The Influence of Loudspeaker Directivity upon the Perception of Reproduced Sound in Domestic Listening Rooms

by

William Evans

Submitted for the degree of Doctor of Philosophy

Institute of Sound Recording
School of Arts
Faculty of Arts & Human Sciences
University of Surrey

June 2013
©2013 William Evans
Abstract

The directivity of a loudspeaker characterises how it radiates sound into the space around it. Loudspeaker directivity is thought to influence the perception of reproduced sound in a room but the types and magnitudes of its effects have yet to be determined. Theory from acoustics and psychoacoustics literature indicates that changes in the characteristics of direct and reflected sound (as a result of changes in directivity, loudspeaker/listener position and room characteristics) will cause small changes in the perceived attributes of reproduced sound. These attributes include timbre, apparent source width, localisation, loudness, envelopment/spaciousness and distance. Tests using loudspeakers with a variety of directivities confirm that all of these attributes are affected.

In order to more fully characterise and quantify these effects in a controlled fashion, an auralisation system is shown to make an appropriate platform for listening experiments. A series of experiments using such a system with a novel elicitation and analysis technique reveals that: changes in loudspeaker directivity interact with changes in room surface absorption to mostly affect perceived width, loudness and reverberence; a narrowing in on-axis directivity is associated with a perceived reduction in width, brightness, closeness and spaciousness; the perceptual changes caused by variations in loudspeaker directivity and in room surface absorption occur in parallel, along one dimension; the magnitude of effects caused by variations in loudspeaker directivity is reduced with the increased presence of reflections; and classical music is an effective signal to highlight the perceptual differences between different loudspeaker directivities. The influence of listener position can be significant, but the degree of its effect is dependent on the range of positions evaluated, and the magnitude of boundary separation has less effect than directivity on the perception of reproduced sound. Finally, loudspeaker directivity does influence listener preference.
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>i</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>xxiii</td>
</tr>
<tr>
<td>Statement of Originality</td>
<td>xxiv</td>
</tr>
<tr>
<td><strong>1 Introduction</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>2 Loudspeaker Directivity and Reproduced Sound</strong></td>
<td>5</td>
</tr>
<tr>
<td>2.1 Loudspeaker Directivity</td>
<td>6</td>
</tr>
<tr>
<td>2.1.1 Acoustic Radiation</td>
<td>6</td>
</tr>
<tr>
<td>2.1.2 Defining Directivity</td>
<td>11</td>
</tr>
<tr>
<td>2.1.3 The Directivity of Loudspeakers</td>
<td>16</td>
</tr>
<tr>
<td>2.1.4 Loudspeaker Directivity Control</td>
<td>22</td>
</tr>
<tr>
<td>2.1.5 Section Summary and Discussion</td>
<td>25</td>
</tr>
<tr>
<td>2.2 Reproduced Sound in Domestic Listening Rooms</td>
<td>26</td>
</tr>
<tr>
<td>2.2.1 Loudspeaker Reproduction in Rooms</td>
<td>26</td>
</tr>
<tr>
<td>2.2.2 The Effect of Changing the Directivity of a Loudspeaker Upon the</td>
<td>30</td>
</tr>
<tr>
<td>Sound Field</td>
<td></td>
</tr>
<tr>
<td>2.2.3 Section Summary and Discussion</td>
<td>32</td>
</tr>
</tbody>
</table>
CONTENTS

2.3 Chapter Summary ................................................................. 33

3 Perception of Reproduced Sound ........................................... 35
  3.1 Auditory Analysis and Perception of Reproduced Sounds in Rooms . . . . 36
    3.1.1 Human Auditory System and Basic Auditory Analysis ................. 36
    3.1.2 Non-Spatial Auditory Analysis ................................................. 38
    3.1.3 Spatial Auditory Analysis ...................................................... 43
    3.1.4 Section Summary and Discussion ........................................... 48
  3.2 Reflections - Considering Adaptation and Preference ....................... 50
    3.2.1 Section Summary and Discussion ........................................... 52
  3.3 Loudspeaker Directivity and the Perception of Reproduced Sound (Previous Research) .......................................................... 52
    3.3.1 Initial Discussion ................................................................. 52
    3.3.2 Previous Studies ............................................................... 54
    3.3.3 Section Summary and Discussion ........................................... 59
  3.4 Chapter Summary ............................................................... 60

4 Preliminary Studies .............................................................. 64
  4.1 Experiment 1: Listening Tests in Surrey, UK ................................... 65
    4.1.1 Introduction/Aims ............................................................... 65
    4.1.2 Experimental Set-up .......................................................... 65
    4.1.3 Stage 1: Elicitation Experiment ............................................ 71
    4.1.4 Choosing the Test Attributes ................................................. 72
    4.1.5 Stage 2: Rating Experiment ................................................ 75
    4.1.6 Listener assessment and screening ........................................ 76
    4.1.7 Results ............................................................................... 81
    4.1.8 Summary/Discussion .......................................................... 87
CONTENTS

6.8.3 Relationship Between Dimension and Attributes .......... 205
6.8.4 Effect of Boundary Conditions and Listener Position ........ 206
6.8.5 Summary ................................................................. 209
6.9 Discussion ................................................................. 210
6.9.1 Dimensionality ....................................................... 210
6.9.2 Relationship Between Perceptual Dimension and Loudspeaker Directivity ........................................... 211
6.9.3 Relationship Between Perceptual Dimension and Attributes . 211
6.9.4 Effect of Boundary Conditions (Absorption and Wall-Separation) . 212
6.9.5 Effect of Listener Position ........................................ 213
6.9.6 Overall Outcome .................................................. 213
6.9.7 Evaluation of Methodology ....................................... 214
6.10 Chapter Summary .................................................... 214

7 Conclusions and Further Work ........................................ 217
7.1 R1: To what extent can loudspeaker directivity be controlled? .......... 218
7.2 R2: In what ways can loudspeaker directivity and room acoustics affect the sound field at the listener? ........................................... 219
7.3 R3: What changes in loudspeaker directivity and boundary characteristics are perceivable, and what are the relative magnitudes of these changes? . 221
7.4 R4: Which perceptual attributes are affected by the combination of loudspeaker directivity and boundary characteristics? . 224
7.5 R5: What part does listener position play in the relationship between loudspeaker directivity and perception? ................................. 225
7.6 R6: What is the best experimental method for this type of investigation? 226
7.7 Research Contribution .................................................. 227
7.8 Limitations of Research .................................................. 228
List of Figures

2.1 Prediction of the acoustic pressure field surrounding a continuous line source. By considering the contribution from a number of 'simple source elements', which radiate in phase with the same power, the pressure as a function of distance and angle can be predicted at a point in the far-field (where r>>L). This approximation demonstrates the directivity of a simple theoretical acoustic source. *Reproduced from Kinsler et al. (2000)*  

2.2 Variation of the *directional factor* of a radiating line source with off-axis angle and kL. This demonstrates the lobing nature of the pressure field at various angles around the source (when reproducing wavelengths that are small with respect to the source itself). Also, at high values of kL, the narrowing of the pressure field, on-axis, can be seen.  

2.3 Variation of the *directional factor* of a radiating circular piston, mounted in an infinite baffle, with off-axis angle and ka.  

2.4 Frame of reference used for directivity definition. The position P is described by the horizontal angle $\theta$, the vertical angle $\phi$, and the distance $r$. The origin of the coordinate system should be the acoustic centre of the loudspeaker.  

2.5 Horizontal directivity pattern of a 115mm diameter moving-coil loudspeaker driver, mounted in a sealed cabinet. The forward-firing nature of the driver is evident, particularly above 4kHz, where there is no radiation to the rear.
LIST OF FIGURES

2.6 Contour plot of the horizontal directivity of a traditional two-way loudspeaker, with a scale in decibels (referenced to the on-axis level). The narrowing of coverage on-axis, with increased frequency, is evident. 15

2.7 Horizontal directivity pattern of a baffled circular piston. This prediction assumes that the baffle is infinite, and, therefore, the piston only radiates into half-space. Reproduced from (Beranek 1954) 17

2.8 Horizontal directivity pattern of a circular piston mounted at the end of a long tube. The effect of edge diffraction is apparent, with an increase in the forward-firing nature of the system. Reproduced from (Beranek 1954) 18

2.9 Horizontal directivity pattern of unbaffled circular piston radiating in free-space. The radiation characteristic is similar to that of a dipole, with the front and rear radiation out-of-phase with each other. Reproduced from (Beranek 1954) 19

2.10 Horizontal directivity plots for two typical 2-way loudspeakers. Measured by Choueiri (2012) at University of Princeton, USA. Both show a narrowing in on-axis directivity with increased frequency. Photograph in (a) taken from www.genelec.com and photograph in (b) taken from www.ascendacoustics.com 21

2.11 'Acoustic Lens' system used by Bang & Olufsen. Sound is radiated upwards from the driver, and reflected from the lens, into the horizontal plane. This is used to achieve an extended (broadened) off-axis directivity (photograph courtesy of www.beophile.com). 23

2.12 Top: Energy Time Curve of a sound radiated in a room. This shows the measured change in acoustic pressure, over time, at a particular position in a room, following a broadband impulse. Bottom: Waterfall plot of same event. This shows the corresponding change in frequency content of the sound field. Reproduced from (Olive and Toole 1989) 27

3.1 Equal Loudness Contours reproduced using iso226.m (Tackett 2010). These contours show the flattening of the response with increased level (with labels in phons where 1phon = 1dB at 1kHz). 42

3.2 This figure was compiled by Toole (2008) and demonstrates the perceived effects of a change in level and delay of a single lateral reflection. 46
LIST OF FIGURES

3.3 Reformatted data from Hidaka et al. (1997), plotted to show the relationship between reflection angle and perception of apparent source width (ASW) / image broadening (Toole 2008) ................................................. 47

3.4 Mean Fidelity ratings for loudspeakers A-d in positions 1-3 averaged across rooms, subpositions, programs, and subjects. Re-edited from (Bech 1994b) to include loudspeaker directivity types. The 95% confidence intervals per loudspeaker range from 0.2 to 0.25 .................................................. 57

4.1 Experimental Set Up in Surrey (not to scale) ............................................. 66

4.2 Loudspeaker 1 (Surrey) Horizontal Directivity Contour Plot .......................... 68

4.3 Loudspeaker 2 (Surrey) Horizontal Directivity Contour Plot .......................... 69

4.4 Loudspeaker 3 (Surrey) Horizontal Directivity Contour Plot .......................... 70

4.5 On-Axis frequency response of loudspeakers used in Experiment 1(Surrey). Normalised to maximum value. .......................................................... 70

4.6 M and S plots for Audio Programme Material. R is the correlation coefficient between left and right channels .................................................. 71

4.7 Specific terms used to describe differences between reproduction systems at Surrey (terms relate to order of reproduction systems - for example, Genelec 1032A has Less LF than Beolab 3) .............................................. 73

4.8 General terms used to describe differences between reproduction systems in Surrey .......................................................................................... 74

4.9 Intra-Listener RMS Difference (the ability of each individual listener to reproduce consistent scores in repeat trials) for the experiment in Surrey. In this case, Listener 2 shows greater than 15% root-mean-square difference for more than two attributes (circled), and as a result will have their data removed from analysis .................................................. 79

4.10 Inter-Listener Correlation - This example is for 'Liking' scores. Each plot shows the listener score (y-axis), against the average of the remaining listener scores (x-axis). The correlation coefficient for each listener is included .................................................. 80
LIST OF FIGURES

4.11 Inter-Listener Correlation Coefficients - This example shows a corresponding bar chart of the correlation coefficients in Figure 4.10 for 'Liking' scores. 81

4.12 Liking(Surrey): Loudspeakers 85

4.13 Loudness(Surrey): Loudspeakers 85

4.14 Timbral Fidelity(Surrey): Loudspeakers vs Recording 86

4.15 Ensemble Width(Surrey): Loudspeakers vs Recording 87

4.16 Experimental Set Up in Denmark (not to scale) 90

4.17 Loudspeaker 1 (Denmark) Horizontal Directivity Contour Plot 92

4.18 Loudspeaker 2 (Denmark) Horizontal Directivity Contour Plot 93

4.19 Loudspeaker 3 - Left Channel (Denmark) Horizontal Directivity Contour Plot 94

4.20 On-Axis frequency response of loudspeakers used in Experiment 2(Denmark). Normalised to maximum value. 95

4.21 Specific terms used to describe differences between reproduction systems at Denmark (terms relate to order of reproduction systems - for example, Genelec 1031A is narrower than Beolab 3) 95

4.22 General terms used to describe differences between reproduction systems in Denmark 96

4.23 Intra-Listener RMS Difference (the ability of each individual listener to reproduce consistent scores in repeat trials) for the experiment in Denmark (averaged over two positions). In this case, no listener shows greater than 15% root-mean-square difference for more than two attributes, and as a result no data is removed from analysis. 98

4.24 Inter-Listener Correlation - Liking (Denmark) 100

4.25 Inter-Listener Correlation - Loudness (Denmark) 101

4.26 Inter-Listener Correlation - Timbral Fidelity (Denmark) 102

4.27 Inter-Listener Correlation - Ensemble Width (Denmark) 103

4.28 Liking(Denmark): Loudspeakers vs Recording 104
# List of Figures

4.29 Loudness (Denmark): Recording .................................................. 105

4.30 Loudness (Denmark): Loudspeaker vs Position .......................... 106

4.31 Timbral Fidelity (Denmark): Loudspeaker vs Recording .......... 106

4.32 Ensemble Width (Denmark): Loudspeaker vs Recording .......... 107

4.33 Liking (Denmark): Listener ...................................................... 108

4.34 Results from informal test where listeners were able to evaluate loudspeakers at any point within a room ........................................ 111

5.1 Thresholds measured in Olive and Toole (1989). Listeners were able to change the relative level of an off-axis loudspeaker reproducing a delayed version of a signal reproduced by a loudspeaker on-axis. ........ 122

5.2 A screenshot taken from CATT Acoustic, indicating the geometry of the simulation created. ......................................................... 123

5.3 In order to maintain a fixed angle of reflection, the boundary had to be rotated as it was moved further from the receiver. ................. 123

5.4 A balloon plot of the 'omnidirectional' directivity pattern used. The scale indicates sound pressure level in decibels. ......................... 124

5.5 A balloon plot of the 'extended omidirectional' directivity pattern used in order to produce reflections that have the same amount of energy as the direct sound (the increase in off-axis sound pressure is indicated by the light red/pink regions at the sides). ................................. 124

5.6 All reflections simulated using the acoustic modelling software are shown here compared to the original thresholds determined by Olive and Toole (1989) ................................................................. 125

5.7 Flow diagram of the binaural head-tracking system, indicating the processing path from creation of BRIRs in CATT Acoustic to reproduction over headphones via the real-time convolution. ................. 126

5.8 The familiarisation interface used in both listening tests and elicitation. Immediate comparison of any of the test stimuli is possible .......... 128

5.9 MATLAB interface used for the initial listening tests. .................... 129
5.10 Results from initial test. Each pie chart shows, for a particular simulation (delay time and level) the proportions of listeners choosing each of the 3 possible responses. ......................................................... 131

5.11 MATLAB listening test interface with refined descriptors. .................. 134

5.12 Results from main test with improved descriptors. Each pie chart shows, for a particular simulation (delay time and level) the proportions of listeners choosing each of the 3 possible responses. ......................... 135

6.1 Variation in parameters for main experiment ................................. 143

6.2 3D representation of the two-boundary test condition (S1) ................. 156

6.3 Main experiment test system diagram .............................................. 158

6.4 A listener with the test system in the listening room at the University of Surrey ................................................................. 160

6.5 Headphones and laptop showing the TCL-LabX software .................. 161

6.6 Boxplots showing the effect of different experimental parameters upon the relationship between directivity and the perceived dimension for the classical excerpt reproduced in condition S1 (two-boundary): (a) absorption, (b) listener position and (c) boundary separation. .......... 165

6.7 Number of occurrence of top antonyms and next highest term in the Width group (two-boundary) ......................................................... 169

6.8 Number of occurrence of top antonyms and next highest term in the Artifacts group (two-boundary) ............................................................. 170

6.9 Number of occurrence of top antonyms and next highest term in the Proximity group (two-boundary) ......................................................... 170

6.10 Number of occurrence of top antonyms and next highest term in the Loudness group (two-boundary) ......................................................... 171

6.11 Number of occurrence of top antonyms (spacious vs boxy/small) and next highest term in the Spaciousness group (two-boundary) .............. 171

6.12 Number of occurrence of top antonyms and next highest term in the Reverberation group ................................................................. 172
6.13 $R^2$ Correlation between the perceptual dimension and brightness for all excerpts in the two-boundary condition (where CS1 is the classical excerpt, PS1 the pop excerpt and SS1 the speech excerpt) ........................................... 177

6.14 Boxplots showing the effect of different experimental parameters upon the relationship between directivity and the perceptual dimension for the classical excerpt reproduced in condition S1 (two-boundary): (b) absorption, (c) listener position and (d) boundary separation. Subfigure (a) shows the $R^2$ correlation of various attributes with the perceptual dimension, the label indicating the direction of a positive change in that dimension (i.e. increase in dimension is well correlated to the stimuli being perceived as more narrow) .......................................................... 179

6.15 Boxplots showing the effect of different experimental parameters upon the relationship between directivity and the perceptual dimension for the pop excerpt reproduced in condition S1 (two-boundary): (b) absorption, (c) listener position and (d) boundary separation. Subfigure (a) shows the correlation of various attributes with the perceptual dimension, the label indicating the direction of a positive change in that dimension (i.e. increase in dimension is well correlated to the stimuli being perceived as wider) .......................................................... 180

6.16 Boxplots showing the effect of different experimental parameters upon the relationship between directivity and the perceptual dimension for the speech excerpt reproduced in condition S1 (two-boundary): (b) absorption, (c) listener position and (d) boundary separation. Subfigure (a) shows the correlation of various attributes with the perceptual dimension, the label indicating the direction of a positive change in that dimension (i.e. increase in dimension is well correlated to the stimuli being perceived as more dry) .................................................. 181
6.17 Boxplots showing the effect of different experimental parameters upon the relationship between directivity and the perceptual dimension for the classical excerpt reproduced in condition S2 (four-boundary): (b) absorption, (c) listener position and (d) boundary separation. Subfigure (a) shows the correlation of various attributes with the perceptual dimension, the label indicating the direction of a positive change in that dimension (i.e. increase in dimension is well correlated to the stimuli being perceived as wider). .................................................. 189

6.18 Boxplots showing the effect of different experimental parameters upon the relationship between directivity and the perceptual dimension for the pop excerpt reproduced in condition S2 (four-boundary): (b) absorption, (c) listener position and (d) boundary separation. Subfigure (a) shows the correlation of various attributes with the perceptual dimension, the label indicating the direction of a positive change in that dimension (i.e. increase in dimension is well correlated to the stimuli being perceived as wider). .................................................. 190

6.19 Boxplots showing the effect of different experimental parameters upon the relationship between directivity and the perceptual dimension for the speech excerpt reproduced in condition S2 (four-boundary): (b) absorption, (c) listener position and (d) boundary separation. Subfigure (a) shows the correlation of various attributes with the perceptual dimension, the label indicating the direction of a positive change in that dimension (i.e. increase in dimension is well correlated to the stimuli being perceived as wider). .................................................. 191

6.20 Boxplots showing the effect of different experimental parameters upon the relationship between directivity and the perceptual dimension for the classical excerpt reproduced in condition S3 (six-boundary): (b) absorption, (c) listener position and (d) boundary separation. Subfigure (a) shows the correlation of various attributes with the perceptual dimension, the label indicating the direction of a positive change in that dimension (i.e. increase in dimension is well correlated to the stimuli being perceived as more narrow). .................................................. 199
### LIST OF FIGURES

6.21 Boxplots showing the effect of different experimental parameters upon the relationship between directivity and the perceptual dimension for the pop excerpt reproduced in condition S3 (six-boundary): (b) absorption, (c) listener position and (d) boundary separation. Subfigure (a) shows the correlation of various attributes with the perceptual dimension, the label indicating the direction of a positive change in that dimension (i.e. increase in dimension is well correlated to the stimuli being perceived as more dry). .................................................... 200

6.22 Boxplots showing the effect of different experimental parameters upon the relationship between directivity and the perceptual dimension for the speech excerpt reproduced in condition S3 (six-boundary): (b) absorption, (c) listener position and (d) boundary separation. Subfigure (a) shows the correlation of various attributes with the perceptual dimension, the label indicating the direction of a positive change in that dimension (i.e. increase in dimension is well correlated to the stimuli being perceived as more reverberant). ........................................ 201

6.23 The correlation of each parameter with the measured perceptual dimension for all conditions (where C1=classical two-boundary, P1=pop two-boundary etc). ......................................................... 209

6.24 An infographic showing a summary of the main outcome of this research (by Alex Hornsby). ................................................................. 210

A.1 Interface of Max/MSP patch used for elicitation tests ..................... 230
A.2 Assessment form for Elicitation Testing ........................................... 231
A.3 Max/MSP GUI for Liking Test ....................................................... 232
A.4 Max/MSP GUI for Timbral Fidelity Test ......................................... 232
A.5 Max/MSP GUI for Ensemble Width Test ....................................... 233
A.6 Max/MSP GUI for Loudness Test ................................................. 233
A.7 Listener instructions (Page 1) ....................................................... 234
A.8 Listener instructions (Page 2) ....................................................... 235
LIST OF FIGURES

B.1 Loudspeaker 1 (Surrey) Vertical Directivity Contour Plot ............... 239
B.2 Loudspeaker 2 (Surrey) Vertical Directivity Contour Plot ............... 240
B.3 Loudspeaker 3 (Surrey) Vertical Directivity Contour Plot ............... 241
B.4 Liking(Surrey): Loudspeakers vs Listeners ................................. 245
B.5 Loudness(Surrey): Loudspeakers vs Listeners ............................ 246
B.6 Loudness(Surrey): Listeners vs Recording .................................. 246
B.7 Timbral Fidelity(Surrey): Listeners vs Recording ........................ 247
B.8 Timbral Fidelity(Surrey): Loudspeakers vs Listeners .................... 247
B.9 Timbral Fidelity(Surrey): Loudspeakers vs Recording .................... 248
B.10 Ensemble Width(Surrey): Loudspeakers vs Listeners ................... 248
B.11 Ensemble Width(Surrey): Loudspeakers vs Recording .................... 249

C.1 Loudspeaker 3 - Right Channel (Denmark) Horizontal Directivity Contour Plot .................................................. 251
C.2 Loudspeaker 1 (Denmark) Vertical Directivity Contour Plot .......... 252
C.3 Loudspeaker 2 (Denmark) Vertical Directivity Contour Plot .......... 253
C.4 Loudspeaker 3 - Left Channel (Denmark) Vertical Directivity Contour Plot ....................................................... 254
C.5 Loudspeaker 3 - Right Channel (Denmark) Vertical Directivity Contour Plot ....................................................... 255
C.6 Liking(Denmark): Loudspeakers vs Listeners ............................ 260
C.7 Liking(Denmark): Loudspeakers vs Recording ........................... 261
C.8 Loudness(Denmark): Loudspeakers vs Listener .......................... 261
C.9 Loudness(Denmark): Loudspeakers vs Position .......................... 262
C.10 Loudness(Denmark): Loudspeakers vs Recording ....................... 262
C.11 Timbral Fidelity(Denmark): Loudspeakers vs Listener ................ 263
C.12 Timbral Fidelity(Denmark): Loudspeakers vs Recording ............... 263
LIST OF FIGURES

C.13 Ensemble Width (Denmark): Loudspeakers vs Listeners ................. 264
C.14 Questionnaire given to listeners in ‘Roaming’ preference test ............ 266

D.1 Main Experiment Listener Instruction Sheet - Page 1 ......................... 269
D.2 Main Experiment Listener Instruction Sheet - Page 2 ......................... 270
D.3 Number of occurrence of top antonyms and next highest term in the Width group (four-boundary) ......................................................... 271
D.4 Number of occurrence of top antonyms and next highest term in the Artifacts group (four-boundary) ......................................................... 271
D.5 Number of occurrence of top antonyms and next highest term in the Proximity group (four-boundary) ......................................................... 272
D.6 Number of occurrence of top antonyms and next highest term in the Loudness group (four-boundary) ......................................................... 272
D.7 Number of occurrence of top antonyms (spacious vs boxy/small) and next highest term in the Spaciousness group (four-boundary) ............ 273
D.8 Number of occurrence of top antonyms and next highest term in the Reverberation group (four-boundary) .................................................... 273
D.9 Number of occurrence of top antonyms and next highest term in the Width group (six-boundary) .............................................................. 274
D.10 Number of occurrence of top antonyms and next highest term in the Artifacts group (six-boundary) ............................................................ 274
D.11 Number of occurrence of top antonyms and next highest term in the Proximity group (six-boundary) .......................................................... 275
D.12 Number of occurrence of top antonyms and next highest term in the Loudness group (six-boundary) ............................................................ 275
D.13 Number of occurrence of top antonyms (spacious vs boxy/small) and next highest term in the Spaciousness group (six-boundary) ............ 276
D.14 Number of occurrence of top antonyms and next highest term in the Reverberation group (six-boundary) .................................................... 276
List of Tables

3.1 A summary of previous research regarding loudspeaker directivity and its effect on perception of reproduction. ........................................ 58

4.1 Categorised terms from elicitation experiment .................................. 73

4.2 Experimental Factors (Surrey) ................................................................. 77

4.3 Significant Factors (Surrey) - this table shows the Degrees of Freedom, $F$-Ratio, Significance, Eta-Squared and Observed Power values for all significant factors and interactions in the preliminary study at Surrey. . . 83

4.4 Further Effects (Surrey) - this table shows the significance of interactions between factor types following post-hoc Bonferroni tests. ............... 84

4.5 Experimental Factors (Denmark) ............................................................... 97

4.6 Significant Factors (Denmark) - this table shows the Degrees of Freedom, $F$-Ratio, Significance, Eta-Squared and Observed Power values for all significant factors and interactions in the preliminary study at Denmark. . 102

4.7 Further Effects (Denmark) - this table shows the significance of interactions between factor types following post-hoc Bonferroni tests. .............. 103

6.1 The three simulated directivity types used for the Main Experiments. The scale is in dB. ................................................................. 141

6.2 Attribute-based sensory evaluation methods ........................................... 146

6.3 Holistic-based sensory evaluation methods ............................................. 147
6.4 Number of resulting dimensions based on the RSQ, and the associated RSQ value in brackets, for each of the MDS analysis types, for the two-boundary experiment. ......................................................... 163

6.5 Groups defined by panel members in verbal analysis session. .................. 167

6.6 Total frequency of words in groups/scales defined by panel members in verbal analysis session (two-boundary). ................................................................. 168

6.7 High-frequency antonyms and the relative percentage of each term within its group (two-boundary) ................................................................. 169

6.8 "Intensity" Qualifiers - Mean Value on a 0-10 scale from Rohrmann (2007) 173

6.9 Frequency of top scaling or 'intensity' words and their perceived position on a 0-10 scale according to Rohrmann (2007) .................. 173

6.10 Weighting scale used to quantify relationship of various antonyms with each stimulus ................................................................. 174

6.11 $R^2$ Correlation between attribute values and the single dimension output from CMDS of the aggregated dissimilarity matrices for the two-boundary condition ................................................................. 175

6.12 $R^2$ Correlation between attribute values (frequency of use in the case of 'Frequency') and the single dimension output from CMDS of the aggregated dissimilarity matrices for the two-boundary experiment. 176

6.13 $R^2$ Correlation between bass increase/decrease and the single dimension output from CMDS of the aggregated dissimilarity matrices for the two-boundary condition. ......................................................... 178

6.14 Significant Factors - Main Experiment (two-boundary) .................. 182

6.15 Further Effects - Main Experiment (two-boundary) .................. 182

6.16 Number of resulting dimensions based on the RSQ, and the associated RSQ value in brackets, for each of the MDS analysis types, for the four-boundary experiment. ......................................................... 185

6.17 Total frequency of words in groups/scales defined by panel members in verbal analysis session (four-boundary). .................. 186
LIST OF TABLES

6.18 High-frequency antonyms and the relative percentage of each term within its group (four-boundary) ................................................................. 187

6.19 $R^2$ Correlation between attribute values (frequency of use in the case of 'Frequency') and the single dimension output from CMDS of the aggregated dissimilarity matrices for the four-boundary experiment. ... 187

6.20 $R^2$ Correlation between bass increase/decrease and the single dimension output from CMDS of the aggregated dissimilarity matrices for the four-boundary condition ................................................. 187

6.21 $R^2$ Correlation between attribute values and the single dimension output from CMDS of the aggregated dissimilarity matrices for the four-boundary condition ................................................. 187

6.22 Significant Factors - Main Experiment (four-boundary) ..................... 192

6.23 Further Effects - Main Experiment (four-boundary) ........................... 192

6.24 Number of resulting dimensions based on the RSQ, and the associated RSQ value in brackets, for each of the MDS analysis types, for the six-boundary experiment. ............................................. 195

6.25 Total frequency of words in groups/scales defined by panel members in verbal analysis session (six-boundary). ................................. 197

6.26 High-frequency antonyms and the relative percentage of each term within its group (six-boundary) ................................................................. 197

6.27 $R^2$ Correlation between attribute values (frequency of use in the case of 'Frequency') and the single dimension output from CMDS of the aggregated dissimilarity matrices for the six-boundary experiment. ... 198

6.28 $R^2$ Correlation between bass increase/decrease and the single dimension output from CMDS of the aggregated dissimilarity matrices for the six-boundary condition ............................................. 198

6.29 $R^2$ Correlation between attribute values and the single dimension output from CMDS of the aggregated dissimilarity matrices for six-boundary ................................................................. 198

6.30 Significant Factors - Main Experiment (six-boundary) ..................... 202

6.31 Further Effects - Main Experiment (six-boundary) ........................... 202
6.32 Attributes that are highest correlated with the single perceptual dimension \(^2\)06

6.33 Associated change in perception of sound reproduction according to change in dimension in same direction as a change in perception caused by reduction in modelled loudspeaker directivity. Changes are presented in order of magnitude. ........................................ \(\)07

6.34 Levels of factors that vary according to a scale ........................................ \(\)08

D.1 Normality Tests for Main Experiment Data ................................................. \(\)68
Acknowledgements

I would like to thank:

- Dr. Tim Brookes and Dr. Russell Mason for their dedicated supervision and guidance;
- Dr. Francis Rumsey and Dr. Slawek Zielinski for introducing me to the IoSR and the research project;
- The IoSR graduate researchers for their continued support - in particular, Chris Hummersone and Laurent Simon for their ears, critical minds and 'unique' sense of humour;
- My friends and family for being there for me on days when I'm not thinking about loudspeakers.
Statement 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.

Some work presented in Chapters 2 and 3 has been published previously in Evans et al. (2009)
Chapter 1

Introduction

The semi-reverberant nature of most domestic listening spaces means that a significant amount of the sound radiated from a loudspeaker is reflected from the walls to the listener. It is the directivity of the loudspeaker that defines the direction of its radiation and, therefore, the amount of sound that will come into contact with particular boundaries in a listening space. The combination of the loudspeaker directivity and the characteristics of the boundaries (their composition, distance from source/listener and angle) will affect the nature of the reflections that arrive at the listener. The nature of these reflections influences the way a listener perceives the sound reproduction within that environment.

Loudspeaker directivity, itself, is the subject of much debate. Whilst some argue that a loudspeaker with a wide horizontal coverage contributes beneficially to the sound field in a listening environment, others feel it is detrimental. For example, when discussing the ideal directivity properties of a conventional loudspeaker, Colloms (2005) writes, ‘outside of this [axial 60°] angle, a rapid reduction in output would be considered a positive advantage, since some reflections from the adjacent walls would be diminished with a consequent reduction in reverberation energy and an improvement in stereo image stability’. On the contrary, Toole (2008) writes, “the results [of several authors] discussed here all point in the same direction: that wide-dispersion loudspeakers, used in rooms that allow early reflections, are preferred by listeners especially, but not exclusively, for recreational listening. There appear to be no notable sacrifices in the ‘imaging’ qualities of stereo reproduction”. This argument has existed for a long period of time and remains unsettled. Often, however, those involved in the argument are concerned
CHAPTER 1. INTRODUCTION

with reproduction situations that differ in terms of several variables, including boundary characteristics and relative source/listener positions. An investigation is, therefore, required in order to establish more specifically, the relationship between loudspeaker directivity, boundary characteristics and the perception of reproduction in domestic rooms for a range of typical listening situations. Results from such an investigation can then be used to inform, more definitively, the predicted influence of a directivity type upon the perception of the sound being reproduced within a particular environment.

Following from the text above, the principal aim of the investigation can be defined as:

**Principal question** How is loudspeaker directivity likely to affect the perception of reproduced sound in domestic listening rooms?

The principal question can be decomposed into several more specific 'initial research questions'. To begin with, it is necessary to understand the limitations of loudspeaker directivity control in domestic listening rooms; it is only of interest to investigate the effect of changes in loudspeaker directivity that can be achieved practically (see R1).

