azimuth and elevation. Listeners were unable to proceed onto the following trial until the values in each number box had changed.

### 7.2.2.4 Listeners

A total of fifteen listeners took part in the experiment and each listener produced results for all four of the response methods. All participants were from the University of Surrey: eight undergraduate Tonmeister students, two undergraduate Music students and five postgraduate members of the Institute of Sound Recording. All listeners were between the ages of 18 and 35 and had no reported hearing problems.

### 7.2.3 Test Procedure

The four response tests were conducted in a single one hour and forty minute block. Each test took approximately 15 minutes with a 5-minute break between tests. Each test consisted of 36 trials in which two stimuli were played from each loudspeaker position. For each response test a familiarisation stage of 10 stimuli was conducted. Data from the familiarisation stage was not used for results analysis. In both the familiarisation and test runs no feedback was given to the
listener as to their LRA. The order of response test was varied randomly between listeners.

The test was conducted in the audio laboratory within the Institute of Sound Recording (IoSR). Listeners were led into the room with their eyes closed and instructed to open them once in the seated listening position. This was to avoid the listener observing the loudspeaker locations before the test began.

For the four response tests, the overall test procedure was the same. Listeners were presented with a test GUI created in MaxMSP. Preceding each test, listeners were asked to calibrate the system. For the verbal and graphical responses this involved sitting in a comfortable listening position and facing a calibration point at 0˚ azimuth and elevation. The program gave a three second verbal countdown until calibration was carried out. Calibration logged the listener’s head position and orientation with six degrees of freedom, averaged over a one second time period. If the listener moved their head more than 2.5˚ in orientation or 2.5 inches in location during this time then they would be instructed to repeat the procedure. For the laser pointer response, the listener was instructed to direct the laser at two points, one at 0˚ azimuth and one at 180˚ azimuth, and a calibration was conducted for both locations.

The actual trial then started automatically, presenting the first trial page. Listeners pressed ‘Play’ to hear the programme item and ‘Stop’ to stop it. Subjects could listen to any programme item as many times as they wished.

For the laser response method, when the listeners had determined the perceived location of the source they stopped the stimulus and pointed their heads at the location. By clicking ‘log’ the listeners could save the head tracker orientation and location. A Polhemus head tracker monitored the listeners’ head movements with six degrees of freedom and each perceived source location, as indicated by the listeners, was calculated using a triangulation technique similar to the method described by Frank et al. [8] and described in detail in Appendix A. The loudspeaker radius from the central listening position was used to create a projected sphere and the perceived location was the point at which the line of the laser pointer pierced this sphere. Following the response, the next
programme item was loaded automatically. For each trial the following set of
data was logged:

- Overall Trial Number
- Movement Condition
- Speaker Number
- Tracker Location
- Tracker Orientation

For the graphical and verbal responses, once the listeners were satisfied with the responses they had input into the GUI, they clicked 'Next' for the next trial. In these trials the data logged were:

- Overall Trial Number
- Movement Condition
- Speaker Number
- Listener Pointed Elevation and Azimuth

7.2.4 Data For Analysis

The calculations described in Appendix A resulted in four dependent variables:

- Reported loudspeaker elevation
- Elevation LRE
- Reported loudspeaker azimuth
- Azimuth LRE

The independent variables for the data set were:

- Actual loudspeaker Elevation
- Actual loudspeaker Azimuth
- Response Method
- Loudspeaker Number
- Listener Number
- Trial Number
In the response method group there were four categories:

- **Movement - Laser pointing method where head movement was permitted.**
- **No Movement - Laser pointing method where head movement was not permitted.**
- **Graphical - Method where two arrow pointers were adjusted on the GUI in MaxMSP.**
- **Verbal – Method where source location was described using world coordinates system and angles typed into GUI.**

### 7.3 Experiment Results

The results for elevation and azimuth LRAs are considered in Sections 7.3.1 and 7.3.2. These sections seek to find the most accurate and consistent localisation response method to be used in subsequent head movement localisation studies.

#### 7.3.1 Elevation LRA

Each loudspeaker was positioned at a known location in both azimuth and elevation with reference to the central listening position. This section studies how the listener responded to the elevation component of this location.

**Absolute Elevation LRE**

The absolute elevation LRE was calculated as the absolute difference in elevation angle between the reported and actual location of the source. The absolute error gives an indication of the LRA.

Parametric tests, such as the ANOVA, require data to be normally distributed. A Kolmogorov–Smirnov test found that half of the conditions were normally distributed (36 of 72). Non-parametric Kruskal-Wallis tests were conducted on the elevation LRE data to check the major findings of the ANOVA.

The dependent variable ‘absolute elevation LRE’ was modelled using an ANOVA, with fixed factors:

- **Response Method**
- **Loudspeaker Number**
and random factor;

- **Listener Number**

The results of this analysis are shown in Table 17.

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>Hypothesis</td>
<td>206582</td>
<td>1</td>
<td>206581.92</td>
<td>101.2</td>
<td>.000</td>
</tr>
<tr>
<td>Speaker</td>
<td>Hypothesis</td>
<td>9480.47</td>
<td>17</td>
<td>557.675</td>
<td>5.006</td>
<td>.000</td>
</tr>
<tr>
<td>Response_Method</td>
<td>Hypothesis</td>
<td>8306.17</td>
<td>3</td>
<td>2768.724</td>
<td>11.91</td>
<td>.000</td>
</tr>
<tr>
<td>Listener</td>
<td>Hypothesis</td>
<td>28587.3</td>
<td>14</td>
<td>2041.949</td>
<td>7.155</td>
<td>.000</td>
</tr>
<tr>
<td>Response_Method * Listener</td>
<td>Hypothesis</td>
<td>9765.88</td>
<td>42</td>
<td>232.521</td>
<td>3.971</td>
<td>.000</td>
</tr>
<tr>
<td>Speaker * Listener</td>
<td>Hypothesis</td>
<td>29513.0</td>
<td>238</td>
<td>111.399</td>
<td>1.903</td>
<td>.000</td>
</tr>
<tr>
<td>Speaker * Response_Method</td>
<td>Hypothesis</td>
<td>20679.6</td>
<td>51</td>
<td>405.482</td>
<td>6.925</td>
<td>.000</td>
</tr>
</tbody>
</table>

*Table 17 - ANOVA Tables for Absolute Elevation LRE*

All primary factors were shown to be significant. The secondary interactions listener*response method (F (42, 2160) = 3.971, p > 0.001, \( \eta^2 = 0.085 \)), loudspeaker*listener (F (238, 2160) = 1.903, p > 0.001, \( \eta^2 = 0.202 \)) and loudspeaker number*response method (F (51, 2160) = 6.925, p > 0.001, \( \eta^2 = 0.164 \)), were also found to be significant. These findings are studied individually in following sections.

A histogram of the ANOVA model's standardised residuals showed a normal distribution (Figure 28) indicating that the ANOVA model provided a good fit to the data.
The listener number factor was shown to be significant \((F (14, 2160) = 7.155, p > 0.001, \eta^2 = 0.623)\). The secondary interactions listener*response method \((F (42, 2160) = 3.971, p > 0.001, \eta^2 = 0.085)\) and listener*loudspeaker \((F (238, 2160) = 1.903, p > 0.001, \eta^2 = 0.202)\) were also shown to be significant. The listener*loudspeaker interaction was only formed of 4 data points and so any significant differences might be a quirk of the data, so it would be dangerous to draw any conclusions from it.

Since the primary aim to find the effect of response method on the LRA, it is necessary to plot the listener*response method interaction. Figure 29 shows a plot of the absolute elevation LRE plotted against loudspeaker number and response method.
The main finding shown by this plot is that some listeners use some response methods more effectively than others. For example, listener 13 is significantly more accurate responding using a laser-guided movement response method than either verbal or graphical methods. There are no listeners for which the elevation LRA of either the verbal or graphical response methods is significantly higher than the elevation LRA for the laser-guided pointing method.

Listener 12 is the only listener for which the no movement laser response method resulted in a higher LRA than the free movement method. For all other listeners the free movement condition resulted in a lower mean absolute elevation LRE than the no movement condition, and for listeners 8, 9, 11 and 13, this difference was shown to be significant. This suggests that the ability to move
the head significantly improved the response of the listeners. The response method factor is studied more in the ‘response method’ section.

To show a general trend in elevation LRA with listener, Figure 30 shows a plot of the absolute elevation LRE plotted against listener number.

The graph shows that there is a large variation in the elevation LRAs of listeners. A Kruskal-Wallis test conducted on the absolute elevation LRE confirmed this finding ($H(14) = 365.342, P < 0.001$). Significant variation in elevation LRAs with listeners has also been shown in previous experiments [Langendijk and Bronkhurst 2013]; some listeners are better than others at using spectral and head movements cues to localise in elevation. This listener dependent effect is more apparent for elevation LRA than azimuth LRA. This trend is apparent in this study because the programme item used in the test was easy to localise, therefore differences in elevation LRA due to the ‘listener’ factor were more apparent. For the studies in which pinna cues were reduced, thereby making
elevation localisation responses more difficult (such as the experiments in Chapters 9, 10 and 12), the 'listener' factor was no longer significant.

**Response Method**

The main aim of this chapter is to find the localisation response method that gives the highest and most consistent three-dimensional LRA. To discover this, it is necessary to study the effect the response method factor had on the elevation LRA. The response method factor was shown to be significant (F (3, 2160) = 11.91, p > 0.001, $\eta^2 = 0.460$) in the ANOVA. The secondary interactions response method*listener (F (42, 2160) = 3.971, p > 0.001, $\eta^2 = 0.085$) and response method*loudspeaker (F (51, 2160) = 6.925, p > 0.001, $\eta^2 = 0.164$) were also shown to be significant. The interaction response method*listener was discussed in the previous section and indicated that the free movement laser pointing response method gave the highest LRA.

A non-parametric Kruskal-Wallis Test was conducted to verify the findings of the ANOVA for the response method factor. This test does not require normal distribution. It is calculated by assigning each test case a rank number and subsequently comparing the mean rank of each group of test cases. A Kruskal-Wallis test was conducted on the absolute elevation LRE with the response method factor and the results are shown in Table 18.

<table>
<thead>
<tr>
<th>Response_Method</th>
<th>N</th>
<th>Mean Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>abs_ele_error Movement</td>
<td>540</td>
<td>919.11</td>
</tr>
<tr>
<td>No Movement</td>
<td>540</td>
<td>1037.33</td>
</tr>
<tr>
<td>Verbal</td>
<td>540</td>
<td>1168.33</td>
</tr>
<tr>
<td>Graphical</td>
<td>540</td>
<td>1197.23</td>
</tr>
<tr>
<td>Total</td>
<td>2160</td>
<td></td>
</tr>
</tbody>
</table>
The results of the Kruskal-Wallis test show that the effect of the ‘response method’ was significant \((H(3) = 68.371, P < 0.001)\), which verifies the significance finding of the ANOVA. The free movement laser pointing method has the lowest mean rank, which suggests that it is the most accurate response method. The mean ranks of the verbal and graphical methods are very similar.

Since the secondary interaction response method* loudspeaker was shown to be significant, this must be investigated before the primary factor ‘response method’. Figure 31 shows a plot of the absolute elevation LRE plotted against loudspeaker and response method.
Loudspeaker 1 is the only loudspeaker for which the verbal response method resulted in a significantly higher LRA than the laser-guided response method. The pointing response method gave a significantly higher elevation LRA than the verbal and graphical response methods for loudspeakers 2, 4, 7, 12, 13 and 15. This suggests that, in general, the elevation LRA was higher for the pointing response method.

To more easily observe the variation in elevation LRA with loudspeaker, Figure 32 plots the absolute elevation LRE plotted against loudspeaker number and separated by response method.
Figure 32 - Absolute elevation LRE plotted against loudspeaker number and separated by response method.

This plot clearly displays large changes elevation LRA with loudspeaker location for the verbal method. For verbal responses, loudspeakers 1, 9 and 17 are all located near 0° elevation and have a high elevation LRA. Loudspeaker 18 has the lowest mean absolute elevation LRE, and so the highest LRA, for this method and its actual location is 90° elevation. It is suggested that perceived source locations tend to cluster around polar locations, and so result in a high LRA, because listeners prefer common 'anchor' angles. For example, if listeners only had an angle resolution of 90°, so could only report sources at 0° and 90°, sources actually located at these points would have a very high elevation LRA for the verbal response method whereas sources in-between would have a significantly reduced elevation LRA. The LRA bias caused by these anchor locations is discussed further in the following 'loudspeaker' section.

In general the laser pointing response methods resulted in a higher response
consistency across location than either the graphical or verbal response methods. All studies that use verbal elicitation, such as Wightman and Kistler [1999], must consider that the listener LRA may not be constant across loudspeaker location.

Loudspeaker 18 resulted in a much higher variance in listener response for both the movement and no movement laser pointing condition, indicated by the larger confidence intervals. This is the only loudspeaker location for which the variance of the movement condition exceeds both the graphical and verbal response methods. During the reported location calculation this loudspeaker location required a correction of 5.7° for the laser pointing response method, approximately five times the average correction angle. The correction angle is included in the calculation of the loudspeaker location to allow for differences between the physical location measured with a tape measure and location of the loudspeaker with respect to the headtracker and pointing response method. The large correction angle suggests that the headtracked pointing response method struggled to report angles of elevation close to 90° and so the error for loudspeaker 18 is due to the response method system rather than a lower listener LRA. It is advised that for further research using the laser pointing method loudspeaker 18 is removed.

**Response Method**

The analysis of the secondary interactions response method*listener and response method*loudspeaker suggested that the laser pointing response method gave the highest elevation LRA and was most consistent across loudspeaker location. The Kruskal-Wallis test, which showed that the movement laser pointing method had the lowest mean rank. The mean ranks of the verbal and graphical methods were very similar. These findings suggest that the laser pointing response method was the most accurate response method. To confirm these findings post-hoc Bonferroni and Tukey-B tests were conducted following the ANOVA (Table 19).
Localisation Response Methods

The tests showed that there were significant differences between all conditions except between the Verbal-Graphical methods. This result confirmed two findings:

- the laser pointing response method resulted in a significantly higher elevation LRA than either the verbal and graphical response methods.
- The ability to move one's head during the playback of the programme item significantly improved the listeners' LRAs using the laser pointing response method.

The degradation in localisation performance caused by the listener being unable to move their head was not as large as the degradation caused by use of the other two pointing methods.

Using verbal and graphical response methods, it is suggested that listener training could significantly reduce elevation LRE. The raw data from the verbal response method shows that in only 8 of the 540 trials did listeners respond with an angle resolution of less than 5°. This suggests that listeners had already decided that their response resolution would be limited. Training may also allow listeners to better perceive their own location in space. There were a number of instances in both graphical and verbal response where listeners responded with an elevation of -90° (a source directly below them). It should have been obvious to the listener that no source could originate from directly below them. Whether this elevation response was actually what they perceived or a misunderstanding of the response method is unclear.

Figure 33 shows the mean absolute elevation LRE and 95% confidence intervals for each response method.

Table 19 - A Tukey-B post-hoc test conducted on the response method factor

<table>
<thead>
<tr>
<th>Response Method</th>
<th>N</th>
<th>Subset</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Tukey B&lt;sup&gt;a,b&lt;/sup&gt; Movement</td>
<td>540</td>
<td>7.2630</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Movement</td>
<td>540</td>
<td>8.5344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graphical</td>
<td>540</td>
<td>11.2004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal</td>
<td>540</td>
<td>12.1204</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Means for groups in homogeneous subsets are displayed.

Based on observed means.

The error term is Mean Square(Error) = 58.550.

Uses Harmonic Mean Sample Size = 540.000.

<sup>a</sup> Alpha =

<sup>b</sup>
The two laser pointing methods have a smaller mean absolute response error than either the verbal or graphical response methods. The confidence intervals for the verbal response are largest, showing that elevation LRA varied most between trials. This suggests that either some listeners were better at using the verbal method than others or some locations were easier to locate, as shown by the loudspeaker and listener results sections. Using a laser pointing localisation response method results in a mean reduction in absolute elevation LRE of approximately $4.5^\circ$ when compared to the verbal and graphical response methods.

**Loudspeaker**

The loudspeaker factor was shown to be significant ($F(17, 2160) = 5.006, \ p > 0.001, \ \eta_p^2 = 0.263$). A Kruskal-Wallis test conducted on the absolute elevation LRE with the loudspeaker factor confirmed this finding ($H(17) = 168.243, \ P <$
0.001). Figure 34 shows the absolute elevation LRE separated by loudspeaker number.

![Figure 34 - Absolute Elevation LRE plotted against loudspeaker number](image)

This plot again shows the ease of response for the anchor locations at 0° elevation (loudspeakers 1, 9 and 17). This effect was shown to only significantly affect the verbal and graphical response methods. Since the subsequent experiments described in this thesis use a laser pointing response method investigation into this effect is not necessary, however, it could be an area of further work. To test whether these anchor locations affected the LRA using the verbal and graphical response methods, it would be necessary to repeat the experiment but rotate and tilt the truncated isocahedron slightly so its loudspeakers did not line up with these polar points. By observing how this rotation affects the LRA it would be possible to find the effect of anchor locations on the elevation LRA.
7.3.2 Azimuth LRA

Each loudspeaker was positioned in a known location in both azimuth and elevation with reference to the central listening position. This section studies the azimuth component of the LRA.

Speaker 18 was removed from all azimuth analysis because, being located at 90° elevation, its azimuth angle is meaningless. Changes of a few degrees in elevation can produce significant changes in the azimuth angle, causing azimuth LRE to be skewed.

**Front-Back Errors**

Before any analysis on the absolute azimuth LRE can be conducted, the front-back errors should be removed from the reported location for each response method. Front-back errors were removed from the azimuth LRE data by calculating $180^\circ - \varphi$ ($\varphi$ = reported azimuth location) for all data exhibiting front-back reversal (azimuth reported in the wrong hemisphere). For the source location at precisely 90° azimuth to the listener the localisation response was left unchanged. Front-back errors are discussed first, followed by analysis of the absolute azimuth LRE data. A pie chart showing the proportion of front-back errors for each response method is shown in Figure 35.
The largest proportion of front-back errors is shown in no movement laser pointing response method. When compared to the free movement laser pointing response method, it can be concluded that being free to move the head significantly reduces the listeners’ front-back error rates. This conclusion is in concurrence with the findings of Wightman and Kistler [1999].

Comparing the three localisation response methods, the free movement laser pointing response method has a smaller proportion of front-back errors than either the verbal or the graphical response methods. It may be that large errors of response that were not caused by front-back errors may have resulted in a front-back confusion correction for the verbal and graphical response methods. Alternatively listeners may have been more inclined to use head movements in

---

**Figure 35 - Proportion of front-back errors for each response method**
the laser pointing response method because they were required to point their heads. These head movements would have reduced the proportion of front-back errors. Either way, the laser-pointing localisation response method resulted in the highest LRA.

Once the front-back errors were removed, the verbal response method still showed a large number of significant errors between the reported and actual loudspeaker locations (>90°). For these responses, the listeners appear to have reported sources in the left hemisphere, even though all sources were located in the right. Left / right confusions are not common in azimuth localisation tests and none were shown for the other response methods. Therefore, the most likely conclusion is that these errors stem from a lack of understanding of the coordinates system used for the verbal response method. Listener training is likely to improve the LRA using the verbal response method.

**Absolute Azimuth LRE**

The absolute azimuth LRE was calculated as the absolute difference in azimuth angle between the reported and actual location of the source. Absolute azimuth LRE gives an indication of the LRA. If a listener gives consistently large errors then it is likely that they are uncertain of the source location in that condition.

A Kolmogorov–Smirnov test found that 28 of the 68 conditions were normally distributed. As a result both ANOVA and Kruskal-Wallis tests were conducted on the azimuth LRE data.

The dependent variable ‘absolute elevation LRE’ was modelled using an ANOVA, with fixed factors:

- *Response Method*
- *Loudspeaker Number*

and random factor:

- *Listener Number*

The results of this analysis are shown in Table 20.
The analysis shows that all single factors, except listener number, had a significant effect upon the absolute azimuth LRE. Response method (F (3, 2040) = 25.537, p < 0.001, \( \eta^2 = 0.646 \)) has a high F value, which shows that it had a strong effect on the absolute azimuth LRE. The partial eta squared value is also largest for the response method indicating that it had the largest effect of all the factors. Loudspeaker number was also found to be significant (F (16, 2040) = 10.693, p < 0.001, \( \eta^2 = 0.433 \)), showing that azimuth LRA varied significantly between loudspeakers.

As with the elevation LRE, there is a significant interaction between loudspeaker number*response method (F (48, 2040) = 6.527, p < 0.001, \( \eta^2 = 0.156 \)). This indicates that the azimuth LRA of the listeners to each loudspeaker location varied depending on the response method used. Furthermore the secondary interaction response method*listener number was also shown to be significant (F (42, 2040) = 3.628, p < 0.001, \( \eta^2 = 0.083 \)). These secondary interactions will be investigated in the following sections.

The primary factor, listener number, was not found to be significant, showing that absolute azimuth LRE did not vary significantly between listeners and so the listeners’ responses were similar.
A histogram of the ANOVA model's standardised residuals showed a normal distribution (Figure 36), indicating that the ANOVA model provided a good fit to the data.

**Figure 36 - Histogram of standardised residuals given by ANOVA**

**Response Method*Loudspeaker**

The main aim of this chapter is to find the localisation response method that gives the highest and most consistent three-dimensional LRA. To discover this, it is necessary to study the effect the response method factor had on the azimuth LRA. The response method factor was shown to be significant (F (3, 2040) = 25.537, \( p < 0.001, \eta^2 = 0.646 \)) in the ANOVA. However, the secondary interaction response method*loudspeaker (F (48, 2040) = 6.527, \( p < 0.001, \eta^2 = 0.156 \)) was also shown to be significant and this must be investigated first.
Figure 37 shows the absolute azimuth LRE separated by loudspeaker number and response method.

For loudspeakers 1 and 17, the verbal and graphical response methods resulted in a significantly higher azimuth LRA than the laser pointing response methods. For all other loudspeaker locations, except loudspeakers 5 and 9, the laser pointing response method resulted in a significantly higher azimuth LRA than either the graphical or verbal response methods. Loudspeakers 1 (0°), 9 (90°) and 17 (180°) were all located at ‘anchor’ locations as discussed in the elevation response chapter. These anchor locations caused the verbal and graphical response methods to have a lower mean absolute azimuth LRE overall.
To view the consistency of the response methods across loudspeaker location, the absolute azimuth LRE was plotted against loudspeaker location and grouped based on response method (Figure 38).

![Figure 38 - Absolute elevation LRE plotted against loudspeaker number for each response method group](image)

The laser pointing response method is more consistent across loudspeaker location than both the verbal and graphical response methods. Verbal and graphical response methods show a much higher LRA for anchor locations for loudspeakers 1, 9 and 17, as discussed previously.

For the no movement condition, the azimuth LRA is reduced as the source is moved behind the listener (as the loudspeaker number is increased). Behind the listener the no movement condition has a much larger variance than the movement condition. This suggests the when the listener is unable to move their head, reporting the location of sources behind them becomes much more difficult. This trend is not apparent in the graphical response method, which
indicates that the listeners were making use of head movements the improve their LRA for rearward sources.

Comparing the movement and no movement response methods, shows two main differences:

1) *In the no movement condition, as the source is moved further in azimuth from the median plane, the azimuth LRA is reduced.*

2) *In the no movement condition there are a number of front-back confusions, while in the movement condition there are none.*

Difference 1 is further evidence that azimuth LRA becomes less accurate as the source moves further from their field of vision. There are three possible reasons for the degradation in azimuth LRA performance. Firstly, when the listener is unable to move their head they are unable to move the source into their most accurate area of localisation. Pinna cues for rearward sources are likely to be less accurate due to its shape and previous studies have suggested that interaural differences are most accurate for frontal sources. In the head movement condition, ability to move the head effectively allows any source to be directly in front of the listener. Secondly, by moving their head, listeners may be able to use dynamic interaural difference cues as hypothesized by Wallach [1938]. Thirdly, the laser pointing method is based on combining visual (where the laser is) and auditory cues. If the sound source is behind the listener then they cannot see it, which means they have to guess the visual part of the pointing method after the sound has finished.

**Response Method*Listener**

The ANOVA showed the secondary interaction response method*listener number to be significant (*F* (42, 2040) = 3.628, *p* < 0.001, *ηp^2^ = 0.083). Figure 39 shows a plot of the absolute azimuth LRE plotted against listener number and response method.
The plot shows that 12 of the 15 listeners had a significantly higher azimuth LRA when using the laser-pointing response method. For no listeners did the LRA of either the verbal or graphical methods significantly exceed the LRA of the laser-pointing response method. Listener 1 showed no significant differences in LRA between any of the response methods. This appears to be due to a significant reduction in LRA using the laser-pointing response method. It may be that this listener was limited by their actual azimuth localisation acuity and this lower accuracy limited the effect of the response method factor.

These findings appear to show that the laser-pointing response method resulted in the highest azimuth LRA. To confirm this, the primary factor ‘response method’ was investigated.
Response Method

The primary factor ‘response method’ was shown to be significant (F (3, 2040) = 25.537, p < 0.001, ηp² = 0.646) in the ANOVA. A non-parametric Kruskal-Wallis Test was conducted to verify the findings of the ANOVA for the response method factor. The test was conducted on the absolute elevation LRE with the response method factor and the results are shown in Table 21.

<table>
<thead>
<tr>
<th>Response Method</th>
<th>N</th>
<th>Mean Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abs_azi_err_fb1</td>
<td>510</td>
<td>713.77</td>
</tr>
<tr>
<td>Movement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Movement</td>
<td>510</td>
<td>1093.10</td>
</tr>
<tr>
<td>Verbal</td>
<td>510</td>
<td>1145.27</td>
</tr>
<tr>
<td>Graphical</td>
<td>510</td>
<td>1129.85</td>
</tr>
<tr>
<td>Total</td>
<td>2040</td>
<td></td>
</tr>
</tbody>
</table>

Table 21 - Results of Kruskal Wallis Test with test variable Absolute Azimuth LRE and fixed factor Response Method

The results of the Kruskal-Wallis test show that the effect of the ‘response method’ was significant (H (3) = 186.495, P < 0.001), which verifies the significance finding of the ANOVA. Once again, the movement laser pointing method has the lowest mean rank, which suggests that it is the most accurate response method. The mean ranks of the verbal and graphical methods are very similar.

The same trend is shown in the plot of mean absolute azimuth LRE and 95% confidence intervals for each response method in Figure 40.
The laser pointing response method is significantly more accurate than either the verbal or graphical response methods. The mean absolute azimuth LRE for the laser pointing response method is more than half the other two methods. Bonferroni and Tukey-B post-hoc tests both confirmed that the laser pointing response method was significantly different from the verbal and graphical response methods (Table 22).

### Table 22 - Tukey-B post-hoc test on absolute azimuth LRE with response method

<table>
<thead>
<tr>
<th>Response_Method</th>
<th>N</th>
<th>Subset 1</th>
<th>Subset 2</th>
<th>Subset 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Movement</td>
<td>510</td>
<td>3.1463</td>
<td>7.6727</td>
<td>8.7920</td>
</tr>
<tr>
<td>Graphical</td>
<td>510</td>
<td>7.7112</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Movement</td>
<td>510</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal</td>
<td>510</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
All of these findings indicate that the Movement method was the most accurate response method. The graphical and verbal response methods all resulted in a significantly higher mean absolute azimuth LRE than the free movement condition.

The no movement method gave a significantly higher mean absolute azimuth LRE than the free movement condition. When the listener was unable to move their head, the azimuth LRA was significantly degraded.

### 7.4 Conclusions

The two goals of this study were:

- To find the most accurate localisation response method
- To measure how the LRA varied when the listener was able move their head.

In this section these two goals will be addressed separately and any further findings will be discussed.

#### 7.4.1 Goal 1 – Response Method Accuracy

The response method significantly affected the listener’s ability to accurately report the location they perceived. The laser pointing method was shown to give a higher LRA than the verbal and graphical methods for both azimuth and elevation localisation. Mean absolute LRE in elevation and azimuth for each response method are summarised in Table 23 (only free movement responses are considered).
### Table 23 - Comparison of Azimuth and Elevation LRE for each response method

<table>
<thead>
<tr>
<th>Response Method</th>
<th>Mean Azimuth LRE (°)</th>
<th>Std. Deviation</th>
<th>Mean Elevation LRE (°)</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Movement</td>
<td>3.14</td>
<td>2.63</td>
<td>7.26</td>
<td>7.38</td>
</tr>
<tr>
<td>Verbal</td>
<td>8.79</td>
<td>9.71</td>
<td>12.12</td>
<td>13.18</td>
</tr>
<tr>
<td>Graphical</td>
<td>7.67</td>
<td>6.45</td>
<td>11.20</td>
<td>9.09</td>
</tr>
</tbody>
</table>

It can be seen that elevation LRE was higher than azimuth LRE for all methods. Better listener localisation consistency in azimuth is indicated by the significance of listener number in the ANOVA of the absolute elevation LRE and non-significance of listener number in the ANOVA of the absolute azimuth LRE. Listeners appear to find elevation localisation more challenging than azimuth localisation and the ability to localise in elevation varies between listeners.

It has been shown that the response method significantly affects the listeners’ abilities to accurately report a location and so localisation studies that use differing response methods must be compared with caution. This is especially true for studies that seek to define the absolute acuity of localisation. If the error of response method is higher than the acuity of localisation, then it is impossible to measure the acuity.

It was suggested that verbal and graphical response methods could be improved by using a longer training stage in which some feedback was given for response angles. However, this training would be time intensive and may bias the listener’s perception by not representing their most natural response. An optimal method of response should be intuitive and should ideally not take intense training.
7.4.2 Goal 2 – Head Movement LRA

Mean absolute LREs in elevation and azimuth for the laser point response method with and without head movements are compared in Table 24.

<table>
<thead>
<tr>
<th>Response Method</th>
<th>Mean Azimuth LRE (°)</th>
<th>Std. Deviation</th>
<th>Mean Elevation LRE (°)</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Movement</td>
<td>3.14</td>
<td>2.63</td>
<td>7.26</td>
<td>7.38</td>
</tr>
<tr>
<td>No Movement</td>
<td>7.71</td>
<td>8.33</td>
<td>8.53</td>
<td>7.76</td>
</tr>
</tbody>
</table>

Table 24 - Comparison of Azimuth and Elevation LRE for movement and no movement conditions

The ability of a listeners to move their heads while the sound source was playing significantly improved their LRAs in azimuth and elevation.

One possible reason for this improvement in azimuth LRA could be that when the listeners could move their heads, they orientated to face the sound source. This meant that the source was presented to them in their area of highest azimuth localisation acuity [Pollack and Rose 1967]. Whether directly in front of a listener is the area of highest elevation localisation acuity is not shown in previous research. Chapters 11 and 12 show how elevation LRA varies on vertical planes of differing azimuth.

A second possible reason for the improved azimuth LRA is that the pointing of the laser is an interactive process between the visual and auditory system. If the listener is unable to move their head while the source is playing then they must guess the location and point the laser at it following the test. This could result in a less accurate laser point.

The ability of a listener to move their head while the source was playing significantly increased the elevation LRA. The mean absolute elevation LRE was significantly higher for the no movement than free movement condition. This is either due to dynamic interaural difference cues given by the head movement or
the head movement improving the quality of static pinna cues. Is the cause of this improvement in elevation LRA better static pinna cues or dynamic interaural cues created by head movement? This question is answered in Chapters 8, 11 and 12.