**R1** To what extent can loudspeaker directivity be controlled?

Having already highlighted that the sound field at the listener position is dependent upon both loudspeaker directivity and the room characteristics, the nature of physical changes expected at the sound field in typical rooms (for a practical region of directivity variation) should be investigated. This leads to R2:

**R2** In what ways can loudspeaker directivity and room acoustics affect the sound field at the listener?

Once the relationship between loudspeaker directivity and boundary characteristics has
been explored, and the effect it has upon the sound field established, it is necessary to consider what perceptual changes might result from alterations to this relationship (see R3).

**R3** What changes in loudspeaker directivity and boundary characteristics are perceivable, and what are the relative magnitudes of these changes?

It is then of interest to find which perceptual attributes, in particular, are affected, and by how much (see R4).

**R4** Which perceptual attributes are affected by changes to loudspeaker directivity and boundary characteristics?

It is realistic to assume that listening within a domestic room occurs at multiple positions, and so a consideration of this variable and its involvement with the directivity/perception relationship is also necessary:

**R5** What part does listener position play in the relationship between loudspeaker directivity, boundary characteristics and perception?

As the principal question here is concerned with measurements that are both physical and psychological in nature, there are likely to be numerous experimental options. Thus, it is also of interest to determine which experimental method is most suitable for this type of study, and why. An additional research question R6 is therefore also considered:
CHAPTER 1. INTRODUCTION

R6 What is the best experimental method for this type of investigation?

This thesis addresses these questions by considering relevant literature and conducting a series of experiments. Chapter 2 introduces the concept of loudspeaker directivity and its influence upon the sound field in a room (answering R1 and R2). Chapter 3 explains how changes in the sound field in a room affect the perception of the reproduced sound (providing initial answers to R3, R4 and R5 based on theory/literature). Chapter 4 includes details of a preliminary test which was designed in order to explore the magnitude of effects caused by differences in the directivity of commercial loudspeakers (providing additional information relevant to R3, R4 and R5, and initial information regarding R6). This test was also used to confirm whether the perceptual attributes thought to be affected by changes in directivity (according to literature) were affected in a real listening set-up or not. Results from this chapter suggest that the perceived changes are likely to be small, and so a test method capable of reproducing and measuring small effects was noted as being necessary (contributing further towards R6). In light of these results, and limited definitive conclusions from the experiments documented in the literature, Chapter 5 considers how to explore this topic further and more effectively answer R3, R4 and R5 - a method for testing the perceived differences between various combinations of loudspeaker directivity and boundary characteristics using a novel auralisation system is proposed, and an experiment performed in order to validate the system. Results from the experiment show that the system is satisfactory for use in further testing, and Chapter 6 describes how this system is used in a series of three further experiments to provide fuller answers to R3, R4 and R5. The final method used is then considered with regards to R6.
Chapter 2

Loudspeaker Directivity and Reproduced Sound in Domestic Listening Rooms

The purpose of this chapter is to:

- introduce loudspeaker directivity;
- identify the range over which it may be controlled;
- establish the relationship between loudspeaker directivity and the listening room characteristics with respect to the reproduced sound field.

By considering these items, an understanding of the differences in the sound arriving at the listener as a result of changing the loudspeaker directivity can be established. If the extent of these physical changes are known, information from perceptual studies can be used to predict the resulting audibility or perceptual magnitude of the differences. Thus, without even carrying out experiments, something about the relationship between directivity and perception of reproduced sound can be defined at this stage.

The chapter begins by introducing the concepts of simple source radiation and directivity. The directivity of loudspeakers is then discussed, followed by an overview of the current scope for loudspeaker directivity control. Finally, the interaction between loudspeakers and rooms, and the subsequent effect on the sound field (as a result of changes in directivity and boundary characteristics), is considered.
CHAPTER QUESTIONS

- What is loudspeaker directivity?
- What factors affect the directivity of a loudspeaker?
- To what extent can loudspeaker directivity be controlled?
- How does loudspeaker directivity influence the reproduced sound field in a room?

2.1 Loudspeaker Directivity

Directivity is a parameter that describes an aspect of the way in which a loudspeaker radiates sound. The following section provides an introduction to loudspeaker directivity and considers the nature of directivity associated with conventional loudspeakers. A further exploration of theory and recent technological developments also allows the scope for directivity control to be assessed, so that by the end of the section, the limitations of directivity control in home audio systems can be established.

2.1.1 Acoustic Radiation

To begin with, consider a loudspeaker comprising a moving-coil driver unit, mounted in a small cabinet. As a first approximation, such a loudspeaker may be modelled as a 'simple source'. A simple source is the most basic theoretical acoustic source, defined as a uniform sphere that is pulsating in free space. The pressure field that a simple source radiates is uniformly spherical and independent of frequency (Kinsler et al. 2000).

According to the concept of acoustic reciprocity (Kinsler et al. 2000), any source that is non-spherical will not radiate as an equivalent simple source when the wavelength it radiates is less than its dimensions. Improving the approximation to take account of this concept, it can be assumed that, at wavelengths smaller than the dimensions of the loudspeaker, the radiation will be non-uniform. This assumption has immediate repercussions upon the topic of the investigation. It suggests that at low frequencies, where wavelengths are large in comparison to the loudspeaker, it can be assumed to radiate as an omnidirectional 'simple source'. Conversely, at high frequencies, where wavelengths become comparable to the dimensions of the loudspeaker, the radiation becomes directional.
In order to predict the pressure field of a non-spherical source, at wavelengths shorter than the source dimensions, the effect can be modelled by replacing the simple source with an arrangement of smaller simple sources. The contribution of each smaller source to the total pressure field may then be calculated at each point in the pressure field (Kinsler et al. 2000).

As an example, consider an acoustic source in the form of a continuous line, with length, $L$. It can be modelled by dividing it into several elements along its length, each representing a simple source. The pressure as a function of radial distance, $r$ and angle $\theta$, can be calculated, for all frequencies, by integrating the contribution of each element over the length of the source (see Figure 2.1).

If it is assumed that the radial distance, $r$, is much greater than the length of the source, $L$ (referred to as being ‘in far field’), then the acoustic pressure amplitude is given by
Kinsler et al. (2000) as:

\[ P(r, \theta) = P_{ax}(r)H(\theta) \]  

(2.1)

where

\[ H(\theta) = \left| \frac{\sin \nu}{\nu} \right| \quad \nu = \frac{1}{2}kL\sin \theta \]  

(2.2)

is the directional factor and

\[ P_{ax}(r) = \frac{1}{2}\rho_0cU_0 \left( \frac{a}{r} \right) kL \]  

(2.3)

is the amplitude of the far field axial pressure.

(a=radius of the cylindrical line source, \( \rho_0c \)=specific acoustic impedance, \( U_0 \)=surface velocity, \( k \)=acoustic wavenumber)

This equation shows that in this case, the pressure field is affected by two key products: \( P_{ax} \) denotes the amplitude of the far field axial pressure and depends only upon the distance from the source, and \( H \) is the directional factor, which depends only on angle. The \( \sin \nu/\nu \) relation of the directional factor is known as the sinc function or zeroth order spherical Bessel function of the first kind and when plotted (Figure 2.2) demonstrates the lobing nature of the pressure field around the source at smaller wavelengths (relative to source size). The function shows that the larger the value of \( kL \) (i.e. the larger the source length compared to the wavelength), the more narrowly directed the major lobe and the greater number of minor lobes in pressure amplitude (Kinsler et al. 2000).

Therefore, further to the initial discovery regarding non-spherical sources at low frequencies, this finding suggests that as the wavelength becomes smaller than the dimensions of the loudspeaker, a narrowing in radiation occurs. This narrowing is focussed towards the on-axis point, and has been shown, in the case above, to increase with frequency.

In order to improve the loudspeaker approximation, a motional rigid circular piston could be modelled. Kinsler et al. (2000) consider such a piston, which has a radius \( a \),
and is 'mounted' in an infinite baffle; this prevents the sound radiated from the front of the piston interfering with sound radiated from the rear of the piston. As with the line source, the pressure field can also be evaluated by dividing the radiating surface into infinitesimal elements, each with the same source strength. And again, assuming a point in the far field, integration across the surface allows for the pressure field as a function of angle and radial distance to be calculated.

For the case of the baffled piston, the acoustic pressure amplitude in the far field is given by Kinsler et al. (2000) as:

\[ |p(r, \theta)| = P_{ax}(r)H(\theta) \quad (2.4) \]

\[ H(\theta) = \left| \frac{2J_1(\nu)}{\nu} \right| \quad \nu = ka \sin \theta \quad (2.5) \]
CHAPTER 2. LOUDSPEAKER DIRECTIVITY AND REPRODUCED SOUND

Variation of Directional Factor with Off-Axis Angle and ka

Figure 2.3: Variation of the directional factor of a radiating circular piston, mounted in an infinite baffle, with off-axis angle and ka.

where \( J_1 \) is a first order Bessel function of the first kind.

As with the line source, this result implies that, for wavelengths much smaller than the piston radius, the pressure field narrows to form a main lobe on-axis, with smaller lobes off-axis (see Figure 2.3). For \( ka \ll 1 \) (wavelengths much larger than the piston radius), the directional factor is nearly unity for all angles, and the piston, therefore, acts as a baffled simple source, radiating equally in all directions (of a hemisphere).

The modelling presented here demonstrates that there are inherent properties of source radiation that are defined by the relationship between the source and the wavelength of sound it is radiating. The modelled outcomes can be directly applied to real loudspeaker sources and so, even with the simple approximations above, the key behavioural characteristics of a loudspeaker at low and high frequencies can be predicted. The findings show that, for any single-driver loudspeaker that radiates in a manner similar to a simple source, the measured pressure field would be expected to be largely omnidirectional at the loudspeaker’s lowest frequency range and then more directional towards on-axis at higher frequencies. This outcome serves as an important introduction to the topic of loudspeaker directivity as it highlights that, fundamentally, the direction
in which a loudspeaker radiates is dependent upon its size and the wavelength of sound that it is reproducing.

In summary:

- a single-driver loudspeaker can be modelled by a simple source at low frequencies - it has an omnidirectional radiation pattern;
- at higher frequencies, where the wavelength becomes comparable with the loudspeaker dimensions, the radiation begins to narrow on-axis and forms off-axis side-lobes; and
- the frequency at which narrowing begins to occur is dependent upon the ratio between the driver radius and the wavelength, \( ka \).

2.1.2 Defining Directivity

It has been shown that not all conventional loudspeakers will radiate sound in the same manner for a given frequency range; it depends upon their size relative to the wavelength they are reproducing. The classification of a loudspeaker's radiation properties will now be considered and the principle of Directivity defined.

Previously, it was shown that traditional single-driver loudspeakers are expected to have a radiation that narrows with increased frequency (relative to the operational bandwidth of the driver-unit). If the radiation of different loudspeakers is not the same, and if the directional nature of their own radiation changes with frequency, then a method for quantifying the radiation properties of a loudspeaker is important.

A common method of classification is called the Directivity Factor, \( Q \), or Directivity Index, \( DI \) (\( Q \) measured in decibels). This is the ratio between the on-axis pressure, and the pressure amplitude for a simple source that generates the same acoustic power (equivalent to the power response). It is given by;

\[
Q = \frac{P_{ax}}{P_0^2}
\]  
(2.6)

\[
DI = 10 \log Q,
\]  
(2.7)
where \( P_{ax} \) is the axial pressure and \( P_s \) is the pressure of a simple source with the same acoustic power.

This measure provides a single value that describes how focused the source is in one direction. This is particularly useful as a general classification for directional sources, which all tend to narrow towards on-axis, but it cannot be used to characterise the off-axis radiation of a source, as two sources with a similar Directivity Index may exhibit very different off-axis radiation.

If the ratio of sound pressure off-axis, to that on-axis, is measured as a function of angle, frequency and distance, then a more complete overview of the source radiation characteristics, in all directions, can be defined. It is this measure that is referred to as Directivity. (In some texts, for example: Kinsler et al. (2000), Directivity is used to mean Directivity Factor; although, the definition provided here is generally accepted by most authors, e.g. Colloms (1985))

Definition

For clarity, a formal definition of directivity follows. This definition is intended to apply to the proceeding work.

Directivity, \( D \), is defined as the complex relationship between the position dependent frequency response and the frequency response of a reference position (see Figure 2.4).

\[
D(j\omega, P(\theta, \phi, r)) = \frac{H(j\omega, P(\theta, \phi, r))}{H(j\omega, P_{ref})} \tag{2.8}
\]

where \( \theta = \) azimuth angle, \( \phi = \) elevation angle.

This equation implies that, at the reference position, directivity, \( D(j\omega, P_{ref}) = 1 \) (\( P_{ref} \), is often chosen as the on-axis position of the loudspeaker). As stated previously, in a far-field condition, the distance \( r \) only affects the level of the sound field. For example, if \( r \) is doubled, the sound pressure level is reduced by 6 dB (in an anechoic environment) (Borwick et al. 1988). Therefore, the horizontal and vertical directivity of a loudspeaker in anechoic far-field is independent of distance.

Consequently, directivity becomes a multivariate complex function of three variables: azimuth, elevation, and frequency. It is difficult to evaluate the entire function graphically using a single plot. It has, therefore, become customary for authors to
Figure 2.4: Frame of reference used for directivity definition. The position $P$ is described by the horizontal angle $\theta$, the vertical angle $\phi$, and the distance $r$. The origin of the coordinate system should be the acoustic centre of the loudspeaker.

plot axial pressure using a two-dimensional polar plot, referred to as the 'directivity pattern' of a source (Colloms 1985)(see Figure 2.5). This type of polar plot is, however, a limited method of representing directivity data, as only pressure maps from discrete frequencies may be plotted effectively. If pressure maps at many frequencies are included on one polar plot, there is a tendency for the data to become illegible. A more useful graphical representation of radiation pattern is found in contour plots (see Figure 2.6), which indicate pressure as a function of angle and frequency. This allows all radiation characteristics over the entire bandwidth of interest to be observed (Pedersen and Munch (2002); Zacharov (1998)). Contour plots are also restricted to a two dimensional plane of reference, and thus the plane must be defined (i.e. the horizontal or vertical angle through which the plane exists) (Beranek 1954).
By using 'directivity patterns' or contour plots, the directivity of a loudspeaker (whether physical or theoretical) may be defined. This provides us with some information about the way in which the loudspeaker will radiate sound into the environment around it, and allows for comparison between different types. The directivity index is also an important tool, providing a single numerical value to describe the narrowing radiation of a sound source. In cases where it would be ideal to assess the changes in the 'narrowness' of a loudspeaker's radiation towards on-axis, and its relationship with other variables (for example, perceptual attributes), this measure alone could be more convenient for interpretation of analysis, rather than attempting to consider the multivariate nature of directivity. It may, however, obscure information regarding off-axis changes which may be important, and so this should be considered if it were to be used.
Figure 2.6: Contour plot of the horizontal directivity of a traditional two-way loudspeaker, with a scale in decibels (referenced to the on-axis level). The narrowing of coverage on-axis, with increased frequency, is evident.

In summary:

- **Directivity Index** is a single-valued variable that quantifies how focussed a sound source is in one direction;
- **Loudspeaker Directivity** is a complex, multivariate function that describes the radiation of a sound source in all directions;
- **Loudspeaker Directivity** has been defined as the anechoic far-field sound pressure as a function of angle, elevation and frequency, normalised to a reference axis (typically on-axis); and
- directivity can be represented graphically using polar plots (at discrete frequency intervals) or with contour plots (continuous frequency intervals). Both are limited to a single reference plane.
2.1.3 The Directivity of Loudspeakers

Following a brief consideration of acoustic theory, it has been demonstrated that single-driver loudspeakers (which have a driver mounted to face on-axis), will tend to become directional at high frequencies, that is, narrow in frontal radiation coverage. A polar plot for such a loudspeaker is presented in Figure 2.5.

Though most commercially available loudspeakers tend to demonstrate directivity characteristics similar to this, the exact variation in directivity with frequency is dependent upon a number of factors related to the design of the loudspeaker. The importance of driver unit size with regards to the frequency at which narrowing occurs has been discussed, but there are a number of additional factors that also affect directivity.

To begin with, it is necessary to consider the effect of adding additional driver units to a loudspeaker. Driver units tend to have a limited range of operation, and near the frequencies where they begin to narrow in directivity, the amount of sound power they radiate decreases. In order to achieve a greater radiation efficiency over a wider bandwidth, additional driver units can be introduced to 'take-over' higher frequency bands before a driver begins to behave inefficiently. The resulting directivity of such a multiple-driver loudspeaker depends on the directivity characteristics of the drivers used and the frequencies at which the drivers are designed to overlap ('cross-over'). The point at which the loudspeaker begins to narrow in coverage would be largely dependent upon the size of the driver used for the highest frequency bands, assuming the other drivers are made to overlap well before they begin to narrow.

The next item to consider is the effect of the cabinet in which the drivers are mounted. Previously, the theoretical example of a piston mounted within an 'infinite-baffle' was considered and although this model is sufficient to demonstrate the 'beaming' effect (see Figure 2.7), it does not demonstrate the additional effects that typically occur as a result of the baffle not being infinite in length, such as edge diffraction.

In order to estimate the effects of a cabinet, it is more appropriate to consult the model of a rigid circular piston placed at the end of a long circular tube. This leads to the inclusion of edge diffraction in the calculation of directivity pattern (Beranek (1954); Colloms (1985)). The resulting radiation pattern (Figure 2.8) shows that a cabinet contributes to the 'forward-firing' nature of the loudspeaker.

The calculation of this radiation pattern is particularly complex and the full derivation
is presented in a 1948 paper by Levine and Schwinger (1948).

In reality, there are additional effects to consider relating to edge-diffraction, which affect any mounted driver (whether it be in a cabinet, or in a finite-baffle). The radiation of sound at boundary edges is non-uniform, and depends upon the characteristics of the boundary termination. At high frequencies, radiated sound is reflected back from the boundary edge and can interfere with the on-axis radiation. This effect is dependent upon the size of the baffle, the relative driver positioning and the type of boundary (i.e. corner, smooth edge, sphere) (Colloms 1985).

Also, cabinet loudspeakers are subject to panel vibrations, which re-radiate sound 'in-sympathy' with the driver motion. This can lead to an interaction between the sound radiated from the driver and that from the panels, affecting the overall directivity pattern.

Some loudspeaker designs, for example 'dipole' loudspeakers, do not have driver units mounted in a cabinet. Instead, the drivers are mounted in a finite baffle, and are allowed
CHAPTER 2. LOUDSPEAKER DIRECTIVITY AND REPRODUCED SOUND

Figure 2.8: Horizontal directivity pattern of a circular piston mounted at the end of a long tube. The effect of edge diffraction is apparent, with an increase in the forward-firing nature of the system. Reproduced from (Beranek 1954)

![Diagram of a circular piston mounted at the end of a long tube with labeled directivity pattern.]

Beranek (1954) and Colloms (1985) comment that the resulting directivity patterns resemble those of an acoustic doublet (two simple sources back to back radiating 180 degrees out-of-phase). Figure 2.9 shows the patterns of a theoretical circular piston in free space, and the ‘figure-of-eight’ directivity is apparent, with characteristic acoustic nulls perpendicular to the driver. These occur as a result of the cancellation between front and rear radiations. Though the fundamental directivity pattern of dipole loudspeakers is different to that of traditional cabinet loudspeakers, a narrowing in the front and rear radiation would still be expected to occur as a result of the narrowing characteristics of moving-coil driver units.
Finally, the transduction mechanism used also affects the directivity of the loudspeaker. Moving-coil drive units show piston-like behaviour up until a point where 'cone-break-up' occurs: above a certain frequency, the motion of the diaphragm becomes non-uniform. Different areas of the diaphragm radiate out-of-phase with each other and, therefore, the pressure field created is altered from that of a theoretical equivalent. The extent of non-uniformity is typically dependent upon the material, its geometry and its rigidity (Colloms 1985).

Other options for transduction mechanisms exist and tend to demonstrate their own unique directivity characteristics (Newell and Holland 2007):

- **Ribbon driver** - Uses a conductive diaphragm mounted between two magnetic poles. It has a directivity that narrows with increasing frequency, although the directivity can be made wider with a horn;

- **Piezoelectric driver** - A voltage is applied to either side of a piezoelectrical material, which causes vibrational motion. Cylindrical designs have a $360^\circ$
horizontal directivity;

- **DML (Distributed Mode Loudspeaker)** - Resonances produced in a vibrating panel give rise to a highly diffuse source. It displays wide bipolar-like directivity, although the radiation in front of and to the back of the diaphragm are typically uncorrelated.

The specific directivity behaviour of all driver types is dependent on the design and materials used by the manufacturer, and two drivers of a similar type can often display differences in directivity.

According to the theory presented here, most traditional 'domestic' loudspeakers, which consist of two or more moving-coil drive units mounted in a cabinet, will have a directivity that narrows in coverage with increased frequency (see Figure 2.10). It is also evident that the specific directivity characteristics of a loudspeaker can be affected by a large number of parameters, meaning that there are likely to be differences in the directivity of different loudspeakers, even if the general nature of the directivity (i.e. narrowing in a conventional loudspeaker) is the same. If the configuration of drivers and the mounting method is changed from that of a cabinet, the resulting directivity pattern can be very different.
CHAPTER 2. LOUDSPEAKER DIRECTIVITY AND REPRODUCED SOUND

(a) Genelec 8030a Horizontal Directivity

(b) Ascend Acoustics CBM-170 Horizontal Directivity

Figure 2.10: Horizontal directivity plots for two typical 2-way loudspeakers. Measured by Choueiri (2012) at University of Princeton, USA. Both show a narrowing in on-axis directivity with increased frequency. Photograph in (a) taken from www.genelec.com and photograph in (b) taken from www.ascendacoustics.com.
In summary:

- the directivity of a multiple-driver loudspeaker is dependent on the drivers used and the frequency band over which they operate;

- loudspeaker cabinets cause a reduction in rear radiation at higher frequencies, contributing to the ‘forward-firing’ nature of these types of loudspeakers;

- cabinet panels can vibrate ‘in-sympathy’ with the drivers and re-radiate sound that interferes with that radiated from the drivers (thus altering directivity);

- edge diffraction effects can cause interference in radiated sound, and thus the directivity; and

- dipole loudspeakers have a ‘figure-of-eight’ directivity, where the radiation at the front is 180° out-of-phase with the radiation at the back.

### 2.1.4 Loudspeaker Directivity Control

Having established that loudspeaker directivity is dependent upon a number of parameters, it would now be of interest to consider ways in which it can be controlled; variation in design and additional measures can be employed to create loudspeakers which demonstrate directivities that differ considerably from the more traditional loudspeakers and dipoles.

It was during the 1950s that devices specifically designed for controlling directivity in loudspeakers first became prominent - so called ‘acoustic lenses’ comprised of a mechanical arrangements of plates which were used to prevent exponential horn loudspeakers from beaming. The plates were used to delay the sound path so that the wave exiting the horn had a greater curvature. The ‘slanted’ and ‘folded’ plate varieties provided only wide dispersion in the horizontal plane, whereas the ‘perforated’ plate lens was designed to spread high frequencies conically (Eargle 2010).

More recently, several designs have been documented which aim to widen the directivity of conventional loudspeakers systems at frequencies where beaming occurs. Given the inherent omnidirectionality at low frequencies, these designs often consider only mid-range and high frequency drivers, and are essentially extensions of the lens techniques employed earlier. In all cases observed, it appears that drivers are positioned so that they are facing upwards. Mounted directly above, is a device that redirects the radiated
sound into the horizontal plane (see Figure 2.11). Berlant (1985) proposed a 360° horn device, whereas later designs utilise a reflector system (Ferralli and Moulton 1986; Pedersen and Munch 2002). With these designs, radiation in the horizontal plane has been shown to have been widened considerably, with a directivity that is either hemi or omni-directional and uniform across a selected frequency range. Increased control in the vertical directivity has also been achieved with these designs, allowing the uniform directivity to extend over a broader range of vertical angles (Ferralli and Moulton 1986).

Such developments are highly notable, but have only been implemented in a small number of commercially available loudspeakers; control of directivity in the home-audio market is not yet widespread.

Despite this, there exists research in the field of digitally-controlled loudspeaker arrays which clearly demonstrate the extent of directivity control achievable. Several research departments have developed loudspeaker systems which can be controlled in real-time to give any directivity desirable (within the restraints defined by driver type, size, number and spacing). They are led primarily by the aim to re-create real-life source directivity characteristics (such as musical instruments) and their prototypes often comprise a large number of independently controlled drive-units mounted in spherical-based cabinets (Avizienis et al. (2006); Misdariis et al. (2001); Warusfel et al. (1997)). By altering the properties of the respective driver signals (often according to spherical-harmonic theory), a large range of overall directivity patterns can be engaged.
CHAPTER 2. LOUDSPEAKER DIRECTIVITY AND REPRODUCED SOUND

More simplistic arrangements of drivers can be also used in order to establish basic control of low-frequency directivity (such as Boone and Ouweltjes (1997)). However, even though this level of control exists, the increased cost and complexity required to achieve it in home audio means that it is not yet commercially viable (though, some intelligent beam-forming arrays using smaller drivers have begun to appear on the market - they are, however, intended to reproduce multi-channel stereophonic sound via beams of sound reflected from the listener's room rather than allowing the listener to vary the type of directivity pattern). Also, very few listeners have been exposed to the concept of variable directivity, and so it will take time for there to be necessary demand in the home-audio market.

This research is concerned with sound reproduction via traditional hi-fi loudspeakers, intended for home use. With this in mind, it places certain limitations upon the type of systems to be investigated and the range of directivities explored. For example, traditional bookshelf loudspeakers often occupy cabinets of volume $\sim 0.01m^3$ (10 litres), with woofers of $\sim 13cm$ or less. The aforementioned designs with multiple subwoofers or large spherical loudspeaker-arrays are impractical and unlikely to be widely adopted in home audio and so the limits of directivity control from traditional-sized units, using an affordable number of drive-units should be considered.

With restrictions on the size of the cabinet, the size of the woofers and the number of woofers, it is likely that directivity control below a particular frequency will be unachievable - the maximum conceivable woofer diameter for a traditional loudspeaker unit would be 30cm, which would be expected to only begin to exhibit forward-facing directivity at $(ka\approx 2)$ at $\sim 700$Hz and have narrow directivity $(ka\approx 4)$ at $\sim 1400$Hz. If we consider the measured driver directivity in Figure 2.5, which narrows at $ka\approx 0.6$, a 30cm driver in a sealed cabinet might even be expected to narrow at $\sim 200$Hz. For this reason, it would be fair to assume that any traditional-sized loudspeaker will be omnidirectional below at least $\sim 200$Hz, and thus for this limitation to be considered in any future tests.

In summary:

- horizontal (and some vertical) directivity control has been achieved in home audio loudspeakers by using reflective devices mounted in front of drive-units to prevent 'beaming';
- effective real-time directivity control in all planes has been achieved by the use
CHAPTER 2. LOUDSPEAKER DIRECTIVITY AND REPRODUCED SOUND

of multiple-loudspeaker arrays and digital signal processing. Such designs are currently too complex to be adapted to the home audio market; and

- it can be assumed that any loudspeakers intended for use in domestic audio, with limited size and number of drivers, will exhibit omnidirectionality below at least 200Hz.

2.1.5 Section Summary and Discussion

Summary

In this section, the concept of loudspeaker directivity has been introduced. Using acoustic theory, it was shown that the radiation of sound sources varies depending on the size of the source and the wavelength of sound being radiated. The nature of this radiation can be classified using a multidimensional variable called directivity, and this indicates the spatial manner in which sound is radiated for various frequencies. Acoustic theory can be used to predict the general directivity characteristics of loudspeakers; however, it was shown that in reality, several parameters affect the specific directivity characteristics. Finally, the extent of loudspeaker directivity control was considered and evidence suggests that the technology for effective real-time control in all directions does exist, despite its confinement to research laboratories. More simplistic methods of control have been adapted for use in home-audio, but are limited to specialist models. It is unlikely that directional control below 200Hz can be achieved with traditional-sized loudspeaker units and so this must be considered in any future tests involving directivity control.

Discussion

The findings in this section answer the first research question, R1 (establish the extent to which loudspeaker directivity could be controlled). Knowing this, the investigation can now be directed to only consider any perceptual effects caused by directivity changes that are known to be practically achievable.
CHAPTER 2. LOUDSPEAKER DIRECTIVITY AND REPRODUCED SOUND

2.2 Reproduced Sound in Domestic Listening Rooms

Loudspeaker directivity describes how a loudspeaker will radiate sound into the space around it. This research is concerned with how different directivity characteristics affect the perception of reproduced sound in a typical domestic room and so it is now necessary to consider how directivity is related to the resulting sound field at the listener position. The following section will consider, firstly, the interaction between a loudspeaker and a room and its effect on the sound field at the listener and then, secondly, the effects of changing directivity upon the sound field.

2.2.1 Loudspeaker Reproduction in Rooms

In a completely free-field acoustical environment, a listener will only hear direct sound from a loudspeaker. However, in a small room, the sound radiated from the loudspeaker will be reflected within a short time interval such that the listener will hear a combination of both direct sound and reflected sound. This may be displayed using an energy time curve (ETC), which shows the arrival of the reflected energy at a point in the room, at some time after the arrival of the direct energy to that same point (See Figure 2.12).

The nature of the reflected sound component is dependent on three key factors. The first factor is the absorption and scattering characteristic of the surfaces. This dictates the spectral content of the reflected sound (amplitude and phase), and is highly dependent on the material properties. The second factor is the relative position of loudspeaker, listener and surfaces. By altering the distance between the three, the time it takes for the reflected components to arrive at the listener will be changed. The third factor is loudspeaker directivity. The directivity defines the direction in which sound is radiated into the room, and, therefore, dictates the relative contributions of the above factors to the reflected sound component. As established earlier, the directivity of a loudspeaker is also frequency-dependent.

It follows from this, that the reproduced sound in a room can only be predicted if the loudspeaker directivity is known. Once known, the relative effects of room geometry and boundary characteristics can be taken into account. The reproduced sound field in a small room for a given directivity will now be investigated, before considering how changing the directivity might affect it. The relationship between directivity and the other factors that affect reflected sound in rooms will also be discussed.
It is convenient to begin by considering the sound field reproduced by a typical loudspeaker. As established in the previous section, such a loudspeaker would be omnidirectional at low frequencies which would cause sound to radiate towards every surface in a room. So long as the walls were appropriately massive, some of the sound energy would be reflected back into the room. *Room modes* occur when the length of a reflected wave path (for example, between two boundaries) is equal to an integer multiple of half wavelengths. The direct and reflected waves interfere such that a fixed pressure/velocity distribution is formed within the space, with regularly distributed nodes and antinodes (points of minimal and maximal pressure/velocity). This leads to a variation in sound pressure level at different points throughout the room.

Modes can take the form of axial modes, which occur between two parallel boundaries,
tangential modes, which are formed between pairs of boundaries forming an edge, and oblique modes, where three boundaries forming a vertex or corner are involved.

The frequencies at which the modes occur, and positions at which nodes and antinodes occur, are dependent upon the geometry of the relative boundaries. Hence, different rooms, with different boundary geometries, would be expected to demonstrate different modal characteristics. The low frequency response perceived or measured at the listener position is dependent upon the position of the receiver, the position of the loudspeaker, the relative distance between boundaries, and the frequency being reproduced by the loudspeaker.

As the reproduced frequency is increased the nodes and antinodes become closer together in space and therefore the loudspeaker response in the room becomes more spatially averaged. This means that the response of the loudspeaker is more consistent across the room, and less defined by specific modal interactions. The Schroeder frequency is the frequency above which the room modes are no longer dominant, although its application to small rooms has been questioned (due to its dependence on diffusivity and meaningful reverberation times (Toole 2008)).

In addition to room modes, the response of conventional loudspeakers, at low frequencies, is subject to adjacent boundary effects. If the loudspeaker is placed close to a boundary, much of the energy that radiates in the direction towards the boundary will be reflected back. This leads to a doubling of pressure (+6dB) at a reference point away from the loudspeaker, in the room. If the loudspeaker is placed closely to more boundaries (between the floor and ceiling, or in a corner), then this doubling in pressure is increased, and so a low frequency peak of up to 18dB (theoretically) can occur at the reference position (Toole 2008) (these effects were investigated thoroughly by Allison (1974) in an AES Journal paper entitled 'The Influence of Room Boundaries on Loudspeaker Power Output'

Also, when the distance between the loudspeaker and the wall is equal to a quarter of the wavelength being reproduced, cancellation between the direct and reflected sound can lead to a large dip in response. This phenomenon is known as ‘Comb Filtering’ (Toole 2008).

These effects occur due to the lack of absorption of low frequency sound in rooms. At high frequencies, much more sound energy is absorbed because the transfer of sound energy to heat is more efficient in porous absorbers and the relative thicknesses of fibrous furnishings (such as rugs, carpets, sofas etc) are more comparable to the
wavelengths, aiding absorption (Smith et al. 1982). Therefore, even if a loudspeaker was omnidirectional at high frequencies, more of the energy would be absorbed in typical domestic rooms than at low frequencies, and so these effects (i.e. those described above) would be reduced.

The absorptive nature of rooms at high frequencies results in short-term high frequency reflections, which may only interact with a few surfaces before being attenuated significantly at the listener position. The exact characteristics of these reflections are determined by the directivity, room geometry (dictating the distance travelled), and boundary composition (absorption/diffusivity).

In order to classify the most dominant reflections, an ordinal system is used. First order reflections are those that reflect from just one surface to the listener, second order are those which reflect from two surfaces before arriving at the listener, and so on. Lower order reflections (first and second) are higher in level and contribute more to the sound field than higher order reflections (higher order reflections are attenuated more as they have lost energy at more surfaces).

In summary:

- loudspeaker directivity specifies the direction of sound radiated by a loudspeaker in a room, and therefore the amount of interaction with room boundaries;

- the distance from the loudspeaker to the boundaries, and the absorption/scattering properties of the boundary, affects the nature of the reflected sound component which arrives at the listener;

- at low frequencies, where conventional loudspeakers are omnidirectional, the sound field in a room is dominated by room modes. It is also affected by adjacent boundary effects, and therefore, at low frequencies, the sound field is highly dependent on room geometry, loudspeaker and listener position; and

- at higher frequencies, modal interactions become more spatially averaged. The furnishings and treatments in rooms are typically more absorptive, and so the sound field becomes dominated by discrete reflections. These reflections are dependent on the directivity of the loudspeaker, the room geometry and the boundary characteristics.
2.2.2 The Effect of Changing the Directivity of a Loudspeaker Upon the Sound Field

Considering previous discussion, a change in loudspeaker directivity alone could change:

- the number of reflections;
- the level of reflections and
- the incident angle of reflections.

The change would be dependent on the boundary characteristics of the room (geometry and absorption) and the listener and loudspeaker positions.

If we were to consider variation in directivity, boundary conditions and listener/loudspeaker position, the following parameters could be changed:

- the number, level, incident angle and arrival time of reflections;
- the level and arrival time of the direct sound field component;
- the spectrum of the sound field at the listener.

By changing the directivity of a loudspeaker in a room, the nature of the direct and reflected sound at the listening position can be changed. Assuming there is some reflection from the boundaries, an increase in the angle of radiation in one plane would cause the sound to be radiated towards more boundaries in that plane. Therefore, there would be an increase in the number of reflections arriving at the listener which would result in the indirect component of the sound field being more diffuse (energy density throughout the space becomes more similar (Kinsler et al. 2000)). Conversely, a narrowing of the directivity in one plane would cause a reduced interaction with the boundaries in that plane, and therefore, the indirect component of the sound field to be less diffuse. The so-called ‘imprint’ of the room upon the resulting sound field is therefore dependent on the nature of directivity, as well as the boundary characteristics which determine how much of the radiated sound is absorbed as opposed to reflected towards the listener. If the boundaries of a room are highly absorbent, it would be expected that a change in directivity would cause little change to the sound arriving at a listener on-axis (if the on-axis response remained unchanged).
CHAPTER 2. LOUDSPEAKER DIRECTIVITY AND REPRODUCED SOUND

Due to the increased absorption of high frequency sound compared to low frequency sound in rooms, it would be expected that a narrowing of directivity towards a listener would cause the high frequency sound field to tend towards that in an anechoic environment; assuming the listener is sat in front of the loudspeaker, sound would likely be reflected only once from the rear part of the room, before decaying (though this depends upon the size of the room). If the same narrow directivity pattern existed for low frequency reproduction, it is likely that the reflections would not be absorbed as effectively, with standing waves created between the front and rear walls. Therefore, the sound field would still be dominated by modes at low frequencies, just limited to a set excluding the side walls and floor-ceiling. Thus, whereas a change in directivity at high frequencies could significantly reduce the imprint of the room upon the sound field at the listener, the same change would not necessarily constitute a reduction in room imprint at low frequencies. This means that at high frequencies, loudspeaker directivity can be changed to make the sound field relatively independent of the room it is in, whereas, at low frequencies, it can only be changed to reduce the number of surfaces involved in the room modes. (If enough low frequency absorption is used, so that it does not reflect at the boundaries, the room imprint may well be reduced).

In summary:

- a uniform increase in off-axis directivity leads to more reflections arriving at the listener in a reflective room;
- a narrowing of directivity reduces the amount of indirect sound arriving at the listener and thus reduces the 'imprint' of the reflective room;
- with increased amounts of absorption, the effect of the directivity change will be lessened, as the 'imprint' tends towards an anechoic environment; and
- it is hypothesised that a narrow directivity would still result in some form of room imprint if the absorption was insufficient to prevent low frequency reflections.
2.2.3 Section Summary and Discussion

Summary

The sound radiated by a loudspeaker can interact in different ways with a room. At low frequencies, loudspeakers are typically omnidirectional, and the wavelength is such that the sound is not easily absorbed. This results in a 'modal' distribution of reflected sound waves, the characteristics of which depend upon the geometry of the room, the placement of the source, and the position of the receiver. At higher frequencies, more sound is absorbed by the room, and so the sound decays more rapidly (modal behaviour is less dominant). As such, the nature of the indirect sound is highly dependent upon the directivity, the absorption characteristic and the distance of each indirect path.

Changing the directivity leads to a change in the interaction between the sound radiated by the loudspeaker and the boundaries of a room. Variation in combinations of the loudspeaker, listener and boundary parameters leads to a number of potential changes in the nature of the direct and indirect sound components (such as arrival time and level, number of reflections and angle of reflections) arriving at a listener in a room. Considering a typical domestic room, in which high frequency sound is absorbed more easily, at high frequencies a narrowing in directivity would reduce the amount of reflections that arrive at the listener from the off-axis surfaces and, therefore, would reduce the influence of the room upon the sound field. At low frequencies, however, the room would still have an influence upon the sound field, despite being narrowly radiated in one direction, as the sound would be reflected between multiple surfaces and demonstrate a modal behaviour dependent on the room geometry.

Discussion

This section has made steps towards answering the second research question, R2: What part does room acoustics play in the relationship between loudspeaker directivity and the sound field at the listener? If the effect that loudspeaker directivity has upon the sound field, and subsequently the perception of reproduction at the listener position is to be considered, the boundary characteristics of the room must also be considered, as they directly affect the level, delay and angle of incidence of any reflected sound with respect to the direct sound. It is, therefore, of interest to investigate the effect that combinations of directivity and boundary characteristics have upon the sound field and
the perception of the reproduced sound.

2.3 Chapter Summary

The purpose of this chapter was to introduce loudspeaker directivity and to establish how it affects the sound field within a listening room.

It was defined, in the first section, to be a measure that describes the way in which a loudspeaker radiates sound into the space around it. Loudspeakers can be manufactured to exhibit different directivity characteristics, and this is largely dependent upon the configuration of driver units and their housing (though, directivity also depends upon more specific parameters such as transducer type and cabinet design). Methods for additional control have been discussed, including mechanical reflector systems, which re-direct radiated sound in order to spread it over a wider area, as well as multiple-loudspeaker arrays, which can be designed to allow full directivity control in real-time.

The latter section of this chapter considered the effect of changes in directivity upon the sound field within a listening room. It was established that any changes caused to the sound field as a result of changing directivity are dependent upon the room boundaries and the loudspeaker/listener position, and so changes in the combination of these variables should be considered. It would be expected that their variation would lead to changes in: reflection number, level and arrival time, direct sound arrival time and level. It would be appropriate, next, to consider how such changes at the listening position affect the perception of the sound reproduced in the room.
CHAPTER CONCLUSIONS

What is loudspeaker directivity?
Loudspeaker directivity is a multivariate function that describes the radiation of a loudspeaker in all directions.

What factors affect the directivity of a loudspeaker?
It is affected by the transduction type, arrangement, mounting and size of the drive mechanisms used.

To what extent can loudspeaker directivity be controlled?
With a sufficient number of drive-units and processing, real-time directivity control can be achieved. The real-time control of more basic configurations is much more limited. Theoretical principles dictate that the directivity of typical two or three-way domestic cabinet loudspeakers cannot be controlled in real-time and that they are all expected to be omnidirectional below around 200Hz, becoming increasingly directional towards the on-axis direction as frequency increases above that.

How does loudspeaker directivity influence the reproduced sound field in a room?
Loudspeaker directivity determines the angle and relative level of sound radiated into the listening space. The geometry of the room, listener position and absorption characteristics of the room subsequently affect the reflections that arrive at the listener.
Chapter 3

Perception of Reproduced Sound

A change in the combination of loudspeaker directivity, listener/loudspeaker position and boundary characteristics can lead to the alteration of a reproduced sound field in a room. More specifically, at the listening position the arrival time and level and the spectrum of the direct sound can be altered, as well as the number, arrival time, level, arriving angles and spectrum of reflections. In order to move closer towards the principal research aim, it is now necessary to consider how such changes might change a listener’s perception of the reproduced sound field. In order to do so, the following questions will be answered:

- What are the basic auditory mechanisms involved in human sound perception?
- How is the perception of reproduced sound affected by the room?

By drawing on the knowledge already established, as well as material from other literature and previous research, the following items will also be considered:

- What is the influence of directivity and boundary characteristics upon the perception of reproduced sound in rooms?
- What changes to directivity and boundary characteristics are perceivable, what might the relative perceived magnitude of the changes be, and which perceptual attributes would be affected?
- What part does listener position play in the relationship between loudspeaker directivity, boundary characteristics and perception of reproduction?
CHAPTER 3. PERCEPTION OF REPRODUCED SOUND

The chapter begins with a section that introduces the human auditory system, and is then divided into two key subsections: spatial and non-spatial auditory analysis. For each, the physiological/cognitive mechanisms involved and how they relate to the subjective perception of reproduced audio signals in the presence of reflections is discussed. Following this, the auditory attributes that are likely to be affected by a change in directivity/boundary characteristics (with consideration of the role of listener position in affecting these attributes) are highlighted. Finally, the contribution of these changes to the overall perceived quality of the listening situation is considered, before an overview of existing research in loudspeaker directivity and the perception of reproduction is presented.

3.1 Auditory Analysis and Perception of Reproduced Sounds in Rooms

3.1.1 Human Auditory System and Basic Auditory Analysis

The human auditory system is the sensory mechanism that allows us to hear and interpret sound. Bregman (1990) provides a thorough description of the auditory system, and this description, therefore, will serve as a basis for the following section.

All sound arriving at the human ear is immediately filtered (as a function of its direction) by the pinnae, the skin-covered flaps of cartilage located at the side of the head. The external fluctuations in air pressure, caused by a vibrating sound source, are then guided toward the tympanic membrane (eardrum) via the ear canal. This initial system is called the outer ear. The succeeding middle ear transmits vibrations from the eardrum to the cochlea, which is located in the inner ear. The cochlea is a complex organ which responds to the vibrations (typically between 20 and 20000 Hz). It does so via the resonances that occur in the basilar membrane. This is located inside the cochlea and occupies its entire length. The positions of the resonances along the basilar membrane correspond to the frequencies of the vibrations. These resonances excite inner hair cells (IHCs) which line the membrane. The IHCs convert their motion into neural activity, which is transmitted to the brain via the auditory nerve.

It is the auditory cortex in the brain that processes the neural information. The cochlea outputs nerve firings based upon frequency and amplitude information - the higher neural centres then evaluate various cues such as Interaural Level Difference (ILD), Interaural
CHAPTER 3. PERCEPTION OF REPRODUCED SOUND

Time Difference (ITD), Amplitude Modulation (AM), Frequency Modulation (FM) and periodicity.

Auditory scene analysis describes the cognitive processes that are required to interpret an acoustic input signal at the ear. It essentially consists of two processes: segmentation, where the signal is deconstructed into Time-Frequency (T-F) regions via the cochlea, and grouping, where these regions are reconstructed into groups that are related to particular sources.

Sounds are either grouped according to an analysis of their structure (Primitive) or based upon pattern-learned recognition (Schema-Based). Primitive grouping takes place initially to distinguish separate auditory sources or events by combining sounds with similar spatial, harmonic, temporal and modulation characteristics. For example, this allows us to perceive that there are different sounds in a room and that they are coming from different sources. Schema-based grouping allows us to classify the separate sounds into objects that we can interpret as meaning something. An example here is the use of memory to deduce that an oscillating swept-sine-wave is a siren or alarm.

With regards to reflected sound in a room, this would suggest that, despite differences in spectral and temporal characteristics compared to the direct sound, the reflections and direct sound would be expected to be perceptually linked via Primitive grouping. Considering the role of listener position upon perception, it is evident that, even at this stage, the direction of a listener relative to the sound source has an effect upon the filtering of the sound arriving at a listener.

In summary:

- an acoustic signal arrives at the ear, where it is filtered by the pinnae. The filtering is dependent upon listener position;
- signal is then guided towards inner ear, where it causes motion of the inner hair cells;
- this motion is converted to neural information which is transmitted to the brain; and
- the brain analyses the signal before attempting to perceive it as a recognisable auditory event.
CHAPTER 3. PERCEPTION OF REPRODUCED SOUND

3.1.2 Non-Spatial Auditory Analysis

Humans are capable of much more than just being able to differentiate between multiple sound sources and recognise different types of sounds - we are sensitive to spectral characteristics of sounds, the level of sounds, the relative size of one sound source compared to another and the positions of sound sources. *Non-Spatial Auditory Analysis* relates to the processes which determine spectral and amplitude cues (i.e. not size or position).

Spectral and Temporal Effects - Timbre

The ability of the auditory system to analyse a signal with regard to its frequency content is fundamental to how a sound is perceived. At low frequencies (below 150Hz), the motion of hair cells triggers single nerve firings relating to the periodicity of the waveform, however, between 150Hz and 4kHz, the brain relies upon a combination of nerve fibre outputs to provide information about the waveform (given that the maximum firing rate of a single nerve fibre is 150Hz) (Rumsey and McCormick 2002). The brain also assesses the position of maximum deflection (resonance) on the basilar membrane to gather frequency information. This provides information about frequencies from 50Hz upwards and is the only mechanism used above ~4kHz.

Most of the sounds that humans are exposed to involve complex spectral and temporal structures and so understanding how subtle differences in sounds are perceived is important to this research.

*Timbre* is a perceptual attribute that describes the way humans perceive the 'colour' of a sound, and it allows the difference between different sound sources to be described. The American Standards Association produced the following definition:

Timbre: *'That attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar.'* ANSI (1960)

The definition adopted by this thesis is that timbre is all attributes distinguishing sounds from one another except from loudness, pitch and spatial factors.
There are two aspects to timbre: time-invariant properties of a sound (such as spectral content) and time-variant properties (such as the amplitude envelope) (Moore 2002). The time-invariant aspect is most commonly discussed; ‘timbre is determined by the number and relative strengths of the instrument’s partials (harmonics)’ (Everest 2001). Thus, though the perceived pitch (dictated by the fundamental frequency) and loudness for two different sources are each perceived to be the same, the nature of the tone reproduced can be ‘coloured’ due to the difference in harmonics produced by the sources.

The time-variant aspect of timbre relates to differences in the onset transients and the temporal structure of the sound envelope. Moore (2002) gives the example that if the sound of a piano note is reversed in time, it sounds more like a harmonium, despite the note being the same, in terms of long-term average spectrum.

In order for the timbre of the sound arriving at the listener to be changed, either the spectral content or the amplitude envelope must be affected.

In a reflective room, the spectral content of a sound arriving at a listener from a loudspeaker can be changed as a result of comb filtering, which follows from the acoustical summation of a sound and a delayed version of itself (for example, the summation of the direct sound from a loudspeaker to a listener and the indirect reflected sound from the wall to the listener). This causes a comb-like response where at some frequencies the two signals sum at the ear and at others they cancel, depending on the relationship between the time delay and frequency. Whilst in many audio-related publications this phenomenon is viewed as detrimental to the listening experience, Toole (2008) provides a defensive outlook upon comb filtering, suggesting that typical curves are wrongly based upon equal-level delays. He also suggests that the well-known high frequency notches are too close together to be audible and that low frequency notches are masked by room resonances, summarising that the overall effects are likely to be inaudible within a reflective environment. Research by Clark (1983) using temporal waterfall plots indicates that, in real reflective environments, the comb-notches that are evident in a synthesised delay are not present and can be shown to be ‘filled’ by the reflective sound component at lower frequencies. This supports another earlier argument of Toole (2008), that wider-dispersion loudspeakers are able to increase reflections and thus neutralise the spectrum by ‘filling the gaps’ which may otherwise occur due to crosstalk from narrow-dispersion loudspeakers. He also elaborates upon the ‘one-toothed comb phenomenon’, whereby the comparison of a real central image, to that of a phantom image, indicates a large dip at 2kHz, which corresponds to the delays encountered via crosstalk. He summarises
that this effect is more prominent in less reflective rooms and thus provides an argument against the removal of reflections in an audio control room, where a 2kHz dip could be detrimental.

Toole (2008) is not the only person to testify against the detrimental nature of comb-filtering, with much cognitive-auditory research also suggesting the ability of the brain to overcome modification of signals arising from comb-filtering. Barron (1971), Case (2001) and Moulton (1995) showed that in the presence of multiple reflections such colourations may be perceptually overlooked completely and Zurek (1979) even showed that some amount of spectral smoothing can occur as a result of multiple reflections.

The excitation of multiple reflections means that several copies of the direct sound will arrive at the listener. Such ‘repetitions’ of the reproduced signal are said to contribute to perceived timbral changes, and research by Toole (2008) supports this; resonances are said to characterise the timbral nuances of the original sound sources, and so Toole aimed to investigate the audibility of resonances in the reproduced audio material, and the influence of reflections upon their detection. Tests were carried out in a reflective hall, an anechoic chamber and in headphones, with results showing that low Q resonances, that is subtle timbral nuances, are more audible with an increased repetition frequency. Thus, in this case, timbral nuances were able to be detected more easily in the presence of reflections. He proposes that the increased number of repetitions allows us to gather more information about the sound we are listening to, including the onset and offset information. It is therefore intuitive that, upon this basis, we would be able to distinguish one instrument from another more clearly in an environment which produced more reflections (compared to an anechoic chamber).

Whilst the above research is concerned with the effects of reflections in the horizontal plane, reflections in real listening rooms are also likely to arrive from all directions (from floor and ceiling etc), and so the consideration of effects of non-horizontal reflections on the perception of a replayed recording is also of importance. Bech (1994a, 1995, 1996) carried out a series of tests as part of the ‘Archimedes Project’ (Bang&Olufsen, Kef and the National Laboratory of DTU (Technical University of Denmark)). These investigated which reflections (within a typical listening environment) were likely to contribute individually to changes in timbre and localisation. Several loudspeakers were set up in an anechoic chamber and used to reproduce simulated reflections. Listeners took part in a range of threshold tests to determine how sensitive they were to changes
in timbre and localisation as a result of changes in the level of each single reflection in the presence of the other reflections (which remained fixed). Overall, after tests with different source directivity / room absorption arrangements, results suggested that changes in first order floor and ceiling reflections, in the presence of other reflections, were most likely to contribute individually to a perceived change in timbre and localisation, and that, for naturally occurring levels, changes in the other reflections would not. This suggests that in a real room, where several reflections in all directions are likely to occur, changes in level of horizontal reflections alone are likely to be less severe in their effect upon perception of sound reproduction than changes in vertical reflections.

The consideration of timbre is important here, as characteristics of the loudspeaker and room will partly determine the spectral and temporal nature of the sound arriving at a listener. It is likely, therefore, that listeners would perceive changes in timbre as a result of changing the directivity of a loudspeaker and the boundary characteristics. Changes in reflection timing (due to position of loudspeaker/boundaries/listener and/or directivity) would contribute to temporal alteration of the sound, and changes in the reflection frequency content (due to frequency-dependence of boundary absorption and/or comb filtering) would change the spectrum of the sound at the listening position.

**Loudness**

Loudness describes how the auditory system interprets the sound pressure level of the acoustical input signal at the ears. The ear is not uniformly sensitive to level, and the level required for a tone to be perceived as uniformly loud over all audible frequencies defines the equal loudness contours. They indicate that with increased level, perceived loudness is more uniform across frequencies (See Figure 3.1). At low levels, low frequencies need to be up to 80dB higher in level and high frequencies need to be up to 20dB higher in level in order to be perceived as being the same loudness as mid-frequency signals. Thus, signals at low levels can sound lacking in 'bass' and 'treble', whereas at high levels, 'bass' and 'treble' will be perceived as louder with respect to the mid-range. It has been also been shown that broadband sounds are louder than those occupying narrower bands and distortions are often louder than the same signal without distortions (Rumsey and McCormick 2002).

Loudness is important to this study, as a change in loudspeaker directivity, depending
Figure 3.1: Equal Loudness Contours reproduced using iso226.m (Tackett 2010). These contours show the flattening of the response with increased level (with labels in phons where 1phon = 1dB at 1kHz).

on the boundary conditions and listener position, could result in a change in perceived loudness. A basic example would be if a loudspeaker was designed to beam off-axis at high-frequencies in an absorbent room, the reproduction would be perceived as quieter to a listener sat on-axis compared to a reproduction of the same signal from a loudspeaker that beamed on-axis, because the level of the sound arriving at the listener would be lower in comparison. The frequency and level dependent nature of loudness perception means that if frequency-specific loudness changes occur because of a change in directivity/boundary condition, they may be more or less noticeable depending upon the frequencies affected and what level playback is at.

Listener position is crucial to the relationship between loudness and changes in directivity/boundary conditions because the change in the level of sound arriving at the listener due to a change in directivity/boundary conditions can be affected by the listener changing position.

In summary:

- the auditory system determines the frequency of a stimulus primarily via reso-
nances on the basilar membrane and timing information from the motion of inner hair cells;

- the basilar membrane acts as a bank of overlapping filters, responding to only a limited bandwidth of frequencies at each position;
- timbre is a perceptual attribute associated with the spectral content and amplitude envelope of a complex sound event;
- the timbre of sound reproduction in a room is expected to change with a change in loudspeaker directivity/boundary characteristics and listener position;
- changes in loudspeaker directivity and boundary characteristics are likely to alter the spectrum of the sound field, though some authors report that it is unlikely that any spectral lumps that are produced would be perceptually significant and that, if anything, a 'smoothing' of the sound field response would result;
- reflections cause repetition, so that timbral nuances are easier to detect;
- vertical reflections can cause changes in timbre; and
- perceived loudness of acoustic sound pressure is frequency and level dependent, as shown by equal loudness contours. Thus, changes in loudness caused by changes in directivity, boundary conditions or listener position are dependent on the frequencies affected and the level of the signal being reproduced.

3.1.3 Spatial Auditory Analysis

Spatial Auditory Analysis relates to the processes used for spatial cues, such as localisation.

Directional Perception

Rumsey (2001) states that, with just one ear (monaural), humans are able to deduce some spatial information about a sound source. This relies on the analysis of spectral changes and level changes, discussed above. Most people, however, have two functioning ears, and by considering the relationship between the acoustical signal arriving at each of them, with regard to amplitude, timing, phase and frequency, can deduce very accurate positional information.
CHAPTER 3. PERCEPTION OF REPRODUCED SOUND

Any source that is located off-axis (i.e. one that does not fall on the median plane) will radiate a signal that arrives at one ear before the other. Therefore, it will be detected by the brain as being in the direction of the ear receiving the early signal (Rumsey 2001). In order to distinguish whether a sound is in front or behind, head movements are often necessary - motion towards a source at the front will lead to the sound arriving later at the nearest ear, whereas motion towards a sound at the back will lead to the sound arriving earlier. Cues resulting from pinna-interaction can also reduce front-back confusion.

At high frequencies, the spectrum of sound arriving at the tympanic membrane can be affected significantly as a result of pinna reflections and resonances, and also as a result of reflections from the shoulders and body. The head itself can even act as a barrier, further aiding directional perception as a result of differences in spectral content at each ear.

Such concepts highlight that the angle of incoming acoustical signals can have an appreciable effect upon the final spectrum of sound arriving at the eardrum, and therefore listener position and angle with respect to the source must be considered carefully when investigating perceptual response to audio.

The interaction of direct sound and sound that is reflected from the head and torso of a listener contributes to what is called the 'Head Related Transfer Function' (HRTF). This describes the spectral and temporal changes to the arriving signal as a function of source position and angle of incidence. The change in HRTF for a given change in source position/angle of incidence is a useful spatial cue.

Directional perception is relevant to this study, because the cues used to perceive the direction of a sound source are related to a number of perceptual phenomena which are discussed later. These cues can be affected by changes in directivity, boundary conditions and position and thus can affect those phenomena.

Envelopment and Spaciousness

Envelopment is a term that is used to 'describe the sense of immersivity and involvement in a (reverberant) sound field' (Rumsey (2001)), whereas spaciousness relates to the perception of space within which the listener perceives to be located. The classification of these terms has been the topic of lengthy debate amongst academics, however, for the purposes of this research, it may be more succinct to simply appreciate that these
attributes are correlated with the amount and type of the reflected sound and the type of recording.

Source Distance

In order for a listener to determine the distance between him/herself and a sound source, a number of cues are utilised. A further source would be quieter with reduced high frequency content due to air absorption. In a reflective environment, a further source would result in a reduced direct-to-indirect sound ratio at the listener, there would be less difference between the direct sound and the first reflection (initial time-delay gap), and the ground reflection would be attenuated. In reflective environments, these cues present an idea of both the distance from the source and the size of the environment, however, in non-reflective environments, the reliance on absolute distance perception using airborne attenuation is known to be unreliable (though, judgements of relative distances are more reliable) (Flanagan and Taylor 1999).

Apparent Source Width, ASW

Before discussing ASW, it is necessary to introduce the precedence effect. The precedence effect means that a delayed version of a sound is likely to be perceptually fused with the original, and the fused sound to have a single perceived source, if the delay time is less than ~50ms (depending on signal type). If it is delayed longer than this, then it is likely to be perceived as a separate source (Haas 1972; Gardner 1968, 1969). For delay times of less than 1ms or so the way the human auditory system perceives sources with this amount of delay is referred to as summing localisation, and the perceived location of the fused sound is between the positions of the original and delayed sound. For delay times longer than this (and up to ~50ms, depending on signal type), a broadening in the perceived sound source width is likely (as well as changes in loudness and timbre for the reasons above). Since a reflection is equivalent to a delayed version of the original sound, reflections arriving from a different direction than the source would be expected to cause these effects.

In 1989, Toole and Olive carried out an experiment investigating the detection of single lateral reflections with speech stimuli reproduced in an anechoic chamber (Olive and Toole 1989). Low-level delays were reported to cause an increase in spaciousness (see below), with higher-level delays causing an 'image shift', where subtle changes in the
primary and secondary images equally loud (Haas, 1972)

second image audible (Meyer & Schodder, 1952)
second image audible (Lochner & Berger, 1958)

shift or spreading of primary image (Olive & Toole, 1989)
detection threshold (Olive & Toole, 1989)

Figure 3.2: This figure was compiled by Toole (2008) and demonstrates the perceived effects of a change in level and delay of a single lateral reflection.

position and size of the image were detected. Research carried out by Meyer and Schodder (1972) and Lochner and Burger (1958) suggest that at higher levels of delay, a second image becomes apparent, perceived to co-exist with the original (although not temporally separated, as in echo). The results of these tests were compiled into a single graph by Toole (2008), and it is presented here in Figure 3.2. Tests were carried out in a free-field environment with the simulated reflection added in front of the listener, but the angle of the direct sound and the simulated reflection were not the same for all tests.

The thresholds determined by this type of research also show some signs of variance as a result of the reflection angle, and it is apparent that reflection angle can contribute significantly to perceptual response at the listening position. Hidaka et al. (1997) produced polar plots relating to the sensation of ASW for different reflection angles. Reflections at around 60° were estimated to generate the greatest perception of ASW and the polar plot is shown in Figure 3.3. Similarly, the direction of the reflection has been shown by Olive and Toole (1989) to determine the nature of the perceptual effect, with vertical reflections causing timbral changes, and lateral reflections causing changes in spaciousness.

Olive and Toole (1989) extended the research concerning single reflections, by repeating the threshold detection test in several environments. These included an anechoic
CHAPTER 3. PERCEPTION OF REPRODUCED SOUND

Figure 3.3: Reformatted data from Hidaka et al. (1997), plotted to show the relationship between reflection angle and perception of apparent source width (ASW) / image broadening (Toole 2008).

chamber, a listening room with attenuated first order reflections, and a ‘live’ listening room. Results indicated that the detection threshold for a single reflection is increased with an increased number of reflections, suggesting that a more reflective environment leads to a reduction in sensitivity to single lateral reflections. This could imply that with increased reflections, the influence of singular prominent reflections (perhaps first order) is reduced, and that the potential for image shift and spaciousness as a result of these prominent reflections is limited. Thus, in an anechoic chamber for example, according to these results, the level of a reflection required to give a perceived shift or spread in image is less than that required for the same perceptual effect to occur within a listening environment that is more reflective.

Much of the aforementioned research, and related research, has been conducted using speech. Barron’s (1971) experiments with music have similar findings to those with speech, although Toole (2008) highlights that the detection thresholds are ‘flatter’ (less variance with delay), owing to the ‘prolongation’ of notes. Similarly to the effect of increased reflections, it would seem that an increased continuity in the sound content allows for lower sensitivity to discrete lateral reflections.

In summary:

- differences in arrival time, and level, at the ears allow directional perception of sound sources to be possible;
- the head and torso interact with sound to cause local reflections and attenuation.
This further aids directional perception, and can alter the spectral content of sound arriving at ears;

- these spectral changes are dependent on the listeners' position, HRTFs and the relative source position/angle;
- lateral reflections can cause apparent source width - an increase in the level of lateral reflections leads to a wider ASW; and
- ASW is dependent on the reflection angle and the reflectiveness of a room - ASW is greatest with a 60° reflection angle or with more prominent singular reflections.

3.1.4 Section Summary and Discussion

When an acoustic signal arrives at a listener, the auditory system responds to variations in level and frequency over time. Using this information, humans are able to perceive:

- Pitch
- Timbre
- Loudness

With schema-based grouping, humans are able to match sounds they hear to sounds they already know by comparing these parameters.

Using differences between the signals arriving at each ear, humans are able to deduce more accurately the localisation of sound sources.

This information highlights that there are a number of perceptual attributes that are affected by basic changes in the sound arriving at the ears. The brain uses changes in the level, frequency and temporal characteristics to deduce spatial and non-spatial information which allows humans to distinguish differences between different types of sound coming from different locations.

Discussion

This section then established that if the level, frequency or temporal characteristics of a sound are changed, some change in the perception of the sound could be evoked in a
CHAPTER 3. PERCEPTION OF REPRODUCED SOUND

listener (depending on the magnitude of the change). This may be in the general 'colour' of the sound, or in its perceived direction. Also, if changes in the reflected sound (in relation to the direct sound) take place, the perceived size and distance of the sound source may change and additional sensations of 'space' or 'immersiveness' may occur.

More specifically, it is possible to link changes in these characteristics to particular auditory attributes (though, many more acoustic-perceptual interactions exist than listed here):

- A change in level contributes towards the perceived loudness of a source.
- A difference between level at the ears contributes towards the localisation of a source.
- A change in the fundamental frequency contributes to perceived pitch.
- A change in the spectral content contributes towards timbre.
- A difference between spectral content at either ear contributes to localisation.
- Changes in level and spectrum over time contribute to our perception of timbre (based on envelope and spectral content) and localisation of moving sources.
- Changes in the amount and type of reflected sound can affect the perceived envelopment and spaciousness.
- A change in direct-to-indirect sound ratio at the listener affects the perceived distance of a sound.
- Changes in the level and delay of a reflected version of a direct sound can affect the apparent source width.

All of these changes (apart from a change in fundamental frequency) could be caused by a variation in directivity, boundary characteristics or loudspeaker/listener position, and so with reference to initial research question R4: 'Which Perceptual Attributes are affected by the combinations of loudspeaker directivity/boundary characteristics', the following list can, at this stage, be compiled:

Attributes expected to be affected by changes in reflections (as a result of changes in directivity, boundary characteristics and loudspeaker/listener position), according to theory and literature:
CHAPTER 3. PERCEPTION OF REPRODUCED SOUND

- Timbre;
- Localisation;
- Loudness;
- Localisation;
- Envelopment/Spaciousness;
- Source Distance;
- Apparent Source Width (ASW).

3.2 Reflections - Considering Adaptation and Preference

The previous section suggests which auditory components may be most influenced by reflections when listening to sound reproduction over loudspeakers. The effects described have been observed in research involving detection thresholds and objective measurement data, but it is still difficult to predict the overall influence of multiple reflections upon the perception of reproduction. Whilst a measured increase in level or spectral change may occur as a result of an increase in the level of reflections, it is still not clear, exactly, how this relates to the overall listening experience, whether one condition is preferred over another or whether such effects are even perceived to change the quality of the reproduction at all.

Our hearing has the ability to adapt to the space that we are in, so that we may focus upon an acoustic source within that space. This adaptation may render any objective changes in sound field (such as those described above) insignificant with regard to the overall perception of quality - if our auditory system automatically re-adjusts in order to reduce the influence of reflections, then exploring the effects of changing a sound field may be irrelevant.

Toole (2008) suggests that, following the initial exposure to a sound source within a space, an impression of the space itself, and some general directional information, is deduced. It is only after some time that the impression of the space will fade, and analysis via the precedence effect reveals the true direction of the sound source.

It follows that with more time particular environments may be 'learned' and committed to memory, and research by Perrot et al. (1989) and Saberi and Perrot (1990) suggests
that, with enough exposure to reflections in listening tests, the precedence effect can be ignored and delayed reflections may be detected as if they were isolated from one another.