A higher elevation and azimuth LRA occurs when the listener is able to move their head. Comparisons of movement and no movement head localisation should be made with caution, as it is unclear whether the improvements are due to a dynamic cue created by the movement (such as the dynamic interaural difference hypothesised by Wallach) or simply a factor of the more accurate pointing as a result of the listener being able to move their head to point at the source. It was also shown that front-back confusions were significantly reduced when the listener was able to move their head.

In both azimuth and elevation, the LRA of the verbal and graphical response methods varied significantly across locations. Sources close to ‘anchors’ were much more accurately reported than intermediate angles. It is suggested that an optimal response method should be as consistent across location as possible. The laser pointing method was the most consistent across azimuth and elevation when head movement was permitted.

Speaker 18 was shown to be the least accurately reported loudspeaker location for the laser pointing method. It was concluded that this error was due to the response method not the listener’s localisation accuracy and so it was suggested that loudspeaker 18 should be removed from further laser pointing method experiments.

Azimuth LRA was shown to be more accurate for sources in front of the listener than for those behind the listener using the no movement laser-pointing response method. It was shown, through analysis of the perceived location, that sources outside of the listener’s field of vision were less accurately localised in the no movement condition than in the movement condition. There were three possible reasons for this change in LRA: head movement moved the source into the listener’s most accurate area of audition; the addition of dynamic interaural cues via head movement improved LRA; the combination of visual and auditory
cues was more effective when head movement were allowed. This finding has implications in the further studies into head movement localisation described in the subsequent chapters. By analysing whether the field of vision has any implication on the LRA, it may be possible to isolate the effect of visual cues on the effect of head movements on localisation; this analysis is conducted on the experimental data in Chapter 8.

Summary

A laser pointing response method was found to give the highest LRA when reporting both azimuth and elevation in this study. Therefore, a laser pointing response method will be used in all the subsequent thesis experiments.

The ability of the listener to move their head while the stimulus played resulted in higher LRA in both azimuth and elevation. It is unclear whether this improvement was due to the cues produced when the listener was moving their head or the more accurate cues present when the listener had moved their head into their best area of audition. Chapter 8 will further investigate the effect of head movements on LRA, while Chapter 11 will try to deduce whether improvements in LRA with head movements are due to static or dynamic cues.
8 The Effect of Head Movements on LRA

The goal of this chapter is to find out how head movements affect a listener's LRA. It has been shown that a listener will not remain motionless while auditioning a sound; instead they will dynamically respond to the source by moving their head. The review of listener head movements in Chapter 2 concluded that yaw movements are the largest and most common head movements. Previous studies give contradicting evidence as to the effect these head movements have on the LRA.

Head movements can be used to resolve front-back confusions [Wightman and Kistler 1999] but whether there exists a further head movement localisation cue that can improve elevation LRA is unverified. Wallach [1938] proposed that dynamic interaural differences created by head movement allow the accurate localisation of both source azimuth and elevation. Ashby et al. [2010] showed that a sphere model of the head could be used to resolve source azimuth and elevation by intersecting the cones of confusion given by multiple microphone pairs. Perrett and Noble [1997] and Wightman and Kistler [1999] came to differing conclusions as to the importance of head movements in localisation, but Chapters 5 and 6 showed that both studies limited their applicability through restricted experimental conditions. Perrett and Noble only studied elevation LRA on the median and left lateral planes, which may have skewed their experimental findings, while Wightman and Kistler did not limit pinna cues and used an inaccurate verbal localisation response method. The presence of pinna cues appeared to significantly affect the impact that head movements had on the elevation LRA [Perrett and Noble 1997]; however, a three-dimensional study of these effects has not been completed. Evans [1998] described how verbal response methods were unintuitive and may result in inaccurate responses. The localisation response method study described in Chapter 7, which compared verbal, graphical and pointing responses, showed that verbal responses were the least accurate, and were prone to biases.

The study described in this chapter was designed to answer three research questions:
• How does head movement affect elevation and azimuth LRA?
• How does the presence of pinna cues affect elevation and azimuth LRA using head movement?
• If LRA is improved by head movement, what factor is the cause of this improvement?

Previous head movement localisation studies described in the preceding literature review have suggested that the final question has three possible answers:

• Dynamic localisation cues - The changing cues created during head movement.
• Improved static localisation cues - The cues created once a listener has moved their head to a new static position.
• Increased pointing accuracy - Changes in LRA due to the integration of the listener's visual and auditory cues.

A localisation experiment was designed that would answer these questions whilst avoiding the limitations of previous head movement experiments. Section 8.1 of this report describes the main experiment setup, which includes loudspeaker placement, programme items, response method, listeners and movement conditions. The results of the experiment are separated into elevation LRA (Section 8.2.2) and azimuth LRA (Section 8.2.3) and the experimental conclusions are given in Section 8.3.

8.1 Experimental Setup and Procedure

The following localisation experiment will investigate the effect of head movements on a listener’s elevation and azimuth LRA, both with and without the presence of pinna cues. This section will detail the setup of the listening environment, loudspeakers, software, programme items and movement conditions and give an overview of the overall test procedure.
8.1.1 Listening Room

The experiment took place in the live room of Studio 2 in the Institute of Sound Recording at the University of Surrey. The room measures 6.55×8.78×4.02m and is a pop studio by design. Although it is not anechoic it has a short RT60 of 235 ms [Coleman et al. 2014], which means the studio can be described as fairly acoustically dead. If the room was excessively reverberant then the localisation cues could be masked or confused. The test required localisation cues to be as distinct and precise as possible, so that any differences due to the test variables could be observed. If the experiment were conducted in an anechoic chamber then the results may not be generalisable to more natural listening conditions.

8.1.2 Loudspeaker Positions

In the experiment, seventeen Genelec 8020b loudspeakers were positioned on a hemisphere to the right of the listener, spread from 0° to 180° in azimuth and −35° to +55° in elevation. Since it has been shown that localisation is unchanged between hemispheres [Oldfield and Parker 1994], loudspeakers were placed only on the right hemisphere.

The loudspeakers were not confined to a single plane as this may bias the localisation test [Makousa and Middlebrook 1990]. Chapter 3 showed that source location significantly affects the LRA; therefore testing a wide range of source locations randomises this variable. The loudspeakers were positioned at the coordinates of the vertices of a truncated icosahedron (Figure 41). Although this shape has 60 vertices, only loudspeaker elevations from -35° to +53° in one hemisphere were used, which reduced the required number of loudspeakers to seventeen. Using a regular shape allowed sources to evenly span locations in azimuth and elevation. The loudspeaker numbers, coordinates and angles used in the experiment are given in Table 40 in Appendix B.
In previous experiments there was an eighteenth loudspeaker located at +90° in elevation. It was shown that this loudspeaker location made pointing responses difficult for the listener and produced inconsistent results. For these reasons, this loudspeaker (previously numbered 18) was excluded from this study.

The loudspeakers were oriented to face the listener so that the frequency response of the loudspeaker did not vary with azimuth and elevation. To avoid any biases due to visual cues, the locations of the loudspeakers were hidden from view by an acoustically transparent but visually opaque curtain.

### 8.1.3 Technical Equipment and Software

The loudspeakers were connected to an RME Fireface 800 audio interface and two 8-channel Presonus Digimax D8 interfaces connected by ADAT to the Fireface to create a total of 24 output channels. A Macbook Pro laptop computer running MaxMSP through the RME Fireface proprietary interface software was
used to play back the sounds. The sound samples were created in MATLAB with a sample rate of 48 kHz and a bit depth of 24 bits.

8.1.4 Programme Item

Pinna cues were suppressed in order to more clearly separate the effects of head movement cues. Chapter 6 considered a number of pinna cue suppression methods, namely pinna occlusion, virtualization and sound source manipulation. Occlusion of the pinna can damage the ear canal and creating pinna moulds can be time consuming. Using virtualization it is difficult to create an externalized source location, especially if spectral variations are removed. Manipulation of the sound source signal was identified as the most practicable method to suppress pinna cues.

In the experiment, three programme items were used: Broadband white noise (N); 2 kHz low pass filtered noise (L); and 2 kHz low pass filtered noise with an additional 5.7-8 kHz bandpass filtered component (B).

The broadband white noise (N) exhibited a flat frequency spectrum up to approximately 20 kHz that allowed pinna cues, which are heavily dependent on frequencies above 4 kHz [Algazi et al. 2004; Perrett and Noble 1997], to be used. The L programme item was intended to suppress the high frequency pinna cues, albeit also suppressing other cues that may be available in the filtered frequency range, such as high frequency dynamic ILD cues created through head movement.

The B programme item was created based on the research of Langendijk and Bronkhorst [2002], who showed that the removal of pinna cues in a half-octave band from a broadband noise signal does not reduce elevation LRA. Therefore, it was hypothesized that the inclusion of a half-octave band of noise above 4 kHz would not enable the listener to use pinna cues, but may be useful for other cues such as dynamic ILDs. In the pilot experiments described in Chapter 6, the B programme item was shown to degrade elevation LRA to a similar level to the L programme item. The B programme allowed the investigation of head movement cues at high frequencies with reduced pinna cues.
Wallach [1938] suggested that dynamic interaural differences given by head movement could be used to localise in azimuth and elevation without the need for pinna cues. Ashby [2010] showed that this was possible using a simple sphere model of the head. The L programme item will test whether listeners use dynamic ITDs to localise effectively in elevation. The B programme item will test whether the addition of ILDs without pinna cues can improve elevation LRA.

8.1.5 Head Movement Conditions

While the listener auditioned the stimuli they were instructed to move their head in one of three ways:

- **Free movement** – *The listener could move their head in any way they wished*

- **No Movement** – *The listener was not permitted to move the head while the sound played.*

- **Forced Movement** – *The listener rotated their head in a regular periodic fashion between ±30° in azimuth.*

The no head movement condition was enforced using a Polhemus head tracker, which measured the listener's head movement with six degrees of freedom, and MaxMSP code. The code muted the stimulus if the listener moved their head beyond ±2.5°, half the value used by Simon *et al.* (5°) [2010] and less than the value chosen by Thurlow and Runge (3°) [1967]. Listeners were also explicitly instructed not to move their head during the trial.

In the forced movement condition each listener was instructed to move their head between two fixed markers placed at ±30° in azimuth and 0° in elevation and that the movement should be conducted in a regular periodic fashion. This resulted in a yaw rotation on the equatorial plane which, as shown in Chapter 2, is the most common head rotation. A speed of between 2 and 6 oscillations during the stimulus length was suggested. The precise speed of oscillation was left to the listener's discretion with the stipulation that it should allow the listener to gain what they perceived as their highest localisation accuracy. The test automatically muted the stimulus if the listener moved their head more than ±35° in azimuth and ±10° in elevation. Listeners were observed to verify that the
instructions were being followed. This forced condition prevented the listener from turning all the way round to make a response for a side/rear stimulus whilst the stimulus was sounding, which removed any improved pointing response for rearward sources.

The forced movement condition allowed dynamic cues resulting from a moving head to be distinguished from those available once the head reached a new static head position, such as improved interaural, pinna or visual cues (improved pointing accuracy using the laser pointing response method). Furthermore, this condition tested whether yaw head movements are the main cause of improved LRA due to head movements.

In the free movement condition listeners were allowed to move their head in any way they wished. It was anticipated that listeners would orient themselves to face the source, as this was the movement most common in previous studies (Chapter 2). Whether this is the most effective movement for the listener to make is unverified in previous research, although it has been shown that facing the source produces the most accurate azimuth localisation [Stevens and Newman 1936].

8.1.6 Experiment Procedure

The listener was asked to report the perceived location of each of a series of signals, each replayed via a single selected loudspeaker. The listener responded using a laser pointing method; this was shown to be the most accurate and consistent localisation response method for a three-dimensional localisation test (Chapter 7).

The laser was attached to the head and positioned to point out from between the eyes via an elasticated band that was wrapped around the listener’s head. A Polhemus head tracker monitored the listener’s head movement to six degrees of freedom and each perceived source location, as indicated by the listener, was calculated using a triangulation technique similar to the method described by Frank et al. [8] and described in Appendix A. The loudspeaker radius from the central listening position was used to create a projected sphere and the
perceived location was the point at which the line of the laser pointer pierced this sphere.

Sixteen listeners auditioned all programme items and movement conditions. All of the listeners were aged between 18 and 35, and had no reported hearing problems. The experiment was split into three sections according to movement condition and each listener conducted all three sections in a single forty-minute session. Each section consisted of 51 trials; programme item and loudspeaker were randomly varied in each section. The movement condition order was randomly varied between listeners and always reversed on repeat.

Listeners could either use the on-screen ‘start’, ‘replay’ and ‘log’ click buttons or keyboard commands to control the playback of the trials. Z was used the play the stimulus, X was used if the listener wished to replay the stimulus and spacebar was used to log the perceived source location once the laser was pointing at the perceived source location. This control method was also used for the experiments in Chapters 9 and 10.

A calibration stage was included to establish the relationship between the laser pointer and the head tracker. The listener was instructed to direct the laser at two points, one at 0° azimuth and +5.5° elevation, and one at 180° azimuth and +5.6° elevation. The calibration was conducted for both locations. The program gave a three second verbal countdown before calibration was carried out. Calibration logged the listener’s head position and orientation with 6 degrees of freedom, averaged over a one second time period. If the listener moved their head more than 2.5° in orientation or 2.5 inches in location during this time then they would be instructed to repeat the procedure. The calibration stage allowed the relationship between the laser pointer and the head tracker to be established and was used in the calculation described in Appendix A.

For each trial the following set of data was logged: Overall Trial Number; Listener Number; Movement Condition; Programme Item; Loudspeaker Number; Tracker Location; Tracker Orientation. The tracker location and orientation were used to calculate the reported and actual loudspeaker locations.
The main metric used to measure LRA was the absolute LRE. This was calculated as:

$$\phi_{ae} = \sqrt{\left(\phi_{reported} - \phi_{actual}\right)^2} \quad (1.)$$

where $\phi_{ae}$ is the absolute LRE, $\phi_{perceived}$ is the reported loudspeaker location angle and $\phi_{actual}$ is the actual loudspeaker location angle.

### 8.2 Results

Section 8.2.1 describes the listener screening that was conducted before the results analysis. The results for elevation and azimuth localisation are discussed in Sections 8.2.2 and 8.2.3 respectively.

#### 8.2.1 Listener Screening

The process of listener screening is detailed in this section. If a listener’s trial data was shown to be erroneous through user error or equipment malfunction then the data was removed from the analysis. This was determined by comparing their absolute response error with other test subjects and observing scatter distributions of responses.

Figure 42 shows the mean and 95% confidence intervals for absolute azimuth LRE separated by listener number. The absolute azimuth LRE was the absolute difference in degrees between the perceived and actual location of the source.
The mean absolute azimuth LRE of Listener 6 is more than three times the overall mean absolute azimuth LRE, while their standard deviation is twice the mean deviation. This shows that Listener 6 was significantly less accurate at localising in azimuth than the other listeners. Furthermore, reviewing the evidence presented in Chapter 3, an absolute azimuth LRE exceeding 30° is unusually high.

Figure 43 and Figure 44 show scatter graphs of the perceived azimuth location against actual azimuth location for the N programme item in the free movement condition, both with and without Listener 6. This combination of variables should have enabled high azimuth LRA due to the many localisation cues created by the programme item. In the plot, high azimuth LRA will be illustrated by a linear relationship between the perceived and actual source location.
The Effect of Head Movements on LRA

Figure 43 - Scatter graph of perceived against actual location for all listeners \((r = 0.973)\)

Figure 44 - Scatter graph of perceived against actual location with Listener 6 removed \((r = 0.997)\)
It can be seen that by removing Listener 6 the majority of erroneous localisation points are removed. A Pearson Correlation can be seen to increase from 0.973 to 0.997.

All erroneous locations for Listener 6 are biased toward loudspeaker 1 at -90° (the listener's starting location for each trial). It is suggested that Listener 6 may have pressed the log location button as they returned to the forward facing position, not when the laser was pointed at the perceived loudspeaker location. This would explain both the forward bias and the erroneous azimuth localisation.

If the listener was incorrectly interpreting static interaural difference cues then it is likely that they were also unable to interpret dynamic interaural cues. Therefore, their inclusion in the results analysis would have increased the noise in the data thereby potentially causing the significant differences to go unnoticed.

Figure 45 shows the mean and 95% confidence intervals for absolute elevation separated by listener number.
Although there is variation in elevation LRA between listeners shown by this plot, this was not deemed large enough to warrant listener removal. Listeners 4 and 6 did exhibit a larger overall elevation LRE than the other listeners. Larger variation in elevation localisation response than azimuth response was expected, as some listeners are better at using spectral cues than others. This trend was shown in the localisation response methods experiment in Chapter 7.

In summary, Listener 6 was removed from all further analysis due to their poor azimuth LRA. A mis-pointing of the laser perhaps combined with an inability to interpret interaural cues may have caused these erroneous responses. Both scenarios suggest that the listener was unable to complete the experimental task.
8.2.2 Elevation LRA

This section describes the analysis conducted on the elevation LRA. The goal of this analysis is to find the effect of programme item and movement condition on the elevation LRA.

8.2.2.1 Normality Checks

To use parametric tests such as ANOVA the data must be normally distributed. The absolute elevation LRE data was checked for normality using a Kolmogorov-Smirnov test and was found to be normal for 147 of the 153 conditions. Each normality check contained 31 data points averaged across the ‘listener’ and ‘repeat’ variables.

8.2.2.2 ANOVA

An ANOVA was conducted on the absolute elevation LRE data with the fixed factors ‘Movement Condition’, ‘Loudspeaker’, ‘Programme Item’, ‘Repeat’ and the random factor ‘Listener’. All interactions up to three-way were modelled. The results of the ANOVA are shown in Table 25.
### Tests of Between-Subjects Effects

**Dependent Variable: abs_ele_error**

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>Hypothesis</td>
<td>658558.087</td>
<td>1</td>
<td>658558.087</td>
<td>200.058</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>46146.262</td>
<td>14</td>
<td>0.018</td>
<td>3291.837</td>
<td>.000</td>
</tr>
<tr>
<td>Speaker</td>
<td>Hypothesis</td>
<td>50238.310</td>
<td>16</td>
<td>3139.894</td>
<td>6.986</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>101216.286</td>
<td>225</td>
<td>2.122</td>
<td>449.422</td>
<td>.000</td>
</tr>
<tr>
<td>Programme_Item</td>
<td>Hypothesis</td>
<td>43545.436</td>
<td>2</td>
<td>21772.718</td>
<td>41.991</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>14558.495</td>
<td>28</td>
<td>0.078</td>
<td>518.505</td>
<td>.000</td>
</tr>
<tr>
<td>Movement_CONDITION</td>
<td>Hypothesis</td>
<td>8959.693</td>
<td>2</td>
<td>4479.847</td>
<td>16.172</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>7866.273</td>
<td>28</td>
<td>0.396</td>
<td>277.012</td>
<td>.000</td>
</tr>
<tr>
<td>Repeat</td>
<td>Hypothesis</td>
<td>988.677</td>
<td>1</td>
<td>988.677</td>
<td>1.693</td>
<td>.216</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>7589.607</td>
<td>13</td>
<td>583.816</td>
<td>135.206</td>
<td>.000</td>
</tr>
<tr>
<td>Listener</td>
<td>Hypothesis</td>
<td>46226.564</td>
<td>14</td>
<td>3301.897</td>
<td>2.909</td>
<td>.007</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>33610.301</td>
<td>29</td>
<td>607</td>
<td>1135.206</td>
<td>.000</td>
</tr>
<tr>
<td>Movement_CONDITION * Listener</td>
<td>Hypothesis</td>
<td>7755.969</td>
<td>28</td>
<td>277.000</td>
<td>.676</td>
<td>.865</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>20107.819</td>
<td>49</td>
<td>104.909</td>
<td>270.889</td>
<td>.000</td>
</tr>
<tr>
<td>Programme_Item * Listener</td>
<td>Hypothesis</td>
<td>14599.919</td>
<td>28</td>
<td>521.426</td>
<td>2.110</td>
<td>.006</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>17936.490</td>
<td>72</td>
<td>583.816</td>
<td>247.117</td>
<td>.000</td>
</tr>
<tr>
<td>Repeat * Listener</td>
<td>Hypothesis</td>
<td>7589.607</td>
<td>13</td>
<td>583.816</td>
<td>1.507</td>
<td>.154</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>16692.312</td>
<td>43</td>
<td>837.444</td>
<td>.019</td>
<td>.183</td>
</tr>
<tr>
<td>Speaker * Listener</td>
<td>Hypothesis</td>
<td>101348.253</td>
<td>224</td>
<td>452.448</td>
<td>1.960</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>85860.165</td>
<td>371</td>
<td>688.889</td>
<td>230.889</td>
<td>.000</td>
</tr>
<tr>
<td>Programme_Item * Movement_CONDITION</td>
<td>Hypothesis</td>
<td>1877.374</td>
<td>4</td>
<td>469.343</td>
<td>3.216</td>
<td>.019</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>8401.301</td>
<td>57</td>
<td>145.908</td>
<td>.019</td>
<td>.183</td>
</tr>
<tr>
<td>Movement_CONDITION * Repeat</td>
<td>Hypothesis</td>
<td>2832.213</td>
<td>2</td>
<td>1416.107</td>
<td>5.080</td>
<td>.014</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>7248.100</td>
<td>10</td>
<td>728.773</td>
<td>.014</td>
<td>.281</td>
</tr>
<tr>
<td>Speaker * Movement_CONDITION</td>
<td>Hypothesis</td>
<td>12862.005</td>
<td>32</td>
<td>401.938</td>
<td>3.843</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>48887.652</td>
<td>465</td>
<td>104.569</td>
<td>230.889</td>
<td>.000</td>
</tr>
<tr>
<td>Programme_Item * Repeat</td>
<td>Hypothesis</td>
<td>922.766</td>
<td>2</td>
<td>461.383</td>
<td>4.476</td>
<td>.021</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>2680.280</td>
<td>10</td>
<td>103.086</td>
<td>.021</td>
<td>.256</td>
</tr>
<tr>
<td>Speaker * Programme_Item</td>
<td>Hypothesis</td>
<td>23150.295</td>
<td>32</td>
<td>723.447</td>
<td>6.159</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>54452.436</td>
<td>463</td>
<td>117.467</td>
<td>.000</td>
<td>.298</td>
</tr>
<tr>
<td>Speaker * Repeat</td>
<td>Hypothesis</td>
<td>2705.585</td>
<td>16</td>
<td>169.099</td>
<td>1.305</td>
<td>.196</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>26957.631</td>
<td>208</td>
<td>129.604</td>
<td>1.305</td>
<td>.196</td>
</tr>
<tr>
<td>Programme_Item * Movement_CONDITION * Listener</td>
<td>Hypothesis</td>
<td>8329.665</td>
<td>56</td>
<td>148.744</td>
<td>2.399</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>165939.998</td>
<td>2676</td>
<td>62.010</td>
<td>62.010</td>
<td>.000</td>
</tr>
<tr>
<td>Movement_CONDITION * Repeat * Listener</td>
<td>Hypothesis</td>
<td>7248.100</td>
<td>26</td>
<td>278.773</td>
<td>4.496</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>165939.998</td>
<td>2676</td>
<td>62.010</td>
<td>62.010</td>
<td>.000</td>
</tr>
<tr>
<td>Speaker * Movement_CONDITION * Listener</td>
<td>Hypothesis</td>
<td>47487.606</td>
<td>448</td>
<td>105.999</td>
<td>1.709</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>165939.998</td>
<td>2676</td>
<td>62.010</td>
<td>62.010</td>
<td>.000</td>
</tr>
<tr>
<td>Programme_Item * Repeat * Listener</td>
<td>Hypothesis</td>
<td>2680.280</td>
<td>26</td>
<td>103.086</td>
<td>1.662</td>
<td>.019</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>165939.998</td>
<td>2676</td>
<td>62.010</td>
<td>62.010</td>
<td>.000</td>
</tr>
<tr>
<td>Speaker * Programme_Item * Listener</td>
<td>Hypothesis</td>
<td>53449.245</td>
<td>448</td>
<td>119.306</td>
<td>1.924</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>165939.998</td>
<td>2676</td>
<td>62.010</td>
<td>62.010</td>
<td>.000</td>
</tr>
<tr>
<td>Speaker * Repeat * Listener</td>
<td>Hypothesis</td>
<td>26957.631</td>
<td>208</td>
<td>129.604</td>
<td>2.090</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>165939.998</td>
<td>2676</td>
<td>62.010</td>
<td>62.010</td>
<td>.000</td>
</tr>
</tbody>
</table>
The Effect of Head Movements on LRA

The main factor ‘repeat’; the repeat interactions, repeat*listener, loudspeaker*repeat and loudspeaker*repeat*programme item; and the 2\textsuperscript{nd} order interaction, movement condition*listener were the only non-significant findings in the ANOVA. All other independent variables and interactions were shown to be significant. A normality plot of the ANOVA’s standardised residuals indicated that the model was a good fit of the data (Figure 46).

Table 25 - ANOVA table of all variables and 1\textsuperscript{st}-3\textsuperscript{rd} level interactions for absolute elevation LRE

<table>
<thead>
<tr>
<th>Source</th>
<th>Hypothesis DF</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Programme_Item * Movement_Condition * Repeat</td>
<td>4</td>
<td>322.783</td>
<td>5.205</td>
<td>.000</td>
</tr>
<tr>
<td>Programme_Item * Movement_Condition * Error</td>
<td>2676</td>
<td>62.010(^3)</td>
<td>.866</td>
<td>.008</td>
</tr>
<tr>
<td>Speaker * Programme_Item * Movement_Condition * Repeat</td>
<td>64</td>
<td>135.645</td>
<td>2.187</td>
<td>.000</td>
</tr>
<tr>
<td>Speaker * Programme_Item * Movement_Condition * Error</td>
<td>2676</td>
<td>62.010(^3)</td>
<td>.000</td>
<td>.050</td>
</tr>
<tr>
<td>Speaker * Movement_Condition * Repeat</td>
<td>32</td>
<td>91.113</td>
<td>1.469</td>
<td>.044</td>
</tr>
<tr>
<td>Speaker * Movement_Condition * Error</td>
<td>2676</td>
<td>62.010(^3)</td>
<td>.000</td>
<td>.017</td>
</tr>
<tr>
<td>Speaker * Programme_Item * Repeat</td>
<td>32</td>
<td>45.243</td>
<td>.730</td>
<td>.866</td>
</tr>
<tr>
<td>Speaker * Programme_Item * Error</td>
<td>2676</td>
<td>62.010(^3)</td>
<td>.000</td>
<td>.009</td>
</tr>
</tbody>
</table>

Figure 46 - Histogram of ANOVA standardised residuals
Further analysis of the findings of the ANOVA will be conducted in sections based each of the main factors. The three main factors of the ANOVA, namely programme item, movement condition and loudspeaker, are discussed individually in the following sections.

### 8.2.2.3 Programme Item

Programme item had the largest significant effect on the elevation LRE ($F(2,4437) = 41.991, p < 0.001, \eta^2 = .749$). The F statistic of the individual factor was much larger than any interaction level, thereby indicating that the significance finding was not compromised by higher order interactions. Figure 47 shows the means and 95% confidence intervals for the absolute elevation LRE separated by programme item.

*Figure 47 - Absolute elevation LRE separated by programme item*

Figure 47 shows that N was the most accurately localised programme item. The N programme item gave the listener a broadband source providing rich pinna
cues. As predicted, by degrading pinna cues, the L and B programme items significantly reduced elevation LRA.

The B and L programme items show similar elevation LREs, which indicates that the addition of the band pass filtered high frequency component did not markedly improve elevation LRA. This supports the findings of Chapter 6 that the addition of a \( \frac{1}{2} \) octave band of high frequency band-pass filtered noise does not stimulate additional pinna cues.

Programme item had the largest effect of any of the factors trialed (\( \eta^2 = .749 \)). This suggests that pinna cues may constitute the most significant factor in elevation localisation. This finding opposes that of Blauert [1983], who stated that ‘if information obtained by means of head movements is evaluated, it overrides information derived from monaural signal characteristics’.

### 8.2.2.4 Movement Condition

The effect of movement condition was significant (F (2, 4437) = 16.172, \( p < 0.001 \), \( \eta^2 = .532 \)). Figure 48 shows the absolute elevation LRE plotted against movement condition.
In general, across programme item and loudspeaker position, both the forced and free movement conditions are significantly different from the no movement condition. Therefore, head movements allowed the listener to localise more accurately in elevation. Furthermore, head movements that are limited to yaw rotations, give the listener the necessary cues to improve LRA with head movement.

Second order interactions movement condition*programme item (F (4, 4437) = 3.216, p = 0.019) and movement condition*loudspeaker (F (32, 4437) = 3.843, p < 0.001) and third order interaction movement condition*programme item*loudspeaker (F (64, 4437) = 2.187, p < 0.001) were also significant. Three-way interactions are difficult to interpret and they have only a small number of data points, which reduces the statistical power of the calculations. However, plotting the three way interactions did highlight some interesting trends in the data. Figure 49 shows the means and 95% confidence intervals for the absolute
elevation LRE separated by loudspeaker, programme item and movement condition.

![Figure 49](image)

**Figure 49 - Absolute elevation LRE separated by loudspeaker, programme item and movement condition**

Loudspeakers 3, 5 and 6 in the B programme item condition exhibit a significantly larger error in the no movement condition than in the forced condition. Loudspeakers 5 and 6 were outside of the field of fixation (the range of positions at which the listener could directly point an eye) for the no movement condition, which suggests that the higher error may have been due to pointing inaccuracies. Furthermore, for loudspeaker 16 there were significant differences between the free and the forced/no movement conditions using the L and N programme items. Loudspeaker 16 was in the field of fixation for the free movement condition and outside it for the no movement and forced
movement conditions. All of these differences are indicative of pointing errors caused by the listener being unable to sight the location of the rearward sources.

8.2.2.5 Movement Conditions & Loudspeaker Groups

It is suggested that the significance of interaction between the movement condition and loudspeaker factors found in the ANOVA was due to pointing accuracy changes with loudspeaker azimuth. The changing pointing accuracy was also highlighted in the response methods experiment in Chapter 7. While auditioning a stimulus, preceding a localisation response, a listener will sight the location at which they want to point. This will allow them to point at their chosen direction more accurately than an unsighted response. Therefore, if a listener cannot fixate on the perceived location of the source whilst it is sounding then they will give a less accurate _pointing_ response. To study this further, the loudspeakers were grouped based on their location within the listener's field of fixation. The field of fixation was defined as anywhere within a 45° radius of the listener's eyes pointing straight ahead [Howard 1982]. Three loudspeaker groups were defined:

1. _Sources within the field of fixation for all movement conditions_.
2. _Sources within the field of fixation for the forced and free movement conditions and not contained within Group 1_.
3. _Sources within the field of fixation for the free movement condition and not contained within Groups 1 and 2_.

Figure 50 shows the means and 95% confidence intervals for the absolute elevation LRE separated by loudspeaker groups and movement condition.
The plot indicates that there are significant differences between the forced and no movement conditions for all loudspeaker groups. A Kolmogorov-Smirnov test showed that only 9 out of the 54 levels were normally distributed using the loudspeaker groups factor and so non-parametric tests were used to verify the significance differences between these groups. A targeted comparison between the no movement and forced movement conditions using a Mann-Whitney test showed significant differences for all loudspeaker groups: Group 1 (U = 29786.0, P = 0.013); Group 2 (U = 28120.5, P = 0.001); and Group 3 (U = 423977.0, P = 0.005). Therefore, forced yaw head movements significantly improve a listener’s
elevation LRA when compared to static listening conditions. Were improved visual or static cues or a dynamic cue the cause of this increased LRA?