Taking this into account, the idea of adaptation should be borne in mind when considering the overall impact of this research upon the overall quality of sound perception. The assessment and measurement of loudspeakers with different directivities may lead to variation in perceptual response but after time and with adaptation, the differences in overall subjective quality of each may be insignificant.

Considering preference with regards to reflections, Ando (1977, 1998) carried out an experiment in an anechoic chamber, where listeners were able to choose the desirable level of an accompanying single reflection of speech. Listeners chose reflections with characteristics in level and time which fall beneath the region where listeners in Meyer and Schodder's (1972) and Lochner and Burger's (1958) tests began to perceive a second image as being present (discussed earlier and presented in Figure 3.2). This suggests that the listeners preferred reflections which increase the perceived spaciousness and image size yet which fall beneath the threshold for which the image splits into two. Experimentation using music was also carried out by Ando (1985). Results showed similar desirability for 'spacious' and not 'image-separating' reflections, but it should be noted that the preferred levels of these singular reflections were far greater than would occur in typical rooms.

Kishinaga et al. (1979) performed listening tests using stereo reproduction which suggested that having absorptive side-walls was preferred for monitoring, and for the analysis of audio products, whereas reflective side walls were preferred when listeners were 'enjoying' the music. The placement of absorption was reported to reduce colouration and localisation of the loudspeakers.

These examples suggest that there is some degree of preference for reflections in casual listening but, in order to establish significant relationships between preference and reflections, extended research would be required.
3.2.1 Section Summary and Discussion

Summary

This section highlights the ability of the human auditory system to adapt to its environmental surroundings over time and that it should be considered in the development of any research in this area. It also presents some evidence to suggest that an increased number/level of reflections is preferred when listening to reproduced sound in a room, specifically for enjoyment (though more significant research would be required to reinforce these findings).

Discussion

The necessity to consider adaptation when carrying out this research is primarily applicable when considering the comparison of different loudspeaker directivities in different rooms - or in one room with changing boundary conditions. It does not apply to a situation where a loudspeaker system is designed which allows the user to instantly change the directivity properties - in this case, it is still important to understand how the changes correspond to the perception of the reproduction.

3.3 Loudspeaker Directivity and the Perception of Reproduced Sound (Previous Research)

Information from previous sections can be used to predict the attributes that should be affected by directivity. Before deciding on how to further explore the topic, it is necessary to consider current opinions and research that has already been carried out with regards to loudspeaker directivity and perception.

3.3.1 Initial Discussion

Various authors have presented arguments for and against particular loudspeaker directivity characteristics. Before reviewing their research, it would be appropriate to consider Allison’s (1995) paper, which presents a key underlying question with regards to loudspeaker directivity. He suggests that when we attend a live musical performance
in a concert hall, accurate localisation of the true sources is compromised in exchange for a greater sense of acoustic inclusion within the environment. In such a case the reverberant field is dominant at most listener positions and, should such a scene be recreated, it would be intuitive to think that the loudspeaker directivity should be wide so as to elicit as many reflections as possible, thus promoting a similar involvement. Whilst the ambience may be improved as a result of a greater level or delay of reflections, he highlights that the accuracy of any intended phantom image may be reduced. This has been discussed previously and is a well-documented consequence - the broadening in apparent source width as a result of increased level/delay of lateral reflections tends to be interpreted as a degradation in image clarity (Queen 1979; Kates 2002; Salmi 1982). Therefore the question arises whether to sacrifice image stability in order to increase a sense of ambience, or to ensure image stability by minimising the room interaction by maintaining a narrow directivity directed towards the listener. It would appear that the shift or spread in image, discussed earlier, as a result of increased level/delay of reflections, could be related to a reduction in image quality and therefore, although an increased sense of spaciousness may be perceived, the overall quality of the spatial audio may be compromised.

As a result, it must be investigated to what extent listeners in casual listening (or listening for enjoyment) feel the need for high precision imaging and, or alternatively, an increased perception of spaciousness. The previous sub-section began to explore this very important question. Allison’s (1995) example relates to the reproduction of concert-hall based recordings, but it is largely applicable to any recordings which intend to ‘involve’ the listener spatially and, consequently, has a significant influence on the choice of loudspeaker directivity and diffusivity of the sound field within the listening environment. This choice also depends on whether or not the listener desires to be ‘involved’ in the recording - for some, the requirement of the reproduction may be to reproduce the recording accurately, so that it faithfully replicates the engineer’s creation, with any interference from the room minimised.

If conventional moving-coil loudspeakers, which narrow in directivity with increased frequency, are considered, both of these spatial characteristics may be achieved, although not within the same frequency range. At low frequencies sound is typically radiated over a wide angle, potentially causing multiple reflections within the listening space, whereas at higher frequencies the directivity often narrows, and so less sound would be directly reflected from the surrounding surfaces (Linkwitz 2007) (that is, assuming that the environment were reflective and not anechoic). Subsequently, image quality could be
expected to be greater at high frequencies, whereas the sense of spaciousness might be greatest at lower frequencies. As such, the sound field could be perceived to change with frequency. In order to maintain a more consistent sound field the radiation would have to be uniform across frequency.

Uniformity of reflections has been widely discussed in a number of texts relating to the topic of loudspeakers (Allison 1995; Salmi 1982; Berlant 1985; Linkwitz 2007). Queen wrote in 1979, “to achieve good imaging and clarity, loudspeaker designs for home music listening rooms must consider directivity not from the standpoint of audience coverage, but from the standpoint of uniformity of the intensity of arriving reflections with respect to frequency” (Queen 1979). Linkwitz (2007) develops this idea, and hypothesises that “reflections generated by the two loudspeakers should be delayed copies of the direct sound to the listener (across all frequencies, with at least 6 ms delay) - under these conditions the direct sound dominates perceptually, the cognitive faculty of the brain better able to separate the static listening room acoustics from the acoustics embedded in the recording.” He comments on the practical requirements necessary for the hypothesis to apply - “the polar response of the loudspeaker must not change over its whole frequency range, the loudspeakers must be placed at least 1m away from adjacent surfaces and the requirement for full spectral content of reflections rules out the use of frequency dependent absorbers on the room surfaces” (Linkwitz 2007). Achieving frequency independent diffusion, by having surfaces which do not change the spectral content of the resulting sound field at the listener, is practically impossible. This research theorises the separation of the room ‘imprint’, as discussed earlier, from the direct sound.

3.3.2 Previous Studies

It is now necessary to consider scientific studies which measure the perceptual response of listeners to changes in directivity or associated parameters. Many authors report their own ‘perceptual’ response to a stimulus as being globally applicable to others. Such conclusions are more reliable when combined with structured and significant results from listening tests and, whilst several informal conclusions exist, supportive listening test data is less widely available. A summary of some findings follows.

Flindell et al. (1991) designed a novel listening test in order to evaluate the stereo image quality, spaciousness and overall impression of material reproduced via various simulated loudspeaker directivities. The outputs of an array of loudspeakers (positioned to simulate the key reflections of a room) were filtered in order to mimic the effects of
changing directional properties of the main stereo pair. Results indicated that there
was no clear consensus of preference, although a tendency in naïve listeners to prefer
more omnidirectional characteristics was noted. Experienced listeners also confirmed
that omnidirectional radiation did not degrade stereo imaging as much as presumed.

Choisel (2005) investigated the effect of loudspeaker directivity on the perceived direction
of sound sources within a stereo image. Listeners were asked to use a laser-pointer in
order to indicate the perceived locations of virtual sound sources, as they were panned
between a pair of loudspeakers. Loudspeakers of varying directivity characteristics were
used (B&W 801 and Beolab 5 with absorbing panels, Beolab 5 with reflecting panels),
but results indicated that perceived direction was not affected by the alteration of this
condition in this case. Zacharov (1998) conducted a group of experiments in order to
assess the effects of loudspeaker directivity in surround systems (five-channel). It is
necessary to consider only the ‘frontal test’ carried out by Zacharov which investigated
the effects of the frontal channels of a five-channel system only (with the rear channels
fixed). Three types of loudspeakers with different directivity characteristics were
compared in the front channel positions (cardioid, horizontal-line and direct radiator),
and listeners were asked to judge the coordination of the audio with a presented image,
the spatial awareness and naturalness. As in the previous examples, results indicated
that little difference was found to exist between the three systems under consideration.

In many experimental cases, even though underlying motivation for testing may not be
published as a directivity study, the recorded listener responses to different loudspeaker
units can provide further insight into the matter. Toole (2008, 1986) published a series of
papers exploring loudspeaker measurements and listener preferences, and in his recent
book these results are discussed to provide a further contribution to the directivity
debate. He concludes that his results, as well as those of other authors, “all point in
the same direction: that wide-dispersion loudspeakers, used in rooms that allow for
eyearly lateral reflections, are preferred by listeners especially, but not exclusively, for
recreational listening”. He also states that “there appear to be no notable sacrifices in
the ‘imaging’ qualities of stereo reproduction” (with such wide-dispersion loudspeakers).
Toole’s work appears to be the most relevant to date, with regard to loudspeaker
directivity effects, but these tests are not definitive. The wide-dispersion loudspeakers
in this case are two and three-way loudspeakers, and are so called in comparison to
the other test loudspeaker, an electrostatic dipole. It would be possible to investigate
loudspeakers which demonstrate more variant directivity characteristics, as well as more
variation in room characteristics, and so while his material provides useful findings, it
could be extended to explore a larger range of directivities and boundary conditions.

It is also possible to consider listening test data from other research, in case the effects of directivity have been measured indirectly. One such example is the work of Bech (1994b). A round-robin arrangement of tests was carried out, whereby listeners judged the perceptual fidelity of monophonic recordings presented over a group of different single loudspeakers, in different rooms. Four loudspeaker directivity types were used: dipole, line source, conventional two-way and spherical. For each type two loudspeakers were used to account for any differences that may occur between loudspeakers of the same type. Thus there were 8 loudspeakers in total, labelled A, a, B, b, C, c, D and d. Within each test room, two loudspeakers of differing types were placed closely together in three different positions (away from walls, centre of back wall, and back corner), and the test was repeated in each room so that all loudspeakers had been compared in different positions. This was repeated for two other rooms. If the resulting data is represented, with the loudspeaker type indicated as illustrated in Figure 3.4, we can further observe listener ratings for different directivity types. It is evident that the dependence of listeners' fidelity ratings on position (and room) is also important. Whilst the dipole is rated as worst in Position 2 (less than 1m from the back wall, central), it is rated as best when moved to Position 1 (over 1m from back and side wall). This suggests that the perceived influence of directivity is dependent on both position and room type, which would be expected following the work discussed earlier. As such, these test results are important with regards to the topic of this investigation.

In order to summarise the nature of listening tests described in the above literature, as well as the results, Table 3.1 is presented. The type of loudspeaker directivity investigated in each test is indicated, as well as other test parameters including environment, loudspeaker/listener position, programme material and the attribute measured.

The table highlights that most tests take place in a listening room environment, with the listener seated at the sweet spot. The number of directivity types investigated in each test is low (typically two or three) and the auditory attributes considered are varied. Results are mostly limited, providing no definitive evidence for relationships between directivity type and the attribute under investigation.
Figure 3.4: Mean Fidelity ratings for loudspeakers A-d in positions 1-3 averaged across rooms, subpositions, programs, and subjects. Re-edited from (Bech 1994b) to include loudspeaker directivity types. The 95% confidence intervals per loudspeaker range from 0.2 to 0.25.

In summary:

• the directivity of conventional loudspeakers is non-uniform with frequency, as are the absorption characteristics of the environment;

• if the directivity and absorption characteristics were controlled, in order to give uniform reflections, the audible influence of the room upon the sound field may be reduced;

• research exploring loudspeaker directivity with regard to perception is limited;

• although some trends have been discovered, there are many unclear or contrary results, and certainly no large-scale investigations with significantly conclusive data (the most notable here is Bech’s (1994b) experiment which shows some interaction between directivity and the room with regard to the perception of fidelity); and

• current opinion suggests that wide-directivity loudspeakers are most preferred in
<table>
<thead>
<tr>
<th>Author(s) (Year)</th>
<th>Loudspeaker / Directivities</th>
<th>Environment</th>
<th>Loudspeaker Position</th>
<th>Listener Position</th>
<th>Programme Material</th>
<th>Attribute(s) Measured</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bauer (1960)</td>
<td>Conventional Dipole</td>
<td>Unclear</td>
<td>Conventional (angled towards listener)</td>
<td>Varied</td>
<td>Unclear</td>
<td>Area of Phantom Image</td>
<td>Area of perception of phantom image widened with dipole</td>
</tr>
<tr>
<td>Linkwitz (2007)</td>
<td>Dipole Omni-directional</td>
<td>Long Living / Listening Room</td>
<td>Conventional</td>
<td>Sweet Spot</td>
<td>Varied</td>
<td>General Timbral and Spatial</td>
<td>Dipole and Omni sound identical in reflective environment if reflections are sufficiently delayed.</td>
</tr>
<tr>
<td>Flindell et al. (1991)</td>
<td>Simulated Directivities (Directional to Omni-directional)</td>
<td>Anechoic Chamber (Standard listening room simulated)</td>
<td>Simulated to be conventional</td>
<td>Sweet Spot</td>
<td>Jazz, Pop, Speech, Classical</td>
<td>Stereo image quality, Spaciousness, Overall Impression</td>
<td>No clear consensus, although a tendency for naive listeners to prefer omni.</td>
</tr>
<tr>
<td>Choisel (2005)</td>
<td>2 Conventional with reflections</td>
<td>Listening Room</td>
<td>Conventional</td>
<td>Sweet Spot</td>
<td>Female Speech, Hand Claps</td>
<td>Sound source location</td>
<td>Perceived direction not seen to change</td>
</tr>
<tr>
<td>Zacharov * (1998)</td>
<td>Cardioid Horizontal-Line Conventional</td>
<td>Listening Room</td>
<td>Conventional</td>
<td>Sweet Spot and Left of Sweet Spot</td>
<td>Commentary, Television, Film</td>
<td>Coordination with Image, Spaciousness Effects, Naturalness</td>
<td>Only marginal differences between systems.</td>
</tr>
<tr>
<td>Bech** (1994)</td>
<td>Dipole Conventional Line Source Omnidirectional</td>
<td>3 Different sized rooms</td>
<td>3 Various</td>
<td>Central Position in Room</td>
<td>Anechoic Speech, Guitar, Brass Section, Orchestra</td>
<td>Fidelity Rating</td>
<td>No striking differences, Dipole rated as poor at centre of back wall</td>
</tr>
<tr>
<td>Toole (2008)</td>
<td>Conventional Dipole</td>
<td>Listening Room</td>
<td>Conventional</td>
<td>Sweet Spot</td>
<td>Choral, Chamber, Jazz, Pop</td>
<td>Sound Quality, Spatial Quality</td>
<td>Differences were evident when loudspeakers compared in mono. In stereo, Conventional loudspeakers which radiate widely preferred although not exclusively.</td>
</tr>
</tbody>
</table>

Table 3.1: A summary of previous research regarding loudspeaker directivity and its effect on perception of reproduction.
casual listening.

3.3.3 Section Summary and Discussion

Section Summary

Here, an overview of research that has been conducted in the area of directivity and perception has been presented. There exists much hypothetical research, proposing the optimal conditions of directivity for increased quality of reproduction, yet information based upon well-structured scientific tests seems limited. Of all experiments carried out, most give indication of the detailed relationship between changes in directivity and the attributes affected, with a trend apparent that wide horizontal directivity loudspeakers may be preferred for casual listening in some situations. However, it was also found that there was an interaction between the loudspeaker directivity and the properties of the room.

Discussion

The outcomes of this section give a clear indication that more specific research is required. Ideally, it should investigate the type and magnitude of perceived auditory changes as a result of variation in directivity/boundary characteristics.

The key attributes that should (according to theory considered so far) be affected by a change in reflections have been identified and it now remains to confirm that they are indeed affected and to measure by how much. The likely changes in timbral fidelity and apparent source width are expected to be subtle as the reflections in typical rooms are likely to be low in level and of short delay, which would cause small changes to the perception of the image and it has been reported that timbral changes caused by small domestic rooms are perceptually insignificant. Thus, the test approach should be such that small differences in perception can be measured adequately.

Therefore a series of specifically-designed scientific experiments (with appropriate variables and sensitivity) should be carried out, in order to answer initial research questions R3 and R4.

At this stage, the following points need to be considered in future experimental work:
• are the attributes identified here definitely affected, and by how much;
• are any additional attributes affected;
• are changes in the attributes related to physical factors as identified (i.e. vertical reflections and timbre);
• what is the overall contribution of these effects to the listening experience?

3.4 Chapter Summary

Having introduced the concept of loudspeaker directivity and its relationship with the sound field in a listening environment, it is necessary to identify the perceptual attributes that are likely to be affected by this relationship.

By linking the physical changes caused by a variation in directivity/boundary characteristics to auditory attributes that are known to be affected by such changes, it was shown that theoretically, for changes in level, time, number and angle of arriving reflections, the perception of timbre, loudness, localisation and apparent source width could be affected.

A brief consideration of the importance of these changes was included - suggesting that despite our ability, as humans, to adapt to any reproduction within a sound field, it is likely that for particular listening situations, listeners would show preferences with regards to the characteristics of reflections accompanying the direct sound of a reproducing source.

Finally, an assessment of other research in this field shows that detailed investigations into this relationship are scarce, and so conducting experiments which look closely at changes caused by changes in directivity/boundary characteristics and their specific contribution to the perception of reproduction, is justified.
CHAPTER 3. PERCEPTION OF REPRODUCED SOUND

CHAPTER CONCLUSIONS

What are the basic auditory mechanisms involved in human sound perception?

Sound causes the hair cells of the cochlea to move. This motion is converted into a neural signal that can be analysed by the brain. Initially, sounds are analysed and grouped according to similarities in spectral and temporal characteristics, before being referred to memory so that they can be recognised.

How do humans perceive reproduced sound?

The mechanisms involved in the perception of reproduced sound are the same as those involved in the perception of natural sound - spectral, temporal and level characteristics can be associated with a number of audible attributes which allow humans to distinguish different types of sound (despite similarities in pitch or content). However, with reproduction of audio material such as speech or music, the exact same source material can be replayed which means that any differences in the reproduction system which affect the way the sound is radiated can be compared. In this case, loudspeakers are of interest and it would be expected that, due to the effect of different components, sizes and arrangements (which cause different spectral and temporal responses), differences in the sound reproduced by different loudspeakers will be perceived and could be labelled using audio attributes.

How is the perception of reproduced sound affected by the room?

Rooms cause the sound reproduced from loudspeakers to be reflected. Therefore, in addition to sound that travels directly between the loudspeaker and the listener, indirect reflected sound arrives at the listener. The interaction of direct sound and delayed versions of itself can cause ‘comb filtering’, where particular frequencies (based on the wave paths) constructively or destructively interfere. This affects the spectral nature of the sound arriving at a listener, and can therefore affect the perception of timbre, with boosts and dips occurring at various frequencies. Room modes also affect perceived timbre and it is thought that timbral nuances of sounds are easier to detect with
additional reflections, due to the repetition. More specifically, reflections in the vertical plane of a room have been linked with timbral changes, whereas lateral reflections are associated with a broadening in apparent source width. Thus, a room can affect the perception of reproduced sound by changing its timbral and spatial properties at a listener position.

What is the influence of directivity and boundary characteristics upon the perception of reproduced sound in rooms?

A change in loudspeaker directivity and/or boundary characteristics would change the way that the reproduced sound interacts with the room before arriving at the listener. The level and spectrum of reflections would be affected, thus affecting the perception of the reproduction. Increased boundary absorption, along with reduced directivity coverage (so that it is narrowed towards the listener) would be expected to result in a much reduced ‘room imprint’ in the reproduction compared to reflective boundaries and wide-coverage directivity loudspeakers. The corresponding levels of timbral and spatial changes would also be expected to be increased for the latter case.

What changes to directivity and boundary characteristics are perceivable, what might the relative perceived magnitude of the changes be, and which perceptual attributes would be affected?

Reflections in domestic rooms are typically low in level with short delays. According to research investigating the perception of single reflections, this would suggest that unless extreme combinations of directivity and boundary characteristics are compared (reflective and wide-directivity vs near-anechoic and narrow), the differences in perceived attributes will be subtle. Additionally, consideration of adaptation would suggest that such differences would only be apparent following immediate comparison.

The types of attributes expected to be affected include: timbre, apparent source width (ASW), loudness, localisation, envelopment/spaciousness, source distance.

What part does listener position play in the relationship between loudspeaker directivity, boundary characteristics and perception of reproduction?
The position of a listener, relative to the loudspeaker, has immediate consequences upon the filtering that takes place (according to the HRTF) and on the level of the signal (and therefore perceived loudness). Perception of changes to the sound field caused by changes in loudspeaker directivity and boundary conditions is dependent on where the listener is - as an example, changes in timbre due to comb filtering are position dependent. Many of the auditory phenomena discussed in this chapter are dependent on the interaction between direct sound and reflections; the time-of-arrival difference between direct sound and reflections is a key factor and is in part determined by the listener position.
Chapter 4

Preliminary Studies

Existing research into loudspeaker directivity and its relationship with the perception of reproduced sound is limited. A study of acoustic/psychoacoustic theory suggests that a number of auditory attributes should be affected. However, most research in this area has failed to validate whether these attributes are affected, the magnitude of effect, and how the effects contribute to the overall listening experience.

Before defining the experimental route that will be undertaken in order to answer the research questions in detail, it is useful to consider a typical listening situation and to measure the effects that reproduction over loudspeakers with different directivities has upon the listening situation. By carrying out such experiments, it is possible to verify that the attributes identified previously, which are expected to be affected by directivity, are affected in real environments. Tests have been carried out in two locations, Surrey (UK) and Struer (Denmark) so as to provide data from different environments - different listeners, loudspeakers and listening test environments were used at each location. An alternative listening position was also included as a factor in the tests in Struer, so that the effect of listener position could be considered. The attributes affected when comparing loudspeakers with different directivity types were elicited before a selection of related attributes were presented to listeners for rating. This provides an initial insight into the types and magnitudes of effects that different loudspeaker directivity types have upon domestic listening.

Tests carried out at this stage are also exploratory and may help to improve future experimental work; additional attributes may be elicited which the literature does not
highlight, the nature of the test and the test method can be assessed and the general
direction of research can be re-considered in light of any results found.

4.1 Experiment 1: Listening Tests in Surrey, UK

4.1.1 Introduction/Aims

This experiment took part in two stages. The first was an elicitation test, where listeners
were asked to identify which attributes were affected when comparing the reproduction
of various audio stimuli via loudspeaker pairs exhibiting different directivities. Using
results from this stage and information from the literature, a set of test attributes was
compiled. In the second stage, the listeners were asked to rate these attributes for each
of the different loudspeaker pairs so that some quantitative data regarding how they
were each perceived was available for analysis.

Aims:

- Identify which attributes are perceived to be affected when comparing the
  reproduction of loudspeakers with different directivities in a listening room;

- Compare the types of attributes elicited with the attributes that were expected to
  be affected according to the literature;

- Based upon these findings, compile a group of test attributes to be rated by
  listeners (thus providing quantitative data with regards to the perception of
  reproduction via different loudspeakers).

4.1.2 Experimental Set-up

The tests were carried out in the listening room (TB07) at the University of Surrey,
which conforms to ITU-R BS 1116 (1997). Several Medium Density Fibre (MDF) panels
were positioned on the side walls of the listening room, in order to emphasize lateral
reflections. It was felt that this would more appropriately mimic a typical domestic
listening area. The test loudspeakers were set-up to form an equilateral triangle with
a single, central listening position, conforming to the ITU-R BS 775-1 (2006) standard
for 2-channel stereophonic reproduction. A diagram of the experimental test set-up is presented in Figure 4.1.

For the passive loudspeaker pairs, remotely located Quad QD4240 amplifiers were used for amplification. The left and right loudspeaker of each pair were level matched to within ±0.25 dBA of each other when reproducing broadband pink noise in the listening room. The sound pressure level was measured at the listening position using an NTI Acoustilyzer AL1 digital sound level meter (with NTI MiniSPL microphone). The level of each pair was then matched to within ±0.25dBA, again using broadband pink noise. A small listening panel (comprising 3 listeners) was then used to adjust the playback level of all audio excerpts to be of the same perceived loudness, and of a comfortable listening level. This was done via a test patch in Max/MSP with all final gain values

Figure 4.1: Experimental Set Up in Surrey (not to scale)
noted and fixed for the following tests.

It should be noted that the loudspeakers were not equalised to have the same frequency response - in order to compensate for the irregularities present, high Q filters would be required which can cause a detrimental effect to performance in the time-domain. It was felt that equalising the frequency response would trade one set of deficiencies (the different on-axis frequency responses) for another (the different time-domain side-effects of the equalising filters) and hence it was decided not to employ any equalisation.

Loudspeaker Selection

Three loudspeaker types, with different loudspeaker directivity characteristics, were chosen to be used:

- **Loudspeaker 1** (*Genelec 1032a*) - selected to represent a typical 2-way loudspeaker, with a directivity that narrows on-axis at higher frequencies;

- **Loudspeaker 2** (*Bang & Olufsen Beolab 3*) - selected to represent a loudspeaker which does not narrow on-axis at high frequencies (it utilises the 'acoustic lens' technology, to re-radiate sound to a wider range of off-axis angles, discussed in Section 2.1.4);

- **Loudspeaker 3** (*Canon S-35*) - selected to represent a loudspeaker with nominally 'broad' horizontal directivity. This model had been specifically designed to have a wide-angle horizontal directivity coverage.

One of each loudspeaker type was measured at the ‘Cube’ facility at Bang & Olufsen, Struer, Denmark. A microphone (positioned at 3m from the loudspeaker) was used to measure the impulse response at 2° angular increments (between 0-360° in the horizontal and vertical plane). A Maximum Length Sequence (MLS) was used to obtain the impulse response, with the measurement system set to average 5 measurements at each angular increment. The sampling rate was 48000 Hz.

Contour plots showing directivity in the horizontal plane are included (Figures 4.2 - 4.4). Vertical directivity contour plots are given in the Appendix (Section B.1). As defined previously, the plots indicate the measured change in sound pressure level as a function of frequency and angle, with each contour line representing the level in dB relative to the on-axis level. The difference between each contour line is 1dB.
Following observation of the contour plots for each loudspeaker type, it can be confirmed that each loudspeaker type demonstrates different horizontal and vertical directivity characteristics. Observing horizontal directivity, the narrowing nature of Loudspeaker 1 (Genelec 1032a) (focussing of directivity towards on-axis with increased frequency) is clear, and it appears that, in fact, Loudspeaker 2 has a more consistent off-axis directivity than Loudspeaker 3, which shows more interaction effects at higher frequencies (identified by the areas of increased and decreased output level, due to summation and cancellation of radiated sound waves).

Audio Programme Material Selection

Three excerpts of audio material were chosen. A small listening panel auditioned 100 audio loops over each reproduction system (loudspeaker pair), and shortlisted 10 loops which were perceived to have clearly identifiable timbral and spatial features. This was done using an ‘Audio Library’ patch, which allows loops in Max/MSP to be reproduced over each loudspeaker pair. Each panel member was able to listen to each loop being
reproduced by the different loudspeaker pairs, and subsequently was able to choose
the audio excerpts which varied most significantly when reproduced over the different
loudspeakers. A smaller listening panel was used to choose the final 3 audio excerpts
from the shortlisted 10, with consideration to what may be appropriate test material;
the criterion in this case was for the test material to emphasize the greatest differences
between systems in either spatial or timbral characteristics.

The three excerpts are all classed as ‘popular’ music, including a range of acoustic
and electrical instrumentation as well as vocals (on Audio 2 and 3). Audio 2 was
monophonic, as demonstrated by the M and S plots (which indicate inter-channel
differences) presented in Figure 4.6. The track details are as follows:

- Audio 1 - Crips - Ratatat (Ratatat, 2004 XL Recordings);
- Audio 2 - Love me do - Beatles (1, 2000 Paralphone);
- Audio 3 - Are you gonna be my girl - Jet (Get Born, 2003 Elektra).
CHAPTER 4. PRELIMINARY STUDIES

Figure 4.4: Loudspeaker 3 (Surrey) Horizontal Directivity Contour Plot

Figure 4.5: On-Axis frequency response of loudspeakers used in Experiment 1 (Surrey). Normalised to maximum value.
CHAPTER 4. PRELIMINARY STUDIES

4.1.3 Stage 1: Elicitation Experiment

Method

The listening panel at Surrey comprised 10 sound recording students, considered to be expert listeners on account of their regular exposure to and interaction with high quality audio, and training in critical listening. The test equipment was as described above.

Each listener was asked to take part in a paired-comparison assessment, whereby each loudspeaker pair was compared with each other, with each reproducing the same piece of audio programme material. For each comparison, the listener was asked to describe (in writing) up to three differences, and to rank the contribution of each difference to their overall perception. Three differences were requested, as it was expected that more than one attribute would have been perceived to change and so this allowed for additional description. An open-elicitation may have caused longer test times and complicated analysis and so three differences was chosen as a sensible limit. This was repeated for all
three pieces of audio used in the rating tests. A test interface was designed in Max/MSP, which selected, at random, two loudspeakers and an audio excerpt to be compared. Each listener was able to switch back-and-forth between each loudspeaker type (of the two pairs of loudspeakers chosen for that comparison), whilst they noted their judgements. The listener was seated behind an acoustically transparent, visually opaque curtain so that there was no bias as a result of seeing the loudspeakers.

The interface used is presented in Figure A.1 (Appendix), and an example of the assessment form used in Figure A.2 (Appendix).

**Results**

Initially, the differences described by listeners were collated and grouped according to similar terms. The terms used were as follows:


The three most frequently-reported differences between each pair of loudspeaker types are shown in Figure 4.7 and it is evident that the perceived width ('Narrower') is the attribute that was most often affected. Other spatial attributes such as 'Closeness' and 'Centrality' were perceived to change, as well as timbral aspects: 'More LF', 'Less LF'.

**4.1.4 Choosing the Test Attributes**

Most of the variables listed above can be condensed into a set of categories (see Table 4.1).
CHAPTER 4. PRELIMINARY STUDIES

Narrower Less LF
Closer
Other Attributes

GENELEC 1032A vs Beolab 3

More central Narrower
Clearer
Other Attributes

GENELEC 1032A vs Canon S-35

More LF Narrower
More central
Other Attributes

Beolab 3 vs Canon S-35

Figure 4.7: Specific terms used to describe differences between reproduction systems at Surrey (terms relate to order of reproduction systems - for example, Genelec 1032A has Less LF than Beolab 3)

<table>
<thead>
<tr>
<th>Categories</th>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timbral</td>
<td>More/less LF, more/less HF, brighter/duller</td>
</tr>
<tr>
<td>Spatial</td>
<td>More/less LF, more/less HF, brighter/duller</td>
</tr>
<tr>
<td>Localisation</td>
<td>Lower/higher, right/left skew</td>
</tr>
<tr>
<td>Width</td>
<td>Narrower/wider</td>
</tr>
<tr>
<td>Envelopment</td>
<td>More central/ambient</td>
</tr>
<tr>
<td>Spatial</td>
<td>More central/ambient</td>
</tr>
<tr>
<td>Distance</td>
<td>Closer/further</td>
</tr>
<tr>
<td>Focus</td>
<td>Focused/diffused</td>
</tr>
<tr>
<td>Others*</td>
<td>Louder/quieter, phase issues, dynamic/less dynamic</td>
</tr>
<tr>
<td></td>
<td>dynamic, lighter/fuller, clearer/muddier</td>
</tr>
</tbody>
</table>

*attributes that are either affected by timbral and spatial features or that are not affected by timbral or spatial features at all

Table 4.1: Categorised terms from elicitation experiment
When observing the overall use of these terms in Surrey, it is clear that the majority are timbral or spatial (see Figure 4.8). These attributes also correspond to the attributes expected to be affected (from the previous chapter):

- Timbre;
- Apparent source width (ASW);
- Loudness;
- Localisation;
- Envelopment / Spaciousness;
- Source distance.
It is apparent from Figure 4.8 that the most commonly elicited categories of terms were Width and Timbre. Hence, these were selected for use in a grading experiment. Perceived width as an attribute may be too vague and thus a more specific width-related attribute should be used - *Ensemble Width* follows from Rumsey's (2002) definition and describes the perceived width of all instruments/sources within a stereophonic image.

It would also be appropriate to choose an attribute that is not based directly upon spatial or timbral changes. Based upon the elicitation, loudness would be a suitable candidate as it was expected to change and featured in the elicitation (although not as prominently as the others) and thus will be used for further investigation.

Finally, testing a hedonic attribute might give some indication as to the relative importance of the differences between loudspeakers. The attribute 'liking' provides an indication as to whether one loudspeaker type is preferred over another. If the results suggest that there is no difference in perceived liking, despite perceived changes in other attributes, then it could be inferred that the differences in reproduction do not contribute towards enjoyment. This would indicate, to some extent, how important these levels of physical changes are to the listening experience.

The final attributes to be used for the rating tests are therefore:

- Ensemble width;
- Timbral fidelity;
- Loudness;
- Liking.

### 4.1.5 Stage 2: Rating Experiment

**Method**

Following the attribute elicitation, each member of the listening panel was invited to take part in a rating test, in order to quantify the perceived differences in attributes as a result of reproduction over the different loudspeakers. In total, 7 of the 10 listeners from the panel took part.

For each of the chosen attributes (ensemble width, timbral fidelity, loudness, liking), each listener was presented with an interface which comprised three scales (each one relevant
to a loudspeaker pair. The interface was designed in Max/MSP, and it randomised the loudspeaker-scale combination that was presented to each listener. The listener was asked to provide scores for the reproduction of the different audio material over each loudspeaker pair. The order of programme material reproduction was also randomised. The test interfaces and listening test instructions given to the listeners are shown in Figures A.3-A.8. It was assumed that the meanings of the test attributes were known by the listeners, although some context was provided in order to assist their judgements.

Each listener was provided with a guide sheet which was intended to help them become more familiar with the attributes that they were asked to rate. Each listener was guided to rate according to the 'highest timbral fidelity imaginable'. The sheet is shown in Figures A.7 and A.8 (Appendix).

The interface used a 100-point scale scoring system (180-points for ensemble width, as listeners were asked to give angle of perceived width in degrees). Each listener was able to compare the reproduction of the given audio excerpt between loudspeaker pairs by using a control mechanism built in to the interface patch which changed which loudspeaker pair the audio signal was routed to. At the same time, they were able to adjust the scores on the scales, which corresponded to the loudspeaker pair that they were hearing. Once all ratings had been given for a particular audio excerpt, the interface was refreshed, with a new audio excerpt, and a different loudspeaker order presented. Each audio excerpt was rated twice by each listener. This allowed assessment of listener consistency (the ability of each listener to reproduce scores for the same stimuli could be checked). This led to a total of 72 ratings per listener (3 loudspeaker pairs x 3 audio excerpts x 2 repeats x 4 attributes = 72). Details of the experimental factors, and their levels, are presented in Table 4.2.

4.1.6 Listener assessment and screening

Before carrying out statistical analysis upon the listeners' scores, the performance of the listeners themselves should be considered. If a listener was unable to provide consistent scores, or if they gave unusual responses to the stimuli, then their results should be assessed separately from the main group, or removed from analysis altogether.
<table>
<thead>
<tr>
<th>Experimental Factor</th>
<th>Level</th>
<th>Description</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loudspeaker Type</td>
<td>1</td>
<td>Genelec 1032a <em>(Loudspeaker 1)</em></td>
<td>Reference 2-way active studio monitor with narrowing directivity at higher frequencies (See Figure 4.2)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>B&amp;O Beolab 3 <em>(Loudspeaker 2)</em></td>
<td>Hi-Fi Loudspeaker with extended horizontal directivity above 2kHz (See Figure 4.3)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Canon S-35 <em>(Loudspeaker 3)</em></td>
<td>'Wide Imaging' Loudspeaker with extended horizontal directivity above 2kHz (See Figure 4.4)</td>
</tr>
<tr>
<td>Programme Material</td>
<td>1</td>
<td>Ratatat <em>(Audio 1)</em></td>
<td>'Crips' - <em>Ratatat</em> (2004, XL) Duration = 19 seconds Correlation $R_{LR} = 0.88016$</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Beatles <em>(Audio 2)</em></td>
<td>'Love me do' - <em>1</em> (2000, Paraphone) Duration = 7 seconds Correlation $R_{LR} = 1$</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Jet <em>(Audio 3)</em></td>
<td>'Are you gonna be my girl' - <em>Get Born</em> (2003, Elektra) Duration = 21 seconds Correlation $R_{LR} = 0.8463$</td>
</tr>
<tr>
<td>Listener</td>
<td>1-7</td>
<td>-</td>
<td>7 expert listeners (all Tonmeister / Institute of Sound Recording postgraduates)</td>
</tr>
<tr>
<td>Repetition</td>
<td>1/2</td>
<td>where 1=Original Score 2=Repeat Score</td>
<td>Each rating-based listening test was carried out twice for each loudspeaker/programme combination</td>
</tr>
</tbody>
</table>

Table 4.2: Experimental Factors (Surrey)
Criteria

Listener scores were assessed based upon the following criteria:

- **Intra-Listener Consistency** - *the ability of each individual listener to reproduce consistent scores in repeat trials*. Data from listeners with more than 15% root-mean-square difference, \( \text{diff}_{\text{rms}} \), (for repeat trial scores) for two or more attributes are removed from the analysis. Previously reported intra-listener difference in listening tests is shown to be 8-12\% (Conetta et al. (2008); George et al. (2008); Rumsey (1998)) and so given that the task involving ensemble width was difficult (listeners had to select angular ranges with no visual indication), a level of 15\% was chosen. The requirement for this level to be reached on two or more attributes was chosen so that if a listener had inconsistent scores for one attribute only, yet satisfactory in others, their data would not be removed.

- **Inter-Listener Consistency** - *the amount of consistency that one listener's scores has with the group as a whole*. If one or more listeners show regular inconsistency with the rest of the group, then the case should be explored, and potentially the scores of those listeners analysed separately.

**Intra-Listener Consistency**

Using scores from repeat trials, the root-mean-square difference, \( \text{diff}_{\text{rms}} \), was calculated for each listener and is shown in Figure 4.9.

**Criterion**: Listeners with above 15\% difference for two or more attributes are screened.

**Outcome**: All results from Listener 2 screened.

**Inter-Listener Consistency**

The correlation between scores for each listener and the average of the remaining listener scores was calculated in order to observe the sample consensus.

Scatter plots (for example see Figure 4.10) of each listener score and bar charts of calculated correlation coefficient (for example see Figure 4.11) were plotted for each
Figure 4.9: Intra-Listener RMS Difference (the ability of each individual listener to reproduce consistent scores in repeat trials) for the experiment in Surrey. In this case, Listener 2 shows greater than 15% root-mean-square difference for more than two attributes (circled), and as a result will have their data removed from analysis.
Figure 4.10: Inter-Listener Correlation - This example is for ‘Liking’ scores. Each plot shows the listener score (y-axis), against the average of the remaining listener scores (x-axis). The correlation coefficient for each listener is included.

**Attribute.**

**Criterion:** Listeners whose scores for all cases are not significantly correlated to average scores are considered for removal.

**Outcome:** 26 out of 28 cases showed significant correlation and thus no scores were removed from the analysis upon this basis.

**Assessment Summary**

- All scores from Listener 2 were removed from the analysis on account of intra-listener inconsistency.
- All remaining listeners performed sufficiently and show a consensus of opinion in
Figure 4.11: Inter-Listener Correlation Coefficients - This example shows a corresponding bar chart of the correlation coefficients in Figure 4.10 for 'Liking' scores.

scoring, according to measured inter-listener correlation.

4.1.7 Results

ANOVA (Analysis of Variance) is a method of data analysis which can be used in order to determine whether different levels of an independent variable (such as loudspeaker type) cause a statistically significant difference in a dependent variable (such as sound quality, rated between 0-100), compared to the random variation that exists due to uncontrolled variables (such as the listeners themselves, the test set-up or the environment). If the independent variable is deemed to be a statistically significant factor, it means that one or more of the variable levels contribute to a difference in the dependent variable that would be expected to occur for the majority of the population sample (as opposed to occurring randomly or due to uncontrolled factors). If an experimental factor is not significant, then an experimenter cannot confidently claim that the different levels of that factor cause a difference in the independent variable.

ANOVA can be used with multiple independent and dependent variables, and the interaction of different variables can be explored. This allows a user to determine whether
factors such as listener, repeat, position etc cause significant differences to the ratings of the dependent variables and also whether their interactions with other independent variables cause significant differences.

In this case, ANOVA is used to identify whether or not each of the independent test variables (loudspeaker, recording, listener) causes a difference to the ratings of the dependent test variables (liking, loudness, timbral fidelity, ensemble width). Once statistically significant differences have been established, the relationships between the different levels of factors can be discussed with greater confidence.

Procedure

Before the analysis was carried out, results were assessed and found to normally distributed and hence the assumptions for ANOVA are satisfied. A univariate analysis of variance (Univariate ANOVA) model was used in order to test the relative statistical significance of each experimental factor. For each attribute at each location, a 'customised' analysis is initially performed in order to include 'repetition' as a variable. If its interaction with other variables is not seen to be significant, then it can be assumed that intra-listener differences in scoring are negligible. If that is the case, a standard full-factorial analysis is carried out - with all combinations of independent test variables and their effect upon the dependent variables considered. Finally, if any experimental factors in the full-factorial analysis are found to be non-statistically significant, a 'reduced' analysis is conducted. This is because it improves the resolution of the tests (Bech and Zacharov 2006).

The analysis procedures and their results are detailed in the Appendix, Sections A.3 and B.2, respectively.

Results Summary

A summary of the significant factors from the 'reduced' UniANOVA model (or from the 'full factorial' model if this could not be reduced) of each attribute at each location is presented in Table 4.3. Further effects are presented in Table 4.4.

The effects of the significant factors are explored, by attribute, in the following sections. Where loudspeaker and/or recording are significant factors, data relating to these factors are plotted here. Data relating to all factors are plotted in Appendix B, Section B.3.
### Table 4.3: Significant Factors (Surrey) - this table shows the Degrees of Freedom, F-Ratio, Significance, Eta-Squared and Observed Power values for all significant factors and interactions in the preliminary study at Surrey.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Sig.Factors/Interactions</th>
<th>df</th>
<th>$F$</th>
<th>Sig.</th>
<th>$\eta^2$</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Liking</strong></td>
<td>Loudspeaker</td>
<td>2</td>
<td>13.952</td>
<td>.001</td>
<td>.736</td>
<td>.985</td>
</tr>
<tr>
<td></td>
<td>Loudspeaker vs Listener</td>
<td>10</td>
<td>20.117</td>
<td>.000</td>
<td>.691</td>
<td>1.000</td>
</tr>
<tr>
<td><strong>Loudness</strong></td>
<td>Loudspeaker</td>
<td>2</td>
<td>17.297</td>
<td>.001</td>
<td>.776</td>
<td>.996</td>
</tr>
<tr>
<td></td>
<td>Listener</td>
<td>5</td>
<td>4.754</td>
<td>.008</td>
<td>.574</td>
<td>.903</td>
</tr>
<tr>
<td></td>
<td>Loudspeaker vs Listener</td>
<td>10</td>
<td>8.653</td>
<td>.000</td>
<td>.539</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Recording vs Listener</td>
<td>10</td>
<td>6.343</td>
<td>.000</td>
<td>.462</td>
<td>1.000</td>
</tr>
<tr>
<td><strong>Timbral Fidelity</strong></td>
<td>Loudspeaker</td>
<td>2</td>
<td>19.916</td>
<td>.000</td>
<td>.799</td>
<td>.999</td>
</tr>
<tr>
<td></td>
<td>Listener</td>
<td>5</td>
<td>3.527</td>
<td>.030</td>
<td>.570</td>
<td>.763</td>
</tr>
<tr>
<td></td>
<td>Loudspeaker vs Recording</td>
<td>4</td>
<td>3.540</td>
<td>.011</td>
<td>.161</td>
<td>.846</td>
</tr>
<tr>
<td></td>
<td>Loudspeaker vs Listener</td>
<td>10</td>
<td>9.736</td>
<td>.000</td>
<td>.568</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Recording vs Listener</td>
<td>10</td>
<td>3.046</td>
<td>.003</td>
<td>.292</td>
<td>.970</td>
</tr>
<tr>
<td><strong>Ensemble Width</strong></td>
<td>Loudspeaker</td>
<td>2</td>
<td>16.249</td>
<td>.001</td>
<td>.765</td>
<td>.994</td>
</tr>
<tr>
<td></td>
<td>Recording</td>
<td>2</td>
<td>46.342</td>
<td>.000</td>
<td>.903</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Listener</td>
<td>5</td>
<td>20.422</td>
<td>.000</td>
<td>.936</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Loudspeaker vs Recording</td>
<td>4</td>
<td>2.628</td>
<td>.041</td>
<td>.124</td>
<td>.709</td>
</tr>
<tr>
<td></td>
<td>Loudspeaker vs Listener</td>
<td>10</td>
<td>2.366</td>
<td>.017</td>
<td>.242</td>
<td>.907</td>
</tr>
</tbody>
</table>
Table 4.4: Further Effects (Surrey) - this table shows the significance of interactions between factor types following post-hoc Bonferroni tests.

Liking

Loudspeaker pairs 1 and 2 (Genelec 1032a and Beolab 3s) were rated as being similarly liked (see Figure 4.12) (yet, ANOVA suggests a significant difference is measured). Loudspeaker 3 (Canon S-35) was rated lower in general, with two listeners disliking them strongly.

Loudness

One listener's scores were positioned much lower on the scale. Overall, general scoring suggests that the Beolab 3 was perceived as louder than the Genelec 1032a, and the Canon S-35 was perceived as being much quieter than both (see Figure 4.13). Some listeners perceived differences in the loudness of the programme material.
CHAPTER 4. PRELIMINARY STUDIES

Figure 4.12: Liking(Surrey): Loudspeakers

Figure 4.13: Loudness(Surrey): Loudspeakers
In general, the timbral fidelity of the Canon S-35s was rated lowest. One listener rated it as very low, which may have affected overall results, by pulling the average rating down even further. The Genelec 1032a and Beolab 3s were rated as more similar, although, trends suggest that the Beolab 3s were perceived to have slightly lower timbral fidelity (a significant difference is measured in ANOVA). Also, the rated timbral fidelity of the Genelec 1032a and Beolab 3s appears to be dependent upon the audio programme material (see Figure 4.14), but this is not reinforced with data from ANOVA.

Ensemble Width

No significant difference in ensemble width was perceived between the Beolab 3s and Genelec 1032as, whereas the Canon S-35s were perceived as having a greater width. The mono recording (Audio 2 Beatles) was perceived as being narrower in ensemble width, as expected, and also the range of perceived widths varied between listeners, and as such, was found to be a significant factor (see Figure 4.15). It should be noted that the positioning of the loudspeakers relative to one another may have contributed to the
difference in perceived width - this is a disadvantage of using multiple loudspeakers for comparison in listening tests.

4.1.8 Summary/Discussion

7 listeners were asked to judge the four chosen attributes (ensemble width, timbral fidelity, loudness, liking). One listener’s results were removed from the analysis based on intra-listener consistencies. Loudspeaker 3 (Canon S-35, wide-horizontal directivity) was rated lower than the others for liking, loudness and timbral fidelity. Loudspeakers 1 and 2 were rated as being similar for these attributes. For ensemble width, Loudspeakers 1 and 2 were rated as similar (despite LS1 having a narrow horizontal directivity and LS2 having a wide horizontal directivity) and Loudspeaker 3 was perceived to give wider reproductions than both.

Liking is often linked to loudness, with louder stimuli being rated higher (Bech and Zacharov 2006). As the loudspeakers were not equalised in spectrum, differences in the response of Loudspeaker 3, which demonstrated a reduced low frequency response and increased high frequency response compared to the others (see Figure 4.5), may have contributed to the reduced perceived loudness; a loudspeaker with increased low
frequency response would be perceived to be louder than one with increased high frequency response, providing it crossed a greater percentage of the equal loudness contours (Moore 2002) in that region. Despite this, width was perceived as much greater, and so some aspect of Loudspeaker 3's radiation did contribute to an increase in perceived width. Why a corresponding increase in width was not perceived with Loudspeaker 2, which demonstrates similar characteristics in directivity pattern, is unknown at this point.

4.2 Experiment 2: Listening Tests in Struer, Denmark

4.2.1 Introduction/Aims

This experiment followed the same format as the experiment conducted in Surrey but with two fixed listening positions for the rating test (so that the effect of listener position on ratings could be considered) and an additional small-scale informal free-roaming test after the rating test. The small-scale test involved listeners choosing their preferred reproduction system when asked to evaluate the audio at all positions in the room - this was carried out to further explore the importance of listener position and explore how non-fixed-position critical listening affected results. This experiment also included one loudspeaker which housed both left and right channels in a single unit, rather than two separate units as with all others involved.

Aims:

- Identify which attributes are perceived to be affected when comparing the reproduction of loudspeakers with different directivities in a listening room;
- Compare the types of attributes elicited with the attributes that were expected to be affected according to the literature;
- Based upon these findings, compile a group of test attributes to be rated by listeners (thus providing quantitative data with regards to the perception of reproduction via different loudspeakers);
- Investigate the effect of listener position upon the perception of these attributes by considering a second fixed position and a small-scale 'roaming' evaluation, where each listener moves around the room to listen.
4.2.2 Experimental Set-up

The tests here were carried out in a listening room at Bang&Olufsen, Denmark, which conforms to IEC 268-13 (1987).

As in Surrey, the test loudspeakers were set up to form an equilateral triangle with a single, central listening position, conforming to the ITU-R BS 775-1 (2006) standard for 2-channel stereophonic reproduction. Loudspeaker 2, which was a single-unit stereo loudspeaker (see description below), was positioned directly in front of the listener, with its loudspeaker drivers forming a narrower triangle. MDF panels were not used in this case, on account of the reflective nature of the side walls within the room. A diagram of the experimental test set-up is presented in Figure 4.16. The second listening position was defined to the rear-left of the central listening position.

The left and right loudspeaker of each pair were level matched to within ±0.25 dB of each other using broadband pink noise at Position 1 in the listening room. The level of each pair was then matched to within ±0.25 dB, again using broadband pink noise in Position 1 in the listening room. A small listening panel was then used to adjust the playback level of all audio excerpts to be of the same loudness and of a comfortable listening level.

The loudspeakers in Denmark were also not spectrally-equalised, for the same reasons as in Surrey.

Loudspeaker Selection

Three loudspeaker types with different loudspeaker directivity characteristics were chosen. Two of the loudspeakers used (Genelec 1031a and Beolab 3) were similar to those used in Surrey (to allow potential comparison) and the other (Array) was different (to increase the range of tested directivity types):

- **Loudspeaker 1 (Genelec 1031a)** - selected to represent a typical 2-way loudspeaker, with a directivity that narrows on-axis at higher frequencies. (This is similar to the 1032a, but smaller);

- **Loudspeaker 2 (Bang&Olufsen Beolab 3)** - selected to represent a loudspeaker which does not narrow on-axis at high frequencies (it utilises the 'acoustic lens' technology, to re-radiate sound to a wider range of off-axis angles);
Figure 4.16: Experimental Set Up in Denmark (not to scale)
• **Loudspeaker 3 (Multi-channel Array (Prototype))** - this is a single-unit stereo loudspeaker, the radiation of which may be steered by using beamforming techniques. Delays in the signals to each of the loudspeaker drivers may be controlled by appropriate programming of the DSP network, allowing a nominated directivity pattern to be implemented. In this case, the array was programmed to radiate sound in a strong beam to the left and right side-wall, with little radiation on-axis. This is a notably different directivity characteristic to those above, which predominantly radiate on-axis.

One of each loudspeaker type was, again, measured at the ‘Cube’ facility at Bang & Olufsen, Struer, Denmark, using the same measurement conditions as previously. Contour plots showing directivity in the horizontal plane are included below in Figures 4.17 - 4.19. The plots clearly indicate the differences in directivity between the loudspeakers: Loudspeaker 1 has a horizontal directivity that narrows at higher frequencies, Loudspeaker 2 demonstrates a wide horizontal directivity across most frequencies, and Loudspeaker 3 is shown to beam significantly off-axis (towards -60°). Only one channel (left) of the array is included in the main report (See Appendix for additional plots). Vertical directivity contour plots are given in the Appendix (Section C.2).

**Audio Programme Material Selection**

The audio programme material used in this experiment was identical to that used in Surrey (see Section 4.1.2).

**4.2.3 Stage 1: Elicitation Experiment**

The listening panel at Denmark comprised 9 engineers from the Acoustics Department at Bang&Olufsen, considered to be expert listeners on account of their regular exposure to and interaction with high quality audio, and training in critical listening. The test procedure was identical to that in Surrey, using the same interface, although listeners repeated the test at the second listening position. The loudspeakers were all placed behind an acoustically transparent curtain so that listeners were not able to see them.

In keeping with the Surrey-based attribute selection (Section 4.1.4), the intention here was to choose one spatial attribute, one timbral, one non-spatial non-timbral, and one
CHAPTER 4. PRELIMINARY STUDIES

4.2.4 Choosing the Test Attributes

The elicited attributes were as follows:


The three most frequently-reported differences between each pair of loudspeaker types are shown in Figure 4.21. The overall elicitation results (see Figure 4.22) show width to have been elicited most often and so this is used as the spatial attribute. With respect to the timbral attribute, there is no specific dominating term and so, as before, timbral
fidelity will be used.

Therefore, similarly to tests in Surrey, the attributes used in ratings tests were as follows:

- Ensemble width;
- Timbral fidelity;
- Loudness;
- Liking.

4.2.5 Stage 2: Rating Experiment

Method

The test procedure in Struer was the same for the tests in Surrey, although listeners had to repeat their ratings for each attribute at a second listening position so that the effect of the listener position could also be investigated. This resulted in 144 ratings per
Listener assessment and screening

The listener assessment procedure and criteria used were the same as for Surrey. Listener scores were assessed based upon the following criteria:

- Intra-Listener Consistency;
- Inter-Listener Consistency.
CHAPTER 4. PRELIMINARY STUDIES

Figure 4.20: On-Axis frequency response of loudspeakers used in Experiment 2 (Denmark). Normalised to maximum value.

Figure 4.21: Specific terms used to describe differences between reproduction systems at Denmark (terms relate to order of reproduction systems - for example, Genelec 1031A is narrower than Beolab 3)
Figure 4.22: General terms used to describe differences between reproduction systems in Denmark
<table>
<thead>
<tr>
<th>Experimental Factor</th>
<th>Level</th>
<th>Description</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loudspeaker Type</td>
<td>1</td>
<td>Genelec 1031a (Loudspeaker 1)</td>
<td>Reference 2-way active studio monitor with narrowing directivity at higher frequencies (See Figure 4.17)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>B&amp;O Beolab 3 (Loudspeaker 2)</td>
<td>Hi-Fi Loudspeaker with extended horizontal directivity above 2kHz (See Figure 4.18)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Array (Loudspeaker 3)</td>
<td>DSP-controlled stereo loudspeaker with directional beams aimed towards the side-walls (See Figure 4.19)</td>
</tr>
<tr>
<td>Programme Material</td>
<td>1</td>
<td>Ratatat (Audio 1)</td>
<td>‘Crips’ - Ratatat (2004, XL) Duration = 19 seconds Correlation $R_{LR} = 0.88$</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Beatles (Audio 2)</td>
<td>‘Love me do’ - 1 (2000, Paralphone) Duration = 7 seconds Correlation $R_{LR} = 1$</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Jet (Audio 3)</td>
<td>‘Are you gonna be my girl’ - Get Born (2003, Elektra) Duration = 21 seconds Correlation $R_{LR} = 0.8463$</td>
</tr>
<tr>
<td>Position</td>
<td>1</td>
<td>‘Sweet-Spot’ (Position 1)</td>
<td>See Figure 4.16</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>‘Back-Left’ (Position 2)</td>
<td>See Figure 4.16</td>
</tr>
<tr>
<td>Listener</td>
<td>1-9</td>
<td>-</td>
<td>9 expert listeners (all Bang&amp; Olufs engineer)</td>
</tr>
<tr>
<td>Repetition</td>
<td>1/2</td>
<td>where</td>
<td>Each rating-based listening test was carried out twice for each loudspeaker/programme combination</td>
</tr>
</tbody>
</table>

Table 4.5: Experimental Factors (Denmark)
Intra-Listener Consistency

Using scores from repeat trials, the root-mean-square difference, $\text{diff}_{\text{rms}}$, was calculated for each listener and is shown in Figure 4.23.

Criterion: Listeners with above 15% difference for two or more attributes are screened.
Outcome: No listeners screened on this basis.
Inter-Listener Consistency

The correlation between scores for each listener and the average of the remaining listener scores were calculated in order to observe the sample consensus.

Scatter plots of each listener score and bar charts of calculated correlation coefficient were plotted for each attribute (see Figures 4.24-4.27).

**Criterion**: Listeners whose scores for all cases are not significantly correlated to average scores are considered for removal.

**Observation**: Correlation coefficients for Liking and Timbral Fidelity Scores show that a broad range of conflicting opinion exists in the results, and that an 'average' score method is not applicable. Several listeners have large negative coefficients, and interestingly some listeners show positive correlation at one position, and negative in the other. More consensus is apparent in Loudness and Ensemble Width scores.

**Outcome**: For liking, loudness and timbral fidelity, many listeners show statistically insignificant correlation, which suggests a general lack of consensus. Correlation for Ensemble Width however is both significant and positive for the majority of listeners in both positions. For these reasons, none of the listeners’ scores is removed.

Assessment Summary

- No listener scores were removed on account of intra-listener inconsistency.

- There is little consensus of opinion in scoring for liking, loudness and timbral fidelity.

4.2.6 Results

Procedure

As for Surrey, a univariate analysis of variance (UniANOVA) model was used in order to test the relative statistical significance of each experimental factor. For each attribute at each location, a ‘customised’, ‘full factorial’, and ‘reduced’ (where possible) model
Figure 4.24: Inter-Listener Correlation - Liking (Denmark)
CHAPTER 4. PRELIMINARY STUDIES

A summary of the significant factors from the ‘reduced’ UniANOVA model (or from the ‘full factorial’ model if this could not be reduced) of each attribute at each location is presented in Table 4.6. Further effects are presented in Table 4.7.

The effects of the above factors are explored, by attribute, in the following sections. Where loudspeaker and/or recording are significant factors, data relating to these factors are plotted. Data relating to all factors are plotted in Appendix C, Section C.4.
Figure 4.26: Inter-Listener Correlation - Timbral Fidelity (Denmark)

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Sig.Factors/Interactions</th>
<th>df</th>
<th>F</th>
<th>Sig.</th>
<th>(\eta^2)</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liking</td>
<td>Recording vs Loudspeaker</td>
<td>4</td>
<td>10.822</td>
<td>.000</td>
<td>.136</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Loudspeaker vs Listener</td>
<td>16</td>
<td>18.273</td>
<td>.000</td>
<td>.515</td>
<td>1.000</td>
</tr>
<tr>
<td>Loudness</td>
<td>Recording</td>
<td>2</td>
<td>7.789</td>
<td>.004</td>
<td>.493</td>
<td>.904</td>
</tr>
<tr>
<td></td>
<td>Loudspeaker vs Position</td>
<td>2</td>
<td>11.799</td>
<td>.000</td>
<td>.083</td>
<td>.994</td>
</tr>
<tr>
<td></td>
<td>Loudspeaker vs Listener</td>
<td>16</td>
<td>18.415</td>
<td>.000</td>
<td>.529</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Position vs Listener</td>
<td>8</td>
<td>4.339</td>
<td>.000</td>
<td>.117</td>
<td>.995</td>
</tr>
<tr>
<td>Timbral Fidelity</td>
<td>Recording vs Loudspeaker</td>
<td>4</td>
<td>6.519</td>
<td>.000</td>
<td>.087</td>
<td>.991</td>
</tr>
<tr>
<td></td>
<td>Loudspeaker vs Listener</td>
<td>16</td>
<td>10.498</td>
<td>.000</td>
<td>.379</td>
<td>1.000</td>
</tr>
<tr>
<td>Ensemble Width</td>
<td>Recording</td>
<td>2</td>
<td>11.925</td>
<td>.001</td>
<td>.598</td>
<td>.983</td>
</tr>
<tr>
<td></td>
<td>Loudspeaker</td>
<td>2</td>
<td>40.013</td>
<td>.000</td>
<td>.833</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Listener</td>
<td>8</td>
<td>4.049</td>
<td>.008</td>
<td>.666</td>
<td>.927</td>
</tr>
<tr>
<td></td>
<td>Loudspeaker vs Listener</td>
<td>16</td>
<td>6.733</td>
<td>.000</td>
<td>.281</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Table 4.6: Significant Factors (Denmark) - this table shows the Degrees of Freedom, F-Ratio, Significance, Eta-Squared and Observed Power values for all significant factors and interactions in the preliminary study at Denmark.
Table 4.7: Further Effects (Denmark) - this table shows the significance of interactions between factor types following post-hoc Bonferroni tests.
Liking scores for each loudspeaker type were dependent on listener, particularly with regard to Loudspeaker 3 (the Array). One listener disliked the Array, whereas others liked the Array the most. Differences between the Genelec 1031a and Beolab 3 were less clear. For the Beolab3 and Array, liking scores were also dependent upon programme material. The Array was liked more for the less-spatial recording (Audio 1 Ratatat) than other recordings, whereas the Beolab 3 was liked least for this recording. Liking scores for the Genelec were independent of programme material (see Figure 4.28). Position did not have a significant effect or interaction effect with regards to liking.

Loudness

In this case, the recording was perceived as being a significant factor with regard to loudness, with the Beatles excerpt being rated as lower than the other two recordings (see Figure 4.29). Loudness scores were dependent upon listener, and it is evident that whilst many did not perceive a difference in loudness between loudspeakers, one listener perceived a large change in loudness, rating the Array as being very loud. This is the
same listener who rated the Array as being highly disliked.

The loudness was also seen to be dependent on position. Whilst the Array was perceived to maintain a high loudness at both positions, the Genelec was perceived as being much quieter in the second position, than the first (see Figure 4.30).

Timbral Fidelity

The ratings for Timbral Fidelity were very similar to Liking.

Ensemble Width

Although dependent upon listener, the perceived ensemble width seems to be generally greater for Loudspeaker 3 (the Array) than Loudspeakers 1 and 2. It is also apparent that, with Loudspeaker 3, listeners could not determine any difference in width between programme items, and that all three were wide (see Figure 4.32).

Position was not seen to have any significant effect on ratings of Ensemble Width.
Figure 4.30: Loudness (Denmark): Loudspeaker vs Position

Figure 4.31: Timbral Fidelity (Denmark): Loudspeaker vs Recording
CHAPTER 4. PRELIMINARY STUDIES

Discussion

9 expert listeners were used, yet it is apparent that there is little general consensus with regards to liking (see Figure 4.33), loudness and timbral fidelity in this group when observing the listener data.

In general, Loudspeakers 1 and 2 were rated similarly for liking, loudness, timbral fidelity and ensemble width. Liking and timbral fidelity scores for Loudspeaker 3 (the Array) were highly dependent on the listener being tested and in one case it was perceived as being much louder than the other loudspeakers. Reproduction via Loudspeaker 3 was perceived to be much wider in ensemble width also.

Listening position had no observable effect, other than on the perceived loudness of Loudspeaker 1 (the Genelec which narrows in on-axis directivity).

Again, despite having very different directivities, Loudspeakers 1 and 2, which demonstrate a narrowing and wide horizontal directivity respectively, were perceived to be similar with regards to the tested attributes. This would suggest that the level of differences in reproduction between these loudspeakers was not significant enough to influence the listening situation with regards to these attributes - even though
Figure 4.33: Liking (Denmark): Listener
the directivities were different, which may have directly affected the perception of reproduction in some circumstances, the resulting sound field at the listener may not have been affected due to the interaction between those loudspeakers and that particular room. With this in mind, future testing should have a reduced number of uncontrollable confounding factors so that the individual contribution of each parameter to the perception of reproduction, whether it be directivity, room characteristics or listening position, can be evaluated.

Loudspeaker 3, which exhibited a particularly unconventional directivity, was found to cause a great variation in listener response. The ensemble width was, however, consistently rated to be greater than that for the other loudspeakers and it is likely that the increased amount of lateral reflections as a result of the directivity caused this. The affect of confounding factors means that the primary factor influencing listeners' judgements with regards to timbral fidelity and liking is not necessarily just directivity and, therefore, there is no conclusive evidence that an increase in indirect sound via reflections, and a corresponding spread in image, is perceived as being beneficial.

The main points that arise here are:

- Using this test procedure, two traditional 2-way loudspeakers, designed to radiate differently, were not perceived to be different in attributes which would be expected to be affected by directivity;

- Such a procedure, using combinations of loudspeakers and rooms, gives limited control over the test parameters. Combinations of different loudspeakers and rooms cause a wide range of differences and those caused by directivity alone cannot be distinguished (this provides important initial information regarding R6: What is the best experimental method for this type of investigation?);

- A loudspeaker with a directivity that radiates sounds towards the walls is shown to cause the perception of image width to increase. Listeners show little consensus when judging the perception of reproduction with regards to timbral fidelity and liking via this type of reproduction. Due to the aforementioned confounding factors, it is unknown to what extent such judgements are related to directivity.
CHAPTER 4. PRELIMINARY STUDIES

4.2.7 Stage 3: Small-Scale Free-Roaming Evaluation

Having listeners limited to testing in two positions may not provide data that correspond to how loudspeakers in the home are typically used; it is expected that many people listen to reproduced audio whilst moving between a range of positions within a room, and so it is of interest to observe how listeners judge the different test loudspeakers when allowed to evaluate at any position within the test room, not just two nominated positions. It should also be noted that, due to loudspeaker-room interaction, the response of a loudspeaker will vary at each position and thus can affect the quality of results from such comparisons.

Method

An informal task was designed which invited listeners to evaluate the three test loudspeaker pairs whilst listening at any position in the room. The test took place in the same listening room in Denmark, with the set-up as described above. One piece of audio programme material was used (Audio 1, since of the three excerpts, this was thought to be the most revealing), and listeners were able to control which reproduction system was used for playback using an interface designed in Max/MSP. Listeners were free to walk around the room in order to evaluate the chosen reproduction. However, loudspeaker control could only be made at two points within the room. The loudspeaker control buttons, and the corresponding loudspeaker pairs, were randomised for each listener.