### 8.2.2.6 Reasons for Improvement with Forced Movement

The forced yaw head movement condition was included in the study to allow improved pointing and static cues to be distinguished from dynamic cues. For the forced yaw movement condition the listener’s head was constantly moving between ±30°. They could not stop their head movement to focus their static cues and higher pointing accuracy on the source location, as they could in the free movement condition. The main cues available in the forced movement condition were dynamic cues created through head movement. Therefore, the significant differences between the forced and no movement conditions indicate that dynamic cues cause the improved LRA with head movement.

It could be argued that the forced yaw movement condition had improved pointing accuracy when compared to the no movement condition for loudspeaker group 2; listeners in the forced movement condition could sight the location at which they were going to point the laser, while the in the no movement condition they could not. However, for group 1 the differences in LRA between the forced and no movement conditions cannot be attributed to pointing accuracy because the loudspeaker locations were within the field of fixation for both movement conditions. Therefore, the differences for group 1 must have been due to an auditory localisation cue gained by head movement. There is a similar difference between the forced and no movement conditions for groups 1 and 2, which indicates that pointing accuracy was not the primary cause of the increased LRA for the forced head movement condition in either group.

The plot also indicates significant differences between the absolute elevation LRE of the free and forced movement conditions for group 3 only. Using a Mann-Whitney test these indications were confirmed: Group 1 (U = 33473.5, P = 0.733); Group 2 (U = 31647.5, P = 0.161); and Group 3 (U = 422858.0, P = 0.004). For group 3 listeners responded more accurately in the free movement condition because they could sight the perceived source location, which resulted in an
increased pointing accuracy. The LRA of the listener in the forced movement condition for group 3 is significantly reduced when compared to the localisation accuracy of the forced movement condition in the other two loudspeaker groups. It is suggested that the listener could no longer sight the perceived loudspeaker location and so their pointing accuracy was reduced.

**8.2.2.7 Movement Condition & Programme Item**

Figure 51 shows the means and 95% confidence intervals for the absolute elevation LRE separated by programme item and movement condition.

![Figure 51](image)

*Figure 51 – Absolute elevation LRE separated by programme item and movement condition*

The same trend in mean error with movement condition is shown for each programme item: movement has the lowest mean error while the no movement condition has the highest. Significant differences in elevation LRE were found between all the free movement and no movement conditions. When the listeners
were able to move their heads their LRA was significantly improved. It has been shown that this improvement was due to higher pointing accuracy, static cues and dynamic cues. The plot for the N programme item shows that, even in the presence of pinna cues, listener head movement still improves LRA.

The plot shows that the forced condition is significantly different from the no movement condition using the B programme item. The B programme item contained both low and high frequency energy with reduced pinna cues. This indicates that the listener was more dependent on the cues given by head movement when pinna cues were not available. Furthermore it indicates that head movement cues are dependent on both dynamic ILD and dynamic ITD cues.

Targeted t-tests were conducted on the L and N programme items to compare the no movement and forced movement conditions. It was shown that there were significant differences between the no movement and forced movement conditions for both the L (t=1.964, p<0.05) and N (t=2.296, p<0.05) programme items. Although these differences were not as large as for the B programme item, they were still shown to be significant.

8.2.2.8 Loudspeaker

By observing how the absolute elevation LRE varies with loudspeaker number, it is hypothesized that the actual loudspeaker elevation caused the loudspeaker factor to be significant in the ANOVA. Figure 52 shows the absolute elevation LRE plotted against loudspeaker elevation.
The most obvious characteristic of this graph is that there is a large variation in response error with loudspeaker elevation; sources located at $+35^\circ$ exhibit a significantly lower elevation LRE than other source locations.

When the listener is faced with a range of possible response angles limited by the location of the acoustically transparent curtain, they might limit the range of perceived elevation responses accordingly. The acoustically transparent curtain concealed locations from approximately $-40^\circ$ to $+90^\circ$ in elevation. For sources at high and low elevations the listener was able to make larger errors (reporting perceived locations as far down as the bottom of the curtain for high sources, and as far up as the top for low sources). Therefore, the lowest value of ‘maximum possible error’ would be found for loudspeakers located halfway between those two points at approximately $+25^\circ$ elevation. A scatter plot of individual listener responses is shown in Figure 53.
Figure 53 - Scatter graph of elevation LRE with actual loudspeaker location

At high and low source locations listeners responded with locations up to 90° in error, while for sources at +35° the largest error was approximately 45°. It is possible that this is, at least in part, a result of listeners being reluctant to report locations near to or outside the edges of the curtain.

There are two general forms of LRE that are grouped together when the absolute error is calculated: the bias and variance of reported location. Plotting the signed error shows both errors separately: the mean signed angle error is the bias and the variance of the signed error is the variance of the perceived location. The signed error is calculated as:

$$\theta_{se} = (\theta_{reported} - \theta_{actual})$$

(7)
where $\phi_{se}$ is the signed localisation angle error, $\phi_{reported}$ is the reported loudspeaker location angle and $\phi_{actual}$ is the actual loudspeaker location angle. This signed error gives an indication of the direction of bias of the localisation responses.

Figure 54 shows the signed elevation LRE for each loudspeaker elevation location.

The mean signed elevation LRE for sources located at +35° in elevation is near zero, which means that the reported and actual locations are similar. For sources further from the median response angle (which appears to be approximately +35°), the mean errors are large. The bias is always towards the median response angle: all loudspeakers located above +35° have a negatively signed error, while all loudspeakers below +35° have a positively signed error. This
compression effect has been noted in previous elevation localisation studies [Perrett and Noble 1997] and is discussed in greater detail in Chapter 11.

8.2.3 Azimuth LRA

This section describes the analysis conducted on the azimuth LRA.

8.2.3.1 Front-Back Errors

Front-back errors are a major source of overall azimuth LRE, and inflate the absolute azimuth response error. Preceding any absolute azimuth LRE analysis, front-back errors were removed and studied separately. Front-back errors were removed by replacing \( \phi \) with \( 180^\circ - \phi \) (\( \phi \) = reported azimuth location) for all data that shows a response in the wrong hemisphere (front-back reversals). For the source location at precisely 90° azimuth to the listener, the localisation response was left unchanged.

Figure 55 shows front-back errors separated by movement condition and programme item.
When the listener was able to move their head, either with a free or forced movement, front-back errors decreased dramatically. For the free movement and the B programme item, front-back errors were removed entirely. Front-back errors were slightly higher for the forced movement condition when compared to the free movement, for the B and N programme items.

In the no movement condition, the B and L programme items have more than double the number of front-back errors of the N programme item. With the N programme item the listener could use pinna cues to resolve front-back errors. This plot shows that pinna cues are less effective than head movements at resolving front-back confusions. When head movements are permitted, trends due to the programme item are not present; head movements reduce the rate of front-back errors regardless of the programme item.

Figure 55 - Pie chart of front-back errors
8.2.3.2 ANOVA

An ANOVA was conducted on the absolute azimuth LRE data with the fixed factors ‘Movement Condition’, ‘Loudspeaker’, ‘Programme Item’, ‘Repeat’ and the random factor ‘Listener’. All interactions up to three-way were modelled. The absolute azimuth LRE data was checked for normality using a Kolmogorov-Smirnov test. The absolute azimuth LRE was found to be normal for 125 of the 153 conditions. Each normality check contained 31 data points averaged across the 'listener' and 'repeat' variables. The results of the ANOVA are shown in Table 26.
Table 26 - ANOVA table of all variables and $1^{st}$-$3^{rd}$ level interactions for absolute azimuth LRE with front-back errors removed
The main factors loudspeaker (F (16, 4437) = 12.911, p < 0.001), movement condition (F (2, 4437) = 97.537, p < 0.001) and programme item (F (2, 4437) = 20.858, p < 0.001) are shown to be significant in the ANOVA. The listener and repeat variables are not shown to be significant.

### 8.2.3.3 Movement Condition & Programme Item

The ANOVA showed the movement condition factor to be highly significant (F (2, 4437) = 97.537, p < 0.001) with a large f value and effect size. The interaction movement condition*programme item was also shown to be significant (F (4, 4437) = 19.868, p < 0.001). Figure 56 shows the means and 95% confidence intervals for the absolute azimuth LRE separated by programme item and movement condition without front-back errors.

*Figure 56 - Absolute azimuth LRE separated by programme item and movement condition without front-back errors*
There are significant differences in error between all movement conditions for each programme item. The azimuth LRA followed the same trend with movement condition for all programme items: the no movement condition was least accurate, while the free movement condition was most accurate. The significant differences shown for the N programme item indicate that even in the presence of pinna cues, the ability to move the head still improves azimuth LRA.

As shown in Chapter 2, a listener will orientate towards the source azimuth when they are allowed free head motion. Furthermore, localisation in azimuth is most accurate on the median plane [Stevens and Newman 1936]. In the forced movement condition the listeners were not able to orientate towards the source location. Therefore, the significant differences between the forced and free movement conditions are likely to have been due to improved static cues.

It is suggested that the significant differences between the no movement and forced movement conditions are caused by a dynamic head movement cue and that when the listener moved their head it allowed them to resolve the source’s location on the cone of confusion. Movement condition had the largest effect on the absolute azimuth LRE of any of the factors trialed ($\eta^2 = .887$).

### 8.3 Conclusions

Previous research has shown that head movements can resolve front-back confusions. However, it was unclear whether head movements could otherwise increase elevation and azimuth LRA. Dynamic interaural differences (without pinna cues) can be used to resolve source location using a spherical head model [Ashby 2010] but whether humans use these cues is unclear from previous research. The goal of this study was to answer the following research questions:

- How does head movement affect elevation and azimuth LRA?
- How does the presence of pinna cues affect elevation and azimuth LRA using head movement?
- If LRA is improved by head movement, what factor is the cause of this improvement?
This conclusions section will show how the study has answered these questions. Elevation LRA and azimuth LRA will be discussed separately.

8.3.1 Elevation LRA

*How does head movement affect elevation LRA?*

For noise programme items, head movements significantly increase elevation LRA.

*How does the presence of pinna cues affect elevation LRA using head movement?*

Pinna cues constitute a more significant factor in elevation LRA than head movement. When high frequency content is removed from noise programme items, thereby suppressing the listener’s pinna cues, elevation LRA is reduced. A half-octave band of white noise 5.7–8 kHz does not provide sufficient pinna cues to allow the listener to report the source location accurately. This finding agrees with those shown in Chapter 6. Whether this is the case for all half-octave bands of noise is unverified. Larger improvements in elevation LRA with head movement were noted for the programme items with suppressed pinna cues. Therefore, the suppression of pinna cues causes head movement cues to become more significant. However, even in the presence of pinna cues, the ability of the listener to move their head significantly improves elevation LRA.

*If elevation LRA is improved by head movement, what factor is the cause of this improvement?*

Using the programme item with impaired pinna and ILD cues (L programme item), both free and forced head movement significantly improved elevation LRA. This shows that dynamic ITD cues alone can be used to improve the elevation LRA. The largest improvement in response with head movement was shown when both ITD and ILD cues were included but pinna cues were impeded (B programme item). Further studies investigating the cues responsible for the improvement in elevation LRA with head movement are required. Perrett and Noble [1997] stated that low frequency energy is required for head movement.
cues to be effective; Chapter 9 will investigate whether head movement localisation is still possible without low frequency ITD cues. Chapter 4 highlighted the crossover region between ITD and ILD cues as an area of low elevation LRA. The experiment in Chapter 9 will also investigate whether this low static elevation LRA causes head movement cues to be less effective.

The results also suggest that head movement cues are based on interaural time and level differences, not spectral cues. Wallach's dynamic localisation cues [1938] are based on ITDs and ILDs, so accurate elevation and azimuth LRA should be possible with extremely narrow bandwidth sources. If head movement cues are shown to be effective at lower bandwidths than pinna cues then this will provide further evidence that head movement cues are independent of pinna cues. Chapter 10 investigates the effect of bandwidth on localisation with head movement.

The three factors outlined in the introduction, namely dynamic cues, increased pointing accuracy and improved static cues, can all improve elevation LRA when the listener moves their head. For rearward sources, when the listener is unable to move their head, reduced pointing accuracy is the main factor that reduces elevation LRA. For sources in the frontal hemisphere dynamic and static cues cause higher elevation LRA with head movement.

Based on this experiment alone it is not possible to conclude whether static or dynamic cues have a larger impact upon elevation localisation with head movement. However, the results do strongly suggest that dynamic cues are the primary cause of improved elevation LRA with head movement.

For sources within the field of fixation for both the forced and free movement conditions, forced yaw movements are as effective as free head movements in resolving source elevation. This indicates that the primary elevation localisation cue given with head movement is caused by a yaw-based rotation. Furthermore, in the forced movement condition listeners were rotating their heads and so (although their heads might have passed through positions where improved static cues would have been available) they were not permitted to statically focus on a specific location. Static cues were more readily available in the free
movement condition yet for frontal sources this did not improve the elevation LRA.

The forced yaw movement condition also gave a higher elevation LRA than no movement conditions for all loudspeaker groups. The forced yaw movement condition was included in the study to attempt to isolate dynamic cues and impede improved static cues. Since the forced yaw movement condition, as well as the free movement condition, was significantly different from the no movement condition, this indicates that some of the improvement in elevation LRA with head movement was due to dynamic cues. However, in the forced yaw movement condition the listener must have stopped moving in order to change direction at the extreme angles of ±30° azimuth. It could be argued that at that instant the listener would have auditioned static cues in the forced movement condition that were different from those auditioned in the no movement condition. Could these different static cues cause an improvement in elevation LRA?

To conclude whether dynamic cues or improved static cues are the main cause of the higher elevation LRA with head movement requires a further experiment. The experiment described in this chapter showed that free head movements improve elevation LRA when compared to no movement for all loudspeaker groups and programme items. Chapter 2 showed that in free head movement conditions listeners will rotate in yaw towards the source location. Therefore if elevation LRA is higher for sources on the median plane then it is likely that static cues are significant in improving elevation LRA with head movement; alternatively if median plane elevation LRA is not significantly higher then static cues do not significantly affect the head movement localisation cue. This is investigated in Chapter 11.

8.3.2 Azimuth LRA

_How does head movement affect azimuth LRA?_
For noise programme items, head movements increase azimuth LRA. The ability of the listener to move their head reduces front-back errors. When the listener is unable to move their head, front-back errors are the main source of azimuth LRE.

*How does the presence of pinna cues affect azimuth LRA using head movement?*

The presence of pinna cues increases azimuth LRA in all movement conditions.

*If azimuth LRA is improved by head movement, what factor is the cause of this improvement?*

When the listener is able to move their head freely they can localise more accurately than when they are forced to move their head in a specified manner. Therefore it is suggested that of the three factors outlined in the introduction, static cues cause the largest improvement in azimuth LRA with head movement. When they are free to move their head they can move the source into their most accurate area of static azimuth localisation, the median plane. Results also show that head movements can also create dynamic cues and increase pointing accuracy causing an improvement in azimuth LRA.
9 Head Movement Localisation Without the Presence of Low Frequency ITDs

Chapters 6 and 8 showed that using a ‘½ octave band of high frequency noise with a low frequency component’ programme item significantly reduced LRA when compared to full bandwidth noise due to the impairment of the listener’s pinna cues. Chapter 8 also showed that yaw based head movements significantly improve the listener’s LRA for this programme item. Therefore, head movements can be effective even when some localisation cues, in this case pinna cues, are impaired.

Chapter 8 showed that head movements improve LRA using a ‘½ octave band of high frequency noise with a low frequency component’ programme item more than for a low pass filtered noise programme item alone. This indicates that the listener may gain extra head movement cues at high frequencies that allow a larger improvement in LRA. The low pass filtered noise predominantly tested ITD cues, while the inclusion of the higher frequencies within the ‘½ octave band of high frequency noise’ allowed ILDs to be used as well.

This experiment will compare the LRAs of free and no head movement conditions when only certain cues are available. All of the programme items used in Chapter 8 contained low frequency ITD cues. The main goal of this chapter will be to find whether head movements can improve LRA without the presence of low-frequency ITDs. The research question this chapter seeks to answer is:

- Is the presence of low-frequency ITDs necessary for improved elevation LRA with head movement?

Since ITDs, ILDs and pinna cues are effective over different frequencies, by studying how head movement LRA varies over frequency one can find out which changing cue is being used by the listener. If head movement improves LRA for programme items containing only high-frequency energy, then one can conclude that ITDs are not necessary for increased LRA with head movement. Band-pass-
filtered noise programme items with differing centre frequencies were chosen to study how head movement cues vary over frequency.

9.1 Experiment Method

The experiment method of the previous head movement studies was replicated in terms of loudspeaker location, response method and movement condition. The only factor that was changed from previous experiments was the programme item. A brief reminder of the experiment method will be given here, followed by an in-depth discussion of the programme item factor.

9.1.1 General Experiment Setup and Method

The experiment used seventeen Genelec 8020a loudspeakers, positioned on a hemisphere to the right of the listener, spread from 0° to 180° in azimuth and -35° to +55° in elevation. The loudspeakers were positioned at the coordinates of the vertices of a truncated icosahedron. The loudspeaker numbers, coordinates and angles used in the experiment are given in Table 40 in Appendix B. To avoid any biases due to visual cues, the locations of the loudspeakers were hidden from view by an acoustically transparent but visually opaque curtain.

Eight listeners auditioned all programme item and movement condition combinations; all were IoSR postgraduate students at the University of Surrey, aged between 18 and 35, with no reported hearing conditions.

The listener was asked to report the perceived location of each of a series of signals, each replayed via a single selected loudspeaker. The listener responded using a laser pointing method. The laser was attached to the head and positioned to point out from between the eyes via an elasticated band that was wrapped around the listener’s head. A Polhemus head tracker measured the listener’s head movement in six degrees of freedom and each perceived source location, as indicated by the listener, was calculated using a triangulation technique similar to the method described by Frank et al. [8]. The radius from the central listening position to the loudspeakers was used to create a projected sphere and the
perceived location was taken to be the point at which the line of the laser pointer pierced this sphere.

9.1.2 Programme Items

The experiment used three band-pass-filtered noise programme items, which were called ‘low-frequency (LF)’, ‘mid-frequency (MF)’ and ‘high-frequency (HF)’. The filter characteristics used to create these programme items were:

LF

- Centre Frequency: 500 Hz
- Passband: 420 - 595 Hz
- Transition bands: 210 - 420 and 595 - 805 Hz
- Stopband: 0 - 210 Hz and 805 - 20000 Hz

MF

- Centre Frequency: 2000 Hz
- Transition bands: 841 - 1682 and 2378 - 3219 Hz
- Stopband: 0 - 841 Hz and 3219 - 20000 Hz

HF

- Centre Frequency: 6000 Hz
- Transition bands: 2522.5 - 5045 and 7135 - 9657 Hz
- Stopband: 0 – 2522.5 Hz and 9657 - 20000 Hz

Each programme item had a half-octave pass band with an octave slope at the lower frequency side to the stopband attenuation. For all programmes, the stopband region was 100dB lower than the passband. The frequency spectrum of the HF programme item is shown in Figure 57.
Spectral energy above 2 kHz is required for pinna cues to be available. The LF programme item contained no perceivable spectral energy above 805 Hz and so there were no pinna cues available to the listener. Furthermore, Begault [2000] states that below 1 kHz ILDs are ineffective as the sound waves can effectively diffract around the head. So it can be concluded that the LF programme item was providing only ITD cues.

The MF programme item was created to test the crossover area between ITDs and ILDs. Mills [1958] found that at 2 kHz the MAA was greatest, indicating that neither ITDs nor ILDs are particularly effective. The MF programme item has also been shown to be the least accurately localised frequency region in azimuth using interaural difference studies [Newman and Stevens 1936]. Since the upper frequency stopband of this programme item was located at approximately 3 kHz, there were no pinna cues to allow more accurate localisation. This region was included to find out whether, when neither interaural cue was working well, head movements could still improve LRA.

The perceivable energy of the HF programme item extended down to approximately 3 kHz. Above approximately this frequency there is a breakdown in phase locking of the inner ear hair cells, which causes ITDs to no longer be perceivable unless the signal varies in amplitude over time, which the signal

Figure 57 - Frequency spectrum of the HF programme item

![Magnitude Response (dB)](image)
used in this experiment did not. Since the bandwidth of the programme item was limited, the pinna cues would be significantly less effective than when using a full bandwidth source, as shown in a previous pilot experiment. Therefore the HF programme item would contain predominantly ILD cues.

These centre frequencies were chosen to highlight the frequency ranges at which the listener’s localisation cues changed. For the LF programme item the listener would predominantly use ITDs; for the HF programme item the listener would predominantly use ILDs and reduced spectral cues; and for the MF programme item the listeners would use both ITDs and ILDs, though static azimuth LRA to these cues is least accurate in this region. Since the HF programme item will contain no low-frequency energy, it will be possible to note whether head movements can improve LRA without the presence of low-frequency ITDs. By studying these frequency ranges, it is possible to resolve which dynamic localisation cue the listener uses most prevalently.

9.2 Results

The localisation results analysis is split into the two location components, elevation LRA (Section 9.2.1) and azimuth LRA (Section 9.2.2). Programme item, movement condition and loudspeaker location were the main factors used to study the listener’s LRA.

The main metric used to measure LRA was the absolute LRE. This was calculated as shown in Equation 2.

\[ \phi_{ae} = \sqrt{\left( \phi_{\text{perceived}} - \phi_{\text{actual}} \right)^2} \]  

Where \( \phi_{ae} \) is the absolute LRE, \( \phi_{\text{perceived}} \) is the reported loudspeaker location angle and \( \phi_{\text{actual}} \) is the actual loudspeaker location angle.

9.2.1 Elevation LRA

In this section the absolute elevation LRE component of the localisation tests is investigated. The goal of this section is to find the effect of programme item on
the elevation LRA with head movement; specifically whether head movements offer an improved elevation LRA without the presence of low-frequency ITDs.

### 9.2.1.1 Normality Checks

To use parametric tests such as ANOVA the data must be normally distributed. Using a Kolmogorov-Smirnov test across the independent variables ‘movement condition’, ‘programme item’ and ‘loudspeaker’, the distribution of the absolute elevation LRE data was found to be normal for all of the conditions. Each normality test contained 16 data points averaged across the listener and repeat variables.

### 9.2.1.2 ANOVA

An ANOVA was conducted on the absolute elevation data with fixed factors ‘movement condition’, ‘loudspeaker’, ‘programme item’, ‘repeat’ and the random factor ‘listener’. All interactions up to three-way were modelled. The results of the ANOVA are shown in Table 27.
9. Head Movement Localisation Without the Presence of Low Frequency ITDs

### Tests of Between-Subjects Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Hypothesis</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>Hypothesis</td>
<td>607118.229</td>
<td>1</td>
<td>607118.229</td>
<td>263.362</td>
<td>.000</td>
<td>.974</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>16136.839</td>
<td>7</td>
<td>2305.263</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Programme_Item</td>
<td>Hypothesis</td>
<td>24725.597</td>
<td>2</td>
<td>12362.848</td>
<td>17.353</td>
<td>.000</td>
<td>.974</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>9974.242</td>
<td>14</td>
<td>712.446</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Movement_Condition</td>
<td>Hypothesis</td>
<td>8012.404</td>
<td>1</td>
<td>8012.404</td>
<td>22.118</td>
<td>.002</td>
<td>.760</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>2535.790</td>
<td>7</td>
<td>362.256</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Speaker</td>
<td>Hypothesis</td>
<td>97206.856</td>
<td>16</td>
<td>6075.429</td>
<td>12.318</td>
<td>.000</td>
<td>.638</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>55239.806</td>
<td>112</td>
<td>493.213</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Repeat</td>
<td>Hypothesis</td>
<td>205.504</td>
<td>1</td>
<td>205.504</td>
<td>1.117</td>
<td>.326</td>
<td>.138</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>1287.668</td>
<td>7</td>
<td>183.952</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Listener</td>
<td>Hypothesis</td>
<td>16136.839</td>
<td>7</td>
<td>2305.263</td>
<td>2.548</td>
<td>.083</td>
<td>.626</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>9628.907</td>
<td>10.643</td>
<td>904.735</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Movement_Condition * Listener</td>
<td>Hypothesis</td>
<td>2535.790</td>
<td>7</td>
<td>362.256</td>
<td>.794</td>
<td>.602</td>
<td>.251</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>7576.743</td>
<td>16.612</td>
<td>456.114</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Programme_Item * Listener</td>
<td>Hypothesis</td>
<td>9974.242</td>
<td>14</td>
<td>712.446</td>
<td>1.438</td>
<td>.207</td>
<td>.445</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>12425.790</td>
<td>25.079</td>
<td>2309.033</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Repeat * Listener</td>
<td>Hypothesis</td>
<td>1287.668</td>
<td>7</td>
<td>183.953</td>
<td>.770</td>
<td>.628</td>
<td>.391</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>2003.700</td>
<td>25.079</td>
<td>2309.033</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Speaker * Listener</td>
<td>Hypothesis</td>
<td>55239.806</td>
<td>112</td>
<td>493.213</td>
<td>1.732</td>
<td>.001</td>
<td>.607</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>35693.822</td>
<td>125.377</td>
<td>284.691</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Programme_Item * Movement_Condition</td>
<td>Hypothesis</td>
<td>687.021</td>
<td>2</td>
<td>343.510</td>
<td>1.058</td>
<td>.373</td>
<td>.131</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>4547.056</td>
<td>14</td>
<td>324.790</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Movement_Condition * Repeat</td>
<td>Hypothesis</td>
<td>89.774</td>
<td>1</td>
<td>89.774</td>
<td>.496</td>
<td>.504</td>
<td>.066</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>1266.315</td>
<td>7</td>
<td>180.902</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Movement_Condition * Speaker</td>
<td>Hypothesis</td>
<td>11212.444</td>
<td>16</td>
<td>700.778</td>
<td>4.138</td>
<td>.000</td>
<td>.371</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>18969.674</td>
<td>112</td>
<td>169.372</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Programme_Item * Repeat</td>
<td>Hypothesis</td>
<td>497.704</td>
<td>2</td>
<td>248.852</td>
<td>1.497</td>
<td>.258</td>
<td>.176</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>2327.121</td>
<td>14</td>
<td>166.223</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Programme_Item * Speaker</td>
<td>Hypothesis</td>
<td>27923.282</td>
<td>32</td>
<td>872.603</td>
<td>3.906</td>
<td>.000</td>
<td>.358</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>50044.231</td>
<td>224</td>
<td>223.412</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Speaker * Repeat</td>
<td>Hypothesis</td>
<td>1855.131</td>
<td>16</td>
<td>115.946</td>
<td>1.046</td>
<td>.415</td>
<td>.130</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>12416.016</td>
<td>112</td>
<td>110.857</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Programme_Item * Movement_Condition * Listener</td>
<td>Hypothesis</td>
<td>4547.056</td>
<td>14</td>
<td>324.790</td>
<td>2.967</td>
<td>.000</td>
<td>.448</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>90864.161</td>
<td>830</td>
<td>109.475</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Movement_Condition * Repeat * Listener</td>
<td>Hypothesis</td>
<td>1266.315</td>
<td>7</td>
<td>180.902</td>
<td>1.652</td>
<td>.117</td>
<td>.014</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>90864.161</td>
<td>830</td>
<td>109.475</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Movement_Condition * Speaker * Listener</td>
<td>Hypothesis</td>
<td>18969.674</td>
<td>112</td>
<td>169.372</td>
<td>1.547</td>
<td>.001</td>
<td>.173</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>90864.161</td>
<td>830</td>
<td>109.475</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Programme_Item * Repeat * Listener</td>
<td>Hypothesis</td>
<td>2327.121</td>
<td>14</td>
<td>166.223</td>
<td>1.516</td>
<td>.098</td>
<td>.252</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>90864.161</td>
<td>830</td>
<td>109.475</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Programme_Item * Speaker * Listener</td>
<td>Hypothesis</td>
<td>50044.231</td>
<td>224</td>
<td>223.412</td>
<td>2.041</td>
<td>.000</td>
<td>.355</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>90864.161</td>
<td>830</td>
<td>109.475</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Speaker * Repeat * Listener</td>
<td>Hypothesis</td>
<td>12416.016</td>
<td>112</td>
<td>110.857</td>
<td>1.013</td>
<td>.450</td>
<td>.120</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>90864.161</td>
<td>830</td>
<td>109.475</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>Hypothesis</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Programme_Item * Movement_Condition</td>
<td>Hypothesis</td>
<td>546.634</td>
<td>2</td>
<td>273.317</td>
<td>2.497</td>
<td>.083</td>
<td>.006</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>90864.161</td>
<td>830</td>
<td>109.475</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Programme_Item * Movement_Condition</td>
<td>Hypothesis</td>
<td>8115.334</td>
<td>32</td>
<td>253.604</td>
<td>2.317</td>
<td>.000</td>
<td>.082</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>90864.161</td>
<td>830</td>
<td>109.475</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Movement_Condition</td>
<td>Hypothesis</td>
<td>1812.991</td>
<td>16</td>
<td>113.312</td>
<td>1.035</td>
<td>.416</td>
<td>.020</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>90864.161</td>
<td>830</td>
<td>109.475</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
<tr>
<td>Programme_Item</td>
<td>Hypothesis</td>
<td>3228.431</td>
<td>32</td>
<td>100.888</td>
<td>1.922</td>
<td>.594</td>
<td>.034</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>90864.161</td>
<td>830</td>
<td>109.475</td>
<td>.000</td>
<td>.713</td>
<td></td>
</tr>
</tbody>
</table>

Table 27 - ANOVA of absolute elevation LRE data
A histogram of the standardised residuals was plotted to check their normality. A K-S test showed the distribution to be non-normal (p = 0.042). However, this distribution is very close to normal (Figure 58) and large sample sizes often cause the K-S test to be significant. Therefore it is suggested that the findings of the ANOVA are reliable.

![Histogram of elevation LRE ANOVA standardised residuals](image)

*Figure 58 - Histogram of elevation LRE ANOVA standardised residuals*

The main factors ‘repeat’ and ‘listener’ were found to be non-significant. The non-significance of the repeat factor means that the listener’s response did not vary significantly when they were asked to repeat the experiment. Thus, there was no improvement in response as the listener became familiar with the task.

The main factors ‘movement condition’, ‘programme item’ and ‘loudspeaker’ were all found to significantly affect the elevation LRA. Further analysis of the
findings of the ANOVA will be conducted in sections based on each of the main factors.

9.2.1.3 Movement Condition

The ANOVA showed the effect of ‘movement condition’ to be significant with the largest effect size and highest F statistic \(F(1, 1632) = 22.118, p = 0.002, \eta^2 = 0.76\).

Movement condition*loudspeaker was the only other interaction including movement condition that was significant with a meaningful effect size and F value \(F(16, 1632) = 4.138, p < 0.001, \eta^2 = 0.371\). Figure 59 shows the mean elevation LRE and 95% confidence intervals separated by movement condition and loudspeaker number.

![Figure 59 - Elevation LRE separated by loudspeaker and movement condition](image-url)
For loudspeakers 6, 9, 11 and 17 the no movement condition actually resulted in a lower mean elevation LRE than the movement condition. All of these loudspeakers were located at between -9° and +9° in elevation. None of these differences were found to be significant but they do indicate that at these elevations the loudspeakers were localised as accurately in the no movement condition as the movement condition. It is suggested that at these low elevations, when a listener was unsure of the elevation location at which they perceived the source, they were likely to respond by pointing their head at the equator. Thus a response where the listener has no idea of source location will be similar to that of an accurately localised response. This would explain the lack of significant variance between the movement and no movement conditions for these locations.