A short questionnaire was given to each listener, which asked them to note the order of the systems according to liking. They were also asked to state where they felt the loudspeaker sounded best, and to describe their typical listening scenario. The questionnaire is given in Figure C.14 in the Appendix. 10 Expert listeners and 4 Naive listeners took part in the test and, as in all previous tests, an acoustically transparent curtain was placed in front of the loudspeakers so that listeners could not see them.

4.2.8 Results and Analysis

The results of the questionnaire are summarised in Figure 4.34. Contrary to results from the formal listening tests in Denmark, Loudspeaker 3 (the Array), was unanimously perceived to be the most liked system when listeners were allowed to listen at any position.
Figure 4.34: Results from informal test where listeners were able to evaluate loudspeakers at any point within a room.
in the room. Most feedback commented that the system exhibited a 'consistency at all positions in the room'. It is interesting to note, also, that a large number of the listeners tested don't primarily listen to music in a sweet-spot position at home. This validates the consideration of the relationship between directivity and perception of reproduced sound in non-fixed listening positions.

Summary/Discussion

This informal test showed that several listeners preferred the reproduction via Loudspeaker 3 (the Array), when allowed to listen at any position within the room.

This simple result suggests that the evaluation of loudspeaker systems is dependent upon the nature of the test, particularly with regards to whether listeners are seated in defined listening positions or not. This research is interested in listening in 'domestic' rooms, and so this acts as justification to consider multiple, non-sweet-spot listening positions as well as sweet-spot listening positions when investigating the influence of loudspeaker directivity.

4.3 Chapter Summary

This chapter documents a series of experiments that were carried out in order to validate earlier work and to provide information to assist the design of future tests. Previously, a number of attributes expected to be affected by changes in loudspeaker directivity were identified, and results from elicitation tests conducted here, using expert listeners and loudspeakers with different directivities in listening rooms, showed that these attributes are indeed affected.

A group of four attributes were chosen to be rated by listeners; these included one spatial attribute (ensemble width), one timbral attribute (timbral fidelity), a non-timbral/spatial attribute (loudness) and an attribute that would indicate changes to the overall enjoyability of the listening experience (liking). In the first experiment, changing the loudspeaker was shown to have a statistically significant effect on these attributes (at least one of the loudspeakers being compared was perceived as different to the others with respect to these attributes), with the reproduction via the wide-horizontal directivity loudspeaker appearing to be perceived as most different from the others. In the second experiment, the reproductions via the test loudspeakers were
perceived as more similar to each other, although there were statistically significant effects on ensemble width, liking and timbral fidelity (for some programme material) and on loudness (for off-centre listening). In the second experiment, the reproduction via the loudspeaker with wall-directed directivity was perceived to be most different from the others.

Despite test results confirming that these attributes were affected by reproductions via different loudspeakers, it is not possible to determine the extent of the effects caused exclusively by directivity - the interactive nature of the room and the loudspeakers, which was uncontrolled, means that any differences caused by directivity could have been overshadowed by other confounding factors. This may explain why, at both test locations, two loudspeakers with markedly different directivities (one had a narrowing horizontal directivity, the other a wide horizontal directivity) were perceived to be similar with regards to the tested attributes. Also, using sets of commercial loudspeakers in listening rooms is limited because the directivity patterns of the loudspeakers are fixed, they cannot be compared at the exact same point in space and significant changes to the room characteristics cannot be compared instantaneously in tests. In light of these limitations, it is necessary that future tests involve greater control over the main contributing test factors.

Finally, an informal test in Denmark highlighted that directivity may become more significant as a contributing factor to the listening experience when considering more than one listening position, thus justifying the inclusion of listener position as a factor in any future testing.
Chapter 5

Investigating the Influence of Loudspeaker Directivity upon Perception

In this chapter, the key findings so far will be summarised. Discussion of the extent to which each of the main research questions has been answered will be followed by the proposal and validation of an experimental route to address the unanswered questions more fully.

5.1 Knowledge Summary

5.1.1 Overview

Directivity has been introduced as a means of defining the directional properties of sound source radiation. It is dependent on the size of the source and the frequency of the sound being reproduced. When considering the directivity of loudspeakers, it is necessary to consider the effects of several parameters, including: driver-arrangement in multiway loudspeakers, cabinets, and transduction mechanism. By altering these parameters, the directivity characteristics of a loudspeaker can be changed and therefore controlled. Real-time control is also possible via multiple-driver DSP-based arrays.

Variation in directivity affects the reproduced sound field within a room - the way in which sound interacts with boundaries is altered, leading to changes in the arriving reflections (in level and spectrum over time) and therefore the direct/indirect sound
The sound arriving at a listener is processed by the human auditory system in such a way that humans are able to perceive spatial and non-spatial changes including: pitch, timbre, loudness and localisation of sources. Attributes most likely to be affected by changes in reflections, according to theory, include: timbre, loudness, localisation, envelopment/spaciousness, source distance and apparent source width (ASW).

The significance of perceived changes in timbre via reflections is debated, and mostly linked to vertical reflections, whereas lateral reflections are linked to changes in ASW. The overall contribution of changes in these attributes with regards to the overall listening situation is likely to be subtle; this is due, at least in part, to the ability of humans to quickly adapt to source/room environments.

Existing directivity-related research does not cover in detail the relationship between reflections and corresponding changes in perceptual attributes. The overall significance of such changes, with respect to the overall listening experience, also remains to be investigated.

Thus, experiments which clarify which attributes are affected, the extent to which they are affected, and their overall contribution to the listening experience are necessary in order to further understanding in this research area.

Preliminary tests show that the attributes expected to change, according to theory, are perceived to change when comparing loudspeakers with different directivities. The changes measured were not large, and it is expected that limitations in the test method (including the use of different loudspeakers in different positions in just two rooms and the uncontrollable interaction of confounding factors) will make this approach unsuitable for further work. Experiments able to affect and measure small changes will require a greater amount of control over a number of parameters (directivity and boundary characteristics/listener position).

Finally, preliminary testing also showed that a particular directivity type was preferred when listeners were asked to evaluate the reproduction all around the room. This provides information that directivity may contribute more significantly to the listening experience when the listener is not seated in a fixed position, but instead engaged in an activity in which they are moving around the room whilst listening.
5.1.2 Current State of Research Questions

This section summarises what has been discussed so far in this thesis in relation to the main research questions.

R1: To what extent can loudspeaker directivity be controlled?

Using traditional transduction mechanisms and low numbers of drive-units, the amount of directivity control is limited to a single directivity-type (typically omnidirectional at low-frequencies, becoming narrow and forward towards on-axis at high-frequencies). Using various mechanical devices and different driver/cabinet arrangements, this directivity can be made to be any type (though its uniformity with frequency depends on the type of pattern and driver size - i.e. a narrow on-axis directivity across all frequencies would require a very large driver). With multiple loudspeaker arrays and DSP, limits are relaxed and real-time pattern control is possible.

It was highlighted that, even with DSP arrays or additional mechanical devices, directivity control would be limited if the size of the unit was restricted so that it was no bigger than typical domestic loudspeaker units. This would result in having no control beneath 200Hz, where it would exhibit omnidirectionality. Therefore, if future work is to consider typical domestic loudspeaker units only, any DSP array/mechanical techniques that are used to vary directivity should be integrated into a comparable size of unit and it should be accepted that control under 200Hz is unachievable. Above 200Hz, depending on the number of drivers and amount of DSP control, a range of directivity patterns may be achieved.

R2: What part does room acoustics play in the relationship between loudspeaker directivity and the sound field at the listener?

The room acoustics wholly define how the radiated sound will be affected before it arrives at the listener. The locations and angles of the boundaries determine the delay and incident angle of any reflected sound (with respect to the direct sound), and the absorption characteristics the level of reflected sound. Therefore, although directivity characterises how the direct sound is radiated, room acoustics affect how the sound is subsequently delivered to the listener.

This information is sufficient to be able to characterise the likely physical changes that
CHAPTER 5. INVESTIGATING DIRECTIVITY/PERCEPTION

will occur in the sound field at a listener as a result of changes in room acoustics. It is from these physical changes that the psychoacoustical changes can be predicted. It is known that differences in the reflected sound, if significant enough, can affect the perceived spatial and timbral qualities of the reproduced sound - delays of high enough level and delay can 'spread' or 'shift' the phantom image, and can also cause comb filtering.

R3: What changes in loudspeaker directivity and boundary characteristics are perceivable, and what are the relative magnitudes of these changes?

Preliminary tests show that there are perceptible differences between the sound fields produced by different loudspeakers when compared in standardised listening rooms. The differences in the measured attributes were subtle and the results were of limited significance (the directivity control was limited and the boundary conditions fixed). In order to answer this question effectively, a higher amount of control over the test parameters is required, yet it is still expected that effects will contribute relatively little to the overall quality of the reproduction, particularly if adaptation is considered.

Further tests that incorporate a high level of parameter control are needed in order to compare the effects of different combinations of loudspeaker directivity and boundary characteristics upon the perception of reproduced sound. It would be desirable to quantify the magnitude of perceptual changes (for specific attributes) that occur with changes in directivity/boundary characteristics.

R4: Which perceptual attributes are affected by the combination of loudspeaker directivity and boundary characteristics

The perceptual attributes affected by loudspeaker directivity, according to theory and preliminary tests are: ASW, timbre, localisation, loudness, envelopment/spaciousness and source distance. It should be noted that envelopment/spaciousness and loudness were not mentioned as predominantly as the others in elicitation experiments, and clarity was perceived to change, despite not being highlighted in the literature.

All of these attributes should be investigated further using the test method with greater parameter control. It would be useful to re-elicit attributes using a new test method and to compare them to those noted here.
R5: What part does listener position play in the relationship between loudspeaker directivity and perception?

In Chapter 3 it was established that listener position is likely to affect the perception of direct sound from a loudspeaker, and the way in which direct and reflected sound interact at the ear. In theory, therefore, listener position is likely to have an effect on the relationship between loudspeaker directivity and perception.

The experiments reported in Chapter 4 revealed this effect when listeners were free to listen from any point within the room, but not when they were restricted to two fixed listening positions. Future experiments relating to the role of listening position in perception of loudspeakers should therefore, perhaps, allow listeners to roam freely. However, the results of free-roaming experiments can be difficult to analyse and interpret because listeners might not all choose the same set of listening positions.

It is possible that the failure of Chapter 4's fixed-positions experiment to reveal a listening position effect was due simply to the particular choice of positions, and that had two other positions been chosen an effect would have been revealed. It is proposed, therefore, that in future experiments two fixed positions (felt likely to reveal a listening position effect) will be used. If no effect of listening position is revealed in these experiments then it is accepted that this might be a result of the choice of positions or a result of the listener being stationary.

R6: What is the best experimental method for this type of investigation?

The requirements of the experimental method have been established and an appropriate method remains to be decided. Results from Chapter 4 give good indication that methods using real loudspeakers in rooms are limited by uncontrollable confounding variables, and so would not be the most suitable.

5.2 Investigation Route

The above knowledge summary shows that a more controlled series of tests is necessary to find out more specifically how loudspeaker directivity affects perception, with variations to speaker directivity (above 200Hz) and to room acoustics independently controlled and with as few other variables as possible.
Varying directivity by switching between alternative loudspeakers is possible but the loudspeakers would have to be placed in separate positions and may exhibit differences in frequency response. Obtaining alternative directivity patterns from a single multiway DSP-based array would offer more control than the switched-loudspeakers method but, in either case, the results would still be limited to the room in which the tests are carried out. The same loudspeaker set-up used in different environments could be compared by listeners, but it is likely that in the time taken to move between one environment and another, their ability to compare the different reproductions will become compromised (see Olive et al. (1995), mentioned earlier). The only way to instantaneously compare a variety of environments based on changes in directivity and boundary characteristics would be to consider auralisation.

Auralisation is the process of rendering virtual sound fields. The impulse response of a loudspeaker within a room can be convolved with an audio signal so that a listener can listen to the signal as if they were listening to it via that loudspeaker in that room.

For this research, auralisation could be used to allow listeners to instantaneously compare the sound fields of various source/room environments in order to assess the effects of directivity/boundary changes upon chosen attributes. In order to compare a large range of environments, a number of impulse response measurements must be made in real rooms using real loudspeakers (a single multi-way DSP-based device could be used here). This method would have an associated limit due to measurement time and the number of suitable acoustic environments. Also, it would lead to a lack of control with regards to room type, as existing rooms would have to be used.

If, instead of using measured binaural room impulse responses, synthesised impulse responses using acoustic modelling software were used, any source/environment combination could be defined, thus giving much greater independent variable control and extending the range of potential data for comparison. The main concern with such a method is its validity, with respect to providing results that are in accordance with real-life - this relies upon the effectiveness of the simulation software, and the transparency (and realism) of the test method. If results of a test using synthesised BRIRs do give the same results as a test conducted in the same real-life conditions, then a synthesis-based system could be used in the main experiments.
5.3 Investigation Method - Validation

Listening tests using the auralisation of real environments (based-on measured BRIRs) have been found to give similar results to those using real sources/environments (Christensen et al. (2005), Hiekkanen et al. (2009)). Tests using synthesised environments are, however, less common and there is no compelling published evidence to suggest that results from synthetic auralisations match those from tests in real acoustics. Therefore, the following study presents an auralisation system that uses synthesised BRIRs, and describes a listening test carried out in order to compare its performance to that of the loudspeaker system used in a previously-published study. If the auralisation system allows listeners to identify the same perceptual changes as documented previously, it will be considered valid for use in further studies involving loudspeaker directivity and perception.

5.3.1 Methodology

Previous Study

A number of previous studies exist, wherein the effect of changing the level and delay of a single off-axis reflection upon the perception of reproduction is investigated. In most cases, tests were carried out in real rooms (often anechoic) with real loudspeakers. In order to validate the proposed auralisation system, and to check that simulations which involve a change in source directivity, listener position and boundary characteristics lead to realistic reproductions, one of these experiments could be replicated; if the acoustic-modelling and binaural reproduction (with head-tracking) is satisfactory, the results of previous experiments should be duplicated.

This study will replicate two parts of an experiment conducted by Olive and Toole (1989), where one loudspeaker was positioned on-axis from the listening position and a second loudspeaker positioned at 65° off-axis. Both loudspeakers were directed towards the listener, were 2m from the listening position and were at 0° elevation. The tests took place in an anechoic chamber.

The same source signal (track 49 from the European Broadcasting Union subjective quality assessment material (EBU-SQAM 1988) - a dry recording of female speech) was reproduced via each loudspeaker. Listeners had a control box which allowed them
to change the level of the signal being reproduced via the off-axis loudspeaker, which therefore simulated a single controllable reflection at 65°.

In the first part of the test to be replicated, listeners were asked to change the level of the off-axis signal until it was perceived to cause 'any audible change in the nature of the sound itself or of the sound field'. They were asked to do this for a range of delays (between 0-80ms). Results were used to determine an absolute threshold, indicating the minimum level and delay for which the 'reflection' caused an audible change in the sound field.

The second part of the test required listeners to reduce the level of the off-axis signal until there 'no longer appeared to be a change in the location or apparent size of the main auditory image'. The results here were used to determine an image-shift threshold. Up until this threshold, it was accepted that other artifacts may be apparent, though not with regards to the perceived location/size of the main image. The absolute threshold and image-shift threshold are presented in Figure 5.1.

Set-up

In order to simulate these parts of Olive and Toole's experiments, it was first necessary to construct a geometrical simulation of the test in CATT Acoustic. A source with frequency-independent level (of 85 dB SPL at 1m) was defined to exist at 2 meters in front of a receiver point. Instead of having a second source off-axis and varying its simulated level and delay with respect to the on-axis source, a reflective boundary was positioned off-axis (see Figure 5.2). Including the boundary instead of a second source meant that the software's ability to render changes in reflections caused by boundary alterations could be tested (future tests would involve changing the boundary conditions like this). By changing the position of the boundary with respect to the source/receiver, the delay of the arriving reflection would be changed, and by changing the absorption of the boundary, the level changed. The boundary also had to be rotated as it was moved further from the listener position so that the reflection angle remained constant (see Figure 5.3). For most reflections, the loudspeaker directivity was defined to be frequency-independent and omnidirectional (see Figure 5.4). For reflections which were high in level (with respect to the on-axis signal), the directivity had to be changed to counteract the natural fall-off in energy over distance that would affect the reflection. The alternative directivity pattern is shown in Figure 5.5.
Figure 5.1: Thresholds measured in Olive and Toole (1989). Listeners were able to change the relative level of an off-axis loudspeaker reproducing a delayed version of a signal reproduced by a loudspeaker on-axis.
CHAPTER 5. INVESTIGATING DIRECTIVITY/PERCEPTION

Figure 5.2: A screenshot taken from CATT Acoustic, indicating the geometry of the simulation created.

Figure 5.3: In order to maintain a fixed angle of reflection, the boundary had to be rotated as it was moved further from the receiver.
Figure 5.4: A balloon plot of the 'omnidirectional' directivity pattern used. The scale indicates sound pressure level in decibels.

Figure 5.5: A balloon plot of the 'extended omidirectional' directivity pattern used in order to produce reflections that have the same amount of energy as the direct sound (the increase in off-axis sound pressure is indicated by the light red/pink regions at the sides).
Figure 5.6: All reflections simulated using the acoustic modelling software are shown here compared to the original thresholds determined by Olive and Toole (1989).

The original experiment used outboard equipment to give continuous control over the delay between the loudspeakers. However, continuous control using synthesised environments would require an extremely large number of BRIRs to have been simulated, so that a listener could cycle through them in the same way. Instead, it was decided that a finite number of simulations would be chosen to best represent the various combinations of level/delay that caused different types of perception. A total of 18 combinations were modelled, with 11 of these used for the tests (see Figure 5.6). The final 11 were chosen as they existed between each of the thresholds (and therefore can be used to test whether the perceived changes are as in the real case), and because they existed in the region that future studies will be limited to considering (between 0-50ms, which is the maximum reflection delay expected in small domestic rooms). For each model, 61 BRIRs were calculated to be used with a head-tracker (to allow the listener to rotate between -30 and 30° either side of the virtual source whilst listening to the reproduction).

The acoustic modelling used an automatically selected number of rays (though, only
the region in time where image-source modelling would have been used was of interest), specular reflections, a 48kHz sample rate and binaural receiver modelling (based on CATT Plain HRTFs) with headphone equalization (based on Sennheiser HD600s).

The BRIRs were formatted into a database so that they could be used with the real-time convolution software. The convolution software convolved the impulse response (relative to the listener's head position as informed by the head-tracker) with an audio signal. In this case, the audio signal was the same as that used by Olive and Toole. A flow-diagram of the system is presented in Figure 5.7.
CHAPTER 5. INVESTIGATING DIRECTIVITY/PERCEPTION

Listening Test - Initial Measurements

8 listeners from the Institute of Sound Recording, University of Surrey took part in the initial listening tests. The listeners were considered to be 'expert listeners', and the tests took place in a small edit suite at the Institute of Sound Recording.

Before each initial listening test was conducted, listeners were asked to take part in a familiarisation process. A playback graphical-user-interface (GUI) was constructed that allowed the listeners to hear all of the reproductions that had been simulated, so that they could identify the range and extent of perceived changes in the test stimuli (see Figure 5.8). The stimuli were presented in a randomised order, but this order was the same for everyone.

Following the familiarisation process, listeners were presented with a three-alternative forced-choice test GUI (see Figure 5.9). The listeners would be presented with a reproduction and asked to choose whether they perceived 'no audible change', 'an audible change not including image-shift' or 'an audible change that included image-shift', when comparing it to the reference, which was a simulation with no boundary (and thus, just the direct sound with no reflection). 18 reproductions were presented in total, at random. These included the 11 chosen reproductions, 4 repeats, a spatial ability check (where the on-axis source was simulated to be on the left, with no reflection) and a timbral ability check (where the on-axis source signal was filtered with a 500Hz high-pass filter and, again, with no reflection). These 'check stimuli' were included to ensure that listeners were able to distinguish between an obvious spatial difference that should constitute image-shift, and an obvious timbral difference that should not. All listeners, bar one, were able to make this distinction. That listener perceived the high-pass filtering to cause an image-shift. As there is some evidence to suggest that a change in spectral content of a signal can affect the perceived characteristics of the image, the listener's results were not removed.

Listeners were asked to provide judgements for some stimuli twice, so that the consistency of their judgements throughout the test could be evaluated. It was decided that a listener's results would be removed from the data set if they were not able to provide the same judgements for at least 75% of the repeated stimuli.

Only one listener was unable to repeat over 75% of their judgements, and so their results were removed. Interestingly, all listeners who showed some inconsistency perceived at least one item to cause image-shift on one occasion, and non-image shift on another.
Figure 5.8: The familiarisation interface used in both listening tests and elicitation. Immediate comparison of any of the test stimuli is possible.
Which of the following options best describes the difference between the stimulus and the reference? (Choose 1 option)

- An audible change including 'image shift' (where the size, position or number of phantom sources is perceived to change)
- An audible change that does not include 'image shift'
- No audible change

Figure 5.9: MATLAB interface used for the initial listening tests.
This suggests that there is some indecision with regards to whether certain stimuli did cause image-shift or not. This only occurred for stimuli that simulated reflections in the region between thresholds (above the absolute and below the 'image-shift' threshold); listeners were certain that they perceived changes above the lower threshold, but they were unsure about whether it was image-shift or not in some cases.

Observation of results (presented in Figure 5.10) from the initial tests indicates that the stimuli based on reflections whose levels were above the image-shift threshold caused a perceived image-shift with this system. The majority of the stimuli that exist in the region between the absolute threshold and the image-shift threshold caused a perceived change that was considered to also be 'image-shift'. All remaining stimuli (based on reflections below the absolute threshold) were perceived to not cause any difference compared to the reference, which did not include a reflection (as in Olive and Toole's experiments).

This suggests that whilst the reproduction of the simulations above and below the thresholds gives results as produced in the real-life reproduction in Olive and Toole's tests, simulations in the region in-between the thresholds do not (in Olive and Toole's tests, listeners did not perceive image-shift, yet here they did). Therefore, either the system is unable to generate similar reproductions in this region, or there is some flaw in the experiment (for example, listeners may be unclear about the definitions provided). In order to shed more light on this outcome, a further investigation was carried out.

Listening Test - Extended Investigation

In order to elicit extended information with regards to the listeners' perception of the stimuli presented in the initial test, particularly those above the absolute threshold, a simple elicitation experiment was carried out. Using the familiarisation interface used previously, the listeners were asked to provide written descriptions of the differences they heard between the reference and a number of selected stimuli. These stimuli included the modelled reflections which caused perceptual effects above the absolute threshold. The remaining stimuli, which were not perceived to be any different from the reference in the first test, were not presented to the listeners on the test sheet.

5 listeners were given test sheets to complete, and were asked to provide as much detail as possible about the perceptual differences, including descriptors and magnitude of effect. The tests took approximately 25 minutes per listener and were carried out in the
Figure 5.10: Results from initial test. Each pie chart shows, for a particular simulation (delay time and level) the proportions of listeners choosing each of the 3 possible responses.
listening room TB07, University of Surrey, which is a well-isolated dry acoustic room conforming to ITU-R BS 1116 (1997).

The stimuli presented to the listeners for further description, and a summary of the descriptions given by all listeners, are presented below:

- **Stimulus 2** (-5dB, 40ms: above Image-shift threshold): Large change - strong echo/split image;
- **Stimulus 3** (-10dB, 10ms: below Image-shift threshold): Small change in source width (broadening);
- **Stimulus 5** (-15dB, 20ms: below Image-shift threshold): Slight change in source width (broadening);
- **Stimulus 6** (0dB, 5ms: above Image-shift threshold): Large source width broadening;
- **Stimulus 8** (10dB, 5ms: below Image-shift threshold): Very slight change in width broadening;
- **Stimulus 11** (-15dB, 40ms: below Image-shift threshold): Very subtle change - spatial blurring / room effect (spaciousness). Not image....;
- **Stimulus 13** (-2dB, 20ms: above Image-shift threshold): Very strong echo / split image.

It appears that there are clear differences perceived between those stimuli that fall above the image-shift threshold and those that fall below, namely in the magnitude of the perceived change and the division of the image. Stimuli that fall beneath the image-shift threshold tend to be perceived as much more subtly different from the reference, and to also include more of a sense of spaciousness. It appears, therefore, that in the initial experiment any slight change in the image, whether it be size, position, clarity or number was interpreted as an 'image-shift' (see Figure 5.10). In Olive and Toole's tests, however, this may have not been so.

**Listening Test - Re-Test with Appropriate Descriptors**

At this point, it is still unclear whether or not the system has produced a similar reproduction to that in the original experiment. If Olive and Toole's text is re-analysed,
and if the descriptors used in conjunction with the term ‘image-shift’ are extracted, it will provide more information on this matter.

The main description from Olive and Toole’s paper was as follows:

‘As the reflection level was reduced, there came a point where the image spreading was no longer significant and the reflection was not separately identifiable. This was the condition required for the image-shift threshold. At this threshold, there were still other artifacts, but they did not affect the location or apparent size of the main auditory image. The principal side effects betraying the presence of the low-level reflection were a slight sense of spaciousness and occasional high-frequency sibilant ‘splashes’, localized at the origin of the lateral reflection.’

From this, it can deduced that,

- for stimuli over the image-shift threshold, the change in image size is ‘significant’, and/or the reflection is ‘separately identifiable’;
- for stimuli below the image-shift threshold, the image spreading is subtle (not significant) and/or there is a slight sense of spaciousness.

This suggests that there is a mismatch between the definition of ‘image-shift’ used in the original paper and that in the minds of listeners used in this test. In light of this, a revalidation experiment, based on a definition of ‘image-shift’ compatible with that in the original paper, should provide more comparable data. Note that the motivation for trying to replicate the results of Olive and Toole’s paper is to validate the auralisation system - even if the original experiment is no longer ‘image shift’ (at least as subjects here understand it), reproducing the original experimental results still validates the system.

The choices available to listeners now become:

- A significant change in image size / separately identifiable reflection;
- A slight sense of spaciousness / subtle image spreading;
- No audible change.

Also, an additional button was included on each test page to allow the listener to hear the stimulus which is deemed to cause the most significant amount of change - a feature of Olive and Toole’s test not included in the initial test here (see Figure 5.11).
The test was carried out using the same set-up as in the previous tests. As before, listeners were asked to partake in a familiarisation phase, before taking the main test. As the new descriptors used are not applicable to the two ‘listener ability’ test signals (a high-pass filtered version to test timbral judgement, and an off-centre source version to test spatial judgement), they were removed from the familiarisation and main test phase, and instead replaced with more repeats of the original stimuli. The removal of the listener ability check caused no problem because all listeners who took the main test proved to be adequate in ability during the initial test. The test was taken by 6 different listeners (all of whom took part in the initial test) in the listening room, TB07, at the University of Surrey, which is a well-isolated dry acoustic room conforming to ITU-R BS 1116 (1997).

Listener consistency was checked as in the initial experiment. All listeners were able to repeat their classification of 75% of the stimuli presented to them more than once and, therefore, can be considered as giving consistent data. As a result, all of the data from this test were kept.
Figure 5.12: Results from main test with improved descriptors. Each pie chart shows, for a particular simulation (delay time and level) the proportions of listeners choosing each of the 3 possible responses.

5.3.2 Results

Results from this test are presented in Figure 5.12

5.3.3 Discussion

Using verbal descriptors that correspond to the definition of 'image-shift' used by Olive and Toole, the majority of simulated reflections using this system were perceived by listeners in the same way as those in Olive and Toole's experiments. The magnitude and nature of the perceptual effects were shown to be dependent on the level and delay of the simulated reflections. This confirms that, by auralising synthesised sound scenes using modelling software and headphone reproduction with head-tracking, it is possible
to affect the perception of reproduction in a realistic way; synthesised changes in source
directivity, absorption and boundary position can be used to reproduce perceptual
changes that have been measured in real environments with real sound sources. This
outcome validates the system's potential use as a tool to investigate perceptual changes
resulting from changes in directivity/room boundary characteristics (and therefore could
contribute to R6).

It is apparent, also, that the changes perceived here were primarily spatial: image,
spaciousness and closeness etc. Timbral/spectral changes are thought to have been
perceived, despite not being reported. It is thought that this is because of their lesser
prominence compared to spatial changes. Future tests using this system to evaluate
both timbral and spatial changes should have test methods that allow listeners to judge
them effectively without spatial changes becoming so dominant that they distract from
timbral judgements.

It is important to note that the test method used here involves a new system, which
convolves simulated impulse responses with audio excerpts in real-time. The additional
real-time head-tracking system (which is used to increase the realism of the auditory
presentation by allowing listeners to rotate their heads within the simulated reproduction
scene) has a high processing cost and, as a result, some issues with the performance of
the system have been identified:

- some 'digital skipping' (i.e. stutter and lag), which appears to occur only when
  opening and closing windows during the time that the system is running;

- some 'popping' in the left channel headphone tends to occur sporadically;

- some minor instability of the image is perceived as the head is turned.

Also, some general ideas about the perception of simulated sources with this system were
considered following informal testing:

- When the source is modelled to be off-axis in a room, it is perceived as being
  external and stable in location;

- When the source is modelled in the free-field, it tends to be perceived as inside-
  the-head. The realism of the auralisation is also affected by head movements; a
  small amount of phantom-image drift occurs when the listener rotates to wider
angles. This may occur because the HRTFs used were non-individualized. It is also felt that such artifacts are more obvious in modelled free-field environments, as non-free field environments introduce reflections which act as a distraction.

5.3.4 Conclusions

This system is able to reproduce results carried out using real loudspeakers in an anechoic chamber. It is therefore validated for use in future tests concerning the perception of reproduced sound and its relationship with loudspeaker directivity/room boundaries; it can provide realistic rendering with appropriate resolution. In its current form, it is not sufficiently stable but this can be remedied with a change of computer hardware.

5.4 Chapter Summary

A 'knowledge summary' has been presented, which reviews the most substantial findings of the preceding chapters of this thesis. This information has been used to answer the specified research questions as effectively as possible, and any unanswered elements have been highlighted. Notably, more experimental data are required in order to answer questions relating to the specific type and magnitude of perceptual differences in reproduced sound caused by changes in loudspeaker directivity.

In order to obtain additional experimental data that furthers knowledge in this area, there is a need for greater parameter control within experiments. The most effective way to control all factors of interest (loudspeaker directivity, boundary characteristics and listener position), and to allow instantaneous comparisons, is to use auralisation. For the most flexibility, the use of acoustic modelling software to create the BRIRs appeared to be favourable. However, its validity with respect to simulating real conditions was questionable.

Therefore, a test was designed in an attempt to validate such a system; a previous well-documented test using real loudspeakers in an anechoic chamber was reproduced by convolving the same audio excerpt with acoustically-modelled simulations that the listener could compare instantaneously. Results show that the system, which used a head-tracker and headphone reproduction, is able to reproduce these results satisfactorily, and can therefore be used for future experiments.
Chapter 6

Main Experiments

Now that the auralisation system has been validated, it is necessary to plan how such a system will be used in experiments to extend this research. It has been identified that the following items should be considered in further experimentation:

- A method that allows control over loudspeaker directivity, boundary conditions and listener position so that a range of combinations of these factors can be directly compared by listeners (within one test environment);
- For differences of these parameters to be ordinal, so that any perceptual differences can be quantified and linked;
- For the perceptual differences to be measured, categorised and labelled, so that the key attributes affected by changes in these parameters can be identified and the magnitude of effect quantified.

6.1 Experiment Aims

To return briefly to the principal aim and the associated research questions, it is of most interest to determine the perceived type and magnitude of changes likely to be caused by varying loudspeaker directivity in a domestic room. It would also be of interest to examine the effect of the listeners' position and the acoustic characteristics of the domestic room. Given that using acoustic modelling and auralisation is the most practical solution to begin to explore this with any scientific rigour, and, that such an
experiment in this area of research hasn't been carried out before, it becomes necessary to consider a 'first-look', simplistic, idealised approach towards selection of the experimental variables and the acoustic environments in which they are tested. A complicated acoustic model of a real room, with comparisons of complicated frequency-dependent directivity patterns would provide very specific results, from which little information could be applied to other situations. Instead, what is most suitable, is to observe the effects of combinations of simple directivity patterns and boundary conditions, such that the test results may be used to predict the likely effects in more complicated situations.

It is necessary to decide which parameters should be compared by listeners in order to satisfy the aim of producing basic informative results. In light of the work in this thesis so far, the following have been chosen as most suitable:

- source directivity (will change the direction of the radiated sound) - this is key to our investigation. A range of simple and realistic directivity patterns should be compared;

- boundary absorption (will change the level of any reflected sound) - the absorption characteristics of a room are thought to play a crucial role in the perception of differences between loudspeaker directivity and so a simple range of absorptive levels of reflective boundaries should be included in the test;

- boundary separation (will change the delay (and level) of any reflective sound) - domestic rooms are different in size in every household and so it is important to consider the effects of changing the separation between boundaries;

- listener position (will change the angle, delay (and level) of reflected sound) - as discussed throughout the thesis, perception of reproduction within a room is known to be dependent on listener position and will be considered in these tests.