For all other loudspeaker locations the free movement condition resulted in a more accurate localisation response than the no movement condition. For loudspeakers 2, 13, 14, 15 and 16 the differences were found to be significant. For rearward sources the differences are larger because the listener has both dynamic localisation cues given by head movement and a higher pointing accuracy, as shown in Chapter 8. Loudspeakers 2 and 13 were both at +35° elevation, while loudspeakers 15 and 16 were at -35° and +52° respectively. Head movements appear to improve LRA more at positions away from the equatorial plane. It is suggested that a larger sample size would have resulted in more significant differences between the free and no movement conditions for loudspeaker locations away from the equatorial plane.

It appears that the significance of differences between the movement and no movement condition depend on the actual elevation of the loudspeaker; loudspeakers closer to the equatorial plane result in no significant differences while loudspeakers away from the plane show significant differences. To analyse this further, loudspeakers were grouped according to their actual elevation angle. The mean absolute elevation LRE and 95% confidence intervals were plotted separated by actual loudspeaker elevation and movement condition (Figure 60).
For the loudspeakers located near the equatorial plane there was no significant difference between the no movement and free movement conditions. However, at elevations far from the equatorial plane, head movement caused a significant improvement in elevation LRA. It is unclear from this experiment whether this improvement was due to dynamic cues or improved static cues. When head movements are allowed listeners commonly orient themselves to face the source. There is no evidence in other research that the cues given on the median plane allow a higher elevation LRA than the cues on any other plane. Furthermore, for the LF and MF programme item there are no pinna cues available, so no known possible cause of improved LRA using static cues on the median plane. Figure 61 shows the mean elevation LRE and 95% confidence
intervals separated by movement condition and loudspeaker number for the low frequency programme item.

![Diagram showing Mean Absolute Elevation LRE separated by actual loudspeaker elevation and movement condition for the low frequency programme item.](image)

**Figure 61 - Absolute elevation LRE separated by actual loudspeaker elevation and movement condition for the low frequency programme item**

With the LF programme item there is no spectral energy above 1 kHz and so, based on previous research (Sections 4.2 and 6), there will be no pinna cues present. Figure 61 shows that, at elevations above the equatorial plane, head movements significantly improve elevation LRA for the LF programme item. Since pinna cues are impeded for this programme item, there are no known static localisation cues that could account for this improvement in elevation LRA with head movement; thus it is likely that this improvement is due to dynamic localisation cues. The question of whether head movement provides static or dynamic cues to elevation is investigated further in Chapter 11.
Why the same trend in improved head movement LRA is not present at elevation angles below the equatorial plane is unclear. It may be that at angles below the equatorial plane there are a large number of up-down confusions. Perrett and Noble [1997] showed that a listener will assume a source is in the upper hemisphere unless given pinna cues to the contrary. As highlighted by Wallach, yaw movements about the equatorial plane do not actually resolve up-down confusions. It may be that the lower LRA for these loudspeakers was due to a bias caused by the position of the acoustically transparent curtain, as described in Chapter 8.

Figure 62 shows the mean absolute elevation LRE and 95% confidence intervals for the two movement conditions. In this graph all loudspeaker locations are combined.

![Figure 62 - Absolute elevation LRE separated by movement condition](image-url)
The free movement condition was localised significantly more accurately than the no movement condition. The movement condition resulted in an overall localisation response improvement of approximately 5°.

The response of the listeners was not significantly affected by the programme item*movement condition interaction. Therefore the improvement given by head movement for the programme item without the presence of low-frequency ITDs was the same as those offered with low-frequency ITDs. This shows that the improvement offered by head movement was significant regardless of the frequency range of the programme item, and so both ITDs and ILDs are used to allow localisation with head movement.

9.2.1.4 Programme Item

The ANOVA showed the main factor programme item (F (2, 1632) = 17.353, p < 0.001, \( \eta^2 = 0.713 \)) and factor interactions programme item*loudspeaker (F (32, 1632) = 3.906, p < 0.001, \( \eta^2 = 0.358 \)), programme item*loudspeaker*listener (F (224, 1632) = 2.041, p < 0.001, \( \eta^2 = 0.355 \)) and programme item*movement condition*loudspeaker (F (32, 1632) = 2.317, p < 0.001, \( \eta^2 = 0.082 \)) to be significant. At the 3-way interaction level for the programme item*loudspeaker*listener interaction there were only 4 data points and so any significant differences might be a quirk of the data, so it would be dangerous to draw any conclusions from this.

Figure 63 shows the mean absolute elevation LRE and 95% confidence intervals for the programme item*loudspeaker interaction.
Significant differences occurred most frequently between the HF and MF programmes. This trend did not occur for 7 of the 17 loudspeakers: namely loudspeakers 1, 5, 9, 10, 13, 14 and 17. Loudspeakers 5, 10 and 13 were located at +35°, while loudspeakers 1, 9 and 17 were located at 0° elevation and loudspeaker 14 was located at +9.8° elevation, all of these loudspeaker locations showed a relatively high elevation LRA for all programme items. It is suggested that the effect of programme item on elevation LRA is dependent on loudspeaker elevation.

A further trend apparent from this graph is the low LRA for loudspeakers 4, 8, 12 and 15 (all located at -35°) for the MF programme item. For all these loudspeakers there were significant difference in LRA between the HF and MF programme items. Loudspeaker 7 and 16 were both located at +53° and for this
location listeners exhibited a high LRA for the HF programme item and low LRA for the LF and MF frequency bands.

The main observation from this plot is that loudspeakers located at the same elevations show the same trends in LRAs based on programme item. From these findings it is suggested that the loudspeakers be grouped based on their elevation. Figure 64 shows the mean absolute elevation LRE and 95% confidence intervals separated by programme item and actual loudspeaker elevation.

![Elevation LRE separated by loudspeaker elevation and programme item](image)

Figure 64 - Elevation LRE separated by loudspeaker elevation and programme item

From the plot it can be seen that there is large variance in accuracy between programme items for loudspeakers located at -35°, -9.8° and +52.9°. For loudspeakers near the median plane, at 0° and +9.8°, there are no significant differences between programme items. Programme item has a large effect on
loudbspeakers far from the equatorial plane and little effect on loudspeakers near the equatorial plane.

Where a significant difference between programme items is observable, the HF programme item is always localised most accurately. For both the -9.8° and +52.9° loudspeaker positions, the HF programme item is localisable with a similar elevation LRA to that seen with the loudspeakers near the equatorial plane, while the low and MF programme items show a significantly reduced LRA. For loudspeakers further from the median response angle there are larger differences between the programme items.

Figure 65 shows the mean absolute elevation LRE and 95% confidence intervals separated by programme item. This plot shows the overall trend of elevation LRA with programme item.

![Elevation LRE separated by programme item](image-url)

*Figure 65 - Elevation LRE separated by programme item*
The HF programme item is the most accurately localised of all the programme items. This programme item gives the listener static ILDs, which should not increase the elevation LRA because they should result in a cone of confusion for each ILD. The higher LRA to this programme item indicates that there may have been some pinna cues still present in the HF programme item.

From previous research it has been shown that using a MF programme item results in the least accurate interaural difference cues. The plot shows MF has lowest LRA both with and without head movements; this suggests that head movement cues result from dynamic interaural differences.

### 9.2.1.5 Loudspeaker

Figure 66 shows the mean absolute elevation LRE separated by loudspeaker number.

![Figure 66 - Absolute elevation LRE separated by loudspeaker number](image-url)

Error Bars: 95% CI
Loudspeakers numbers 4, 8, 12 and 15 were all located at -35°, the lowest loudspeaker elevation tested. This group of loudspeakers shows a significantly higher absolute elevation LRE than any other loudspeaker locations. Loudspeakers 2, 5, 10 and 13, all located at +35°, show the lowest absolute elevation LREs. Loudspeaker 13 has a slightly higher absolute elevation LRE than the other loudspeakers in this group probably because it is in the rearward quadrant and so has a higher pointing error than the frontal quadrant. The plot suggests that the main variation in absolute elevation LRE with loudspeaker is caused by the loudspeaker's elevation location. Therefore the loudspeakers were grouped based on their actual elevation location.

Figure 67 shows the mean absolute elevation LRE and 95% confidence intervals for each actual loudspeaker location in elevation.

![Graph showing mean absolute elevation LRE separated by actual loudspeaker elevation](image_url)
As shown in previous experiments, the most accurately reported source location is $+35^\circ$. This trend may have been due to the positioning of the acoustically transparent curtain or merely due to an increased elevation localisation acuity at this location; this is discussed further in Chapters 11 and 12.

A second ANOVA was conducted with the loudspeaker number group replaced by a loudspeaker elevation group. Having a loudspeaker elevation group combined the loudspeakers that were located at the same elevation angle. The ANOVA showed the loudspeaker elevation group to be significant ($F(5, 1632) = 15.841, p < 0.001, \eta^2 = 0.694$), which substantiates the suggestion that it was the loudspeaker elevation that caused the significance of the ‘Loudspeaker’ factor. A non-parametric Kruskal Wallis test was checked to verify this finding and also showed significant differences between groups ($H(5) = 273.622, p < 0.001$).

### 9.2.2 Azimuth LRA

In this section the azimuth component of the localisation tests is investigated. This will show how the head movement azimuth LRA is affected by the different programme items.

Front-back errors were removed from the azimuth LRE data by replacing $\varphi$ with $180^\circ - \varphi$ ($\varphi = \text{reported azimuth location}$) for all data exhibiting front-back reversal (azimuth reported in the wrong hemisphere). For a source location at precisely $90^\circ$ azimuth to the listener the localisation response was left unchanged. Front-back errors are discussed first, followed by analysis of the absolute azimuth LRE data.

#### 9.2.2.1 Front-Back Errors

Figure 68 shows a pie chart highlighting the proportion of trials that exhibited front-back confusions separated by movement condition and programme item.
In the movement condition all programme items exhibit a similar front-back error ratio approaching zero. Dynamic cues given by head movement are the main front-back discrimination cue and so when head movements are present the effects of spectral cues are negligible. Head movements reduce front-back errors to a similar level regardless of programme item.

In the no movement condition front-back errors are reduced as the centre frequency of the programme item is increased. The LF programme item shows an almost doubling in front-back errors when compared to the HF programme item. When head movements are not available the listener depends heavily on the high frequency content of the source. Pinna cues have previously been shown to reduce front-back errors and this plot substantiates that finding.
9.2.2.2 Absolute Azimuth LRE

The following sections analyse the absolute azimuth LRE with front-back confusions corrected. To use parametric tests such as ANOVA the data must be normally distributed. Using a Kolmogorov-Smirnov test across the independent variables ‘movement condition’, ‘programme item’ and ‘loudspeaker’, the error was found to be normal for 98 of the 102 conditions. Each normality check contained 16 data points averaged across the listener and repeat variables.

9.2.2.3 ANOVA

An ANOVA was conducted on the absolute azimuth LRE data with fixed factors ‘movement condition’, ‘loudspeaker’, ‘programme item’, ‘repeat’ and the random factor ‘listener’. All interactions up to three-way were modelled.
### Tests of Between-Subjects Effects

**Dependent Variable:** FB_Corrected_Absolute_Azimuth_Error

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>Hypothesis</td>
<td>118376.332</td>
<td>1</td>
<td>118376.332</td>
<td>459.154</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td>1804.698</td>
<td>7</td>
<td>257.814</td>
<td>4.247</td>
<td>.036</td>
<td>.378</td>
</tr>
<tr>
<td>Programme_Item</td>
<td>Hypothesis</td>
<td>856.795</td>
<td>2</td>
<td>428.398</td>
<td>4.247</td>
<td>.036</td>
</tr>
<tr>
<td>Error</td>
<td>1412.231</td>
<td>14</td>
<td>100.874</td>
<td>.000</td>
<td>.966</td>
<td></td>
</tr>
<tr>
<td>Movement_Condition</td>
<td>Hypothesis</td>
<td>30706.393</td>
<td>1</td>
<td>30706.393</td>
<td>201.076</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td>1068.972</td>
<td>7</td>
<td>152.710</td>
<td>.000</td>
<td>.966</td>
<td></td>
</tr>
<tr>
<td>Speaker</td>
<td>Hypothesis</td>
<td>15153.061</td>
<td>16</td>
<td>947.066</td>
<td>9.558</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td>11097.344</td>
<td>112</td>
<td>99.083</td>
<td>.000</td>
<td>.966</td>
<td></td>
</tr>
<tr>
<td>Repeat</td>
<td>Hypothesis</td>
<td>16.720</td>
<td>1</td>
<td>16.720</td>
<td>.112</td>
<td>.748</td>
</tr>
<tr>
<td>Error</td>
<td>1046.929</td>
<td>7</td>
<td>149.561</td>
<td>.000</td>
<td>.966</td>
<td></td>
</tr>
<tr>
<td>Listener</td>
<td>Hypothesis</td>
<td>1804.698</td>
<td>7</td>
<td>257.814</td>
<td>1.723</td>
<td>.401</td>
</tr>
<tr>
<td>Error</td>
<td>328.639</td>
<td>2.197</td>
<td>149.588</td>
<td>.000</td>
<td>.966</td>
<td></td>
</tr>
<tr>
<td>Movement_Condition</td>
<td>Hypothesis</td>
<td>1068.972</td>
<td>7</td>
<td>152.710</td>
<td>.751</td>
<td>.634</td>
</tr>
<tr>
<td>Listener</td>
<td>Error</td>
<td>10190.949</td>
<td>88.000</td>
<td>115.806</td>
<td>.000</td>
<td>.966</td>
</tr>
<tr>
<td>Programme_Item</td>
<td>Hypothesis</td>
<td>1412.231</td>
<td>14</td>
<td>100.874</td>
<td>.557</td>
<td>.867</td>
</tr>
<tr>
<td>Listener</td>
<td>Error</td>
<td>3569.801</td>
<td>19.722</td>
<td>181.004</td>
<td>.000</td>
<td>.966</td>
</tr>
<tr>
<td>Repeat * Listener</td>
<td>Hypothesis</td>
<td>1046.929</td>
<td>7</td>
<td>149.561</td>
<td>1.596</td>
<td>.297</td>
</tr>
<tr>
<td>Error</td>
<td>540.642</td>
<td>5.770</td>
<td>93.701</td>
<td>.000</td>
<td>.966</td>
<td></td>
</tr>
<tr>
<td>Speaker * Listener</td>
<td>Hypothesis</td>
<td>11097.344</td>
<td>112</td>
<td>99.083</td>
<td>.856</td>
<td>.783</td>
</tr>
<tr>
<td>Error</td>
<td>10190.949</td>
<td>88.000</td>
<td>115.806</td>
<td>.000</td>
<td>.966</td>
<td></td>
</tr>
<tr>
<td>Programme_Item</td>
<td>Hypothesis</td>
<td>1225.687</td>
<td>2</td>
<td>612.844</td>
<td>4.766</td>
<td>.026</td>
</tr>
<tr>
<td>Movement_Condition</td>
<td>Error</td>
<td>1800.254</td>
<td>14</td>
<td>128.590</td>
<td>.000</td>
<td>.966</td>
</tr>
<tr>
<td>Movement_Condition</td>
<td>Hypothesis</td>
<td>13.382</td>
<td>1</td>
<td>13.382</td>
<td>.155</td>
<td>.705</td>
</tr>
<tr>
<td>Repeat</td>
<td>Error</td>
<td>602.821</td>
<td>7</td>
<td>86.117</td>
<td>.000</td>
<td>.966</td>
</tr>
<tr>
<td>Movement_Condition</td>
<td>Hypothesis</td>
<td>8494.761</td>
<td>16</td>
<td>530.923</td>
<td>5.306</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td>11207.053</td>
<td>112</td>
<td>100.063</td>
<td>.000</td>
<td>.966</td>
<td></td>
</tr>
<tr>
<td>Programme_Item</td>
<td>Hypothesis</td>
<td>155.301</td>
<td>2</td>
<td>77.650</td>
<td>.998</td>
<td>.393</td>
</tr>
<tr>
<td>Repeat</td>
<td>Error</td>
<td>1089.012</td>
<td>14</td>
<td>77.787</td>
<td>.000</td>
<td>.966</td>
</tr>
<tr>
<td>Programme_Item</td>
<td>Hypothesis</td>
<td>8013.284</td>
<td>32</td>
<td>250.415</td>
<td>2.914</td>
<td>.000</td>
</tr>
<tr>
<td>Speaker</td>
<td>Error</td>
<td>19251.900</td>
<td>224</td>
<td>85.946</td>
<td>.000</td>
<td>.966</td>
</tr>
<tr>
<td>Speaker * Repeat</td>
<td>Hypothesis</td>
<td>893.517</td>
<td>16</td>
<td>55.845</td>
<td>1.358</td>
<td>.176</td>
</tr>
<tr>
<td>Error</td>
<td>4604.902</td>
<td>112</td>
<td>41.115</td>
<td>.000</td>
<td>.966</td>
<td></td>
</tr>
<tr>
<td>Programme_Item</td>
<td>Hypothesis</td>
<td>1800.254</td>
<td>14</td>
<td>128.590</td>
<td>2.310</td>
<td>.004</td>
</tr>
<tr>
<td>Movement_Condition</td>
<td>Error</td>
<td>46197.029</td>
<td>830</td>
<td>55.659</td>
<td>.000</td>
<td>.966</td>
</tr>
<tr>
<td>Movement_Condition</td>
<td>Hypothesis</td>
<td>602.821</td>
<td>7</td>
<td>86.117</td>
<td>1.547</td>
<td>.148</td>
</tr>
<tr>
<td>Repeat * Listener</td>
<td>Error</td>
<td>46197.029</td>
<td>830</td>
<td>55.659</td>
<td>.000</td>
<td>.966</td>
</tr>
<tr>
<td>Movement_Condition</td>
<td>Hypothesis</td>
<td>11207.053</td>
<td>112</td>
<td>100.063</td>
<td>1.798</td>
<td>.000</td>
</tr>
<tr>
<td>Speaker * Listener</td>
<td>Error</td>
<td>46197.029</td>
<td>830</td>
<td>55.659</td>
<td>.000</td>
<td>.966</td>
</tr>
<tr>
<td>Programme_Item</td>
<td>Hypothesis</td>
<td>1089.012</td>
<td>14</td>
<td>77.787</td>
<td>1.398</td>
<td>.147</td>
</tr>
<tr>
<td>Repeat * Listener</td>
<td>Error</td>
<td>46197.029</td>
<td>830</td>
<td>55.659</td>
<td>.000</td>
<td>.966</td>
</tr>
<tr>
<td>Programme_Item</td>
<td>Hypothesis</td>
<td>19251.900</td>
<td>224</td>
<td>85.946</td>
<td>1.544</td>
<td>.000</td>
</tr>
<tr>
<td>Speaker * Listener</td>
<td>Error</td>
<td>46197.029</td>
<td>830</td>
<td>55.659</td>
<td>.000</td>
<td>.966</td>
</tr>
<tr>
<td>Speaker * Repeat</td>
<td>Hypothesis</td>
<td>4604.902</td>
<td>112</td>
<td>41.115</td>
<td>.739</td>
<td>.978</td>
</tr>
<tr>
<td>Listener</td>
<td>Error</td>
<td>46197.029</td>
<td>830</td>
<td>55.659</td>
<td>.000</td>
<td>.966</td>
</tr>
</tbody>
</table>

**Table 28 - ANOVA of absolute azimuth LRE data**

---

9. Head Movement Localisation Without the Presence of Low Frequency ITDs
A histogram of the standardised residuals was plotted to check their normality. The plot shows a distribution that is slightly leptokurtic, but very close to normal.

![Histogram of absolute azimuth LRE ANOVA standardised residuals](image)

**Figure 69 - Histogram of absolute azimuth LRE ANOVA standardised residuals**

The main factors ‘repeat’ and ‘listener’ were found to be non-significant. The non-significance of the listener factor shows that azimuth LRA did not vary significantly between listeners. The non-significant repeat factor shows that listeners did not improve between the two tests.

The main factors ‘movement condition’, ‘loudspeaker’ and ‘programme item’ were all shown to significantly affect the azimuth LRA. As with elevation, further analysis of the findings of the ANOVA will be conducted in sections based on each of the main factors.
9.2.2.4 Movement Condition

The ANOVA showed the effect of movement condition to be significant with the largest effect size and highest F statics ($F(1, 1632) = 201.076$, $p < 0.001$, $\eta^2 = 0.966$). Figure 70 shows azimuth LRE plotted against movement condition.

![Figure 70 - Absolute azimuth LRE separated by movement condition](image.png)

This plot shows that azimuth LRA for the free movement condition was significantly higher than the no movement condition. Free movement more than halves the error when compared to the no movement condition. It is suggested that the main components of this higher LRA were improved interaural static cues once the listener had moved their head, improved pointing cues and dynamic cues given by listener head movement. Chapter 8 indicated that although all factors are significant, improved static cues have the largest effect on the azimuth LRA with head movements.
9.2.2.5 **Loudspeaker**

The main factor ‘loudspeaker’ was shown to significantly affect the azimuth LRA (F (16, 1632) = 9.558, p < 0.001, η² = 0.577). The second order interaction loudspeaker*movement condition was also shown to be significant (F (16, 1632) = 5.306, p < 0.001, η² = 0.431). This interaction was not discussed in the movement condition section as the f value of the main factor, movement condition, was much larger than this interaction f value. Figure 71 shows azimuth LRE plotted against loudspeaker number and movement condition.

![Graph showing Mean Absolute Azimuth LRE (degrees) against Loudspeaker number with Error Bars: 95% CI for Movement and No Movement conditions.]

*Figure 71 – Absolute azimuth LRE separated by movement condition and loudspeaker*

Only loudspeakers 1 and 17 resulted in non-significant differences in azimuth LRE between movement conditions. It is suggested that because these loudspeakers were located directly behind the calibration points the calibration points provided an anchor at which the listener could direct the pointer. This
suggests that the azimuth responses for these loudspeaker locations were biased. For all other loudspeaker locations the movement condition resulted in a significantly more accurate localisation response than the no movement condition.

9.2.2.6 Programme Item

The main factor ‘programme item’ was shown to significantly affect the absolute azimuth LRE ($F(2, 1632) = 4.247, p = 0.036, \eta^2 = 0.378$). The significance finding of this factor was much less than the other main factors in the ANOVA. Furthermore the second order interactions programme item*movement condition ($F(2, 1632) = 4.766, p = 0.026, \eta^2 = 0.405$), programme item*loudspeaker ($F(32, 1632) = 2.914, p < 0.001, \eta^2 = 0.294$) and third order interaction programme item*movement condition*loudspeaker ($F(32, 1632) = 4.484, p < 0.001, \eta^2 = 0.147$) were shown to be more significant. Figure 72 shows azimuth LRE plotted against programme item, movement condition and loudspeaker number.
In the movement condition there are overlapping confidence intervals between programme items for all loudspeaker locations. It is suggested that by moving the head the listener nullified any effects related to the programme item.

In both movement conditions, loudspeakers 1 and 17 show the localisation accuracy bias due to the position of the calibration points. In the no movement condition, there is no frontal source that shows significant differences between the programme items (although loudspeaker 5 is very close to having significant difference between then low and medium frequency programme items).

For loudspeaker 16, the HF programme item has a much higher error than the other two programme items. This loudspeaker location was at the highest elevation tested and behind the listener. All loudspeaker locations had front-
back errors removed so the differences cannot be contributed to these. When viewing the raw data file it was seen that listeners 1, 2 and 8 all made errors of greater than 60° in azimuth for this programme item. Interestingly for these trials the listeners gave extremely accurate elevation responses with a mean elevation LRE for the 5 erroneous responses of 2.35°. As stated previously small errors in pointing are accentuated when the source is located at a higher elevation. Alternatively it may be that the listeners pressed to 'log location' button as they returned to their starting position and not when their head was pointed at the source.

Figure 73 shows that absolute azimuth LRE separated by programme item.

![Figure 73 - Absolute azimuth LRE separated by programme item](image)

There is no significant difference in absolute azimuth LRE between the LF and MF frequency programme items. The HF programme item appears to show a slightly higher mean error than the other two programmes. It is hypothesised that this difference was caused by the unusual response in the no movement
condition to loudspeaker 16 that was highlighted in the previous section. When the overall data are plotted with loudspeaker 16 excluded (Figure 74), all programme items show a very similar response (there are no significant differences between programme items). Therefore, it can be concluded that, overall, the azimuth LRA was unaffected by the frequency range of the programme item.

![Graph showing mean absolute azimuth LRE separated by programme item with loudspeaker 16 excluded](image)

**Figure 74 - Absolute azimuth LRE separated by programme item with loudspeaker 16 excluded**

### 9.3 Conclusions

A localisation experiment was conducted that compared static and free head movement LRA for band-limited programme items in three frequency ranges, LF (centred at 500 Hz), MF (centred at 2 kHz) and HF (centred at 6 kHz). These frequency ranges were used to show whether head movement localisation was
possible without the presence of low frequency ITDs and, more generally, to show the localisation cues that affect LRA with head movement.

9.3.1 Elevation LRA

In elevation, significant improvements in LRA were shown with head movement for the HF programme item. This shows that the presence of low frequency ITDs is not necessary for improved LRA with head movement. This finding opposes that of Perrett and Noble [1997] described in Chapter 5.

In elevation, head movements improve LRA for all the programme items tested. Furthermore, head movements improve LRA to a similar degree for all programme items. This was shown by the absence of a significant programme item*movement condition interaction. This finding shows that head movement cues are consistent over frequency and do not appear to have an optimum frequency range. Thus head movements can be used when other cues, such as pinna cues, fail.

Without the presence of pinna cues, using the LF programme item, head movements still improve LRA. Therefore, head movement cues are not solely dependent on pinna cues or some kind of dynamic spectral variation. At least some head movement cues must take the form of dynamic interaural time and level difference cues, as suggested by Wallach [1938].

9.3.2 Azimuth LRA

Head movements significantly improve azimuth LRA. By moving their head, the listener moves the source into their most accurate area of localisation, thus improving azimuth LRA. It is suggested that dynamic interaural cues given by head movement allow a further increase in LRA.

Head movements improve LRA regardless of source azimuth. However, there are larger differences in error with the movement condition for sources outside of the listener's field of fixation because of pointing accuracy differences.
Head movements are the main factor in front-back discrimination. When head movements are allowed, the proportion of front-back errors are greatly reduced, regardless of the programme item being used. However, when the listener is not permitted to move their head, the presence of high frequency content reduces front-back errors. This is because there are some spectral cues present in the signal for the high frequency programme item. If head movement is not allowed, azimuth LRA, excluding front-back confusions, is unaffected by the frequency range of the source programme item.

### 9.3.3 Further Work

Further study into elevation LRA is necessary to find out whether the improved LRA with head movement is due to the static cues given when a listener reaches a new listening azimuth or the dynamic cues created as a listener moves their head. This will show whether the improvement due to head movement results from a static or dynamic cue. When localising a source, listeners rotate their head towards the source location (Chapter 2). By studying elevation LRA on vertical planes at various azimuth angles in relation to the listener’s head position it will be possible to show how static localisation cues vary with head movement. A study investigating static elevation localisation is described in Chapters 11 and 12.
The effect of programme item bandwidth on the localisation with head movement

Chapters 6 and 8 showed that using a ‘1/2 octave band of high frequency noise with a low pass filtered noise component’ programme item significantly reduces a listener’s pinna cue LRA when compared to the localisation of broadband white noise. It was also shown that head movements improve LRA more using this programme item than they do when using broadband noise. These findings lead to a number of questions concerning the bandwidth of the programme item:

- **What is the minimum bandwidth of the programme item that is required to allow improvement in LRA due to head movement?**
- **How does LRA change as the bandwidth of the programme item is reduced and how does allowing a listener to move their head affect this change in LRA?**

These questions will be answered by comparing the movement and no movement LRAs of listeners for various bandwidth noise programme items.

10.1 Experiment Setup

The setup of the previous head movement experiments was replicated in terms of loudspeaker location, response method and movement condition. The only factor that was changed from previous experiments was the programme item. A brief reminder of the general experiment setup will be given here, followed by an in-depth discussion of the programme item factor.

10.1.1 General Experiment Setup and Method

The experiment used seventeen Genelec 8020a loudspeakers, positioned on a hemisphere to the right of the listener, spread from 0° to 180° in azimuth and -35° to +55° in elevation. The loudspeakers were positioned at the coordinates of the vertices of a truncated icosahedron. The loudspeaker numbers, coordinates and angles used in the experiment are given in Table 40 in Appendix
B. To avoid any biases due to visual cues, the locations of the loudspeakers were hidden from view by an acoustically transparent but visually opaque curtain.

Eight listeners auditioned all programme item and movement condition combinations. There were 7 IoSR postgraduate students and 1 undergraduate music student at the University of Surrey; all were aged between 18 and 35 and had no reported hearing conditions.

The listener was asked to report the perceived location of each of a series of signals, each replayed via a single selected loudspeaker. The listener responded using a laser pointing method. The laser was attached to the head and positioned to point out from between the eyes via an elasticated band that was wrapped around the listener's head. A Polhemus head tracker monitored the listener's head movement to six degrees of freedom and each perceived source location, as indicated by the listener, was calculated using a triangulation technique similar to the method described by Frank et al. [8]. The radius from the central listening position to the loudspeakers was used to create a projected sphere and the perceived location was taken to be the point at which the line of the laser pointer pierced this sphere.

10.1.2 Programme Items

In the test, four programme items of varying bandwidths were used: three bandpass filtered noise programme items, which were called ‘Octave’, ‘Half-Octave’ and ‘Quarter-Octave’, and one ‘Sine Tone’ programme item. All programmes were centred at 6 kHz. The filter characteristics used to create these programme items were:

Octave

- **Passband**: 4243 - 8485 Hz
- **Transition bands**: 4063 - 4243 and 8485 - 8861 Hz
- **Stopband**: 0 - 4063 Hz and 8861 - 20000 Hz
10. The effect of programme item bandwidth on the localisation with head movement

Half-Octave

- **Passband**: 5045 - 7135 Hz
- **Transition bands**: 4831 - 5045 and 7135 - 7451 Hz
- **Stopband**: 0 - 4831 Hz and 7451 - 20000 Hz

Quarter-Octave

- **Passband**: 5502 - 6543 Hz
- **Transition bands**: 5269 - 5502 and 6543 - 6833 Hz
- **Stopband**: 0 - 5269 Hz and 6833 - 20000 Hz

In each case the stopband region was 100dB lower than the passband. The frequency spectra of the 'Octave' and 'Half-Octave' programme items are shown in Figure 75 and Figure 76.

![Figure 75 - Frequency spectrum of 'Octave' programme item](image-url)
The effect of programme item bandwidth on the localisation with head movement

The tonal programme item was used to test Wallach’s assertion that pure ILDs used in conjunction with head movement can be used to allow listeners to localise in elevation. Previous studies have suggested that the quarter-octave and half-octave bandwidth programme items do not excite the listener’s pinna cues [Langendijk and Bronkhorst 2002] and the findings of Chapters 6 and 8 support this suggestion. These programme items were, therefore, also used to test whether head movements improve elevation LRA when using interaural difference cues alone.

It was anticipated that elevation LRA without head movement would remain unchanged as the bandwidth of the programme item was increased until pinna cues were present, at which point elevation LRA would improve dramatically. It has been shown that filtering out a half-octave portion of a full bandwidth white noise programme item between 4 kHz and 16 kHz does not significantly reduce the listener’s static LRA [Langendijk and Bronkhorst 2002], while filtering out an octave portion of the programme item in the same frequency range significantly reduces LRA. This suggests that an octave programme item gives vital pinna cues and so the dramatic increase in LRA would occur between the octave and half-octave regions. The sine wave, quarter-octave noise and half-octave noise programme items should all produce the same poor static elevation LRA, while the octave bandwidth noise might show higher LRA.
With head movements it was anticipated that a significant improvement in LRA would be shown when compared to the static movement condition for some of the programme items. If Wallach’s dynamic interaural differences assertion is correct then head movements should significantly improve LRA for all programme items down to a sine wave. However, a real room might lower LRA for the narrower bandwidth programme items. For example, since the room is non-anechoic, room resonances are likely to confound the ILDs and ITDs used to localise the sine wave programme item. The effects of reverberant rooms on ITD and ILD localisation was described in Chapter 4. If head movements are seen to improve LRA at a bandwidth providing negligible pinna cues, then it can be asserted that pinna cues and head movement cues are distinct, and that increased LRA with head movement is not due to improved static pinna cues.

10.2 Results

The localisation results analysis is split into the two location components, elevation (Section 10.2.1) and azimuth (Section 10.2.2). Programme item, movement condition and loudspeaker location were the main factors used to study the listeners’ LRA.

10.2.1 Elevation LRA

In this section the elevation LRA component of the test is investigated by analysing the absolute elevation LRE of the listener. The goal was to find out how the bandwidth of the programme item affected the improvement in elevation LRA offered by head movement.

10.2.1.1 Normality Checks

To use parametric tests such as ANOVA the data should be normally distributed. Using a Kolmogorov-Smirnov test across the independent variables ‘movement condition’, ‘programme item’ and ‘loudspeaker’, the distribution of the absolute elevation LRE data was found to be normal for 109 of the 136 conditions. Each normality test contained 16 data points averaged across the listener and repeat
variables. It is suggested that for each significant factor in the ANOVA, a secondary Kruskal-Wallis test should be conducted to confirm the findings.

### 10.2.1.2 ANOVA

Table 29 shows the results of an ANOVA conducted on the absolute elevation data with fixed factors 'movement condition', 'loudspeaker', 'programme item', 'repeat' and the random factor 'listener'. All interactions up to three-way were modelled.

<table>
<thead>
<tr>
<th>Source</th>
<th>Hypothesis</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td>1311058.78</td>
<td></td>
<td>2173.999</td>
<td>603.063</td>
<td>.000</td>
<td>.989</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>7</td>
<td></td>
<td>2173.999</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker</td>
<td>Hypothesis</td>
<td>411090.446</td>
<td>16</td>
<td>25693.153</td>
<td>41.502</td>
<td>.000</td>
<td>.856</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>112</td>
<td></td>
<td>619.087</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Programme_Item</td>
<td>Hypothesis</td>
<td>58274.804</td>
<td>3</td>
<td>19424.935</td>
<td>25.682</td>
<td>.000</td>
<td>.786</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>21</td>
<td></td>
<td>756.360</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movement.Condition</td>
<td>Hypothesis</td>
<td>3351.913</td>
<td>1</td>
<td>3351.913</td>
<td>4.722</td>
<td>.066</td>
<td>.403</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>7</td>
<td></td>
<td>709.920</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repeat</td>
<td>Hypothesis</td>
<td>514.203</td>
<td>1</td>
<td>514.203</td>
<td>1.565</td>
<td>.251</td>
<td>.183</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>7</td>
<td></td>
<td>328.544</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Listener</td>
<td>Hypothesis</td>
<td>15217.996</td>
<td>7</td>
<td>2173.999</td>
<td>1.531</td>
<td>.230</td>
<td>.415</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>15.079</td>
<td></td>
<td>1420.192</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movement.Condition * Listener</td>
<td>Hypothesis</td>
<td>4969.440</td>
<td>7</td>
<td>709.920</td>
<td>1.340</td>
<td>.276</td>
<td>.284</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>23.609</td>
<td></td>
<td>529.933</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Programme_Item * Listener</td>
<td>Hypothesis</td>
<td>15883.553</td>
<td>21</td>
<td>756.360</td>
<td>1.800</td>
<td>.071</td>
<td>.569</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>28.680</td>
<td></td>
<td>420.105</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repeat * Listener</td>
<td>Hypothesis</td>
<td>2299.806</td>
<td>7</td>
<td>328.544</td>
<td>1.865</td>
<td>.559</td>
<td>.334</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>12.058</td>
<td></td>
<td>379.964</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker * Listener</td>
<td>Hypothesis</td>
<td>69337.788</td>
<td>112</td>
<td>619.087</td>
<td>1.598</td>
<td>.006</td>
<td>.608</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>115.609</td>
<td></td>
<td>387.396</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Programme_Item * Movement_Cond</td>
<td>Hypothesis</td>
<td>6634.673</td>
<td>3</td>
<td>2211.558</td>
<td>9.528</td>
<td>.000</td>
<td>.576</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>21</td>
<td></td>
<td>232.114</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movement_Condition * Repeat</td>
<td>Hypothesis</td>
<td>67.253</td>
<td>1</td>
<td>67.253</td>
<td>.279</td>
<td>.614</td>
<td>.038</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>7</td>
<td></td>
<td>241.265</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker * Movement_Condition</td>
<td>Hypothesis</td>
<td>5539.716</td>
<td>16</td>
<td>346.232</td>
<td>1.060</td>
<td>.401</td>
<td>.132</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>112</td>
<td></td>
<td>326.593</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Programme_Item * Repeat</td>
<td>Hypothesis</td>
<td>665.557</td>
<td>3</td>
<td>221.852</td>
<td>.828</td>
<td>.498</td>
<td>.106</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>21</td>
<td></td>
<td>267.962</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker * Programme_Item</td>
<td>Hypothesis</td>
<td>42534.178</td>
<td>48</td>
<td>886.129</td>
<td>4.662</td>
<td>.000</td>
<td>.400</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>336</td>
<td></td>
<td>190.067</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker * Repeat</td>
<td>Hypothesis</td>
<td>2392.374</td>
<td>16</td>
<td>149.523</td>
<td>1.062</td>
<td>.400</td>
<td>.132</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>112</td>
<td></td>
<td>140.775</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Programme_Item * Movement_Cond * Listener</td>
<td>Hypothesis</td>
<td>4874.398</td>
<td>21</td>
<td>232.114</td>
<td>1.719</td>
<td>.023</td>
<td>.029</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>1189</td>
<td></td>
<td>135.019</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movement_Condition * Repeat * Listener</td>
<td>Hypothesis</td>
<td>1688.856</td>
<td>7</td>
<td>241.265</td>
<td>1.787</td>
<td>.086</td>
<td>.010</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>1189</td>
<td></td>
<td>135.019</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker * Movement_Condition * Listener</td>
<td>Hypothesis</td>
<td>3657.8364</td>
<td>112</td>
<td>326.593</td>
<td>2.419</td>
<td>.000</td>
<td>.186</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>1189</td>
<td></td>
<td>135.019</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Programme_Item * Repeat * Listener</td>
<td>Hypothesis</td>
<td>5627.204</td>
<td>21</td>
<td>267.962</td>
<td>1.985</td>
<td>.005</td>
<td>.034</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>1189</td>
<td></td>
<td>135.019</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker * Programme_Item * Listener</td>
<td>Hypothesis</td>
<td>6386.2548</td>
<td>336</td>
<td>190.067</td>
<td>1.408</td>
<td>.000</td>
<td>.285</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>1189</td>
<td></td>
<td>135.019</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker * Repeat * Listener</td>
<td>Hypothesis</td>
<td>15766.790</td>
<td>112</td>
<td>140.775</td>
<td>1.043</td>
<td>.367</td>
<td>.089</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>1189</td>
<td></td>
<td>135.019</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The main factors ‘loudspeaker’ (F (16, 2176) = 41.502, \( \eta^2 = 0.856 \)) and ‘programme item’ (F (3, 2176) = 25.682, \( \eta^2 = 0.786 \)) were shown to be significant. These factors are discussed under individual headings in the following sections.

As shown in previous experiments, the ‘listener’ and ‘repeat’ factors were non-significant. This shows that across all programme items, listeners produced a similar elevation LRE. Interestingly the movement condition factor was non-significant (F (2, 2176) = 25.682, \( p = 0.066, \eta^2 = 0.403 \)). It is suggested that the variance cause by the loudspeaker and programme item factors masked any overall difference in variance between movement conditions. The programme item*movement condition interaction was significant (F (3, 2176) = 9.528, \( p < 0.001, \eta^2 = 0.576 \)) and this will be discussed further in the programme item section. A normality plot of the ANOVA's standardised residuals indicated that the model was a good fit of the data (Figure 77).

<table>
<thead>
<tr>
<th>Source</th>
<th>Hypothesis</th>
<th>328.981</th>
<th>3</th>
<th>109.660</th>
<th>( \eta^2 )</th>
<th>.812</th>
<th>.487</th>
<th>.002</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error</td>
<td>160537.783</td>
<td>1189</td>
<td>135.019</td>
<td>( q )</td>
<td>1189</td>
<td>160537.783</td>
<td>( q )</td>
<td>1189</td>
</tr>
<tr>
<td><strong>Programme_Item * Movement_Condition * Repeat</strong></td>
<td>Hypothesis</td>
<td>768.164</td>
<td>16</td>
<td>48.010</td>
<td>.356</td>
<td>.991</td>
<td>.005</td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>160537.783</td>
<td>1189</td>
<td>135.019</td>
<td>( q )</td>
<td>1189</td>
<td>160537.783</td>
<td>( q )</td>
<td>1189</td>
</tr>
<tr>
<td><strong>Programme_Item * Movement_Condition * Speaker</strong></td>
<td>Hypothesis</td>
<td>10042.702</td>
<td>48</td>
<td>209.223</td>
<td>1.550</td>
<td>.010</td>
<td>.059</td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>160537.783</td>
<td>1189</td>
<td>135.019</td>
<td>( q )</td>
<td>1189</td>
<td>160537.783</td>
<td>( q )</td>
<td>1189</td>
</tr>
</tbody>
</table>

*Table 29 - ANOVA of absolute elevation LRE*
The results of the ANOVA indicated that the loudspeaker factor had the most significant effect on the elevation LRE data. A Kruskal-Wallis non-parametric test confirmed the significance of the loudspeaker factor given by the ANOVA (H (16) = 793.069, p < 0.001). Figure 78 shows a plot of the absolute elevation LRE separated by loudspeaker number.
Loudspeakers 4, 8, 12, and 15 have a significantly higher elevation LRE than all other loudspeaker locations. These loudspeakers were all located at -35°, the lowest loudspeaker elevation in the test. Loudspeakers 2, 5, 10 and 13 are the four loudspeakers with the lowest absolute elevation LRE shown in the plot. These loudspeakers were all located at +35°. The plot shows that the main cause of the significant variance shown by the ANOVA for the factor 'loudspeaker number' was the elevation of the loudspeakers. Previous experiments agree that the elevation of the loudspeakers is the relevant element of the factor 'loudspeaker number'. Figure 79 shows a graph of the absolute elevation LRE plotted against the loudspeaker elevation.
Figure 79 – Absolute elevation LRE separated by actual loudspeaker elevation

Loudspeakers were grouped together by elevation to give a ‘loudspeaker elevation number’ factor. A second ANOVA was conducted with the ‘loudspeaker’ factor replaced by loudspeaker elevation number. The ANOVA showed that the loudspeaker elevation number group was highly significant ($F (5, 1632) = 67.113, p < 0.001, \eta^2 = 0.904$), which substantiates the suggestion that it was the loudspeaker elevation that was the relevant element of the ‘Loudspeaker’ factor. A non-parametric Kruskal-Wallis test verified this finding and also showed significant differences between groups ($H (5) = 779.464, p < 0.001$). To find the cause of the ‘loudspeaker elevation number’ dependent elevation LRE a further experiment would be required, but such an experiment is tangential to the head movement focused research presented in this thesis.

Here are three hypotheses that may explain why the elevation LRA is greatest for the loudspeakers located at +35°:
1. The location of the acoustically transparent curtain biased the listeners’ responses upwards from the equatorial plane.

2. The listener’s area of highest elevation localisation acuity occurs when the loudspeaker is located at approximately +35°.

3. The ‘pitch height effect’, whereby narrow bandwidth sources are localised in elevation based on their pitch and not their actual location, caused sources to be located at +35°. The pitch height effect does suggest that sources of 6 kHz should be localised at approximately +35°. However, the upward bias has also been noted for the full bandwidth sources, which this hypothesis does not explain.

10.2.1.4 Programme Item

As discussed, the programme item factor was shown to be significant in the ANOVA. A Kruskal-Wallis non-parametric test confirmed the significance of the programme item factor given by the ANOVA (H (3) = 114.031, p < 0.001). However the factor interaction programme item*movement condition interaction was also shown to be highly significant (F (3, 2176) = 9.528, p < 0.001, ηp² = 0.576) with a high F value and large effect size. Figure 80 shows the absolute elevation LRE separated by programme item and movement condition.
The effect of programme item bandwidth on the localisation with head movement

For the octave noise programme item there is a significant difference between the movement and no movement conditions. This finding was confirmed by a targeted comparison between the no movement and forced movement conditions using a non-parametric Mann-Whitney test: Octave Noise (U = 27840.0, P < 0.001). For all other programme items the difference between movement conditions is not significant. By allowing head movement the mean absolute elevation LRE for the octave programme item is nearly halved. This suggests that for head movement to be effective requires a minimum programme item bandwidth of one octave.

There are two possible explanations for why there was a halving in error for the octave bandwidth noise:

Figure 80 – Absolute elevation LRE separated by programme item and movement condition
• By moving their head the listener gained dynamic cues that allowed them to localise more accurately.
• By moving their head the listener moved the source into a more accurate area of static elevation localisation.

This finding requires further investigation; Chapter 12 describes an experiment that separates static and dynamic cues for elevation localisation for reduced-bandwidth sources.

In the no movement condition all the noise programme items produced a lower elevation LRE than the tonal programme item. The room was not anechoic and while listening to the tonal programme item the resonances of the room were pronounced. It is suggested that these room resonances may have confounded ILDs. It is suggested that this effect may also be apparent in the azimuth localisation section and that the effect may be more pronounced.

In the no movement condition, there were no significant differences between the noise programme items. The listener was able to report elevation just as accurately using the quarter-octave noise as the octave noise. This finding opposes the hypothesis stated in Section 10.1.2 that pinna cues might allow higher LRA for the octave programme item; pinna cues were equally degraded for all programme items. Chapters 6 and 8 showed that adding a half-octave band of 6 kHz white noise to a low-pass-filtered noise programme did not significantly improve elevation LRA. Since this experiment shows that there is no difference in elevation LRA between any of the noise programmes, it can be concluded that, in the no movement condition, pinna cues require bandwidth of greater than an octave to allow an improvement in elevation LRA.

Once head movement was permitted, significant differences appeared between the listeners’ elevation LRA for octave noise and that for the other two noise programme items. This finding shows that head movement cues do not improve LRA for noise based programme items with a half-octave bandwidth or less. Furthermore it shows further evidence that head movement cues are independent of pinna cues.
10.2.2 Azimuth LRA

In this section the absolute azimuth LRE component of the test is investigated. The goal of this section is to find the effect of programme item bandwidth on the azimuth LRA, considering both front-back errors and azimuth LRA with front-back errors removed.

Front-back errors were removed from the azimuth LRE data by replacing $\varphi$ with $180^\circ - \varphi$ ($\varphi = \text{reported azimuth location}$) for all data exhibiting front-back reversal (azimuth reported in the wrong hemisphere). For a source location at precisely $90^\circ$ azimuth to the listener the localisation response was left unchanged. Front-back errors are discussed first, followed by analysis of the absolute azimuth LRE data.

10.2.2.1 Front-Back Errors

Front-back errors were analysed separately from the absolute azimuth LRE results. Figure 81 shows the proportion of front-back errors for each trial separated by programme item and movement condition.
The free movement condition is shown to significantly reduce the number of front-back confusions. The tone programme item has the highest number front-back confusions in the free movement condition; however, the number of front-back confusions for the tone programme item in the movement condition is smaller than the octave programme item in the no movement condition (which has the lowest number of front-back confusions for the no movement condition). This finding indicates that head movements are the most significant factor in reducing front-back confusions. For the noise programmes, head movements reduce front-back confusions to a similar level regardless of bandwidth.

In the no movement condition for the tonal programme item, front-back confusions occur approximately 50% of the time. This finding indicates that for this programme item, listeners are completely confused as to whether the source is in front of or behind them. For this programme item the listener gained no
head movement or spectral cues and so had no front-back discrimination cues. As the bandwidth of the noise programme item is increased, the front-back error rate is reduced.

10.2.2.2 Absolute Azimuth LRE

Following the removal of the front-back errors, the azimuth LRA was analysed. To use parametric tests such as ANOVA the data must be normally distributed. Using a Kolmogorov-Smirnov test across the independent variables ‘movement condition’, ‘programme item’ and ‘loudspeaker’, the distribution of the absolute azimuth LRE data was found to be normal for 99 of the 136 conditions. Each normality test contained 16 data points averaged across the listener and repeat variables. It is suggested that for each significant factor in the ANOVA, a secondary Kruskal-Wallis test should be conducted to confirm the findings.

10.2.2.3 ANOVA

An ANOVA was conducted on the absolute azimuth data with fixed factors ‘movement condition’, ‘loudspeaker’, ‘programme item’, ‘repeat’ and the random factor ‘listener’. All interactions up to three-way were modelled. Table 30 shows the significant factors given by the ANOVA.
The main factors movement condition, programme item and loudspeaker were all shown to be significant. Each of these factors will be discussed individually in the following analysis sections. A plot of the standardised residuals was examined and the ANOVA was shown to be a good fit of the data.

Table 30 – ANOVA of absolute azimuth LRE

<table>
<thead>
<tr>
<th>Source</th>
<th>Hypothesis</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>Hypothesis</td>
<td>418187.071</td>
<td>1</td>
<td>418187.071</td>
<td>1125.179</td>
<td>.000</td>
<td>.994</td>
</tr>
<tr>
<td>Error</td>
<td>2601.640</td>
<td>7</td>
<td>371.663</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movement_Condition</td>
<td>Hypothesis</td>
<td>56571.786</td>
<td>1</td>
<td>56571.786</td>
<td>205.026</td>
<td>.000</td>
<td>.967</td>
</tr>
<tr>
<td>Error</td>
<td>1931.478</td>
<td>7</td>
<td>275.925</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker</td>
<td>Hypothesis</td>
<td>13253.071</td>
<td>16</td>
<td>828.317</td>
<td>2567</td>
<td>.002</td>
<td>.268</td>
</tr>
<tr>
<td>Error</td>
<td>36147.069</td>
<td>112</td>
<td>322.742</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Programme_Item</td>
<td>Hypothesis</td>
<td>40472.633</td>
<td>3</td>
<td>13490.878</td>
<td>46.162</td>
<td>.000</td>
<td>.868</td>
</tr>
<tr>
<td>Error</td>
<td>6137.316</td>
<td>21</td>
<td>292.253</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repeat</td>
<td>Hypothesis</td>
<td>13.546</td>
<td>1</td>
<td>13.546</td>
<td>.120</td>
<td>.739</td>
<td>.017</td>
</tr>
<tr>
<td>Error</td>
<td>789.230</td>
<td>7</td>
<td>112.747</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Listener</td>
<td>Hypothesis</td>
<td>2601.640</td>
<td>7</td>
<td>371.663</td>
<td>1.573</td>
<td>.443</td>
<td>.847</td>
</tr>
<tr>
<td>Error</td>
<td>470.832</td>
<td>1.993</td>
<td>236.212</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movement_Condition</td>
<td>Hypothesis</td>
<td>1931.478</td>
<td>7</td>
<td>275.925</td>
<td>.595</td>
<td>.753</td>
<td>.153</td>
</tr>
<tr>
<td>Listener</td>
<td>Hypothesis</td>
<td>10727.296</td>
<td>23.147</td>
<td>463.438</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>9361.344</td>
<td>23.957</td>
<td>390.759</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Programme_Item</td>
<td>Hypothesis</td>
<td>6137.316</td>
<td>21</td>
<td>292.253</td>
<td>.748</td>
<td>.748</td>
<td>.396</td>
</tr>
<tr>
<td>Listener</td>
<td>Hypothesis</td>
<td>789.230</td>
<td>7</td>
<td>112.747</td>
<td>.632</td>
<td>.720</td>
<td>.372</td>
</tr>
<tr>
<td>Error</td>
<td>1331.841</td>
<td>7.463</td>
<td>178.451</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker</td>
<td>Hypothesis</td>
<td>36147.069</td>
<td>112</td>
<td>322.742</td>
<td>1.204</td>
<td>.166</td>
<td>.555</td>
</tr>
<tr>
<td>Listener</td>
<td>Hypothesis</td>
<td>29032.230</td>
<td>108.322</td>
<td>268.017</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>7155.333</td>
<td>21</td>
<td>340.730</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repeat</td>
<td>Hypothesis</td>
<td>5.847</td>
<td>1</td>
<td>5.847</td>
<td>.038</td>
<td>.850</td>
<td>.005</td>
</tr>
<tr>
<td>Error</td>
<td>1068.997</td>
<td>7</td>
<td>152.714</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movement_Condition</td>
<td>Hypothesis</td>
<td>8582.653</td>
<td>16</td>
<td>536.416</td>
<td>2.626</td>
<td>.002</td>
<td>.273</td>
</tr>
<tr>
<td>Speaker</td>
<td>Hypothesis</td>
<td>22874.810</td>
<td>112</td>
<td>204.239</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>1194.889</td>
<td>3</td>
<td>398.296</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Programme_Item</td>
<td>Hypothesis</td>
<td>25692.046</td>
<td>48</td>
<td>535.251</td>
<td>3.321</td>
<td>.000</td>
<td>.322</td>
</tr>
<tr>
<td>Error</td>
<td>7155.333</td>
<td>21</td>
<td>340.730</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repeat</td>
<td>Hypothesis</td>
<td>2182.567</td>
<td>16</td>
<td>136.410</td>
<td>.997</td>
<td>.466</td>
<td>.125</td>
</tr>
<tr>
<td>Error</td>
<td>15328.913</td>
<td>112</td>
<td>136.865</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movement_Condition</td>
<td>Hypothesis</td>
<td>7155.333</td>
<td>21</td>
<td>340.730</td>
<td>2.909</td>
<td>.000</td>
<td>.049</td>
</tr>
<tr>
<td>Listener</td>
<td>Hypothesis</td>
<td>139258.673</td>
<td>1189</td>
<td>117.123</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>1068.997</td>
<td>7</td>
<td>152.714</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repeat</td>
<td>Hypothesis</td>
<td>22874.810</td>
<td>112</td>
<td>204.239</td>
<td>1.744</td>
<td>.000</td>
<td>.141</td>
</tr>
<tr>
<td>Listener</td>
<td>Hypothesis</td>
<td>139258.673</td>
<td>1189</td>
<td>117.123</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>2585.461</td>
<td>21</td>
<td>123.117</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Programme_Item</td>
<td>Hypothesis</td>
<td>139258.673</td>
<td>1189</td>
<td>117.123</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>139258.673</td>
<td>1189</td>
<td>117.123</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker</td>
<td>Hypothesis</td>
<td>54148.778</td>
<td>336</td>
<td>161.157</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Listener</td>
<td>Hypothesis</td>
<td>139258.673</td>
<td>1189</td>
<td>117.123</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>139258.673</td>
<td>1189</td>
<td>117.123</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repeat</td>
<td>Hypothesis</td>
<td>15328.913</td>
<td>112</td>
<td>136.865</td>
<td>1.169</td>
<td>.120</td>
<td>.099</td>
</tr>
<tr>
<td>Listener</td>
<td>Hypothesis</td>
<td>139258.673</td>
<td>1189</td>
<td>117.123</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>219.444</td>
<td>3</td>
<td>73.148</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The effect of programme item bandwidth on the localisation with head movement was examined and the ANOVA was shown to be a good fit of the data.
The main factors ‘repeat’ and ‘listener’ were found to be non-significant. The non-significance of the repeat factor means that the listener's response did not vary significantly when they were asked to repeat the experiment. Thus, there was no improvement in response as the listener became familiar with the task. These findings concur with previous localisation test results.

### 10.2.2.4 Movement Condition

The movement condition factor was shown to be highly significant with a large F factor and partial eta squared value \( F(1,4437) = 205.026, p < 0.001, \eta^2 = .967 \). The significant differences between the absolute elevation LRE of the free and no movement conditions were verified using a Mann-Whitney test \( U = 349811.5, P < 0.001 \).

From the ANOVA it is suggested that movement condition is the most important factor in azimuth LRA. The absolute azimuth LRE is plotted against movement condition in Figure 82.
When the listener is free to move their head their azimuth LRE is significantly reduced. The much higher F value of the movement condition factor suggests that its effect overrides any factor interactions. This suggests that the increase in azimuth LRA with head movement is not due to pointing accuracy as this would be highlighted by a high F value in the loudspeaker*movement condition. Steven and Newman [1936] showed that azimuth localisation was most accurate on the median plane. Furthermore it was observed that most listeners point their heads towards the source, putting the source on their median plane. Once again this experiment suggests that the observed improvement in azimuth LRA with head movement is due to static cues created when the listener moves the source into their most accurate area of localisation.

10.2.2.5 Programme Item

The main factor ‘programme item’ was shown to be significant using the ANOVA ($F (2,4437) = 46.162, p < 0.001, \eta^2 = .868$). A Kruskal-Wallis non-parametric
test confirmed the significance of the programme item factor given by the ANOVA \((H (3) = 176.52, p < 0.001)\). The absolute azimuth LRE is plotted against movement condition in Figure 83.

![Figure 83 - Absolute azimuth LRE separated by programme item](image)

The tonal programme item is shown to have approximately twice the azimuth LRE of the three noise programmes. Following the test, listeners suggested that their LRA was reduced by room modes. The room modes caused the listener’s interaural level differences, the main azimuth localisation cue at high frequencies, to be altered and to no longer indicate the source azimuth. It is suggested that this is the main reason for the reduced azimuth LRA for the tone programme item.

A Bonferroni post-hoc test showed significant differences between the absolute azimuth LRE of the octave and quarter-octave bandwidth noise programme items. From the plot it can be seen that this improvement was approximately 1-
2°. From this plot it is unclear what cue may have caused this improvement in LRA. Figure 84 shows the absolute azimuth LRE plotted against programme item and movement condition.

![Plot showing mean absolute azimuth LRE](image)

**Figure 84 - Absolute azimuth LRE plotted against programme item and movement condition**

The no movement condition shows no significant differences between the noise programme items; it is only when the listeners were free to move their heads that significant differences occur between the noise programmes. This indicates that this significant difference is head movement dependent. It is unclear whether this improvement is due to dynamic cues or improved static cues.

**10.2.2.6 Loudspeaker**

The loudspeaker factor was shown to be significant in the ANOVA. A Kruskal-Wallis non-parametric test confirmed the significance of the loudspeaker factor
given by the ANOVA (H (16) = 117.973, p < 0.001). Factor interactions were also shown to be significant, namely loudspeaker*movement condition; loudspeaker*movement condition*programme item; loudspeaker*programme item*listener; and loudspeaker*movement condition*listener. Figure 85 shows the absolute azimuth LRE plotted against loudspeaker number, movement condition and programme item.

Figure 85 - Absolute azimuth LRE separated by movement condition, loudspeaker number and programme item

In the no movement condition, LRA for the noise programme items follows a definite trend with loudspeaker number: azimuth absolute error is lower for loudspeakers located near ‘anchor’ locations such as loudspeakers 1 (−90°), 9 (0°) and 17 (90°); in this experiment 0° azimuth is directly to the right of the listener, while -90° is directly in front of the listener. It is suggested that these anchor locations are not more accurately localised by the listeners but that, when the
listeners are less certain of a location, they are more likely to respond with an anchor location, i.e. this trend is caused by a pointing error.

The movement condition appears to remove any trend created by loudspeaker azimuth. This is unsurprising as head movement tends to cause all loudspeakers to be moved into the most accurate area of azimuth localisation, directly in front of the listener. In the movement condition the listener is more certain of the source azimuth and so will be less likely to guess an anchor location, as in the case of the no movement condition. This effect is highlighted when the tonal programme item is removed from the analysis and the response from all noise programme items is plotted against actual source azimuth, as shown in Figure 86. In this plot 0° is to the right of the listener while -90° is directly in front of them.

![Figure 86 - Absolute azimuth LRE plotted against actual loudspeaker azimuth and movement condition](image-url)
There is a slightly higher azimuth LRE for sources behind the listener than for those in front in the no movement condition. This indicates that pointing errors, caused by sources being outside the listener's field of fixation, are affecting the absolute azimuth LRE data.

10.3 Conclusions

A localisation experiment was conducted that compared static and free head movement LRA for sources of varying bandwidth. The goal of the experiment was to find the bandwidth at which head movement cues improve LRA and to see whether this improvement with head movement occurs at lower bandwidths than static pinna cues. The two primary research questions were:

• What is the minimum bandwidth of the programme item that is required to allow improvement in LRA due to head movement?
• How does LRA change as the bandwidth of the programme item is reduced and how does allowing the listener to move their head affect this change in LRA?

These research questions will be answered for azimuth and elevation in sections 10.3.1 and 10.3.2 respectively.

10.3.1 Elevation LRA

What is the minimum bandwidth of the programme item that is required to allow improvement in LRA due to head movement?

Head movements significantly improve elevation LRA when compared to a static head position for programme items with an octave bandwidth centred at 6 kHz. For noise programme items, the pairing of an octave noise programme item and the free head movement condition was the only combination that resulted in significantly reduced elevation LRE. All other combinations of movement condition and noise programme item resulted in non-statistically significant differences. For programme items of half-octave bandwidth or less, head movement cues do not improve LRA.
How does LRA change as the bandwidth of the programme item is reduced and how does allowing the listener to move their head affect this change in LRA?

In static listening conditions there is no improvement in LRA as the bandwidth is increased from quarter octave noise to octave noise. This shows that bandwidths of greater than an octave are required for pinna cues to be effective.

The improvement offered by head movement to octave bandwidth sources shows that head movement cues can be used to localise at narrower bandwidths than pinna cues. Furthermore, it may indicate that head movement cues are not based on pinna cues, as this would cause both types of cue to fail simultaneously. However, it is possible that yaw head movements only serve to move the source into an area of higher LRA, where pinna cues can be used more effectively, as is the case with interaural difference cues in azimuth. Whether the improvement in LRA offered by head movement is due to dynamic or static cues is an area that is investigated further in Chapters 11 and 12.

In elevation, a noise programme item is more accurately localised than a tonal programme item. This was true in both the no movement and free movement conditions and for bandwidths of noise down to ¼ octave. It is suggested that for the tonal programme item, the listeners were confounded by room resonances and that, were the experiment to be repeated in anechoic conditions, the LRA of the listener to the ¼ octave noise and tone programme items would be more similar.

10.3.2 Azimuth LRA

What is the minimum bandwidth of the programme item that is required to allow improvement in LRA due to head movement?

Head movements significantly improve azimuth LRA when compared to a static head position for 6 kHz centred noise programme items down to ¼ octave bandwidth and 6 kHz tonal programme items. Whether this is the case for other frequency regions is unclear; since listeners will use ILD cues for all frequencies
above approximately 2 kHz, it is likely that this conclusion is applicable for all frequencies above 2 kHz.