These parameters become the independent variables for the main experiment - listeners will be asked to compare the reproduction of audio based on different combinations of them, contributing data towards the principal aim of this thesis. In order to acquire a wide range of data, which considers the effects of first reflections and reverberation, the modelled acoustic environments tested will include a simple two-boundary configuration, modelling a side-wall only scenario, a four-boundary configuration, modelling a situation with no reflections from floor or ceiling and a six-boundary situation, modelling a
more typical domestic room. This also enables the change in effects resulting from the introduction of additional reflective boundaries to a space to be monitored.

**Summary of Main Experimental Aims**

- design and carry out a ‘first look’ experiment to acquire data that can be applied to a wide range of listening situations;
- compare auralised reproductions of modelled loudspeakers/environments which differ in directivity, boundary absorption, boundary separation and listener position;
- determine which perceptual attributes are affected;
- measure how much the attributes are affected.

### 6.2 Stimulus Selection

#### 6.2.1 Independent Variables

The initial simulations chosen for the main experiment involve a single loudspeaker positioned off-centre, in-between two boundaries (as discussed previously, more boundaries are to be added in subsequent tests). A single loudspeaker is chosen because two loudspeakers operating simultaneously within a reflective environment cause effects that are considered to overly complicate the investigation at this stage. It is positioned off-centre because most listening situations do not involve a single source reproducing sound from the middle of the room, and asymmetrical reflections are likely to cause different perceptual changes (from those caused by a central loudspeaker position, which could produce perfectly symmetrical reflections).

With just one simulated off-centre source and two boundaries, the effect of asymmetrical reflections arriving from the two walls and their effect upon perception can be investigated.

The modelled directivities of the simulated loudspeaker will be omnidirectional, cardioid and narrow-beam on-axis. These have been selected as they span an extreme range of horizontal directivities, between full off-axis coverage and no off-axis coverage (thus, maximal to minimal interaction with side walls). Table 6.1 details the different patterns used.
### Table 6.1: The three simulated directivity types used for the Main Experiments. The scale is in dB.

<table>
<thead>
<tr>
<th>3D Balloon Plot</th>
<th>Description</th>
<th>Directivity Index</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Omni Plot" /></td>
<td>Omni</td>
<td>0 dB</td>
</tr>
<tr>
<td><img src="image" alt="Cardioid Plot" /></td>
<td>Cardioid</td>
<td>5.3 dB</td>
</tr>
<tr>
<td><img src="image" alt="Narrow On-Axis Plot" /></td>
<td>Narrow On-Axis</td>
<td>15.3 dB</td>
</tr>
</tbody>
</table>
The boundary models will be separated by 2m and 6m, which represent the extremes of the range found in most typical British living rooms/bedrooms. These will lead to the earliest reflections arriving at around 0.3 to 9ms after the direct sound, and are thus likely to cause effects as discussed and measured earlier (Chapters 3 and 5) in the region relating summing localisation and precedence effect.

The modelled boundary absorptions should span a wide range of absorption conditions and should be representative of typical domestic rooms. The highly absorbent condition will have an absorption coefficient of 0.5 (an extreme case, equivalent to medium-weight folded curtains and carpeted floor (Smith et al. 1982)). The highly reflective condition will have a coefficient of 0.05 (equivalent to floor tiles and brickwork (Smith et al. 1982)).

Finally, the simulated listener positions cover two typical listening situations: on-axis and off-axis of a centrally-pointing loudspeaker. It is expected that the sound field at these two different positions will reflect the types of changes that would have been heard had the test been able to include many positions or a 'free-roaming' element as earlier - reflections arriving at each position would be different in level and time-of-arrival, and the coverage at each position different for narrowing directivity patterns (to clarify - each additional listening position would require another 12 stimuli to be compared and 'free-roaming' auralisation is not possible using our system). The distance between loudspeaker and listener positions will be 2.5m, again representing a typical distance in a domestic situation.

Variation in all of these parameters means that for every audio excerpt used in the first test stage of each experiment there are 24 combinations to be tested. This is demonstrated in Figure 6.1.

### 6.2.2 Audio Excerpts

Three audio excerpts will be used:

- Speech;
- Dry pop recording;
- Reverberant classical recording.

Speech is used in most of the key tests in the literature. It is expected that perceived changes with less reverberant recordings will be different from those that are more
Figure 6.1: Variation in parameters for main experiment
reverberant, and therefore a dry pop recording and reverberant classical recording will be used. (Previous literature highlighted the that the effects of reflections are less audible with continuous material or amongst other reflections, than with dynamic material).

6.2.3 Summary

The aim of the following experiments is:

*to identify the magnitude and type of perceptual effects that are likely to occur in a real listening situation, as a result of changing loudspeaker directivity, room characteristics and listener position.*

The simulated listening environments include three different loudspeaker directivities (omnidirectional, cardioid and narrow on-axis beam), two wall-separation distances (2m and 6m), two absorption levels (absorption coefficient of 0.05 and 0.5) and two listener positions (centre and left-of-centre). This results in the listeners having to compare audio programme material (of which there will be 3 excerpts: speech, dry music and reverberant music) in 24 different simulated environments, with all combinations of the aforementioned parameters. The comparison will be carried out using the auralisation system discussed previously in Chapter 5.

6.3 Experimental Method Selection

6.3.1 Attribute vs Holistic methods

With the stimuli prepared, it is necessary to consider the measurement techniques that may be employed to determine which attributes are perceived to be different with different simulations and by how much.

There are two clear methodology types in traditional sensory evaluation: attribute and holistic testing. Attribute testing refers to the type of test where specific sensory attributes are provided for, or elicited from, subjects to then rate with respect to the presented stimuli. Holistic testing refers to an alternative, where subjects are asked to carry out simple comparisons or groupings of stimuli based on their holistic properties
and are not required to rate with respect to specific elicited or provided attributes. The relationships measured between the stimuli are used to subsequently map the stimuli onto representative perceptual spaces, marked by key dimensions or components in which they are shown to differ. Despite elicitation of attributes not being required to derive a perceptual map, its interpretation can be aided by the consideration of attributes. This can occur on a number of levels: the experimenter may use his own knowledge to label the key dimensions, he may consult the subjects for assistance in labelling, or he may combine data from separate attribute rating tests and then observe the correlation between the stimuli variation and dimensions/attributes. An overview of these methods are provided in Table 6.2 and 6.3.

Berg and Rumsey (1999) similarly categorise the various methods for arriving at sound attribute scales into three groups: (i) those that aim at a common set of attributes for grading by all panel members, (ii) those that are based on free categorisation of individualised scales and (iii) those which use some form of multidimensional analysis based on non-semantic similarity/difference relationships between stimuli. They subsequently remark that (i) is advantageous to the experimenter, (ii) has a lack of bias and allows personal reflection and (iii) also shows a lack of bias but is problematic with regards to interpretation.

Berg elaborates on this concern for bias in provided attribute scales by suggesting that the subject is constrained to responding in a way defined by the experimenter and goes on to quote Kjeldsen (1998): “you only get an answer to what you ask”. The idea that subjects may want to use descriptions that the experimenter has not permitted them to use is considered by many in sensory evaluation (particularly in the study of foods) to be considerably limiting and to not allow the true response to be captured. Berg and Rumsey (1999) state additional issues with basic scaling of attributes as being: whether an attribute definition is clear and unambiguous, whether it is understood in the same way by all subjects, whether it was agreed with the subjects in the context of the task at hand and whether the subject had any influence over the definitions. By allowing subjects to elicit their own attributes for ratings (as in Quantitative Descriptive Analysis (QDA), Repertory Grid Technique (RGT) and Free Choice Profiling (FCP), all of which can be found in Table 6.2), results are arguably more reliable, but are still questioned because of the generality of terms, the influence of the panel leader, the influence of discussion between subjects and differences in their use of terms/scales (Bech 1999). Koivuniemi and Zacharov (2001) reinforce these limitations and even suggest that correlation between grading scales can affect results.
<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provided Attribute Scaling</td>
<td>Subjects required to rate stimuli according to attributes chosen by the experimenter.</td>
<td>Considered by many to be limiting and not to allow the true response to be captured - ‘subject is constrained to responding in a way defined by the experimenter’ (Berg and Rumsey 1999)</td>
</tr>
<tr>
<td>Quantitative Descriptive Analysis (QDA)</td>
<td>'Conventional Profiling' is employed, where the panel discusses and determines attributes in a given sensory area (i.e. defining a common language). Attributes from the language which are agreed by the panel to be affected by test stimuli are agreed before rating scales are designed and used (Pagès (2005), Neher (2004), Bech (1999), Berg and Rumsey (1999), Koivuniemi and Zacharov (2001))</td>
<td>Assumes that the intensity of the described sensation can be assigned to a value along a scale (Bech 1999). Audio Descriptive Analysis and Mapping (ADAM) was a variation of QDA involving naïve listener preference test (via pairwise comparison) w/PCA and QDA w/ANOVA carried out by experts. Results of each were then correlated to give a predictive model (Koivuniemi and Zacharov 2001).</td>
</tr>
<tr>
<td>Repertory Grid Technique (RGT)</td>
<td>Triads of stimuli are presented to subjects and they are asked to describe how two of the stimuli differ from the other. This is repeated until the subject produces no new answers. A grid is then constructed with which a subject can rate the stimuli according to his/her own attributes (Neher (2004), Bech (1999), Koivuniemi and Zacharov (2001), Berg and Rumsey (1999)).</td>
<td>Had been criticised for being influenced by the experimenter in comparison to FCP, but tests show that no differences are found (Bech 1999).</td>
</tr>
<tr>
<td>Free Choice Profiling (FCP)</td>
<td>Subjects gather descriptions they’ve used to describe the stimuli, remove synonyms and antonyms and define scale bounds. Samples are then scored according to their own descriptors (Pagès 2005). Subjects can use as many terms as necessary and Procrustes analysis is typically used to produce a perceptual map (Neher 2004). No training, experimenter interaction or discussion between subjects.</td>
<td>Results assumed to be more general than DA (Bech 1999). Suggested to be the best method for obtaining additional semantic data (for MDS) (Neher et al. 2006). Versus DA it had higher agreements between subjects (Bech 1999). Appealing but analysis is complex (Koivuniemi and Zacharov 2001).</td>
</tr>
</tbody>
</table>

Table 6.2: Attribute-based sensory evaluation methods
### Type

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pairwise Comparisons</strong></td>
<td>Subjects assess similarity/dissimilarity between all stimuli. Stimuli are presented in pairs and the number of trials for n stimuli is n(n-1)/2 (Næs and Risvik 1996).</td>
</tr>
<tr>
<td><strong>Conditional Rank Ordering</strong></td>
<td>Each stimulus is used as a standard and the panelist ranks the remaining stimuli according to the similarity to the standard. This method requires special models for analysis (Næs and Risvik 1996).</td>
</tr>
<tr>
<td><strong>Sorting</strong></td>
<td>Stimuli are sorted into mutually exclusive groups based on similarity. Subjects are often told that they must sort into no fewer than two groups (Næs and Risvik 1996).</td>
</tr>
<tr>
<td><strong>Projective Mapping (Napping ®)</strong></td>
<td>Subjects arrange stimuli on a two-dimensional surface (typically a tablecloth or sheet in food/drink analysis). Stimuli placed close together are deemed to be similar, and stimuli placed far apart, dissimilar. In some cases, labels are added to aid with analysis (Pagès 2005).</td>
</tr>
</tbody>
</table>

Table 6.3: Holistic-based sensory evaluation methods
The style of attribute rating tests has also been questioned by King et al. (1998) who implore that data collection procedures should be easy to understand and structured so that the process doesn't become laborious. Also, with elicitation-based methods, it can take a very long time to establish attributes and scales - for example, Koivuniemi and Zacharov (2001) took approximately 30 hours (over 5 weeks).

This all contributes to the notion that holistic test methods, which allow subjects to freely express the relationships between stimuli without being confined to predisposed categories, as in rating tests based on provided or elicited attributes, would be ideal in this type of study, where the exercise is exploratory. However, it should be noted that some means of interpreting the outcome of such tests will need to be considered, which may lead to its own difficulties.

6.3.2 Pairwise Comparison vs Sorting vs Projective Mapping

The type of holistic method for the evaluation of the prepared stimuli must now be chosen. Options, as outlined in Table 6.3, include pairwise comparisons, conditional rank order, sorting and projective mapping. Conditional rank order is the least common method and also requires special models for analysis and therefore the remaining methods will be explored as a priority.

Pairwise Comparison, whilst perhaps seeming like the most straightforward test method from a listener point-of-view, can involve an excessive number of evaluations (for example, 7 stimuli = 21 pairs, 10 stimuli = 45 pairs, 25 stimuli = 300 pairs) (Næs and Risvik 1996). This can cause sensory fatigue, adaptation and potential lapse of attention, and therefore this method is little used in sensory evaluation with large sample sizes (particularly with multidimensional scaling (MDS), discussed later) (Nestrud and Lawless 2010).

A Pairwise Comparison of the 24 simulated environments here (considered to be a large sample size) would require 276 trials per audio excerpt used and so it will not be considered as an appropriate test method for this experiment. An experiment was carried out to see if the number of stimuli could be reduced, by removing any stimuli that did not cause a noticeable difference compared to an anechoic reference stimulus, but results showed that all stimuli were perceivably different. This experiment, although documented, was not included in the thesis as it would cause a digression.

The two remaining holistic methods available are sorting and projective mapping. Both
methods are free from the systematic rigour of attribute-scaling or pairwise comparison, allowing subjects to collate and arrange stimuli based on similarities. Both have also been used successfully in numerous sensory studies.

**Sorting**: In sorting tasks, subjects are typically exposed to all stimuli at once and are asked to sort them into mutually exclusive groups based upon similarity (normally at least two groups). Following the grouping stage, the number of times any two samples are placed in the same group is calculated to derive a similarity matrix, where pairs found to be grouped together more often indicate a greater similarity than pairs not found to be grouped together often. This similarity matrix can then be used for analysis (discussed later) (Bonebright 1996).

**Projective Mapping** (or Napping®): Napping® is a term derived from the French ‘nappe’, meaning ‘tablecloth’, and it is used to describe a test type whereby subjects arrange test samples (originally cheese or wine) directly on a tablecloth, with the distances between them on the 2-dimensional surface representing the magnitude of their similarities/dissimilarities. The distances between each sample, for each test subject, are noted and used to construct a similarity matrix which can then be analysed (Pages 2005). This method can be adapted to compare any stimuli on a 2-dimensional surface (for example, audio clip comparison on a computer-based GUI).

This technique has been used successfully for the sensory analysis of wines (Pages 2005), evaluation of digital cameras (Ramsgaard et al. 1994), breads and yoghurts (Pfeiffer and Gilbert 2008).

A paper by King et al. (1998) compares the two methods. An overview of sorting examples in previous sensory studies suggests that whilst there are instances where it has provided reliable results (such as for aroma categorization) (and in Bonebright (1996) vs Paired-Comparisons), some loss of information is seen to occur, particularly in small differences between individual samples. An experiment is carried out to compare the sorting and mapping of snack bars (following analysis via MDS) and results show that projective mapping leads to better differentiation between bar type (meaningful
information is provided in an additional dimension) and also requires fewer iterations to reach the convergence criteria. A better goodness of fit is also seen. Nestrud and Lawless (2010) compared the methods (using different analysis techniques for each) and found that, although both maps were similar, more information could be drawn from the projective mapping data when Cluster Analysis was utilised in addition.

With regards to the ease-of-use, King et al. (1998)'s study also compared this and results showed no significant difference. Panellists, however, did find it easier to change their minds using projective mapping.

A more recent study by Vowels (2012) using auditory stimuli known to vary in three distinct dimensions showed that a projective mapping task outperformed corresponding pairwise and sorting tasks in allowing the recovery of the three dimensions, and took less time than the pairwise task for listeners to complete (but took a similar time for sorting).

In light of this, Projective Mapping would seem like an appropriate choice. It has been shown to be robust and effective in comparison to other methods (Risvik et al. (1997), Nestrud and Lawless (2008)) and should benefit listeners with respect to time (when compared to pairwise comparison) and performance, in this case. One concern with this method that has been has raised is whether the limitation of the two dimensional response space constrains the respondents and therefore the resulting perceptual maps (Nestrud and Lawless (2008), Nestrud and Lawless (2010)). He remarks that it remains an open question, but the aforementioned study by Vowels (2012) indicates that more than 2 dimensions can be identified via projective mapping tasks.

6.3.3 Multidimensional Scaling Analysis

Similarity data from projective mapping experiments may be analysed via a number of multivariate statistical techniques including Multidimensional Scaling (MDS), Principal Component Analysis (PCA), Cluster Analysis, Partial Least Squares Analysis and Procrustes Analysis. Existing examples show that Mapping has been successfully paired with Multiple Factor Analysis (similar to Principal Component Analysis) (Pages (2005); Nestrud and Lawless (2008)), MDS (Barcenas et al. (2004); King et al. (1998)) and Procrustes Analysis (King et al. (1998); Nestrud and Lawless (2008)), with MDS presenting more meaningful configurations than GPA in King et al. (1998). Given that MDS has been used successfully with projective mapping in the past and was found
to present more meaningful data than GPA by King et al. (1998), it will be used for analysis in the main experiment here. The technique is described in full below.

'Multidimensional Scaling (MDS) is used to construct a spatial representation of the similarity among objects with the purpose of discovering relationships or patterns' (Næs and Risvik 1996)

It uses a 'dissimilarity matrix' (comprising the measured differences between each stimulus in a stimulus set) to place each stimulus at a position (relative to the other stimuli) within a multidimensional space.

A typical analogy used to describe MDS is re-creating a geographical map from the measured distances between different cities - this would result in a perfect 2 dimensional solution, corresponding directly to the original geographical map.

The 'goodness-of-fit' (i.e. how well the data can be fit into the multidimensional space), tends to improve as the dimensionality of the solution is increased. It has two key measures:

- S-Stress - how well the map fits the data (how good the model is) (Neher et al. 2006). Low stress indicates a good fit (data has been mapped into object space with low compromise). High stress means either data input error, or not enough dimensions. Below 0.1 indicates a good fit (Davison 1983). Stress above 0.2 represents a poor fit (Kruskal 1964);

- RSQ - the squared correlation index i.e. the proportion of variance explained by the model (Neher et al. 2006). 0.6 is acceptable and 0.95 is a very confident solution. An RSQ improvement of below ~0.05 for an increase in dimensionality suggests that improvement is negligible (Davison 1983).

The number of dimensions that the data can be mapped onto is limited by the number of stimuli. Kruskal and Wish (1991) state that the number of stimuli should be at least four times larger than the number of dimensions to the solution, whereas Schiffman and Knecht (1993) suggest it is preferable to use 12 stimuli for two dimensional solutions and 18 for three-dimensional solutions.

The best solution often occurs where there is no improvement in stress/RSQ for an increase in dimensionality. If stress/RSQ is plotted against dimensionality, this is indicated by the 'knee' point.
With large numbers of stimuli, it is possible to derive high dimensionality solutions. However, Næs and Risvik (1996) suggest that usually two or three spatial dimensions are sufficient to reveal the most important relationships among the objects.

6.3.4 Verbal Elicitation

The multidimensional data/plots derived from MDS analysis may be interpreted in a number of ways: via the experimenter's own judgement, via suggestions from panelists (after having shown them the spatial configuration), via lists of the criteria upon which panelists made their judgements, or via a separate study conducted to generate additional information, such as an attribute rating experiment (Næs and Risvik 1996).

Regression of attributes/physical measurements with the dimensions of the MDS space can be done (to find how numerical changes in an attribute or physical factor are correlated with a dimension). This typically involves multiple regression which can be used to, one at a time, regress attributes against co-ordinates of the MDS dimensions. Vectors can be plotted, where angle indicates correlation with horizontal/vertical dimension and magnitude represents the magnitude of correlation (Næs and Risvik 1996).

Given the previous concerns regarding rating experiments, it would be preferable to use a method for this study that does not require the listeners to rate specific attributes. It would therefore remain that the MDS data could be evaluated either by the experimenter's own judgement, by the panelists themselves, or by the experimenter linking verbal data from the panelists to the MDS data. Having the experimenter decide what attributes are associated with the dimensions carries the same risk of bias as discussed earlier, and having the panelists interpret the data themselves would be expected to give more truthful results, yet would require further experimentation that would be complex and costly of time. If verbal data collected by the experimenter in the projective mapping stage of the test could be managed and used to interpret the MDS data in some meaningful way, it would mean that there is a) no experimenter bias, b) no group bias and c) no additional testing required beyond the projective mapping stage. It has been shown previously (for example in Nestrud and Lawless (2010)) that having subjects label the test stimuli, and/or identify groups, at the mapping stage provides verbal data that can be successfully combined with quantitative data in these types of tests (in that case subjects were asked to write down onto the mapping sheet words or attributes that described samples of cheese and apple).
CHAPTER 6. MAIN EXPERIMENTS

In the experiments outlined below, this approach will be attempted.

6.3.5 Verbal Analysis

The collection and preparation of verbal data in listening experiments tends to vary from case to case and there exists little in the way of specific recommended guidelines. If the approach mentioned above is to be taken, then a form of analysis that allows listeners' descriptions/labelling of stimuli, or the differences/similarities between them, to be compiled, assessed and quantified so that they may be used to interpret the MDS data, is required.

Attribute or description elicitation is common in listening tests, and may be carried out jointly by a panel of experts, as in QDA, or by individuals as in RGT and FCP. Triadic elicitation, where listeners are asked to describe how one of three sounds is different to the other two, and how the two are similar, can be used to derive a pair of terms which should implicitly be opposite in meaning. Pairwise elicitation, where listeners simply describe the difference between two sounds using a pair of opposite words or expressions is a more explicit way of deriving opposing terms. Finally, single term elicitation, where a subject provides one descriptor to a stimulus, or group of stimuli, can also be used (Nestrud and Lawless (2010)).

These methods assume that there is a connection between the sensation and the verbal description. However, as discussed in Choisel and Wickelmaier (2005), this is problematic because firstly the terms elicited are dependent upon there being an adequate label in the subject's lexicon and secondly it cannot be ensured that a listener actually experiences a sensation that they describe - it can only be assumed. Despite this, many experiments in this field have relied upon listeners describing sensations or nominating associated attributes and the results used to draw meaningful conclusions (for example: Williams (2010); Lee and Rumsey (2004)).

Typically, such as in the aforementioned methods (QDA, RGT and FCP), quantitative data regarding the elicited attributes is derived by rating stimuli on scales which are designed using the attributes as scale ends or labels - in Lee and Rumsey (2004), attributes elicited from the comparison of mono and phantom images were used to then measure the effects of various time and intensity differences, and source type, by having listeners rate stimuli according to the attributes. It is, however, also possible to measure the frequency of use in order to derive quantitative data - for example, the
number of times a stimulus has been labelled with a particular attribute (this technique featured in Williams (2010), where the occurrence of terms used to describe changes along a measured perceptual dimension were used to identify an appropriate label for that dimension).

This data can then be used to quantify the relationship between the stimuli and the elicited attributes.

Details of the method used in this experiment will be presented in the main experimental write-up, below, but it can be mentioned at this stage that a system based on frequency of occurrence, along with some smaller panel sessions to help categorise terms, will be used to analyse simple verbal data gathered at the end of the projective mapping stage.

6.3.6 Section Summary and Discussion

Summary

In discussing the multitude of experimental methods that can be used for listening experiments, the following parameters were considered to be important and evaluated where possible: risk of experimenter/listening panel bias, test fatigue and data-collection constraints. It was decided that a projective mapping technique in conjunction with MDS analysis (with an integrated labelling phase during the projective mapping) would be most appropriate; according to previous research, it should be time effective, beneficial to listener performance and non-inhibitive of the listener's opinion. This decision will be assessed with regard to R6 following analysis and discussion of the main experimental outcomes.

Discussion

This combination of test methodology and analysis has not (as far as the author is aware) been used before in any listening experiments. Therefore there is a limitation that the results cannot be directly compared to other tests for reference. However, the methods used have been evaluated and compared with reference to previous research and this suggests that results will be valid.

Any potential methodological improvements that become apparent during the following experimentation will be noted.
6.4 Experiment Set-Up

6.4.1 Test Design

Based on discussion in Section 6.3.2, it has been decided to utilise the projective mapping / Napping® technique for the main listening tests. The resulting data would then be used as input for MDS analysis.

As mentioned previously, a number of simulated listening environments, including three different loudspeaker directivities, two wall-separation distances, two absorption levels and two listener positions (constituting 24 different combinations), would be compared in order to identify the magnitude and type of perceptual attributes affected. Each of the 24 stimuli would be compared using three different audio excerpts: speech, dry music and reverberant music. The following excerpts were chosen:

- Speech - A dry recording of female speech - Track 49 from the European Broadcasting Union subjective quality assessment material (EBU-SQAM 1988);
- Dry music - By The Way - Red Hot Chilli Peppers (By The Way, 2002 Warner Bros.);
- Reverberant music - Violin Concerto No.1 in A Minor BWV 1041 (Andante) - Bach (Bach Violin Concertos, 1989 Naxos).

The speech excerpt had been used previously, the Red Hot Chilli Peppers excerpt a very dry, time-variant recording, and the Bach excerpt continuous, slow and reverberant - all were thought to be suitable for highlighting perceptual differences between reproductions.

Again, as described earlier, the simulated environments initially comprised two reflective side-walls, a source and a receiver; the distance between these walls and their absorption, the directivity of the source and the position of the speaker, the variables to be adjusted (see Figure 6.2 for graphical representation of two-boundary condition). The comparison of stimuli was repeated under two more conditions: with the addition of a front and rear reflective wall (four walls total) and with the addition of front, rear walls, floor and ceiling (six boundaries total). Construction details are as follows:
Figure 6.2: 3D representation of the two-boundary test condition (S1)

The Two-Boundary Condition

- CATT-Acoustic™ settings: diffuse surface reflections, 10'000 rays/octave, 1000ms ray truncation
- Exported impulse response: length = 1486ms at 48000 kHz

The four-boundary condition

- CATT-Acoustic™ settings: diffuse surface reflections, 10'000 rays/octave, 1000ms ray truncation
- Exported impulse response: length = 1486ms at 48000 kHz

The six-boundary condition

- Two side walls measuring 4m x 2.4m*
- Front and back walls measuring 2m (or 6m) x 2.4m
CHAPTER 6. MAIN EXPERIMENTS

- Floor and ceiling measuring 2m (or 6m) x 4m
- CATT-Acoustic™ settings: diffuse surface reflections, 10'000 rays/octave, 1000ms ray truncation
- Exported impulse response: length = 1486ms at 48000 kHz

*Ceiling height - 2.4m is the 'standard' UK domestic ceiling height.

Adaptation for this condition - some reflective situations required an extension of ray truncation time (≈1400ms) and longer outputted BRIRs (≈2000ms). Also, for 6m wall separation with reflective conditions, a ray truncation of 2000ms was used, and the BRIR length set to 3000ms at 48kHz.

6.4.2 Test System

For the three conditions, 61 BRIRs were outputted for each of the 24 environmental combinations (forming 24 BRIR 'databases' for use with the auralisation system, which convolves these impulse responses with a nominated audio excerpt). As before, the 61 BRIRs were based on 61 receiver angles (30 to -30 degrees with respect to the source, on-axis) and used so that the head tracker could be worn and used to select the relevant BRIRs based on the listener's head position. This was intended to aid realism within the virtual environment.

Gaillard (2009) has recently developed specialised software based upon the projective mapping technique, whereby the experimenter may assign different audio samples to playable/movable objects which can be moved within a two-dimensional space (on a computer). Results are automatically saved at set time intervals and can be saved easily to be used for MDS analysis. Previously, listeners had used test GUIs designed in MATLAB, which directly selected the BRIR database. However, with the availability of the software recently developed by Gaillard (2009), called TCL-LabX, it seemed appropriate to design a solution that allowed the projective mapping software to control the selection of BRIR database, so that listeners could compare and audition the 24 stimuli easily in one interface.

The TCL-LabX software is designed to play back an unlimited number of audio stimuli - users must import stereo .wav files to be assigned to the playback buttons. These are triggered when the listener presses the corresponding button.
Figure 6.3: Main experiment test system diagram
CHAPTER 6. MAIN EXPERIMENTS

The solution involves importing a range of sinusoidal tones into the TCL-LabX software. By evaluating the audio output from a computer running the software, the BRIR database chosen for convolution can be made to be dependent on the frequency of the tone triggered, and thus, listeners are able to control the auralised simulation via the software. This process is made possible using MATLAB's data acquisition tool box, which can be used to carry out real-time frequency analysis on incoming signals.

Now, listeners would be able to carry out the projective mapping procedure whilst utilising the benefits of the auralisation system (see Figure 6.3 for system diagram).

6.5 Two-Boundary Test (Condition 1)

6.5.1 Description

On arrival, the listener was presented with an instruction sheet to read before the test began (See Appendix D, Figures D.1 and D.2).

The 24 BRIR databases were loaded into the convolution software by the experimenter via a MATLAB script, which attributes a random number between 1-24 to each. The random number assigned to each of the databases was then saved.

Having been invited into the room being used, the listener is presented with the TCL-LabX software interface on the test laptop. 24 clickable items exist, which, when double-clicked on by a listener, produces a tone. The audio output of the laptop is routed to the main test computer's sound card, where a real-time FFT function in MATLAB analyses the incoming signal. As mentioned previously, specific tones were defined to be associated with certain stimuli and so the listener pressing different items in the projective mapping interface triggers different databases to be convolved. Figure 6.4 shows the test environment, and Figure 6.5 shows the main test setup.

Listeners were instructed to listen to all stimuli before carrying out the projective mapping task, in order to have a prior awareness of the range over which the stimuli are perceived to differ. They then proceeded to listen, compare and position the items relative to each other based upon similarity, within the space. Once the listeners were satisfied with the relative positions of the stimuli, they were able to add descriptive comments to the items (either individually or in groups) and, if more than one was labelled with the same comments, the most typical of that group could be defined. When
all items had been described or associated with a description, the listener could end the test. At this point, the experimenter saved the listener’s results (stimuli positions and descriptions) for decoding and analysis.

For the first condition (two-boundary model), 7 listeners, considered to be expert on account of their involvement with research of high quality audio, were invited to take part in the tests (several projective mapping experiments have had panel numbers in this range: Pagès (2005), Perrin et al. (2008), Pfeiffer and Gilbert (2008), Ramsgaard et al. (1994)). Tests took part in an edit suite at the University of Surrey and typically lasted between 25-45 minutes per audio excerpt (beyond 45 minutes, the convolution engine sometimes crashed, but was simply restarted with all of the listener’s projective mapping data preserved, and the original random number allocation restored). The auralisation system headphone output was set at a level deemed to be ‘comfortable’ by a small listening panel before the tests began, and remained at that level throughout all tests. It should also be noted that there was no loudness alignment between stimuli because it was expected that this was an important factor that would have changed according to the different combinations of independent variables.
Figure 6.5: Headphones and laptop showing the TCL-LabX software
6.5.2 Analysis

Quantitative Data from Projective Mapping

The 'Napped' positions of the stimuli were used to determine dissimilarity matrices for each excerpt/listener. This was achieved by calculating the Euclidean distance between each stimulus and entering the resulting value into the relevant element of the matrix. Here, stimuli which were perceived to be similar would have been placed closely together, with a small Euclidean distance between them and, therefore, a low amount of dissimilarity. Those that were far apart would represent a higher amount of dissimilarity.

At this point, two forms of MDS analysis can be utilised: Classic MDS (CMDS) which requires two-way data (just one dissimilarity matrix with rows and columns), and Weighted MDS (WMDS) which requires three-way data (multiple matrices with rows and columns that vary according to an additional factor, typically subject). WMDS is able to derive weightings and thus reflect the relative importance of each dimension to the subject (Næs and Risvik 1996). This is a particularly useful extension beyond CMDS as it can highlight variation amongst the test panel.

Also, both metric and non-metric forms of analysis were used. Metric analysis assumes that the data has been measured at an interval or ratio level, and is more quantitative than non-metric, where the dissimilarity data is treated as ordinal (only rank order is used to determine the spatial configuration). Research has shown that non-metric analysis of dissimilarities is sufficient to derive a spatial configuration that closely matches that based on metric analysis (Næs and Risvik 1996) and some consider metric as unsuitable for measurements made in perceptual space, as subject's proximities are unlikely to be absolute (Wickelmaier 2003).

A CMDS analysis was carried out for each excerpt, using a single aggregated dissimilarity matrix of all listeners' scores (determined by calculating the mean values across all listeners). The dissimilarity matrices for each listener were then used as input to a WMDS for each excerpt.