*How does LRA change as the bandwidth of the programme item is reduced and how does allowing the listener to move their head affect this change in LRA?*

When the listener is able to move their head, the number of front-back discrimination errors is reduced regardless of programme item. When the listener is unable to move their head, the bandwidth of the source programme item has a significant effect on the number of front-back errors. For a tonal programme item at 6 kHz in the no movement condition, the proportion of front-back errors is close to 50%. This indicates that the listener has no localisation cues that enable them to discriminate a frontal and rearward source location.

In the no movement condition there are no significant differences in azimuth LRE between noise programmes. As the bandwidth of programme is increased, azimuth LRA also increases for free head movement listening conditions. Head movements either cause this increased LRA by moving the source into a higher area of azimuth acuity, which allows the small differences in acuity between different bandwidth programme items to be apparent; or provides dynamic cues, which are more accurate for wider bandwidth sources.

The azimuth LRA is significantly higher for noise programme items than tonal programme items for both the no movement and free head movement conditions. It is suggested that, for the tonal programme item, the room resonances altered interaural differences, thus rendering the listener's cues inaccurate.
11. Do head movements improve elevation LRA by way of static or dynamic cues?

Chapter 8 showed that yaw head movements significantly improve elevation LRA. For rearward sources, it was suggested that some of this improvement was due to the pointing response method. However, the main improvement in LRA with head movement was due to an auditory cue; it was unclear which of two auditory factors caused the improved LRA with yaw head movements. On the one hand, it is possible that, as the head is moved, the listener detects the variations in localisation cues during this movement, and makes use of this additional information to determine the location of the sound source. Alternatively, it was shown in Chapter 2 that when listeners are free to move their heads, they rotate their heads in azimuth towards the source location. Furthermore, it is known that the azimuth LRA is best when the sound source is on the median plane, so this head movement optimizes the azimuth cues [Pollack and Rose 1967]. If the same is true for elevation, then head movement would allow the head to be positioned such that the stimulus lies in the region of highest acuity, and improved elevation LRA might result from the improved cues available at this position, rather than from additional cues available during movement. Within this thesis, these two factors have been referred to as 'dynamic' cues (obtained during head movement) and 'static' cues (obtained when the head has reached a new stationary position, facing the source).

The primary research question for this chapter is:

- **Do head movements improve elevation LRA by way of static or dynamic cues?**

This will be found by determining whether elevation LRA is higher towards the median or lateral planes. Chapter 2 showed that when free to move their head, listeners make a yaw movement towards the source azimuth; moving the source towards their median plane. If elevation LRA is higher on the median plane than other planes then it is possible that head movement improves elevation LRA due
to static cues. On the other hand, if the LRA is similar or worse on the median plane, it is likely that the static cues do not play a role in the observed improvement resulting from head movement. This would show that dynamic cues were the primary cause of improved elevation LRA with head movement. In the experiment, loudspeakers were positioned on four vertical planes of differing azimuths and the elevation LRAs for the planes were compared.

A similar study conducted by Butler and Humanski [1992] suggested that the locations of sources on the lateral plane were reported more accurately than those on the median plane. However, Perrett and Noble [1995] showed that the numbered loudspeaker response method used by Butler and Humanski had significantly increased the lateral-plane LRA by resolving the listener’s cone of confusion, which threw their conclusions into question. The experiment described in this chapter also investigated elevation LRA but used a laser pointing response method, which did not provide any additional visual location cues.

Section 11.1 details the setup used in the experiment, including the loudspeaker position, programme item, and test procedure. Section 11.2 goes on to describe the results, analyzing the elevation LRA using the absolute elevation LRE, signed elevation LRE and correlation of the actual and reported loudspeaker elevations. Finally Section 11.3 gives the conclusions, describing how elevation LRA changes with plane and thus suggesting whether dynamic or static cues are the main contributor to the improved LRA with head movement.

**11.1 Experiment Setup**

Loudspeakers were placed at a variety of locations in elevation and azimuth and the listeners were asked to report the perceived loudspeaker position. The LRA was then calculated by comparing the actual location of the loudspeaker against the reported location given by the listener.
11. Do head movements improve elevation LRA by way of static or dynamic cues?

11.1.1 Listening Room

The experiment was conducted in 'VISLAB' listening room at the University of Surrey. The room measures $6.93 \times 7.81 \times 3.98$ m and has an RT60 of 217 ms averaged over 500, 1000 and 2000 Hz octave bands [Coleman 2014]. It is a large acoustically dead space that allows the subtle cues necessary for localisation to be perceived, while also showing a listener’s response to a real room.

11.1.2 Loudspeaker Positions

The experiment used 32 Genelec 8020a loudspeakers located on 4 vertical planes at $0^\circ$, $36^\circ$, $72^\circ$ and $108^\circ$ azimuth. The loudspeakers were concealed behind an acoustically transparent but visually opaque curtain, and were placed at a radius of 1.68 m and at elevations ranging from $-55^\circ$ to $+81^\circ$ at intervals of approximately $15^\circ$. Table 31 shows the elevation angles of the loudspeakers on each plane and designates a ‘Loudspeaker Elevation Number’ to each loudspeaker group. The layout of the vertical planes in azimuth is shown in Figure 87.

<table>
<thead>
<tr>
<th>Loudspeaker Elevation Number</th>
<th>Plane 1 ($0^\circ$)</th>
<th>Plane 2 ($36^\circ$)</th>
<th>Plane 3 ($72^\circ$)</th>
<th>Plane 4 ($108^\circ$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$-55$</td>
<td>$-42$</td>
<td>$-55$</td>
<td>$-41$</td>
</tr>
<tr>
<td>2</td>
<td>$-30$</td>
<td>$-30$</td>
<td>$-29$</td>
<td>$-30$</td>
</tr>
<tr>
<td>3</td>
<td>$-14$</td>
<td>$-14$</td>
<td>$-11$</td>
<td>$-13$</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>14</td>
<td>13</td>
<td>17</td>
<td>13</td>
</tr>
<tr>
<td>6</td>
<td>32</td>
<td>28</td>
<td>32</td>
<td>29</td>
</tr>
<tr>
<td>7</td>
<td>44</td>
<td>56</td>
<td>44</td>
<td>58</td>
</tr>
<tr>
<td>8</td>
<td>65</td>
<td>79</td>
<td>66</td>
<td>81</td>
</tr>
</tbody>
</table>

*Table 31 - Loudspeaker elevation number and loudspeaker elevations in degrees for each vertical plane*
Do head movements improve elevation LRA by way of static or dynamic cues?

The overall loudspeaker number (ranging from 1-32) can be calculated as:

\[
SN = SEN + (P - 1) \times 8
\]  

where SN is loudspeaker number, SEN is loudspeaker elevation number and P is plane number.

The head movement experiment in Chapter 8 indicated that visual cues might affect the listeners’ laser pointing accuracy. In this experiment, every effort was made to remove visual cues. As well as the loudspeakers being concealed behind an acoustically transparent curtain, the experiment also was conducted in total darkness (except the laser pointer and calibration point LED) so that no visual cues would have been present to skew the listener’s response.
11. Do head movements improve elevation LRA by way of static or dynamic cues?

11.1.3 Technical Equipment and Software

The loudspeakers were connected to two Fireface 800 audio interfaces (16 analogue outputs) and two 8-channel Presonus Digimax D8 audio interfaces connected by ADAT to the Fireface to create a total of 32 output channels. A Macbook Pro laptop computer running MaxMSP through the RME Fireface proprietary interface software was used to playback the sounds. The sound samples were created in MATLAB with a sample rate of 48 kHz and a bit depth of 24 bits.

11.1.4 Programme Item

A 4.5 second train of 150ms white noise bursts alternating with 300ms of silence was used as the programme item. This broadband amplitude-modulated source was chosen to provide ample scope for the creation of ITD, ILD and spectral localisation cues. It was hoped that by using an easily localisable programme item, the differences in LRA between planes would be most apparent.

11.1.5 Head Tracker

A Polhemus Patriot headtracker was used to monitor the listener’s movements during the test. This was used to measure the perceived sound source location and limit the listener’s head movement. The coordinates system of the tracker was used to define the loudspeaker positions, central listening position and calibration point as described in Appendix A.

11.1.6 Experiment Procedure

Listeners auditioned the stimuli whilst keeping their heads stationary, pointed at a calibration position. Head movements were limited to a maximum movement of 0.39° in azimuth, elevation or tilt, which is less than the smallest MAA shown in previous research [Perrott and Saberi 1990]. This was achieved using the Polhemus Patriot tracker and muting the sound if the head was moved by a greater angle.
A calibration point was positioned on the median plane at +9.85° in elevation to allow the relationship between the headtracker and laser to be established. It was suggested in previous experiments that positioning the calibration point at 0° in elevation may have biased the listener’s response. By moving the calibration point away from the location of a loudspeaker it was thought this bias could be reduced.

At the start of each trial the listener aimed the laser at the calibration point. Once the stimulus had finished playback the listener could move their head to point the head-mounted laser at the perceived loudspeaker position; the tracker was used to determine the direction in which it was pointed. The listener was allowed to repeat the playback of the programme item during a trial.

The listener was not able to see a user interface during the test. Instead they were shown the keyboard commands that would control the playback of the trials. In line with previous experiments, Z was used the play the stimulus, X was used if the listener wished to replay the stimulus and spacebar was used to log the perceived source location once the laser was pointing at the perceived source location.

Each stimulus was repeated once, resulting in 64 trials per listener. A familiarity session of 10 trials was included at the start of the test. The test took approximately 20-30 minutes to complete. Seventeen listeners aged between 18 and 35 undertook the test, all with no reported hearing problems.

11.2 Results

In order to gain a sufficient impression of the overall trends of the listeners’ elevation responses and the effect of source azimuth, a number of metrics are examined. The absolute elevation LRE, signed elevation LRE and the correlation of reported and actual source elevation are analysed in the following results sections. The goal of this section is to find out whether there are significant differences in elevation LRA between vertical planes of differing azimuth.
11.2.1 Absolute Elevation LRE

An ANOVA was conducted on the absolute elevation LRE data with the fixed factors 'loudspeaker' and 'repeat' and the random factor 'listener'. Validation tests on the absolute elevation LRE data preceding the ANOVA showed 50 of the 64 factor comparisons to be normally distributed. The results of the ANOVA are shown in Table 32.

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>Hypothesis</td>
<td>99559.617</td>
<td>1</td>
<td>99559.617</td>
<td>444.713</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td>Hypothesis</td>
<td>3581.981</td>
<td>16</td>
<td>223.874</td>
<td>.000</td>
<td>.405</td>
</tr>
<tr>
<td>Speaker</td>
<td>Hypothesis</td>
<td>26180.169</td>
<td>31</td>
<td>844.522</td>
<td>.000</td>
<td>.405</td>
</tr>
<tr>
<td>Error</td>
<td>Hypothesis</td>
<td>38420.568</td>
<td>496</td>
<td>77.461</td>
<td>.000</td>
<td>.405</td>
</tr>
<tr>
<td>Repeat</td>
<td>Hypothesis</td>
<td>10.624</td>
<td>1</td>
<td>10.624</td>
<td>.231</td>
<td>.014</td>
</tr>
<tr>
<td>Error</td>
<td>Hypothesis</td>
<td>736.207</td>
<td>16</td>
<td>46.013</td>
<td>.009</td>
<td>.371</td>
</tr>
<tr>
<td>Listener</td>
<td>Hypothesis</td>
<td>3581.981</td>
<td>16</td>
<td>223.874</td>
<td>.009</td>
<td>.371</td>
</tr>
<tr>
<td>Error</td>
<td>Hypothesis</td>
<td>6059.991</td>
<td>63.132</td>
<td>95.989</td>
<td>.051</td>
<td>.051</td>
</tr>
<tr>
<td>Repeat * Listener</td>
<td>Hypothesis</td>
<td>736.207</td>
<td>16</td>
<td>46.013</td>
<td>.167</td>
<td>.048</td>
</tr>
<tr>
<td>Error</td>
<td>Hypothesis</td>
<td>13632.603</td>
<td>496</td>
<td>27.485</td>
<td>.278</td>
<td>.083</td>
</tr>
<tr>
<td>Speaker * Listener</td>
<td>Hypothesis</td>
<td>38420.568</td>
<td>496</td>
<td>77.461</td>
<td>.218</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td>Hypothesis</td>
<td>13632.603</td>
<td>496</td>
<td>27.485</td>
<td>.218</td>
<td>.000</td>
</tr>
<tr>
<td>Speaker * Repeat</td>
<td>Hypothesis</td>
<td>1234.869</td>
<td>31</td>
<td>39.834</td>
<td>1.449</td>
<td>.058</td>
</tr>
<tr>
<td>Error</td>
<td>Hypothesis</td>
<td>13632.603</td>
<td>496</td>
<td>27.485</td>
<td>.218</td>
<td>.000</td>
</tr>
</tbody>
</table>

Table 32 - ANOVA of absolute elevation LRE with factors speaker, repeat and listener

The factors 'loudspeaker' (F (31, 1088) = 10.903, p < 0.001, \( \eta^2 = 0.405 \)) and 'listener' (F (16, 1088) = 2.332, p = 0.009, \( \eta^2 = 0.371 \)) were shown to be significant. This shows that each listener's LRA was consistent during the test.

A histogram of the ANOVA model's standardised residuals showed a normal distribution (Figure 88), indicating that the ANOVA model provided a good fit to the data.
Do head movements improve elevation LRA by way of static or dynamic cues?

Figure 88 - Histogram of ANOVA Standardised Residuals
11.2.2 Listener

Figure 89 shows the absolute elevation LRE separated by listener number.

![Figure 89 - Mean absolute elevation LRE separated by listener number](image)

It is likely that the inter-listener variance shown in the graph is due to differences in the shape and size of the listeners' pinnae, which will change the spectral cues that they receive and so, potentially, their elevation LRA. Previous studies have noted that some listeners are better than others at localising with pinna cues [Langendijk and Bronkhurst 2013]. Other factors such as experience with localisation tasks and mental stimulation (concentration/ boredom) may also have affected their response. Although there was significant variance in the response of listeners, no single listener responded significantly more or less accurately than all the others and so there is no indication that any significant errors in response occurred.
11.2.3 Loudspeaker Number

Based on the ANOVA results, the effect of the loudspeaker location on the results was investigated. Figure 90 shows the absolute elevation LRE plotted against loudspeaker number.

![Loudspeaker Number Diagram]

*Figure 90 - Absolute elevation LRE separated by loudspeaker number*

Loudspeaker 1 has a significantly mean higher elevation LRE than all other loudspeaker locations. A single errant response might have biased the overall mean for this loudspeaker location. A boxplot is a good method used to highlight outliers in the response; Figure 91 shows a box plot of the absolute elevation LRE separated by loudspeaker.
Do head movements improve elevation LRA by way of static or dynamic cues?

Figure 91 - Box plot of absolute elevation LRE plotted against loudspeaker number

Loudspeaker 1 has the highest median error of all loudspeakers and the highest number of outliers. This loudspeaker was located on the median plane at -55° elevation, the lowest loudspeaker elevation tested.

Response number 25, a significant outlier for loudspeaker 1, may have been an up-down confusion. Alternatively it may have been an entirely erroneous response due to an operation error or because the listener found the loudspeaker location to be very difficult to judge. Both hypotheses justified removing response 25 from further analysis. With response 25 removed the ANOVA was repeated and the significance of the loudspeaker variable remained unchanged (p < 0.001) while the significance of the listener variable was increased (p = 0.001).

For Loudspeaker 1 there were five further outliers. For these responses the listener reported a location of approximately 0° elevation, i.e. straight ahead. The minimum absolute elevation LRE (i.e. the most accurate single response by any
listener) for loudspeaker 1 for all of the cases measured was 11.85°. Therefore, on no occasion did a listener respond to this loudspeaker location with an accuracy of greater than the overall mean elevation LRE of the test (9.6°). The mean error of loudspeaker 1 (28.5°) was much larger than that for loudspeaker 17 (17.6°), which was located at the same elevation but on Plane 3. One explanation for the low LRA for Loudspeaker 1 is that elevation localisation of this loudspeaker was dependent on monaural spectral cues; no interaural cues were present because it was on the median plane. Loudspeaker 17 allowed the listener both interaural and spectral cues. For this reason loudspeaker 1 appears to be the most difficult loudspeaker to localise in the test.

11.2.4 Loudspeaker Plane

In this section the variance in elevation LRA due to loudspeaker plane is investigated. Figure 92 shows the absolute elevation LRE data separated by the loudspeaker plane and plotted against the actual loudspeaker location.
Figure 92 - Mean absolute elevation LRE with 95% confidence intervals plotted against actual loudspeaker elevation separated by plane number

The trend for all four planes is similar: for loudspeakers between $-40^\circ$ and $+60^\circ$ the response error was fairly consistent on all planes; for loudspeakers outside these bounds the LRA decreased dramatically. At the lowest elevations this may have been due to boundary biases caused by the edge of the acoustically transparent curtain. However, at high elevations there was no boundary as the curtain extended all the way over the listener's head. End effects could still be apparent for the top of the curtain, forcing all responses below $90^\circ$ and causing the mean reported elevation to be reduced. The shift in LRA appears to be quite sudden, with listeners displaying a dramatic decrease. Error bars are much tighter for elevations between $-40^\circ$ and $+60^\circ$, indicating that listeners are more consistent for these regions.

An ANOVA was conducted in which the 'loudspeaker number' factor was replaced by 'loudspeaker elevation number' and 'plane number' (these variables
are summarised in Table 31). This allows one to separate the variance in LRE due to the changes in loudspeaker elevation from the variance due to the loudspeaker azimuth (i.e. loudspeaker plane). The results of the ANOVA are shown in Table 33.

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>Hypothesis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>1</td>
<td>97203.376</td>
<td>516.766</td>
<td>.000</td>
<td>.970</td>
</tr>
<tr>
<td>Plane</td>
<td>Hypothesis</td>
<td>16</td>
<td>188.099</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>16</td>
<td>632.700</td>
<td>7.760</td>
<td>.000</td>
<td>.327</td>
</tr>
<tr>
<td>Speaker_Elevation</td>
<td>Hypothesis</td>
<td>7</td>
<td>1866.919</td>
<td>12.055</td>
<td>.000</td>
<td>.430</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>7</td>
<td>139.936</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repeat</td>
<td>Hypothesis</td>
<td>1</td>
<td>.001</td>
<td>.000</td>
<td>.994</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>1</td>
<td>.001</td>
<td>.000</td>
<td>.994</td>
<td>.000</td>
</tr>
<tr>
<td>Listener</td>
<td>Hypothesis</td>
<td>16</td>
<td>188.120</td>
<td>1.103</td>
<td>.366</td>
<td>.176</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>16</td>
<td>170.540</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plane * Listener</td>
<td>Hypothesis</td>
<td>48</td>
<td>81.558</td>
<td>1.640</td>
<td>.014</td>
<td>.357</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>48</td>
<td>48.724</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repeat * Listener</td>
<td>Hypothesis</td>
<td>16</td>
<td>16.763</td>
<td>.729</td>
<td>.748</td>
<td>.240</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>16</td>
<td>17.743</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker_Elevation * Listener</td>
<td>Hypothesis</td>
<td>112</td>
<td>140.020</td>
<td>3.096</td>
<td>.000</td>
<td>.623</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>112</td>
<td>23.006</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plane * Repeat</td>
<td>Hypothesis</td>
<td>3</td>
<td>13.402</td>
<td>.596</td>
<td>.621</td>
<td>.036</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>3</td>
<td>22.500</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plane * Speaker_Elevation</td>
<td>Hypothesis</td>
<td>21</td>
<td>362.323</td>
<td>8.104</td>
<td>.000</td>
<td>.336</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>21</td>
<td>44.709</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker_Elevation * Repeat</td>
<td>Hypothesis</td>
<td>7</td>
<td>23.261</td>
<td>1.293</td>
<td>.260</td>
<td>.075</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>7</td>
<td>17.995</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plane * Repeat * Listener</td>
<td>Hypothesis</td>
<td>48</td>
<td>22.501</td>
<td>1.286</td>
<td>.107</td>
<td>.156</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>48</td>
<td>17.491</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plane * Speaker_Elevation *</td>
<td>Hypothesis</td>
<td>336</td>
<td>44.745</td>
<td>2.558</td>
<td>.000</td>
<td>.720</td>
</tr>
<tr>
<td>Listener</td>
<td>Error</td>
<td>336</td>
<td>17.491</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker_Elevation * Repeat *</td>
<td>Hypothesis</td>
<td>112</td>
<td>17.995</td>
<td>1.029</td>
<td>.417</td>
<td>.256</td>
</tr>
<tr>
<td>Listener</td>
<td>Error</td>
<td>112</td>
<td>17.491</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plane * Speaker_Elevation *</td>
<td>Hypothesis</td>
<td>21</td>
<td>21.640</td>
<td>1.237</td>
<td>.217</td>
<td>.072</td>
</tr>
<tr>
<td>Repeat</td>
<td>Error</td>
<td>21</td>
<td>17.491</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 33 - ANOVA of absolute elevation LRE with factors plane, loudspeaker elevation, repeat and listener

The ANOVA showed both the factors ‘loudspeaker elevation number’ (F (7, 1088) = 12.055, p < 0.001, \( \eta^2 = 0.430 \)) and ‘plane’ (F (3, 1088) = 7.76, p < 0.001, \( \eta^2 = 0.327 \)) to be significant. The interaction loudspeaker elevation number*plane was also significant and will be investigated first. Figure 93 shows the absolute elevation angle plotted against plane and loudspeaker elevation number.
11. Do head movements improve elevation LRA by way of static or dynamic cues?

Figure 93 - Mean absolute elevation LRE with 95% confidence intervals separated by plane number and loudspeaker elevation number

The elevation LRAs for loudspeaker elevation numbers 1 and 8 are significantly worse than those for other loudspeaker elevation numbers in the majority of cases, the main exception being loudspeaker elevation number 1 on plane 4. These loudspeakers were located at below -40° and above +60° respectively; they were the highest and lowest loudspeaker locations on each plane. One can conclude that when the listener is unable to move their head, sources at elevations further from the equatorial plane are harder to localise. It is hypothesised that the differences between loudspeaker elevation numbers 1 and 8 and the other loudspeaker locations caused the significance of the Loudspeaker Elevation Number in the ANOVA. To investigate further the region in which
elevation LRA was consistent across elevation (between -40° and +60°), a second ANOVA was conducted excluding loudspeakers with elevation numbers 1 and 8.

This ANOVA found that the loudspeaker elevation number \( (F(5, 1088) = 1.184, p = 0.325, \eta_p^2 = 0.069) \) was no longer significant, while the variable plane \( (F(3, 1088) = 3.202, p = 0.031, \eta_p^2 = 0.167) \) was still slightly significant. The interaction of plane*loudspeaker elevation number was not significant. Figure 94 shows the absolute elevation LRE separated by plane number and loudspeaker elevation number.

![Figure 94: Mean absolute elevation LRE with 95% confidence intervals separated by plane number and loudspeaker elevation number (excluding speaker elevation numbers 1 and 8)](image)

Loudspeaker 5 on plane 1 shows a significantly lower elevation LRE than four of the other loudspeakers on the same plane. It is suggested that the increased LRA
for loudspeaker 5 was due to the positioning of the calibration point. Loudspeaker 5 was located at $+14^\circ$, close to the calibration point at $+9.85^\circ$. The angle between these two points, $4.15^\circ$, was lower than the mean absolute LRE for that location and half the mean absolute LRE for that plane. If listeners had responded with the calibration point on every occasion then the mean absolute LRE would have been more accurate. It is suggested that the calibration point caused the improved LRA for this location. It is hypothesised that, had the calibration point not been present, elevation LRA for this loudspeaker would have been similar to others on plane 1.

There is a general trend of slightly increasing LRA from Planes 1 to 3, i.e. as the loudspeaker locations get further from the median plane. This trend can be more easily seen in Figure 95, which shows the mean absolute elevation LRE plotted against plane number.

![Figure 95 - Mean absolute elevation LRE with 95% confidence intervals separated by plane number](image-url)
Significant differences can be seen between plane 3 and plane 1 and between plane 4 and plane 1. A Kruskal-Wallis test, conducted to verify this assertion, showed the ‘plane’ variable to be significant \((H (3) = 12.771, P = 0.005)\). The mean ranks showed the same order of error as the graph in Figure 95, with plane 3 being most accurate and plane 1 being least accurate.

The differences between plane 3 and plane 4 were found to be non-significant. This shows that moving a source from the front right quadrant to the rear right quadrant but at the same angle from the median plane did not significantly reduce elevation LRA.

### 11.2.5 Reported and Actual Angle Correlation

A Pearson’s Correlation test, conducted on the listener’s reported elevation data and actual elevation data, showed that the two variables were closely correlated for all the planes (Table 34). All correlations were shown to be significant.

<table>
<thead>
<tr>
<th>Plane Number</th>
<th>Pearson Correlation</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.938</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>2</td>
<td>0.963</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>3</td>
<td>0.969</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>4</td>
<td>0.968</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

*Table 34 - Pearson’s correlation test results for each plane*

This is an unsurprising finding: as the actual loudspeaker elevation was increased, the reported loudspeaker elevation increased. However, there is also a trend of increasing correlation as the azimuth angle from the median plane is increased. It was suggested that the inaccuracy of response to loudspeakers 1 and 8 on certain planes might have reduced the correlation on the median plane.
However, the test was repeated using only loudspeaker 2 – 7 and the results were similar (Table 35).

<table>
<thead>
<tr>
<th>Plane Number</th>
<th>Pearson Correlation</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.929</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>2</td>
<td>0.946</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>3</td>
<td>0.956</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>4</td>
<td>0.957</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 35 - Pearson’s correlation test results for each plane excluding loudspeakers 1 and 8

This is a further indication that elevation LRA was increased as the azimuth angle from the median plane was increased.

**11.2.6 Response Bias**

The signed elevation LRE of the listener, plotted in Figure 96, gives an indication of the elevation response bias.
As with previous studies all listener responses were biased towards a central location. The point at which no bias is present occurs when the scatter plot is centred on 0° actual elevation LRE. In this test this appears to be for loudspeaker elevation numbers 5 and 6, which are located above the equatorial plane, between approximately +30° and +50°.

Since the experiment was conducted in darkness, it is unlikely that visual cues were responsible for this bias. One possible reason for the bias was that the calibration point was located at +9.85°. However, one would assume that this would cause a distribution of response around the calibration point. A second possible reason was that the listener may have pointed their head, rather than the laser pointer, at the source location. However, every effort was made preceding the test to align the listener’s eye-line with the laser pointer and the listener was specifically instructed to point the laser at the perceived loudspeaker location.

**Figure 96 - Scatter plot of signed elevation LRE against loudspeaker number**
Whilst the cause of the bias is unknown, it seems that when a loudspeaker is positioned below +30°, listeners perceive it to be higher than it actually is, and when a loudspeaker is positioned above +30°, listeners perceive it to be below its actual location. Furthermore it is likely that listeners produce the highest elevation LRA for source located at approximately +30° because that is the source elevation they can localise most accurately.

11.3 Discussion and Conclusions

The primary research question for this chapter was:

- Do head movements improve elevation LRA by way of static or dynamic cues?

From the absolute elevation LRE analysis it can be concluded that when the listener is unable to move their head, for loudspeakers between -40° and +60° elevation, elevation LRA gets higher as the source is located further in azimuth from the median plane. The mean absolute error for sources located 72° laterally (plane 3) was shown to be 2° lower than the mean absolute error for sources located directly in front (plane 1). Correlations between the actual and reported elevation data also showed this trend of increasing elevation LRA as the source was located further in angle from the median plane.

For angles away from the median plane, the listener has both monaural spectral cues and interaural difference cues available to them, which may explain the increased elevation LRA. Furthermore, as the source location gets further from the median plane, the cone of confusion associated with the interaural difference cues gets smaller, creating a reduced range of possible responses.

As shown in Chapter 2, when attempting to localise a sound source, listeners move their heads toward the source azimuth; therefore moving the source onto their median plane. This experiment has shown that positioning the head in this way will reduce elevation LRA, indicating reduced efficacy of elevation localisation cues when facing the source. However, Chapter 8 showed that head movements increase overall elevation LRA. Therefore, dynamic cues, created
11. Do head movements improve elevation LRA by way of static or dynamic cues?

through the movement of the head, must be the cause of this increase in elevation LRA.

11.3.1 Additional Conclusions

In this experiment it is suggested that the positioning of the calibration point created a LRA bias that caused a single loudspeaker location to be reported significantly more accurately than it should have. Visual cues can override auditory cues when a listener is attempting to localise a stimulus and so tight control of all visual stimuli must be maintained while running a localisation test.

A response bias was noted that loudspeakers below approximately +30° elevation were reported above their actual location, while loudspeakers above approximately +30° elevation were reported below their actual location for sources on the median plane. Furthermore loudspeakers positioned at +30° elevation were shown to produce the smallest absolute elevation LRE.
12 Static and Dynamic Cues with Head Movement for Band-Limited Noise Sources

Chapter 8 showed that head movements significantly improved elevation LRA for both full and reduced bandwidth noise sources. Chapter 11 showed that, for full bandwidth sources, when no head movements were permitted, elevation LRA was higher for planes located further in azimuth from the median plane. This means that dynamic cues created through the movement of the listener’s head caused the improvement in elevation LRA with head movement noted in Chapter 8. The experiment ‘Localisation of Band-pass Filtered Noise with Head Movement’, described in Chapter 10, showed that head movements had no effect on the elevation LRA of band-pass filtered noise programme items with a bandwidth of less than an octave. For an octave bandwidth programme item head movements were shown to significantly improve elevation LRA. There was a change in the effect that head movements had on the elevation LRA between these two bandwidths. It was hypothesised that this may have been due to dynamic or static cues becoming ineffective at the bandwidths below an octave.

The aim of this chapter is to find whether, for reduced bandwidth sources, elevation LRA is higher for sources further in azimuth from the median plane. This will show whether, for reduced bandwidth sources, improvements in LRA with head movements, shown in chapters 8, 9 and 10, were due to dynamic or static cues. Furthermore, studying both half-octave and octave programme items might highlight whether the change in LRA shown in Chapter 10 was due to changing static or dynamic cues.

The experiment described in this chapter was designed to answer one primary research question:

- For reduced bandwidth sources, is elevation LRA higher for vertical planes further in azimuth from the median plane?

The most common head movement when attempting to localise a source is towards the source azimuth. If static elevation LRA is equal across planes or higher for planes to the side of the listener, then it can be concluded that head
movement does not improve the listener's static cues. Therefore this experiment will show whether the improved elevation LRA noted for an octave bandwidth programme item with head movement was due to static or dynamic cues. Comparing the programme item results might also show why there was an improvement in LRA with head movement for the octave programme item and not for the half-octave programme item.

The experiment setup and procedure was similar to the ‘Elevation Localisation on Vertical planes of Differing Azimuth’ experiment in the previous chapter. Listeners were asked to localise loudspeakers located on planes of differing azimuths. The primary difference between the experiments was that reduced bandwidth noise sources were used as the programme items.

### 12.1 Experiment Setup

This section describes the experiment parameters, including programme items, loudspeaker positions, movement conditions and number of trials.

#### 12.1.1 Programme Items

In the experiment described in Chapter 10, significant changes in elevation LRA with head movement occurred between the octave and half-octave programme items. For the octave noise, head movement improved LRA; for the half-octave noise, head movement had no effect on LRA. Therefore, the two programme items used in this test were:

- *Octave band-limited white noise*
- *Half-Octave band-limited white noise*

The experiment used programme items with a single centre frequency at 6 kHz. This allowed a direct comparison with the previous bandwidth localisation test, which used the same centre frequency.
12.1.2 Loudspeaker Positions

The loudspeakers were placed at the same locations as the loudspeakers in the experiment in Chapter 11. This allows one to more easily compare between the results of this trial and the results of the previous experiment.

There were 4 vertical planes located in the listener’s right hemisphere, at 0°, 36°, 72° and 108° in azimuth. Each plane had 8 loudspeakers spread evenly in elevation from -54° to +81°. Table 36 summarises the loudspeaker positions.

<table>
<thead>
<tr>
<th>Loudspeaker Elevation Number</th>
<th>Plane 1 (0°)</th>
<th>Plane 2 (36°)</th>
<th>Plane 3 (72°)</th>
<th>Plane 4 (108°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-53</td>
<td>-44</td>
<td>-54</td>
<td>-44</td>
</tr>
<tr>
<td>2</td>
<td>-28</td>
<td>-31</td>
<td>-28</td>
<td>-31</td>
</tr>
<tr>
<td>3</td>
<td>-13</td>
<td>-15</td>
<td>-13</td>
<td>-14</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>18</td>
<td>14</td>
<td>18</td>
<td>14</td>
</tr>
<tr>
<td>6</td>
<td>33</td>
<td>31</td>
<td>33</td>
<td>32</td>
</tr>
<tr>
<td>7</td>
<td>46</td>
<td>57</td>
<td>46</td>
<td>57</td>
</tr>
<tr>
<td>8</td>
<td>65</td>
<td>81</td>
<td>67</td>
<td>81</td>
</tr>
</tbody>
</table>

Table 36 - Loudspeaker elevation number and loudspeaker elevations in degrees for each vertical plane

The overall loudspeaker number (ranging from 1-32) can be calculated as:

\[ SN = SEN + (P - 1) \times 8 \]  \[ (4) \]

where \( SN \) is loudspeaker number, \( SEN \) is loudspeaker elevation number and \( P \) is plane number.
12.1.3 Movement Conditions

The listeners were instructed not to move their heads while the programme item was replayed. Listeners auditioned each stimulus whilst keeping their heads stationary, pointed at a reference position. Head movements were limited to a maximum movement of 0.39° in azimuth, elevation or tilt, less than the smallest minimum audible angle. This was achieved by mounting a Polhemus Patriot tracker on the listener's head and muting the sound if the listener's head moved by a greater angle.

In previous experiments it has been suggested that the position of the calibration point may have biased the listener's response. In this experiment the calibration point was moved from above the equatorial plane to below the equatorial plane. The calibration point was at -6.29°, approximately located between loudspeakers 3 and 4. Any changes in response caused by the change in the location of the calibration point can be noted by comparing the results of this experiment with the previous experiment. Furthermore, the previous experiment showed that sources located below the equatorial plane were less accurately reported than those above. This may have been caused by the location of the calibration point and so the repositioning of this point may cause it to have a smaller impact upon the listeners' responses.

12.1.4 Number of Trials

The experiment was split into two tests, one for each programme item. Each test had four loudspeaker planes; eight loudspeaker positions on each plane; one movement condition; and one repeat, resulting in 64 trials per listener. Each test took approximately ½ an hour to complete. Each listener was required to leave at least 4 hours rest between the two tests.

12.1.5 Listeners

There were 19 listeners for each programme item test; 16 of the listeners completed both programme item tests, while 6 listeners completed only one of
the programme item tests. Therefore a total of 22 listeners aged between 18 and 35 undertook the test, all with no reported hearing problems.

For the listeners who completed both tests, the order of the tests was varied: half the listeners auditioned the octave programme test first, while the other half auditioned the half-octave programme. The results for the two programme items were combined into an overall results database. A familiarity session of 10 trials was included at the start of the test.

### 12.2 Results

An ANOVA was conducted on the absolute elevation LRE data with fixed factors ‘programme item’ and ‘loudspeaker’ and random factor ‘listener’ (Table 37). Validation tests on the absolute elevation LRE data preceding the ANOVA showed that 42 of the 64 loudspeaker factor comparisons were normally distributed. Therefore, non-parametric tests will be conducted to confirm the major findings of the ANOVA.

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>Hypothesis</td>
<td>2248193</td>
<td>1</td>
<td>2248193.40</td>
<td>1005.87</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>46996.3</td>
<td>21.027</td>
<td>2235.073</td>
<td>a</td>
<td></td>
</tr>
<tr>
<td>Programme_Item</td>
<td>Hypothesis</td>
<td>2144.61</td>
<td>1</td>
<td>2144.606</td>
<td>9.673</td>
<td>.007</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>3325.60</td>
<td>15</td>
<td>221.70</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td>speaker</td>
<td>Hypothesis</td>
<td>57179.1</td>
<td>31</td>
<td>18444.875</td>
<td>20.692</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>582553</td>
<td>653.511</td>
<td>891.422</td>
<td>c</td>
<td></td>
</tr>
<tr>
<td>listener</td>
<td>Hypothesis</td>
<td>47209.3</td>
<td>21</td>
<td>2248.062</td>
<td>2.647</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>131437</td>
<td>154.738</td>
<td>849.416</td>
<td>d</td>
<td></td>
</tr>
<tr>
<td>Programme_Item * speaker</td>
<td>Hypothesis</td>
<td>18210.4</td>
<td>31</td>
<td>587.433</td>
<td>2.194</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>124497</td>
<td>465</td>
<td>267.736</td>
<td>e</td>
<td></td>
</tr>
<tr>
<td>Programme_Item * listener</td>
<td>Hypothesis</td>
<td>3325.60</td>
<td>15</td>
<td>221.707</td>
<td>.828</td>
<td>.646</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>124497</td>
<td>465</td>
<td>267.736</td>
<td>e</td>
<td></td>
</tr>
<tr>
<td>speaker * listener</td>
<td>Hypothesis</td>
<td>582935</td>
<td>651</td>
<td>895.445</td>
<td>3.345</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>124497</td>
<td>465</td>
<td>267.736</td>
<td>e</td>
<td></td>
</tr>
<tr>
<td>Programme_Item * speaker * listener</td>
<td>Hypothesis</td>
<td>124497</td>
<td>465</td>
<td>267.736</td>
<td>1.648</td>
<td>.000</td>
</tr>
</tbody>
</table>

Table 37 – ANOVA of absolute elevation LRE separated by programme item, loudspeaker number and listener number

The primary factors ‘loudspeaker’ (F (31, 2432) = 20.692, p < 0.001, ηp² = 0.495), ‘programme item’ (F (1, 2432) = 9.673, p = 0.007, ηp² = 0.392) and
‘listener’ \( (F (21, 2432) = 2.647, p < 0.001, \eta^2 = 0.264) \) were shown to be significant. Secondary interactions loudspeaker*listener \( (F (651, 2432) = 3.345, p < 0.001, \eta^2 = 0.824) \) and programme item*loudspeaker \( (F (31, 2432) = 2.194, p < 0.001, \eta^2 = 0.128) \) were also shown to be significant. A histogram of the ANOVA model’s standardised residuals showed an approximately normal distribution, indicating that the ANOVA model provided a good fit to the data (Figure 97).

![Histogram of standardised residuals given by the ANOVA](image)

**Figure 97 - Histogram of standardised residuals given by the ANOVA**

### 12.2.1 Loudspeaker

The ANOVA showed a large f value for the loudspeaker factor and so this factor was investigated first. To check the significance given by the ANOVA a non-parametric Kruskal-Wallis test was conducted. It was shown that the loudspeaker number was highly significant \( (H (31) = 711.079, p < 0.001) \). Figure 98 shows of absolute elevation LRE plotted against loudspeaker number.
Referring to both the graph in Figure 98 and the table of loudspeaker positions (Table 36), it can be seen that the significant differences highlighted by ANOVA for the factor 'loudspeaker number' were due to both the plane on which the loudspeaker was located and its elevation location. On each plane the elevation LRA follows a similar trend: sources at higher and lower elevations are reported less accurately than those at elevations nearer the equatorial plane. The variance in LRA also increases as the loudspeaker is located further from the equatorial plane, i.e. the listeners’ responses vary more for loudspeakers further from the equatorial plane. It would also appear that the absolute elevation LRE is reduced as the plane on which the loudspeakers are located is moved further in azimuth from the median plane.

12.2.2 Plane Number and Loudspeaker Elevation Number

To further investigate the trends caused by the loudspeaker's location in azimuth and elevation, the loudspeaker number was divided into plane number and
loudspeaker elevation number. An ANOVA was conducted on the absolute elevation LRE data with fixed factors programme item, plane number and loudspeaker elevation number and random factor listener. The results of the ANOVA are shown in Table 38.

### Table 38 - ANOVA of absolute elevation LRE separated by programme item, plane number, loudspeaker elevation number and listener number

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>Hypothesis</td>
<td>2248193.40</td>
<td>1</td>
<td>2248193.40</td>
<td>1005.870</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>46996.254</td>
<td>21.027</td>
<td>2235.073</td>
<td>.000</td>
</tr>
<tr>
<td>Programme_Item</td>
<td>Hypothesis</td>
<td>2144.606</td>
<td>1</td>
<td>2144.606</td>
<td>9.673</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>3325.603</td>
<td>15</td>
<td>221.709</td>
<td>.000</td>
</tr>
<tr>
<td>Plane</td>
<td>Hypothesis</td>
<td>44479.734</td>
<td>3</td>
<td>14826.578</td>
<td>18.169</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>51687.114</td>
<td>63.339</td>
<td>816.043</td>
<td>.000</td>
</tr>
<tr>
<td>speaker_ele_no</td>
<td>Hypothesis</td>
<td>434723.785</td>
<td>7</td>
<td>62103.398</td>
<td>24.669</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>370803.583</td>
<td>147.294</td>
<td>2517.441</td>
<td>.000</td>
</tr>
<tr>
<td>listener</td>
<td>Hypothesis</td>
<td>47209.312</td>
<td>21</td>
<td>2248.062</td>
<td>.835</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>314872.555</td>
<td>116.932</td>
<td>2692.773</td>
<td>.000</td>
</tr>
<tr>
<td>Programme_Item * Plane</td>
<td>Hypothesis</td>
<td>184.637</td>
<td>3</td>
<td>61.546</td>
<td>.180</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>15368.354</td>
<td>45</td>
<td>341.519</td>
<td>.000</td>
</tr>
<tr>
<td>Programme_Item * speaker_ele_no</td>
<td>Hypothesis</td>
<td>9719.856</td>
<td>7</td>
<td>1388.551</td>
<td>3.540</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>41185.858</td>
<td>105</td>
<td>392.246</td>
<td>.000</td>
</tr>
<tr>
<td>Programme_Item * listener</td>
<td>Hypothesis</td>
<td>3325.603</td>
<td>15</td>
<td>221.707</td>
<td>.428</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>33068.784</td>
<td>63.830</td>
<td>518.073</td>
<td>.000</td>
</tr>
<tr>
<td>Plane * speaker_ele_no</td>
<td>Hypothesis</td>
<td>92587.609</td>
<td>21</td>
<td>4408.934</td>
<td>12.241</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>160064.225</td>
<td>444.396</td>
<td>360.184</td>
<td>.000</td>
</tr>
<tr>
<td>Plane * listener</td>
<td>Hypothesis</td>
<td>51603.552</td>
<td>63</td>
<td>819.104</td>
<td>1.682</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>38039.397</td>
<td>78.119</td>
<td>486.943</td>
<td>.000</td>
</tr>
<tr>
<td>speaker_ele_no * listener</td>
<td>Hypothesis</td>
<td>372079.267</td>
<td>147</td>
<td>2531.151</td>
<td>4.708</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>81434.709</td>
<td>151.459</td>
<td>537.670</td>
<td>.000</td>
</tr>
<tr>
<td>Programme_Item * Plane * speaker_ele_no</td>
<td>Hypothesis</td>
<td>8305.937</td>
<td>21</td>
<td>395.521</td>
<td>1.834</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>67943.033</td>
<td>315</td>
<td>215.692</td>
<td>.000</td>
</tr>
<tr>
<td>Programme_Item * Plane * listener</td>
<td>Hypothesis</td>
<td>15368.354</td>
<td>45</td>
<td>341.519</td>
<td>1.583</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>67943.033</td>
<td>315</td>
<td>215.692</td>
<td>.000</td>
</tr>
<tr>
<td>Programme_Item * speaker_ele_no * listener</td>
<td>Hypothesis</td>
<td>41185.858</td>
<td>105</td>
<td>392.246</td>
<td>1.819</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>67943.033</td>
<td>315</td>
<td>215.692</td>
<td>.000</td>
</tr>
<tr>
<td>Plane * speaker_ele_no * listener</td>
<td>Hypothesis</td>
<td>159252.106</td>
<td>441</td>
<td>361.116</td>
<td>1.674</td>
</tr>
<tr>
<td></td>
<td>Error</td>
<td>67943.033</td>
<td>315</td>
<td>215.692</td>
<td>.000</td>
</tr>
<tr>
<td>Programme_Item * Plane * speaker_ele_no * listener</td>
<td>Hypothesis</td>
<td>67943.033</td>
<td>315</td>
<td>215.692</td>
<td>1.328</td>
</tr>
</tbody>
</table>

The single factors ‘programme item’ (F(1, 2432) = 9.673, p = 0.007, ηp² = 0.392), ‘plane’ (F(3, 2432) = 18.169, p < 0.001, ηp² = 0.463) and ‘loudspeaker elevation number’ (F(7, 2432) = 24.669, p < 0.001, ηp² = 0.540) are shown to be significant.
The secondary interactions plane*loudspeaker elevation number (F(21, 2432) = 12.241, p < 0.001, \( \eta^2 = 0.366 \)), programme item*loudspeaker elevation number (F(7, 2432) = 3.540, p < 0.001, \( \eta^2 = 0.191 \)) and listener*loudspeaker elevation number (F(7, 2432) = 4.708, p < 0.001, \( \eta^2 = 0.820 \)) and tertiary interaction programme item*loudspeaker elevation number*plane (F(21, 2432) = 1.834, p = 0.015, \( \eta^2 = 0.109 \)) were all shown to be significant. A histogram of the ANOVA model’s standardised residuals showed an approximately normal distribution, indicating that the ANOVA model provided a good fit to the data (Figure 99).

![Histogram of the standardised residuals given by the ANOVA](image)

**Figure 99 - Histogram of the standardised residuals given by the ANOVA**

### 12.2.3 Listener Number

The single factor listener was not shown to be significant, indicating a generally consistent response to each stimulus between listeners. However, the secondary interaction listener*loudspeaker elevation number (F(105, 2432) = 4.708, p < 0.001, \( \eta^2 = 0.820 \)) had a high f value and effect size so the cause of this was
investigated. A plot showing the absolute elevation LRE against listener number separated by loudspeaker elevation number is shown in Figure 100.

![Figure 100 - Absolute elevation LRE plotted against listener number and separated by loudspeaker elevation number. Please refer to Table 1 for the elevation locations of the loudspeaker elevation numbers.](image)

Each data point is only formed of 16 responses and so individual differences between points could just be random variation within the data. However, general trends between listeners for each loudspeaker elevation number can be noted: there is a higher elevation LRE and a reduced inter-listener consistency for loudspeaker elevation numbers 1 and 2 (loudspeakers located below −29° elevation), and 8 (loudspeakers located above +65° elevation), while loudspeaker elevation numbers 4, 5 and 6 (loudspeakers located between +1° and +32° elevation) give a lower absolute elevation LRE and a more consistent listener response. It is the reduced inter-listener consistency for some loudspeaker locations that was highlighted by the ANOVA.
12.2.4 Plane*Loudspeaker Elevation Number

Both single factors ‘plane’ and ‘loudspeaker elevation number’ had a high significance and f value in the ANOVA. The plane*loudspeaker elevation number factor will be investigated first. Figure 101 shows the absolute elevation LRE separated by plane number and loudspeaker elevation number.

![Figure 101 - Absolute elevation LRE separated by plane number and loudspeaker elevation number. Please refer to Table 36 for the elevation locations of the loudspeaker elevation numbers.](image)

On planes 3 and 4 the most accurately reported ‘loudspeaker elevation number’ was loudspeaker position 4. This loudspeaker was located approximately on the equatorial plane. However, for plane 1, loudspeaker elevation numbers 5 and 6 have a significantly lower mean response than loudspeaker elevation number 4.
For plane 2, loudspeaker elevation numbers 4, 5, and 6 were most accurately reported. As loudspeakers are moved from the median plane to the right lateral plane, the most accurately reported loudspeaker elevation appears to shift from between approximately +32° and +15° down to approximately 0° or the equatorial plane. The upwards bias for sources located on the median plane has been reported previously in Chapters 8 and 11.

Wallach [1938] stated that if a listener is unsure of a source location they tend to place it above rather than below the equatorial plane. It is certainly true that for narrow bandwidth sources on the median plane, the listener will be more uncertain of the source location, as there will be fewer useable localisation cues. However, why there is an upwards bias for sources on the equatorial plane is unclear. It may be that a 6 kHz narrow bandwidth noise source highlights a frequency that the listener has learned, from their pinna response, to attribute to that elevation position i.e. for a position slightly above the equatorial plane listeners have learned that there is a maximum at 6 kHz. Evidence to substantiate this hypothesis is shown by Roffler and Butler [1968], who showed that sinusoidal pulses at 4.8 kHz are perceived at approximately +12° elevation regardless of the loudspeaker elevation.

The effect of increasing error as the source is moved from the equatorial plane is more pronounced for loudspeakers located below the equatorial plane. Listeners find it more challenging to localise sources below the equatorial plane than above. Listeners are less familiar with sources appearing to come from below them so they are less likely to respond with those angles if they are unsure. Following the test one listener said that they had speculated at the dimensions of the room and assumed that the loudspeakers would be mounted on the walls. This led them to presume that no loudspeakers were below them and since there were no distinct cues to oppose this view, they reported no loudspeaker locations at low elevations. A second listener stated that they could localise lower elevations more accurately at the side than in front of them. This statement is backed up by Figure 101, which shows that, for loudspeakers below the equatorial plane, the absolute elevation LRE is reduced as the loudspeakers are located further in azimuth from the median plane.
It can be seen that the most accurately localised loudspeakers overall were located near the equatorial plane on planes 3 and 4. This result contrasts the previous full bandwidth experiment in which the position of a loudspeaker on plane 1 (the median plane) was reported as accurately as the loudspeakers on planes 3 and 4. It was suggested in the previous experiment report that the positioning of the calibration point caused this higher LRA for the median plane loudspeaker. In this experiment the calibration point was moved to -6.3° (below the equatorial plane). It is suggested that this movement of the calibration point reduced the response bias.

Planes 3 and 4 were located at ±72° azimuth from the median plane. Therefore, it was anticipated that the LRAs for loudspeakers on plane 3 would be similar to those for loudspeakers on plane 4. However, it can be seen that for loudspeaker elevation number 1 there is a significant difference between the listeners’ responses on these two planes. One reason for this significant difference was the lower location of this loudspeaker on plane 3 (-53°) than on plane 4 (-44°). This agrees with the earlier finding that, the further from the equatorial plane the loudspeaker is located, the harder it is to localise. A targeted T-test was used to find if there are significant differences between planes 3 and 4 when loudspeaker 1 was removed; there were not shown to be significant differences between then planes, \( t = 0.646 \) \( p = .519 \).

For all frontal planes, loudspeaker 1 was significantly more difficult to localise than all other loudspeaker locations. This further indicates that loudspeakers below the equatorial plane are more difficult to localise than those above. On plane 4 there is no significant difference between the elevation LRE reported for loudspeaker elevation numbers 1 and 8. It would appear that for rearward sources, there is less of a difference in accuracy between sources above and below the equatorial plane.

**12.2.5 Plane Number**

Plane number was shown by the ANOVA to significantly affect the absolute elevation LRE data. To most directly answer the research question posed in the introduction, it is simplest to look at the plane number variable alone. A Kruskal-
Wallis non-parametric test confirmed that plane number significantly affects the absolute elevation LRE \( (H (3) = 45.430, p < 0.001) \). Figure 102 shows the absolute elevation LRE separated by plane number.

![Figure 102](image)

*Figure 102 - Absolute Elevation LRE plotted against plane azimuth measured from the median plane*

As the loudspeakers are located further in azimuth from the median plane, the absolute elevation LRE is reduced. This means that listeners are more accurate at reporting the location of loudspeakers on planes that are further in azimuth from the median plane. It can also be seen that the confidence intervals get smaller as the plane number increases. This is caused by a reduced variation in response with elevation for the planes further from the median plane.

**12.2.6 Programme Item**

The programme item factor was shown by the ANOVA to be significant \( (F (1, 2432) = 9.673, p = 0.007, \eta^2 = 0.392) \). However the secondary interaction
programme item*speaker elevation number was also shown to be significant with a high f value ($F (7, 2432) = 3.540$, $p = 0.002$, $\eta^2 = 0.191$) and so this interaction will be investigated first. Figure 103 shows the absolute elevation LRE plotted against loudspeaker elevation number and programme item.

![Figure 103](image)

**Figure 103 - Absolute elevation LRE separated by loudspeaker elevation number and programme item**

Figure 103 shows that for loudspeaker elevation numbers 1, 2 and 3 (below the equatorial plane) the octave programme item has a lower mean response error than the half-octave programme item. Using a targeted t-test for loudspeakers below the equatorial plane (loudspeaker elevation number < 4) it was shown that the difference between the octave and half-octave programme items was significant ($p = 0.019$). Interestingly for loudspeaker elevation numbers 4, 5 and 6 the half-octave programme item actually shows a lower mean absolute
response error, though a targeted t-test for these loudspeaker elevation numbers did not show a significant difference (p = 0.081).

The results of the ANOVA for the single factor ‘programme item’ were checked using a non-parametric Mann-Whitney Test (Table 39).

<table>
<thead>
<tr>
<th>Programme Item</th>
<th>N</th>
<th>Mean Rank</th>
<th>Sum of Ranks</th>
</tr>
</thead>
<tbody>
<tr>
<td>abs_ele_err</td>
<td>1216</td>
<td>1180.19</td>
<td>1435107.50</td>
</tr>
<tr>
<td>Octave</td>
<td>1216</td>
<td>1252.81</td>
<td>1523420.50</td>
</tr>
<tr>
<td>Half Octave</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2432</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test Statistics²</th>
<th>abs_ele_err</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>695171.500</td>
</tr>
<tr>
<td>Wilcoxon W</td>
<td>1435107.50</td>
</tr>
<tr>
<td>Z</td>
<td>-2.550</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.011</td>
</tr>
</tbody>
</table>

² Grouping Variable: Programme_Item

Table 39 - Results of Mann-Whitney non-parametric test

The results also show a significant variation in absolute elevation LRE based on programme item (p = 0.011). The difference in mean absolute elevation response error between the programme items was very low (1.8°) and an absolute elevation response error of approximately 30° for both cases indicates that listeners were uncertain of the source location.

In order to find how the static elevation LRA varied on each vertical plane, the absolute elevation LRE was plotted against plane number for each programme item in Figure 104.
This plot and the ANOVA show that, within each plane, the octave and half-octave programme items are not significantly different. Therefore the static cues given by the octave and half-octave programme items on each plane are not significantly different. Thus, these non-significant static cues cannot be the cause of the significant differences between the octave and half-octave programme items shown in the previous head movement experiment in Chapter 10. Furthermore the difference in absolute elevation LRE between the two programme items in the previous bandwidth experiment was approximately 10°, much larger than the 1.8° mean improvement shown here. Therefore a dynamic cue, created through the movement of the head, must contribute to the improvement in LRA with head movement.
This plot also shows that, for each programme item, there are significant differences in absolute elevation LRE between planes. It is shown that as the loudspeaker is moved further in azimuth from the median plane the elevation LRA is increased.

12.3 Conclusions and Further Discussion

The primary goal of this experiment was to answer the following question:

- For reduced bandwidth sources, is elevation LRA higher for vertical planes further in azimuth from the median plane?

For both the octave and half-octave bandwidth programme items, static elevation LRA is significantly higher for sources further in azimuth from the median plane. This finding corroborates the finding of the previous study on full bandwidth sources. When combined with the statement that listeners move their heads to the source azimuth when they perceive a sound, it can be concluded that: for both the octave and half-octave programme items, the static cues created through yaw head movement do not improve elevation LRA. Therefore, a dynamic cue, created through movement of the head, must cause the improved elevation LRA for reduced bandwidth sources.

12.3.1 Effect of Calibration Point on LRA

For the loudspeakers located at 72° and 108° azimuth from the median plane, the elevation location of loudspeakers at elevations far from the equatorial plane are less accurately reported than those nearer. On the median plane, the area of highest accuracy shifts up to centre on loudspeakers 5 and 6, which are located at +14° and +32° respectively. This upwards bias was also shown in the previous full bandwidth experiment. The locations of sources in front of the listener are more accurately reported if they are located slightly above the equatorial plane. From previous experiments it was thought that the positioning of the calibration point might have caused this bias but in this experiment, with the calibration point moved to below the equatorial plane, there was still an upwards localisation bias for median plane sources. Previous experiments have shown
that tonal sources are positioned in elevation based on their frequency not the position of the loudspeaker. Roffler and Butler showed that a tonal source with a centre frequency between 4.8 kHz and 7 kHz would be positioned between approximately +12° and +21°. It may be that some pitch-based elevation localisation cues are affecting the listener’s response due to the narrow bandwidth of the sources used in this test.

For both the octave and half-octave sources, the most accurately localised loudspeaker was located at 0° elevation on planes 3 and 4. Both planes were located the furthest in azimuth from the median plane in this study. This finding is in contrast to the full bandwidth study, in which the most accurately reported loudspeaker was on plane 1. It is suggested that the higher LRA shown for plane 1 in the previous study was created by the position of the calibration point. In this study the calibration point was repositioned and loudspeaker 5 on plane 1 was no longer the most accurately reported.

12.3.2 Effect of Loudspeaker Elevation on Elevation LRA

In general, the further the source was located from the equatorial plane the less accurately its location was reported. On each vertical plane, loudspeaker elevation number 1 (located between −53° and −44° elevation) showed a significantly higher absolute elevation LRE than all other loudspeakers. This substantiates the suggestion that listeners find it more difficult to localise sources below the equatorial plane than above it.

The elevation LRA was higher for loudspeakers on plane 4 than those on plane 3. This finding was initially surprising as both planes were located at the same distance in azimuth from the median plane and it was anticipated that sources in front of the listener would be most accurately localised. However, the loudspeakers on plane 3 were displaced further below the equatorial plane than the loudspeakers on plane 4 and it has been shown that loudspeakers below the equatorial plane are less accurately reported than those above.
13 Conclusions

A number of research questions regarding head movements in localisation were stated in Chapter 1. This chapter will answer each of these research questions and then give a summary of the original contributions to knowledge given by the thesis. The final section will detail the impact of these findings on localisation models and sound reproduction systems, as well as indicating the areas for further research.

13.1 Research Questions

The primary research question posed in Chapter 1 was:

*How do yaw based head movements affect a listener’s elevation LRA?*

In order to answer this question fully, a number of smaller research questions were formulated. These research questions are answered below.

- *How do listeners move their heads when localising a sound source?*

Chapter 2 showed that head movements can be described in terms of yaw, pitch and roll rotation angles. Yaw head movements cover the largest displacement angle and are the most frequently occurring rotational head movement. Pitch and roll movements do occur but they are most often used in conjunction with yaw movements. Yaw movements are fundamental to Wallach’s dynamic localisation cue as they produce changing interaural differences, which can be used to calculate both source azimuth and elevation. Pitch movements produce no changing interaural cues because the ears remain stationary during the movement. The findings of Chapter 2 directed the focus of the thesis onto yaw-based head movements.

Chapter 2 also showed that, when free to move their heads, listeners are most likely to orient toward the source location. This statement provides the basis for the experiments described in Chapters 11 and 12, which study whether these head movements improve static LRA, and subsequently whether overall improvements in LRA with head movements are due to dynamic or static cues.
• *How do the experiment conditions affect the LRA of the listeners?*

Chapters 2 to 7 show that a listener’s LRA can be dramatically affected by the use of different experimental methods. Visual cues, response methods, loudspeaker location, programme item, and head movement conditions, are all factors that can affect the outcome of the experiment and care must be taken when selecting an experimental method.

In this thesis, the effects of experimental conditions were limited by controlling certain experiment factors. To limit the effect of visual cues, all loudspeaker locations were hidden behind an acoustically transparent but visually opaque curtain. Following the response methods experiment, all experiments used a laser pointing localisation response method. Using a consistent response method allowed the results for the different factors and experiments to be compared. Loudspeakers were positioned to cover a variety of elevation and azimuth locations, to randomise the effect of the loudspeaker location factor.

• *What are the static localisation cues?*

Chapter 4 showed that interaural time differences (ITDs) and interaural level differences (ILDs) are the primary localisation cue for sources on the equatorial plane. These differences occur because the ears are located on opposite sides of the head; this difference in location causes time and level differences in the signals that arrive at each ear, which are used by the listener to calculate source location. When extended to three dimensions, interaural differences give a whole series of possible source positions, called the cone of confusion. Spectral modifications, primarily from the pinna, can resolve the cone of confusion in some cases. However, knowledge *a priori* of the sound source spectrum is required for spectral cues to be used effectively.

• *What is the current evidence that head movements are an elevation localisation cue?*

Chapter 5 showed that whether head movements significantly improve elevation LRA is unclear from previous literature. Indications that dynamic cues given
through head movements affect elevation localisation were given in Wallach [1938], Thurlow and Runge [1967] and Perrett and Noble [1997]; however the extensive study of Wightman and Kistler [1999] could find no trace of the dynamic cue affecting elevation localisation. This led to the yaw head movements in localisation experiment described in Chapter 8.

• How can the effect of pinna cues be suppressed so that head movement cues can be studied more effectively?

Chapter 6 showed that low-pass filtered noise reduces static elevation LRA by suppressing pinna cues. A novel finding given by the pilot study in this chapter was that low-pass filtered noise with a narrowband high frequency component reduces elevation LRA to the same degree as low-pass filtered noise. Thus, the addition of the narrowband high frequency component does not stimulate a listener’s pinna cues. The results of Chapter 8 also confirmed this finding. Use of this programme item allowed the study of head movement cues with both ITD and ILD cues present but pinna cues suppressed.

Initial efforts to suppress pinna cues showed that use of broadband spectrally swept noise does not reduce a listener’s elevation LRA. This indicates that pinna cues may be integrated over time rather than calculated instantaneously. Further investigation based on this hypothesis is beyond the scope of this thesis but may be an area for further work.

• Which localisation response method most accurately measures the perceived source location of the listener?

Chapter 7 showed that a laser pointing response method is more accurate and consistent than either a graphical or verbal method for eliciting the reported source location from a listener in three-dimensions. In elevation, verbal and graphical responses tend to pool around anchor locations causing the location of some loudspeakers to be reported more accurately than others. This result informed the listening tests conducted for this thesis; all tests used a laser pointing response method.
• Do head movements improve elevation LRA?

Chapter 8 showed that yaw based head movements improve both elevation and azimuth LRA for noise programme items. When pinna cues are impaired, the significance of these head movement cues increases. This finding opposes that of Wightman and Kistler [1999], who state that no improvement in elevation LRA is given by head movement.

For sources within the listener’s field of fixation, forced yaw head movements improve elevation LRA to the same degree as free head movements. For sources outside the field of fixation there is a small improvement in LRA with free movement due to increased pointing accuracy.

Yaw based head movements also significantly reduce front-back errors for noise programme items; a finding that agrees with previous research [Thurlow and Runge 1967; Pollack and Rose 1967; Perrett and Noble 1997; Wightman and Kistler 1999]. When head movements are not permitted, the bandwidth of the programme item significantly affects the number of front-back errors: as the bandwidth of the programme item is reduced, the listener’s pinna cues are degraded, which results in a higher proportion of front-back errors. Forced yaw and free head movements both remove this bandwidth dependent effect.

• Is the improvement in elevation LRA with yaw head movements due to static or dynamic cues?

Chapter 2 showed that, when free to move their heads, listeners orient towards the direction of sound source location. Chapter 3 showed that, in azimuth, this allows listeners to move the source into a more accurate area of audition (azimuth LRA is highest for sources directly in-front of the listener, i.e. on the median plane), and so free head movement improves their azimuth LRA. Therefore the improvements in azimuth LRA with head movement are primarily due to improved static cues created by turning to face the source.

Chapter 11 showed that, when a listener is unable to move their head, the further the loudspeaker is placed in azimuth away from the median plane, the
more accurate the elevation LRA. When coupled with the finding that, when free to move their head, listeners will orient to face the source, it can be concluded that yaw based head movements actually reduce static cue elevation LRA. Therefore, dynamic cues created during head movement are the primary cause of increased elevation LRA with head movement.

- **What cue, created by head movement, causes an improvement in elevation LRA?**

Chapter 9 showed that free head movements improve elevation LRA to a similar degree for bandpass (one octave wide at -10 dB) noise signals centred at 0.5 kHz, 2 kHz and 6 kHz; this finding shows head movement cues are dependent on both dynamic ILDs and dynamic ITDs. Furthermore there is a similar improvement offered by both interaural cues. This result contradicts the findings of Perrett and Noble [1997], who state that energy below 2 kHz is required for head movement cues to be effective.

Chapter 10 showed that head movements do not improve the elevation LRA for programme items with less than an octave bandwidth. In static listening conditions, localisation of octave-band noise at 6 kHz was not significantly different from half-octave and quarter-octave sources. Therefore, pinna cues do not improve the elevation LRA for sources of an octave bandwidth or less. The octave noise programme item gained a significant improvement in LRA when dynamic listening conditions were allowed. This indicates that dynamic cues are dependent on ITD and ILD cues, and independent of pinna cues.

- **Are head movement cues for narrowband programme items dependent on static or dynamic cues?**

Chapter 12 showed that, for both the octave and half-octave bandwidth programme items, static elevation LRA is significantly higher for sources further in azimuth from the median plane. Furthermore the small error differences in accuracy given by the octave and half-octave programme items do not account for the much larger differences in elevation LRA shown when comparing the free and no movement conditions in the bandwidth experiment in Chapter 10. Both
of these findings indicate that dynamic cues given through the yaw movement of the head increase elevation LRA for octave bandwidth programme items.

13.2 Original Contributions to Knowledge

The original contributions to knowledge given by this thesis are listed below. These contributions are grouped into three categories: head movements in localisation, elevation localisation and localisation response methods. All contributions were drawn from experiments conducted by the author and the chapters in which these conclusions were reached are also given.

13.2.1 Head Movements in Localisation

- **Yaw head movements improve both the elevation and azimuth LRA of listeners to noise based programme items** (Chapter 8).
- **Yaw head movements improve elevation LRA by creating dynamic interaural difference cues** (Chapters 8 and 11).
- **Forced yaw head movements improve LRA to the same degree as free head movements for sources within the field of fixation** (Chapter 8).
- **When pinna cues are impeded (using the low pass filtered noise or low pass filtered noise with a narrow band high frequency programme items) the significance of yaw head movement cues increases** (Chapter 8).
- **Yaw head movements are a more significant localisation cue than spectral cues in resolving front-back confusions** (Chapters 8, 9 and 10).
- **When head movements are impeded the significance of pinna cues to the resolving of front-back confusions increases. In this situation, high frequency content significantly decreases the presence of front-back confusions.** (Chapters 8, 9 and 10)
- **Head movements reduce front-back confusions for both noise and tonal programme items.** (Chapter 10)
- **Dynamic ITDs and ILDs both improve elevation LRA with head movement.** (Chapter 9)
- **Head movements do not improve elevation LRA for programme items with less than an octave bandwidth.** (Chapter 10)
13.2.2 Elevation Localisation

- End effects significantly reduce the static elevation LRA to extreme elevation locations (towards ±90° elevation) (Chapters 8 and 11)
- Between bounds of -40° and +60° elevation, static elevation LRA does not vary significantly. (Chapter 11)
- On vertical planes close in azimuth to the lateral plane (the side of the listener) the source elevation position of highest acuity is 0° elevation, while for sources on the median plane the source elevation position of highest acuity is +30° elevation for the octave programme and +15° for the half-octave programme. There is an upward trend in the area of highest acuity as the source is moved towards the median plane. (Chapters 11 and 12)
- The location of loudspeakers positioned below the equatorial plane are less accurately reported than those positioned an equivalent elevation angle above the equatorial plane. (Chapters 11 and 12)
- Visual cues can profoundly change the response of a listener to an auditory stimulus. Positioning a visual cue, such as a calibration point, can significantly bias a listener’s response. (Chapters 7, 8, 11 and 12)
- Mean response bias of elevation localisation response is always toward a central response location. (Chapters 7-12)
- Low pass filtered noise reduces elevation LRA by limiting pinna cues (Chapters 6 and 8)
- Low pass filtered noise with a narrowband high frequency component reduces elevation LRA to the same degree as low pass filtered noise. Thus, the addition of the narrowband high frequency component does not stimulate a listener’s pinna cues. (Chapter 8)

13.2.3 Localisation Response Methods

- A laser pointing response method is more accurate and consistent than either a graphical or verbal method for eliciting the reported source location from a listener. (Chapter 7)
13. Conclusions

- *The responses from Verbal and Graphical methods tend to pool around anchor locations, causing the location of some loudspeakers to be reported more accurately than others. (Chapter 7)*

13.3 Applications

The findings of this thesis can inform the development of an improved head-movement-aware localisation model, providing a more accurate representation of the human localisation system. Since head movements significantly improve three-dimensional LRA, head movement cues should be incorporated into localisation models. Yaw based head rotations create dynamic localisation cues that improve elevation LRA. Therefore localisation models based on dynamic cues created through head movement, such as the multiple-microphone sphere model developed by Ashby et al. [2010], have some physiological basis. A physiologically valid localisation model will incorporate head movement cues. Since head movements are the primary front-back localisation mechanism, it is extremely important that these be included in a localisation model. Localisation models that use only pinna cues to resolve front-back confusions are physiologically invalid and so are less useful when creating a perceptually based localisation model.

Laser pointing methods are more accurate and consistent than either graphical or verbal response methods; this finding can be used to justify pointing techniques in further localisation studies. Furthermore it guides the analysis of previous experiments, allowing the relative accuracies of the response methods to be compared. For example, if a localisation study uses a verbal response method then it is likely that the listeners’ LRAs will be inconsistent across location and this must be considered when conclusions are drawn.

Low-pass filtered noise can be used in further experiments where a researcher desires to reduce the effects of pinna cues. Furthermore low-pass filtered noise with a half-octave 6 kHz bandpass component can be used to reduce pinna cues when high frequency content is to be retained. This programme item can also be
used to increase the effects of head movement cues so they can be studied more directly.

Elevation LRA on the median plane is highest for sources located at +30°. Therefore, it is important to be aware that if, for example, a loudspeaker is located far below the equatorial plane, then the listeners will find it difficult to localise. If loudspeakers are to be accurately localised then they should not be located significantly below or above the optimum response angle. Furthermore, to give the impression of a loudspeaker located below the equatorial plane it may be necessary to emphasise the loudspeaker’s elevation displacement away from the optimum response angle to account for the response bias back towards it. For example, a sound that is intended to be perceived at -30° may need to be reproduced from a loudspeaker positioned below -30°. The bandwidth of the source programme item should also be considered when selecting the loudspeaker location for a source, if a higher LRA is desired.

Elevation LRA is higher for loudspeakers located on planes further in azimuth from the median plane. Therefore if the location of a loudspeaker is to be most accurately reported in elevation then it should not be placed on median plane. Once again, this is especially true for narrower bandwidth sources. This finding will affect the placement of loudspeakers in three-dimensional loudspeaker arrays; for a higher LRA, loudspeakers should be placed away from the median plane.

This thesis provides further evidence that both ITD and ILD cues should be dynamically varied with head movement for the listener to experience immersive three-dimensional sound. Binaural reproduction systems that do not include head movements will reduce the listener’s LRA, resulting in an unnatural spatial experience. For some reduced bandwidth sources dynamic cues given by head movement are the only major elevation localisation cue.

Head movements allow a subject to more accurately understand their orientation and location in a space, as well as the orientation and location of other objects. Models such as KEMAR and HATS that do not incorporate head movement cues limit their application and extension. As elevated sources are
becoming common in spatial reproduction systems, being aware of head movement as a primary elevation localisation cue allows one to conduct physiologically valid, accurate and fast analysis of these systems.

13.4 Further Work

This section will highlight the areas of further work that have been noted during the experiments in the thesis.

13.4.1 The Calculation of Pinna Cues Over Time

Chapter 4 stated that listeners require \textit{a priori} knowledge of the source spectrum to allow high elevation LRA using pinna cues when head movements are not permitted. The results of Pilot Experiment 1 described in Chapter 6 suggested that if the spectrum of the source is varied over time then pinna cues can be averaged to allow the underlying source spectrum to be detected. This method would allow accurate elevation LRA without \textit{a priori} knowledge of the source spectrum. The primary research question arising from these conclusions is:

- \textit{Does a source with a time-varying spectrum allow a higher elevation LRA using pinna cues?}

13.4.2 Bias of Visual Cues on LRA using Verbal and Graphical Response Methods

The Localisation Response Methods Experiments in Chapter 7 showed that the verbal and graphical response methods do not give a constant azimuth or elevation LRA across loudspeaker location. Loudspeakers located at 'anchor' locations (i.e. directly in-front of (0º,0º) or behind the listener (180º,0º), directly above (0º,90º) or to the right (90º,0º)) were reported significantly more accurately than intermediate loudspeaker locations. This effect was not shown in the laser-pointing response method, and was a characteristic of the response method, not the listeners’ localisation acuities. The primary research question arising from these conclusions is:
• Why do verbal and graphical response methods result in higher LRA at anchor locations?

It was suggested that this effect occurred because the listeners had coarse response precision using these two methods; they were unlikely to respond in increments of less than 5°. At anchor locations this coarse precision lined up with an actual loudspeaker locations, thus resulting in an higher LRA. To test whether the improved LRA for anchor locations was due to this course response precision, the experiment should be repeated but the locations of the loudspeakers rotated slightly so that no loudspeaker lines up exactly with anchor locations. It is suggested that this will decrease the LRA and improve the localisation response consistency across loudspeaker location for the verbal and graphical response methods.

13.4.3 Elevation LRA with Loudspeaker Elevation

Chapters 6, 8, 11 and 12 noted that when localising on the median plane without head movements, the loudspeaker location reported with the highest elevation LRA was positioned at approximately +30° elevation. Chapter 12 showed that for sources closer to the lateral plane, the loudspeaker location reported with the highest elevation LRA was positioned on the equatorial plane. Furthermore Chapter 9 and 10 all showed that the position of highest elevation LRA was above the equatorial plane, both with and without head movements. The primary research question arising from these conclusions is:

• Why do loudspeakers located at approximately +30° elevation on the median plane result in the highest elevation LRA?

One would anticipate that sources on the equatorial plane might be reported most accurately, as these are the source locations with which listeners would be most familiar. It was suggested in Chapter 6 that this trend may been due to the positioning of the acoustically transparent curtain. Subsequent analysis of the results suggested this was unlikely however conclusive proof was not given. A further experiment is suggested to find the effect of the location of the acoustically transparent curtain on the listeners’ localisation responses.
13.4.4 The Effects of Pitch and Roll Head Movements on Elevation LRA

It has been shown that static elevation cues are not improved by yaw head movements towards the source location. Further research could be conducted investigating the effect of pitch and roll head movements on static LRA. It may be that rotating the pitch of the listener's head toward the source location improves elevation LRA. The primary research question arising from these conclusions is:

- What is the effect of pitch and roll head movements on elevation and azimuth LRA?

Higher acuity with pitch movement was indicated by the higher elevation LRA for sources closer to the equatorial plane shown in Chapters 11 and 12; however it may be that this was caused by a response bias i.e. if the listener is uncertain of a source location then they are likely to report one closer to the equatorial plane. An experiment similar to Chapter 11 is advised, but where the listeners head is offset in elevation at the start of each trial. By comparing localisation responses using each elevation offset it should be possible to ascertain whether the higher LRA with pitch movement is due to a localisation response bias or higher localisation acuity for sources nearer the equatorial plane.
Appendix A - Reported Location Calculation Method

This section shows how the source location perceived by the listener was calculated from the head-tracker data. The output of the Polhemus headtracking system and the setup of the experiment are described and the overall steps required to get from the raw data to the reported source location are outlined. The terms *central listening position, actual loudspeaker location* and *reported location* are defined and calculated.

A.1 Head Tracking System and Experiment Setup

For each test stimulus, the headtracker output its location using six degrees of freedom; three Cartesian coordinates \((x, y, z)_t\), and three orientation angles \((\varphi, \theta, \tau)_t\).

The Cartesian coordinates were measured in inches from the transmitter to the receiver. The transmitter’s position was kept constant during all trials, fixed atop a microphone stand. The receiver was positioned on the listener’s head using a headband. Care was taken to place both transmitter and receiver as far as possible from interferers.

Orientation angles were calculated as the difference in orientation between the transmitter’s and receiver’s coordinates systems. Orientation angles were calculated in the following order: rotation about the z-axis, azimuth angle \((\varphi)\); rotation about the y-axis, elevation angle \((\theta)\); and rotation about the x-axis, tilt angle \((\tau)\). In this section, all angles will be given in degrees in accordance with the headtracker’s output. The orientation of the receiver with respect to the listener’s laser pointer varied between listeners. To correct for this a calibration stage was included (described in Section A.6).

Figure 105 shows how the tracker was set up during the localisation response methods experiment described in Chapter 7. The setup for all experiments was similar; only the precise distances between objects varied slightly for each
The z-axis output from the tracker was positive in the downward direction. Elevation was measured from the equatorial plane with upwards movement being the positive elevation angle displacement, while azimuth was measured from the x-axis, with rightwards movement being the positive azimuth angle displacement.

The x-axis was not positioned to point forward with respect to the listener because the tracker has two distinct hemispheres, one extending positively and one negatively in terms of x. As stated in the Patriot Guide, “only half of the total spatial sphere surrounding the source can be utilized at any one time for unambiguous position measurement”. Therefore, the listener had to remain within one hemisphere for the duration of the task. This was only possible using the experimental setup shown below if the x-axis were to extend to the right of the listener. Therefore, -90° in azimuth was directly in front of the listener, while +90° was directly behind the listener.

*Figure 105 - Setup of tracker for the experiment in Chapter 7*
A.2 Coordinate Systems

Two different locations were used as origins in this calculation: the Polhemus transmitter location and the central listening position.

The Polhemus transmitter location was used as the origin for the raw data given by the Polhemus head tracker system. In this calculation, coordinates given with respect to the Polhemus transmitter will be denoted as \((x, y, z)\).

The central listening position, \((x, y, z)_c\), was the point from which all loudspeaker locations were measured. It was located at \((19.7 \ 0 \ 0)\) from the head tracker base for the localisation response method experiment. The location of the listener's head varied during the test but the central listening position remained constant. Any location given with respect to the central listening position will be denoted as \((x', y', z')\).

Some Cartesian coordinates given in the calculation do not reference either origin. These coordinates are called difference vectors and describe the change in location between two given points.

A.3 Calculation Overview

Figure 106 and Figure 107 show the main locations described in the calculation. The location from which the listener hears the sound originate is called the reported location. The loudspeaker location is the point at which the loudspeaker is actually located. A comparison of these two points allows the LRA to be resolved.
Figure 106 – Main locations used to calculate LRA (elevation view)

Figure 107 - Main locations used to calculate LRA (azimuth view)
A.4 Actual loudspeaker Angle

The Cartesian coordinates of each loudspeaker location \((x', y', z')_a\), measured from the central listening position, were derived from the vertices of a truncated icosahedron in Chapters 7 to 10. The Cartesian coordinates of each loudspeaker location are given in Appendix B.

The actual loudspeaker angle was defined as the azimuth and elevation angle from the central listening location to the loudspeaker location. \((x', y', z')_a\) was used to calculate the radial distance to source, \(r'_a\), the actual source elevation, \(\theta'_a\), and the actual source azimuth, \(\varphi'_a\).

\[
\begin{align*}
    r'_a &= \sqrt{x'_a^2 + y'_a^2 + z'_a^2} \\
    \theta'_a &= \cos^{-1}\frac{z'_a}{r'_a} \\
    \varphi'_a &= \tan^{-1}\frac{y'_a}{x'_a}
\end{align*}
\]

This radial distance will be used as the radius to loudspeaker in the calculation of reported location in Section A.7.

A.5 Coordinates of the Head Centre

To find the reported source location, angles must be calculated from the centre of the listener’s head. The headtracker was located on top of the listener’s head and so it was necessary to convert the coordinates given by the tracker into the coordinates of the listener’s head centre.
The relationship between the listener’s head centre and the receiver was assumed to be constant, i.e. the receiver didn’t move during each test. The calculations for this stage were described by Kim [2009]. The radius of the head was assumed to be 13cm, in accordance with the dimensions given by Buckhard and Sachs [1975]. Other measurements, such as the thickness of headband and distance to receiver magnetic centre, resulted in a total radial distance of 14cm or 5.51in. This calculation resulted in the Cartesian coordinates of the centre of the listener’s head, \((x, y, z)_h\).

The orientation angles of the centre of the listener’s head were the same as the orientation angles of the top of the listener’s head. Therefore, the transform calculation did not affect the orientation angles.

**A.6 Calibration**

In the localisation response methods experiment, preceding every trial listeners were instructed to direct the laser pointer towards two calibration points, one in front of them, A, and one behind, B. These calibration points were used to establish the relationship between the laser pointer and the head tracker.
orientation angles. The calibration points, \((x, y, z)_n\), were at Cartesian coordinates A and B. The calculation described below was conducted for both calibration points, A and B.

![Diagram of calibration point calculation](image)

**Figure 109 - Angles used in calibration point calculation**

When the listener was pointing the laser at the calibration point, the Cartesian coordinates \((x, y, z)_t\), and orientation angles \((\varphi, \theta, \tau)_t\) of the tracker were logged. The location of the listener’s head, \((x, y, z)_h\) was given by the head centre calculation of Section A.5. The difference vector, \((x, y, z)_m\) was calculated:

\[
(x, y, z)_m = (x, y, z)_n - (x, y, z)_h
\]  

(9)

This was then used to calculate the radius, \(r_m\), elevation, \(\theta_m\), and azimuth, \(\varphi_m\), of the calibration points.

\[
r_m = \sqrt{x_m^2 + y_m^2 + z_m^2}
\]

\[
\theta_m = \cos^{-1}\frac{z_m}{r_m}
\]

(10)

\[
\varphi_m = \tan^{-1}\frac{y_m}{x_m}
\]

The difference between the elevation, \(\theta_m\), and azimuth, \(\varphi_m\), given by this calculation, and the elevation, \(\theta_t\), and azimuth, \(\varphi_t\), given by the headtracker defined the relationship between the head tracker and the laser. These were called the laser offset angles, elevation, \(\theta_l\), and azimuth, \(\varphi_l\).
The laser offset angles were averaged between points A and B. The mean differences between the laser offset angles for points A and B were 1.12° and 1.82° in azimuth and elevation respectively for the localisation response methods experiment in Chapter 7.

A.7 Reported Location

In the experiment, the listener to loudspeaker angle varied depending on the position of the listener’s head. It would be inaccurate to compare orientation angles of elevation and azimuth without considering the position of the listener’s head during the response. In this section the reported location is defined and its derivation detailed.

The reported location is the point at which a line extending from the listener’s head, defined by the listener’s head location and orientation angles, pierces a hypothetical sphere surrounding the central listening position at the same radial distance as the loudspeaker location. This situation is shown in two dimensions in Figure 110.
In Section A.4 the radial distance from the CLP to each loudspeaker, $r_a$, was calculated. During each trial the system logged the head tracker position $(x, y, z)_t$ and orientation $(\phi, \theta, \tau)_t$ when the listener clicked ‘Log’. These were used to calculate the coordinates of the centre of the listener’s head $(x, y, z)_h$ in Section A.5.

The tracker angles can be corrected for the laser offset angle described in Section A.6. The corrected elevation, $\theta_f$, and azimuth, $\varphi_f$, angles give the angle of the laser rather than the head tracker.

$$\theta_f = \theta_t + \theta_l$$  \hspace{1cm} (12)$$

$$\varphi_f = \varphi_t + \varphi_l$$
Any point on a line extended out from the listener’s head starting from a point 
\((x', y', z')_e\) at a given azimuth, \(\varphi'_e\), and elevation, \(\theta'_e\), at a distance of \(t'\) can be found by:

\[
x'_p = x'_e + t'\cos\varphi'_e\sin\theta'_e \\
y'_p = y'_e + t'\sin\varphi'_e\sin\theta'_e \\
z'_p = z'_e + t'\cos\theta'_e
\]

To use these equations the corrected elevation, \(\theta_f\), must be converted from an 
equatorial reference to an angle measured from the z-axis, while azimuth 
remains unchanged:

\[
\theta'_e = 90 - \theta_f \\
\varphi'_e = \varphi_f
\]

Since the sphere is to be centred at the central listening position the initial 
coordinates used in the line equation must reference this central listening 
position. This creates the intermediate stage, \((x', y', z')_i\):

\[
(x', y', z')_i = (x, y, z)_h - (x, y, z)_c
\]

To use Equation 11 the z-axis must also be positive in the upward direction:

\[
x'_e = x'_i \\
y'_e = y'_i \\
z'_e = -z'_i
\]

The formula for a sphere of radius, \(r\), centred at \((0,0,0)\) is:

\[
x^2 + y^2 + z^2 = r^2
\]
By replacing \((x, y, z)\) with the line equation we can find the point at which the line and sphere intersect:

\[
(x' e + t \cos \varphi' e \sin \theta'_e)^2 + (y' e + t \sin \varphi' e \sin \theta'_e)^2 + (z' e + t \cos \theta'_e)^2 = r^2
\]  
(18)

Multiplying out the brackets:

\[
(x'^2 + t^2 \cos^2 \varphi' e \sin^2 \theta'_e + 2x' e t \cos \varphi' e \sin \theta'_e) + ... \\
(y'^2 + t^2 \sin^2 \varphi' e \sin^2 \theta'_e + 2y' e t \sin \varphi' e \sin \theta'_e) + ... \\
(z'^2 + t^2 \cos^2 \theta'_e + 2z' e t \cos \theta'_e) = r^2
\]  
(19)

and grouping like terms to create a quadratic in terms of \(t\):

\[
t^2 (\cos^2 \varphi' e \sin^2 \theta'_e + \sin^2 \varphi' e \sin^2 \theta'_e + \cos^2 \theta'_e) + ... \\
t'(2(x' e \cos \varphi' e \sin \theta'_e + y' e \sin \varphi' e \sin \theta'_e + z' e \cos \theta'_e)) + ... \\
(x'^2 + y'^2 + z'^2 - r^2) = 0
\]  
(20)

The coefficient of \(t^2\) can be simplified using simple trigonometric identities:

\[
\cos^2 \varphi' e \sin^2 \theta'_e + \sin^2 \varphi' e \sin^2 \theta'_e + \cos^2 \theta'_e \\
= \sin^2 \theta'_e (\cos^2 \varphi' e + \sin^2 \varphi' e) + \cos^2 \theta'_e \\
= \sin^2 \theta'_e + \cos^2 \theta'_e \\
= 1
\]  
(21)

The quadratic in terms of \(t\) can be solved using the quadratic formula:

\[
t' = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}
\]  
(22)
with,

\[ a = 1 \]

\[ b = 2(x'_e \cos \phi'_e \sin \theta'_e + y'_e \sin \phi'_e \sin \theta'_e + z'_e \cos \theta'_e) \quad (23) \]

\[ c = x'_e^2 + y'_e^2 + z'_e^2 - r^2 \]

The radius of the sphere for each loudspeaker location, \( r_a \), was found in Section A.4 and for each localisation trial, the Cartesian coordinates with reference to the central listening position \((x', y', z')_e\) and corrected orientation angles \((\phi'_e, \theta'_e)\) are calculated.

Once \( t' \) is found, the reported location Cartesian coordinates \((x', y', z')_p\) can be calculated using Equation 22.

\[ x'_p = x'_e + t' \cos \phi'_e \sin \theta'_e \]

\[ y'_p = y'_e + t' \sin \phi'_e \sin \theta'_e \quad (24) \]

\[ z'_p = z'_e + t' \cos \theta'_e \]

**A.8 Reported Loudspeaker Angle**

\((x', y', z')_p\) can be used to calculate the reported azimuth and elevation from the central listening position.

\[ r'_p = \sqrt{x'_p^2 + y'_p^2 + z'_p^2} \]

\[ \theta'_p = \cos^{-1} \frac{z'_p}{r'_p} \quad (25) \]

\[ \phi'_p = \tan^{-1} \frac{y'_p}{x'_p} \]
A.9 Angle Error

The angle error can be calculated as the difference in degrees between the reported and actual loudspeaker angles. The reported and actual loudspeaker angles both reference the central listening position so the angles can be directly compared. The mean absolute error is one indication of the listener’s LRA. By finding the signed error one can discover the offset of each reported location in relation to the actual location of the loudspeaker.

A.10 Further Calibration

For the localisation response methods experiment, following the final trial a further calibration stage was carried out. The laser was pointed at each loudspeaker in turn and its location logged, then the reported location method was used to find its location. Ideally there should be zero error between the reported and actual loudspeaker locations using this test. However, due to measurement errors that occurred while physically placing the loudspeakers at specific Cartesian coordinate locations, there were variations in this error. The mean correction was 1.08° and 1.16° in azimuth and elevation respectively. Since all angles were calculated from the central listening position, these variations could be subtracted from the reported angle before any LRE was calculated.

A.11 Verbal and Graphical Location Calculation

In the localisation response methods experiment described in Chapter 7, the overall location calculation for verbal and graphical responses had only minor variations from the method described above. The reported location calculation was the same, using the raw elevation and azimuth angles from the listener. The listener’s head location was only taken at the beginning of the test and used as the reference location throughout. This method corrected for variations in listener height.
Appendix B – Loudspeaker Locations for Experiments 7 - 10

The coordinates for the loudspeaker locations used in Chapters 7, 8, 9 and 10 are given in Table 40. These points, given in cm, form the vertices of a truncated icosahedron. All measurements are given from the CLP.

<table>
<thead>
<tr>
<th>Loudspeaker Number</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Azimuth (˚)</th>
<th>Elevation (˚)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0</td>
<td>-172.0</td>
<td>0.0</td>
<td>-90.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2</td>
<td>27.7</td>
<td>-129.4</td>
<td>-93.5</td>
<td>-77.9</td>
<td>35.3</td>
</tr>
<tr>
<td>3</td>
<td>93.5</td>
<td>-129.4</td>
<td>27.7</td>
<td>-54.1</td>
<td>-9.8</td>
</tr>
<tr>
<td>4</td>
<td>93.5</td>
<td>-93.5</td>
<td>93.5</td>
<td>-45.0</td>
<td>-35.3</td>
</tr>
<tr>
<td>5</td>
<td>93.5</td>
<td>-93.5</td>
<td>-93.5</td>
<td>-45.0</td>
<td>35.3</td>
</tr>
<tr>
<td>6</td>
<td>129.4</td>
<td>-93.5</td>
<td>-27.7</td>
<td>-35.9</td>
<td>9.8</td>
</tr>
<tr>
<td>7</td>
<td>93.5</td>
<td>-27.7</td>
<td>-129.4</td>
<td>-16.5</td>
<td>53.0</td>
</tr>
<tr>
<td>8</td>
<td>129.4</td>
<td>-27.7</td>
<td>93.5</td>
<td>-12.1</td>
<td>-35.3</td>
</tr>
<tr>
<td>9</td>
<td>172.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>10</td>
<td>129.4</td>
<td>27.7</td>
<td>-93.5</td>
<td>12.1</td>
<td>35.3</td>
</tr>
<tr>
<td>11</td>
<td>129.4</td>
<td>93.5</td>
<td>27.7</td>
<td>35.9</td>
<td>-9.8</td>
</tr>
<tr>
<td>12</td>
<td>93.5</td>
<td>93.5</td>
<td>93.5</td>
<td>45.0</td>
<td>-35.3</td>
</tr>
<tr>
<td>13</td>
<td>93.5</td>
<td>93.5</td>
<td>-93.5</td>
<td>45.0</td>
<td>35.3</td>
</tr>
<tr>
<td>14</td>
<td>93.6</td>
<td>129.4</td>
<td>-27.7</td>
<td>54.1</td>
<td>9.8</td>
</tr>
<tr>
<td>15</td>
<td>27.7</td>
<td>129.4</td>
<td>93.6</td>
<td>77.9</td>
<td>-35.3</td>
</tr>
<tr>
<td>16</td>
<td>27.7</td>
<td>93.5</td>
<td>-129.4</td>
<td>73.5</td>
<td>53.0</td>
</tr>
<tr>
<td>17</td>
<td>0.0</td>
<td>172.0</td>
<td>0.0</td>
<td>90.0</td>
<td>0.0</td>
</tr>
<tr>
<td>18</td>
<td>0.0</td>
<td>0.0</td>
<td>172.0</td>
<td>0.0</td>
<td>90.0</td>
</tr>
</tbody>
</table>

Table 40 – Loudspeaker numbers, coordinates of loudspeaker locations (cm), and loudspeaker angles from CLP
Appendix C – List of Publications


This paper used Wallach’s study [1938] as a basis to create a head-movement-aware localisation model using a multiple microphone sphere. This paper prompted the research described this thesis, to find out whether the head-movement-aware spatial localisation model was physiologically valid.


This paper primarily described the work presented in Chapter 8, finding the effect of head movements on three-dimensional LRA. Literature described in Chapters 2 to 5 was referenced and the laser guided localisation response method studied in Chapter 7 was used.


This paper described the work presented in chapter 11 of this thesis, with the goal of finding how elevation LRA changed on vertical planes of differing azimuth.


This was a summary paper that described the experiments in Chapters 6 to 11 of this thesis.
Data Archive

The data underlying the findings presented in this thesis are available from doi: 10.5281/zenodo.19034. Further project information can be found at http://iosr.uk/elevation.
References


Hartmann, W. M. (1990). Localization of a source of sound in a room. AES 8th International Conference, New York, USA, May 1


Kim, C., Brookes, T., & Mason, R. (2010). Validation of a Simple Spherical Head Model as a Signal Capture Device for Head-Movement-Aware Prediction of Perceived Spatial Impression. AES 40th International Conference, Guildford, UK, October 8


Rayleigh (1907) 'On Our Perception Of Sound Direction', Philosophical Magazine, 8, 214-232


Young, P. (1931). The rôle of head movements in auditory localization. Journal of Experimental Psychology, 14(2), 95–1