Finally, given that listeners were able to group stimuli together during the labelling stage of the test, the number of times that each pair of stimuli were grouped together (i.e., described by the same label) was calculated for each excerpt. These numbers could then be used to construct a similarity matrix. This was re-coded as a dissimilarity matrix.
CHAPTER 6. MAIN EXPERIMENTS

<table>
<thead>
<tr>
<th>Method</th>
<th>Analysis</th>
<th>Pop</th>
<th>Classical</th>
<th>Speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mapping</td>
<td>Metric WMDS</td>
<td>2</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.6699)</td>
<td>(0.88885)</td>
<td>(0.83278)</td>
</tr>
<tr>
<td></td>
<td>Non-Metric WMDS</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.73082)</td>
<td>(0.78863)</td>
<td>(0.86572)</td>
</tr>
<tr>
<td></td>
<td>Aggregated Metric CMDS</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.95854)</td>
<td>(0.9769)</td>
<td>(0.98225)</td>
</tr>
<tr>
<td></td>
<td>Aggregated Non-Metric CMDS</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.97217)</td>
<td>(0.93742)</td>
<td>(0.98908)</td>
</tr>
<tr>
<td>Sorting</td>
<td>Metric CMDS</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.83657)</td>
<td>(0.91115)</td>
<td>(0.91303)</td>
</tr>
<tr>
<td></td>
<td>Non-Metric CMDS</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.83542)</td>
<td>(0.91004)</td>
<td>(0.86623)</td>
</tr>
</tbody>
</table>

Table 6.4: Number of resulting dimensions based on the RSQ, and the associated RSQ value in brackets, for each of the MDS analysis types, for the two-boundary experiment.

(by subtracting each cell value from the highest value in the matrix) for each and then analysed via classic MDS (again, metric and non-metric).

In all cases, MDS analysis was carried out in order to determine the number of dimensions in which the data can best be represented spatially. The best solution in terms of number of dimensions for each analysis type (based on when negligible improvement occurs), is presented in Table 6.4 for the first condition (two-boundary).

Results summarised in this table suggest that:

- a good-fit, one dimensional solution exists for all excerpts with a non-metric Classic MDS analysis of listeners' aggregated data;
- generally, a reduction in solution dimensionality is seen with a non-metric analysis compared to the related metric analysis;
- solution dimensionality is generally highest for the sorted data analysis;
- Both metric and non-metric Weighted MDS, which take into account individual listener scores, generally show higher dimensionality solutions than the corresponding CMDS of aggregated matrices and also lower goodness of fit.

Carrying out a CMDS of aggregated dissimilarity data has been described as typical
in multidimensional analysis - in Næs and Risvik (1996), they consider a hypothetical average dissimilarity matrix to be analysed via CMDS in one of his examples (entries represent average dissimilarity ratings for the pairs of samples). Whether or not their solution is more or less valid than a WMDS with individual matrices is unclear. Gilbert and Heymann (1995) carried out a test with apple essences and analysed the average dissimilarity ratings via CMDS. Næs and Risvik (1996) re-analysed this data via 2/3 dimensional non-metric WMDS and stress/RSQ indicates a worse fit compared to the average group analysis. A wide variation in one of the dimensions was shown to exist between listeners - perhaps the individual differences are obscured by the aggregated-data CMDS. Which is 'correct' is still not clear though. As introduced at the start of this chapter, a key experimental aim is to acquire data that is applicable to a range of situations, and indeed listeners, and so the use of CMDS is justifiable on the basis that it will present an overview of all listeners' responses, which is more useful here than obtaining specific details about individual listeners.

It is possible, at this stage, to plot the stimuli against the one dimensional solution from the CMDS analysis of aggregated listener matrices as in Figure 6.6. This demonstrates how the different levels of each factor (directivity, absorption, position and wall separation) relate to the dimension for that particular excerpt and condition (note that only the plots for the Classical excerpt with two walls are given here to serve as an example - this presentation style is developed for the main presentation of results, later on). They were created in SPSS by taking the aggregated CMDS dimension value for each stimulus and plotting according to the directivity and one other variable of interest (the box plot data indicates the spread of results for that combination, across the remaining variables).

It is of interest to determine which attribute(s) are associated with this dimension (there may be overlapping, parallel dimensions) and their quantitative relationships with the dimension before presenting this data in full.

Qualitative Data

During the second-phase of the tests, listeners were asked to provide comments describing the projectively mapped stimuli (as in Nestrud and Lawless (2010)). Comments could be attached to singular stimuli, or to groups of stimuli. This meant that, by the end of testing, each stimulus, for each listener, condition and excerpt, would have some form of verbal description attached to it. In total, this presented 1512 descriptions, some of which
Figure 6.6: Boxplots showing the effect of different experimental parameters upon the relationship between directivity and the perceived dimension for the classical excerpt reproduced in condition S1 (two-boundary): (a) absorption, (b) listener position and (c) boundary separation.

contained multiple sentences (7 listeners x 3 excerpts x 3 conditions x 24 stimuli). These descriptions were compiled into a large spreadsheet, which, if analysed, could be used to identify the key attributes used and the relationship between their usage and the stimuli. Meaningful data would allow us to interpret the spatial configurations without having to conduct an additional attribute rating test - thus, successfully carrying out quantitative
and qualitative assessment of audio stimuli using just the projective mapping method with MDS.

Analysis

To begin with, the text from the large spreadsheet (comprising all descriptions used in the main experiments) was input into an online text analysis software (Voyeur http://voyeur.hermeneuti.ca/). The frequencies of all words (598 used in total) were calculated and ranked, which is common for elicited terms (see Martens and Sungyoung (2007)). The experimenter then removed non-sensical singular letters (like ‘a’, ‘s’, ‘f’), before selecting only terms relating to audio/sound (i.e. removing any non-contextual words, such as descriptive terms like ‘fairly’, ‘very’ or linking words such as ‘is’, ‘and’, ‘with’ etc).

A small panel comprising 3 listeners who took part in the listening test were then asked to take part in a verbal analysis session in a seminar room at the University of Surrey. They were asked by the experimenter firstly to go through each of the chosen audio/sound-related words and to group them according to association, whilst at the same time finding any synonyms that existed (such as ‘close’ and ‘near’, or ‘high frequency’ and ‘hf’). The words were printed out onto paper so that the panel could easily move and group them.

For any synonyms found, the panel were asked to select the most appropriate term to be used, and this was noted so that the original descriptions could be updated by the experimenter. This third-party attribute-reduction is a variation of the method used by Williams (2010), who used a single expert to group elicited terms in this way.

Finally, the panel was asked to place as many of the elicited terms as possible into groups based on if they thought the terms described the same perceptual experience or that they were related to one perceptual effect (for example, ‘wide’, ‘width’ and ‘narrow’ would be grouped together on the basis that they were connected to image width). If it was possible to establish a rank order between the terms in that group, for example, from one extreme of an effect to another (i.e. wide to narrow) then the panel did so. The results of this are shown in Table 6.5, with indication as to whether the terms were able to be ordered or not.

Once terms had been assigned to groups/scales, the total usage frequency for all of the terms in each group or scale was calculated. This gave an indication of ‘strength’ of that group/scale with respect to the overall experimental condition and is a variation of
### Table 6.5: Groups defined by panel members in verbal analysis session.

<table>
<thead>
<tr>
<th>Group</th>
<th>Words</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Timbral Descriptors</strong> (Order Attempted)</td>
<td>timbre, timbral, tonal, tone, air, harsh, splaszy, sharp, splashier, bright, thin, crisp, brighter, tinny, crispy, brightness, thinner, nasal, telephone-like, punchy, warm, muffled, muddy, muted, grungy, boomy, thud, booming, dull, duller</td>
</tr>
<tr>
<td><strong>Width (Ordered)</strong></td>
<td>broad, broader, widest, wider, wide, widened, widths, medium-wide, mid-width, narrowish, narrower, narrow</td>
</tr>
<tr>
<td><strong>Filtering Terms</strong></td>
<td>filtered, filtering, filter, ringing, band-passed, ringy, band-pass, resonances, band, limited, resonance, resonant, notch, hollow, band-limited, comb-filtered, boost, boosts</td>
</tr>
<tr>
<td><strong>Clarity (Ordered)</strong></td>
<td>clarity, detail, focused, clear, distinct, defined, precise, direct, wash, diffuse</td>
</tr>
<tr>
<td><strong>Artifacts (Ordered)</strong></td>
<td>processing, artefact, colouration, colour, normal, natural, real, neutral, sibilance, sibilant, flanger, phasey, chorus, chorussy, fake, coloured, cheap, distorted, unnatural, strange</td>
</tr>
<tr>
<td><strong>Pleasantness</strong>       (Ordered)</td>
<td>pleasant, enjoyable, unpleasant</td>
</tr>
<tr>
<td><strong>Proximity</strong> (Ordered)</td>
<td>location, distance, distance-wise, further, far, distant, mid-distance, close-mic’d, closer, close, intimate</td>
</tr>
<tr>
<td><strong>Frequency Response</strong> (Ordered)</td>
<td>spectrum, spectrally, spectral, frequency, frequencies, response, eq, low-frequency, low-pass, bass, lf-heavy, bass-heavy, bottom, mid-low, lower-mid-frequency, mid-frequencies, midly, mid, mid-range-focused, mid-hf, high-mids, upper-mid, mid-high, high, high-frequencies, high-frequency, topyp, top, treble-heavy, treble, top-end</td>
</tr>
<tr>
<td><strong>Flatness of Frequency Response (Ordered)</strong></td>
<td>homogenous, balance, equal, balanced, level, flatter, flat, smoother, tilt, tilted</td>
</tr>
<tr>
<td><strong>Loudness</strong> (Ordered)</td>
<td>loudness, loudest, louder, loud, quiet, quieter</td>
</tr>
<tr>
<td><strong>Prominence</strong> (Unordered)</td>
<td>foremost, prominent, forward, dominant, pronounced</td>
</tr>
<tr>
<td><strong>Room Description</strong> (Unordered)</td>
<td>environment, space, rooms, room, room-affected, stadium, arena, hall, bathroom</td>
</tr>
<tr>
<td><strong>Spaciousness</strong> (Ordered)</td>
<td>spaciousness, spatial, larger, spacious, large, roomy, medium-sized, small, boxy</td>
</tr>
<tr>
<td><strong>Reverberation</strong> (Ordered)</td>
<td>echo, reverb, reverberation, echoic, reflections, reflective, reflection, reverberant, wetter, wet, echoes, absorption, unreflective, anechoic, dead, dry</td>
</tr>
<tr>
<td><strong>Envelopment</strong> (Unordered)</td>
<td>envelopment, immersive, engulf, enveloping</td>
</tr>
</tbody>
</table>

*a closely tied to Spaciousness*
Table 6.6: Total frequency of words in groups/scales defined by panel members in verbal analysis session (two-boundary).

<table>
<thead>
<tr>
<th>Group</th>
<th>Total Frequency of Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency Response (Ordered)</td>
<td>214</td>
</tr>
<tr>
<td>Spaciousness (Ordered)</td>
<td>196</td>
</tr>
<tr>
<td>Timbral Descriptors (Order Attempted)</td>
<td>175</td>
</tr>
<tr>
<td>Artifacts (Ordered)</td>
<td>172</td>
</tr>
<tr>
<td>Width (Ordered)</td>
<td>122</td>
</tr>
<tr>
<td>Proximity (Ordered)</td>
<td>113</td>
</tr>
<tr>
<td>Reverberation (Ordered)</td>
<td>94</td>
</tr>
<tr>
<td>Room Description (Ordered)</td>
<td>76</td>
</tr>
<tr>
<td>Clarity (Ordered)</td>
<td>67</td>
</tr>
<tr>
<td>Filtering Terms (Unordered)</td>
<td>62</td>
</tr>
<tr>
<td>Loudness (Ordered)</td>
<td>32</td>
</tr>
<tr>
<td>Flatness of Frequency Response (Ordered)</td>
<td>30</td>
</tr>
<tr>
<td>Prominence (Unordered)</td>
<td>18</td>
</tr>
<tr>
<td>Envelopment (Unordered)</td>
<td>13</td>
</tr>
<tr>
<td>Pleasantness (Ordered)</td>
<td>5</td>
</tr>
</tbody>
</table>

* closely tied to Spaciousness

the technique used by Williams (2010) to measure ‘overall prominence’ of term groups. Results are shown in Table 6.6.

In observing the frequency of terms, it was noticeable that, for most groups/scales, a pair of highly used antonyms existed. For example, for the scale labelled ‘Width’, the highest used terms are ‘Wide’ (34) and ‘Narrow’ (68), and these are used much more than any other terms within that group. These antonyms are identified in Table 6.7 and their prominence compared to other terms used in that group is shown in Figures 6.7 to 6.12. Note that, in each figure, the highest ranked term after the top antonym pair is also included to demonstrate the relative occurrence of the top antonyms compared to other highly used terms. In this case, for two boundaries, the results confirm the clear dominance of wide/narrow, close/distant and loud/quiet. The other top antonyms are less dominant compared to other terms.
Table 6.7: High-frequency antonyms and the relative percentage of each term within its group (two-boundary)

<table>
<thead>
<tr>
<th>Antonym 1</th>
<th>Freq</th>
<th>%</th>
<th>Antonym 2</th>
<th>Freq</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wide</td>
<td>34</td>
<td>28</td>
<td>Narrow</td>
<td>68</td>
<td>56</td>
</tr>
<tr>
<td>Natural</td>
<td>49</td>
<td>28</td>
<td>Unnatural</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>Distant</td>
<td>42</td>
<td>37</td>
<td>Close</td>
<td>61</td>
<td>54</td>
</tr>
<tr>
<td>Loud</td>
<td>15</td>
<td>47</td>
<td>Quiet</td>
<td>9</td>
<td>28</td>
</tr>
<tr>
<td>Spacious</td>
<td>59</td>
<td>30</td>
<td>Small</td>
<td>61</td>
<td>31</td>
</tr>
<tr>
<td>Spacious</td>
<td>59</td>
<td>30</td>
<td>Boxy</td>
<td>48</td>
<td>24</td>
</tr>
<tr>
<td>Reverberant</td>
<td>9</td>
<td>10</td>
<td>Dry</td>
<td>34</td>
<td>36</td>
</tr>
</tbody>
</table>

Figure 6.7: Number of occurrence of top antonyms and next highest term in the Width group (two-boundary)
Figure 6.8: Number of occurrence of top antonyms and next highest term in the *Artifacts* group (two-boundary)

Figure 6.9: Number of occurrence of top antonyms and next highest term in the *Proximity* group (two-boundary)
CHAPTER 6. MAIN EXPERIMENTS

Figure 6.10: Number of occurrence of top antonyms and next highest term in the Loudness group (two-boundary)

Figure 6.11: Number of occurrence of top antonyms (spacious vs boxy/small) and next highest term in the Spaciousness group (two-boundary)
If the text from each experimental case is analysed according to these group terms, some way of measuring the relationship between each stimulus and group/scale term may be possible - a simple solution would be to count the occurrences of each antonym and to assign them a positive and negative score, respectively. The difference in total scores for each stimulus could then be observed - a stimulus considered to be at one end of the scale, say 'Wide', would have a high positive score, and a stimulus considered to be at the other end, say 'Narrow' would have a high negative value.

It is likely, however, that listeners used varying magnitudes of the antonyms, and so if occurrences of the various levels are assigned different weightings, a more accurate solution can be implemented. The scaling words (i.e. ‘quite’, ‘very’, ‘fairly’ etc) from the text were identified and ranked and it was found that the most frequently used were 'quite', 'very', 'fairly' and 'slightly'.

Each of these featured in a paper by Rohrmann (2007), who carried out a number of experiments in order to present quantitative information about commonly used English scale labels (see Table 6.8). These words featured in the ‘Intensity’ category, and their mean chosen values along a 0-10 scale were given (see Table 6.9).

As pairs of antonyms have been determined, each of which may be described as being
<table>
<thead>
<tr>
<th>Term</th>
<th>0-10 scale Mean Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>a little</td>
<td>2.5</td>
</tr>
<tr>
<td>average</td>
<td>4.8</td>
</tr>
<tr>
<td>completely</td>
<td>9.8</td>
</tr>
<tr>
<td>considerably</td>
<td>7.6</td>
</tr>
<tr>
<td>extremely</td>
<td>9.6</td>
</tr>
<tr>
<td>fairly</td>
<td>5.3</td>
</tr>
<tr>
<td>fully</td>
<td>9.4</td>
</tr>
<tr>
<td>hardly</td>
<td>1.5</td>
</tr>
<tr>
<td>highly</td>
<td>8.6</td>
</tr>
<tr>
<td>in-between</td>
<td>4.8</td>
</tr>
<tr>
<td>mainly</td>
<td>6.8</td>
</tr>
<tr>
<td>medium</td>
<td>4.9</td>
</tr>
<tr>
<td>moderately</td>
<td>5.0</td>
</tr>
<tr>
<td>not</td>
<td>0.4</td>
</tr>
<tr>
<td>not at all</td>
<td>0.0</td>
</tr>
<tr>
<td>partly</td>
<td>3.5</td>
</tr>
<tr>
<td>quite</td>
<td>5.9</td>
</tr>
<tr>
<td>quite a bit</td>
<td>6.5</td>
</tr>
<tr>
<td>rather</td>
<td>5.8</td>
</tr>
<tr>
<td>slightly</td>
<td>2.5</td>
</tr>
<tr>
<td>somewhat</td>
<td>4.5</td>
</tr>
<tr>
<td>very</td>
<td>7.9</td>
</tr>
<tr>
<td>very much</td>
<td>8.7</td>
</tr>
</tbody>
</table>

Table 6.8: "Intensity" Qualifiers - Mean Value on a 0-10 scale from Rohrmann (2007)

<table>
<thead>
<tr>
<th>Term</th>
<th>Total Frequency</th>
<th>0-10 scale Mean Value (Rohrmann, 2007)</th>
</tr>
</thead>
<tbody>
<tr>
<td>quite</td>
<td>251</td>
<td>5.9</td>
</tr>
<tr>
<td>very</td>
<td>243</td>
<td>7.9</td>
</tr>
<tr>
<td>fairly</td>
<td>158</td>
<td>5.3</td>
</tr>
<tr>
<td>slightly</td>
<td>153</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Table 6.9: Frequency of top scaling or ‘intensity’ words and their perceived position on a 0-10 scale according to Rohrmann (2007)
Weighting (Scale Value) | Term | Example
--- | --- | ---
7.9 | very ‘+ve antonym’ | Very Wide
5.9 | quite ‘+ve antonym’ | Quite Wide
5.3 | fairly ‘+ve antonym’ | Fairly Wide
5 | ‘+ve antonym’ | Wide
2.5 | slightly ‘+ve antonym’ | Slightly Wide
0 | Neither antonym | Neither Wide nor Narrow
-2.5 | slightly ‘-ve antonym’ | Slightly Narrow
-5 | ‘-ve antonym’ | Narrow
-5.3 | fairly ‘-ve antonym’ | Fairly Narrow
-5.9 | quite ‘-ve antonym’ | Quite Narrow
-7.9 | very ‘-ve antonym’ | Very Narrow

Table 6.10: Weighting scale used to quantify relationship of various antonyms with each stimulus

‘slightly’, ‘very’ etc, using these ‘intensity’ qualifiers, it would be appropriate to use these as positive and negative scale values. Hence, occurrence of neither term has a value of zero (i.e. neither ‘wide’ nor ‘narrow’), and the occurrence of ‘very’ has a value of 7.9 or -7.9, depending on which term it precedes. Based on Rohrmann’s weightings, it would seem logical to associate the term itself with a value of 5 on the scale. The scales used are shown in Table 6.10.

The design of these scales is an adaptation of typical scale construction methods found in Martens and Sungyoung (2007), for example, where the most frequently used elicited antonyms form the scales used for the rating of stimuli.

For each experimental case (of which there were 9: Classical, Pop and Speech excerpts reproduced in between two, four and six boundary conditions), the total value for antonym and magnitude term across listeners were summed per stimulus. This meant that, for a given stimulus, a numerical value relating to each antonym and therefore attribute scale could be achieved (i.e. for a classical excerpt reproduced via stimulus 23 in the two-boundary condition, a single value for ‘Width’ could be calculated).

The attribute values for each stimulus could now be plotted against the singular dimension values found earlier (using an aggregated dissimilarity matrix input to CMDS). This allows for the correlation between each attribute and the singular dimension to be calculated, thus providing information as to what the true dimension label may be. Table 6.11 shows the correlation values for the attributes and dimension
CHAPTER 6. MAIN EXPERIMENTS

Table 6.11: $R^2$ Correlation between attribute values and the single dimension output from CMDS of the aggregated dissimilarity matrices for the two-boundary condition

<table>
<thead>
<tr>
<th>Excerpt</th>
<th>Wide / Narrow</th>
<th>Natural / Unnatural</th>
<th>Distant / Close</th>
<th>Loud / Quiet</th>
<th>Spacious / Small</th>
<th>Spacious / Boxy</th>
<th>Reverb. / Dry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical</td>
<td>0.81</td>
<td>0.02</td>
<td>0.57</td>
<td>0.2</td>
<td>0.3</td>
<td>0.57</td>
<td>0.02</td>
</tr>
<tr>
<td>Pop</td>
<td>0.84</td>
<td>0.07</td>
<td>0.29</td>
<td>0.69</td>
<td>0</td>
<td>0.77</td>
<td>0.29</td>
</tr>
<tr>
<td>Speech</td>
<td>0.34</td>
<td>0</td>
<td>0.08</td>
<td>0.56</td>
<td>0.02</td>
<td>0.18</td>
<td>0.76</td>
</tr>
</tbody>
</table>

for each case and it appears that width is highly correlated to the dimension for classical and pop excerpts, spaciousness for the pop excerpt and reverberance for speech. Note that, in this study, 'high correlation' relates to a correlation of 0.7 or more - this is in line with a study by Mattila (2002), who referred to correlations between MDS dimensions and direct attribute data in the same way (he described 0.9 and above as 'strong', around 0.7 as 'high', around 0.6 as 'moderate' and around 0.5 as 'low').

At this point, it should be reiterated that the analysis method used comprises a range of techniques that, as far as the author is aware, have not been combined in such a way before. There are also limitations in that the antonym pairs that formed the scales were selected by the experimenter, the scoring was indirect (values were not assigned to the attributes directly by listeners) and the general nature of the analysis is based on inferred outcomes. However, the techniques used have been justified with reference to other research in this field and the test is not affected by the bias of rating tests, discussed earlier. It is felt that the method used, although novel, will lead to an insight into the main perceptual effects which is as robust as one where prescribed attributes and scales were used - as earlier, the aim here is to obtain information that is applicable and useful to a range of situations, and it is expected that this method is capable of presenting that type of data.

Consideration of Timbre and Spectral Differences

Despite the 'Timbral Descriptors' and 'Frequency Response' group accounting for a substantial number of the words used by listeners, analysis using the method above was more difficult. In the 'Timbral Descriptors' group, the two terms that could be used as key antonyms were less clear: bright was a highly used term in that group (18 times), and three possible antonyms: muffled, boomy and dull were all used similar amounts of times (29, 12 and 13 respectively). The usages here are much lower than width/narrow
CHAPTER 6. MAIN EXPERIMENTS

<table>
<thead>
<tr>
<th>Excerpt</th>
<th>Bright/ Muffled</th>
<th>Bright / Boomy</th>
<th>Bright/ Dull</th>
<th>Bright / Muffled, Boomy and Dull</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical</td>
<td>0.57</td>
<td>0.02</td>
<td>0.39</td>
<td>0.36</td>
<td>0.01</td>
</tr>
<tr>
<td>Pop</td>
<td>0.66</td>
<td>0.39</td>
<td>0.35</td>
<td>0.56</td>
<td>0.3</td>
</tr>
<tr>
<td>Speech</td>
<td>0.71</td>
<td>0.12</td>
<td>0.13</td>
<td>0.57</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Table 6.12: $R^2$ Correlation between attribute values (frequency of use in the case of 'Frequency') and the single dimension output from CMDS of the aggregated dissimilarity matrices for the two-boundary experiment.

(34/68) and distant/close (42/61) and thus the language used to describe these effects was clearly less common. The 'Frequency Response' group contained technical terms only (i.e. bass (54) and high-frequencies (4), and typically required an accompanying descriptor, for example: 'lots of bass' or 'boosted high-frequencies'. The variation of the descriptors used makes it very difficult to analyse using the method above; 'very bass', 'quite bass' or 'slightly high-frequency', clearly doesn't work.

Despite these difficulties, some analysis was attempted. Using the same method as above, the correlation between the position of each stimuli along the dimension and the weighting for each stimuli with regards to bright/muffled, bright/boomy, bright/dull and bright/muffled and boomy and dull was measured (see Table 6.12). Results showed notable correlation (more than $r$ squared = 0.57) for each excerpt here with the antonyms bright/muffled. This is shown in Figure 6.13.

An initial analysis of 'Frequency Response' was carried out by counting the total use of 'bass' and 'high-frequencies' for each stimuli and correlating that with the dimension. This showed little notable correlation (see Table 6.12). An extended analysis was then carried out, whereby the experimenter assessed the descriptor attached to each stimulus by each listener for each case with respect to whether it suggested a perceived bass increase (descriptions such as 'boomy', 'bass-heavy', 'bassy') or decrease (descriptions such as 'toppy'). A description thought to be related to bass increase was attributed with a -5 on the scale, no mention was attributed a zero and a bass decrease attributed a 5. Changes according to these parameters also showed little correlation to the difference in the dimension for each stimuli (see Table 6.13).
Detailed Analysis Using a Combination of Quantitative and Qualitative Results

Correlation between the variation in these attributes, and the dimensional output from MDS carried out earlier, provides insight into the relationship between the different stimuli and their effect upon the listeners' perception of reproduction.

So that the initial research questions may be considered, it is necessary to observe how the specific combination of different directivities, wall absorption/separation conditions and listener positions relate to the various attributes found to correlate to the dimensions.

Boxplots showing the changes along each dimension for directivity versus wall absorption, wall separation and listener position are presented, along with plots indicating the correlation of each attribute to the dimension (Figures 6.14, 6.15 and 6.16). This allows the relative differences between each variable with respect to the dimension to be compared, and for the attributes likely connected to that dimension to be considered. The correlation of each attribute is labelled according to the antonym associated with the positive end of the dimension (i.e. ‘Narrow’ with an r-squared value of 0.8 suggests that a positive change in the dimension is highly correlated with the stimuli being perceived
Excerpt | Bass increase/decrease
---|---
Classical | 0.0007
Pop | 0.41777
Speech | 0.03711

Table 6.13: $R^2$ Correlation between bass increase/decrease and the single dimension output from CMDS of the aggregated dissimilarity matrices for the two-boundary condition.

as more narrow).
(a) $R^2$ Correlation of attributes with perceptual dimension

(b) Absorption - classical, two-boundary

(c) Position - classical, two-boundary

(d) Wall-separation - classical, two-boundary

Figure 6.14: Boxplots showing the effect of different experimental parameters upon the relationship between directivity and the perceptual dimension for the classical excerpt reproduced in condition S1 (two-boundary): (b) absorption, (c) listener position and (d) boundary separation. Subfigure (a) shows the $R^2$ correlation of various attributes with the perceptual dimension, the label indicating the direction of a positive change in that dimension (i.e. increase in dimension is well correlated to the stimuli being perceived as more narrow).
CHAPTER 6. MAIN EXPERIMENTS

180

Attribute Correlation - Pop Two Walls

Wide Unnatural Close Loud Spacious Spacious Reverb.

(a) Correlation of attributes with perceptual dimension

(b) Absorption - pop, two-boundary

Figure 6.15: Boxplots showing the effect of different experimental parameters upon the relationship between directivity and the perceptual dimension for the pop excerpt reproduced in condition S1 (two-boundary): (b) absorption, (c) listener position and (d) boundary separation. Subfigure (a) shows the correlation of various attributes with the perceptual dimension, the label indicating the direction of a positive change in that dimension (i.e. increase in dimension is well correlated to the stimuli being perceived as wider).
CHAPTER 6. MAIN EXPERIMENTS

181

Figure 6.16: Boxplots showing the effect of different experimental parameters upon the relationship between directivity and the perceptual dimension for the speech excerpt reproduced in condition S1 (two-boundary): (b) absorption, (c) listener position and (d) boundary separation. Subfigure (a) shows the correlation of various attributes with the perceptual dimension, the label indicating the direction of a positive change in that dimension (i.e. increase in dimension is well correlated to the stimuli being perceived as more dry).
Chapter 6. Main Experiments

<table>
<thead>
<tr>
<th>Excerpt</th>
<th>Sig. Factors/Interactions</th>
<th>df</th>
<th>F</th>
<th>Sig.</th>
<th>$\eta^2$</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical</td>
<td>Directivity</td>
<td>2</td>
<td>68.985</td>
<td>.000</td>
<td>.885</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Absorption</td>
<td>1</td>
<td>7.281</td>
<td>.015</td>
<td>.288</td>
<td>.723</td>
</tr>
<tr>
<td></td>
<td>Wall Separation</td>
<td>1</td>
<td>28.070</td>
<td>.000</td>
<td>.609</td>
<td>.999</td>
</tr>
<tr>
<td>Pop</td>
<td>Directivity</td>
<td>2</td>
<td>43.089</td>
<td>.000</td>
<td>.827</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Absorption</td>
<td>1</td>
<td>7.447</td>
<td>.015</td>
<td>.293</td>
<td>.733</td>
</tr>
<tr>
<td></td>
<td>Wall Separation</td>
<td>1</td>
<td>20.118</td>
<td>.000</td>
<td>.528</td>
<td>.989</td>
</tr>
<tr>
<td>Speech</td>
<td>Directivity</td>
<td>2</td>
<td>61.754</td>
<td>.000</td>
<td>.873</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Absorption</td>
<td>1</td>
<td>21.791</td>
<td>.000</td>
<td>.548</td>
<td>.993</td>
</tr>
<tr>
<td></td>
<td>Wall Separation</td>
<td>1</td>
<td>17.457</td>
<td>.001</td>
<td>.492</td>
<td>.997</td>
</tr>
</tbody>
</table>

Table 6.14: Significant Factors - Main Experiment (two-boundary)

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Factor</th>
<th>Post-Hoc Bonferroni</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical</td>
<td>Directivity</td>
<td>Omni.</td>
</tr>
<tr>
<td></td>
<td>Omni.</td>
<td>-.607</td>
</tr>
<tr>
<td></td>
<td>Cardioid</td>
<td>-</td>
</tr>
<tr>
<td>Pop</td>
<td>Directivity</td>
<td>Omni.</td>
</tr>
<tr>
<td></td>
<td>Omni.</td>
<td>-.241</td>
</tr>
<tr>
<td></td>
<td>Cardioid</td>
<td>-</td>
</tr>
<tr>
<td>Speech</td>
<td>Directivity</td>
<td>Omni.</td>
</tr>
<tr>
<td></td>
<td>Omni.</td>
<td>-.117</td>
</tr>
<tr>
<td></td>
<td>Cardioid</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 6.15: Further Effects - Main Experiment (two-boundary)

ANOVA

A UniANOVA was carried out on the data used for the boxplots (before the analysis was carried out, results were assessed and found to normally distributed and hence the assumptions for ANOVA are satisfied - see D.1 in Appendix D). Significant factors and interactions for each audio excerpt are presented in Table 6.14 and further effects in 6.15.
Observations of the boxplots and results from the ANOVA suggest that:

_for the classical excerpt (two-boundary condition),_

- the perceptual dimension was found to be most highly correlated with the following attributes: width (correlation of 0.81), distance (0.57), and spaciousness (0.57);
- the directivity had a large effect on the perceptual dimension, and hence is correlated with the terms above (a narrowing in directivity may be related to a perceived narrowing in width, increase in distance and reduction in spaciousness);
- higher absorption leads to the reproduction being perceived as higher on the perceptual dimension (which may relate to it being narrower, farther and more boxy) compared to lower absorption;
- listener position has no significant effect;
- greater wall separation leads to the reproduction being higher on the perceptual dimension (which may relate to it being narrower, farther and more boxy) for the omni and cardioid loudspeakers, compared to lesser wall separation.

_for the pop excerpt (two-boundary condition),_

- the perceptual dimension was found to be most highly correlated with the following attributes: width (correlation of 0.84), loudness (0.69), and spaciousness (0.77);
- the directivity had a large effect on the perceptual dimension, and hence is correlated with the terms above (a narrowing in directivity may be related to a perceived narrowing in width, decrease in loudness and reduction in spaciousness);
- for omni and cardioid directivities, higher absorption leads to the reproduction being perceived as lower on the perceptual dimension (which may relate to it being narrower, quieter and more boxy) compared to lower absorption;
- according to the boxplots, position has little effect on the difference between reproductions via omni and cardioid directivities. However, for reproduction via narrow beam directivities, it appears that the left position is perceived to be lower on the perceptual dimension (which may relate to it being narrower, quieter and more boxy) than the central position. ANOVA results suggest that position has no statistically significant effect;
• greater wall separation leads to the reproduction being perceived as lower on the perceptual dimension (which may relate to it being narrower, quieter and more boxy) for the omni and cardioid loudspeakers, compared to lesser wall separation.

for the speech excerpt (two-boundary condition),

• the perceptual dimension was found to be most highly correlated with the following attributes: reverberance (correlation of 0.76) and loudness (0.56);
• the directivity had a large effect on the perceptual dimension, and hence is correlated with the terms above (a narrowing in directivity may be related to a perceived reduction in reverberance and decrease in loudness);
• for omni and cardioid directivities, higher absorption leads to the reproduction being perceived as higher on the perceptual dimension (which may relate to it being less reverberant and quieter) compared to lower absorption;
• listener position has no significant effect;
• greater wall separation leads to the reproduction being perceived as higher on the perceptual dimension (which may relate to it being less reverberant and quieter), compared to lesser wall separation. This effect is most evident when comparing the reproduction between narrow beam directivities.

across all excerpts (two-boundary condition),

• a narrowing in directivity (increased directivity index) corresponds to a change in the perceptual dimension;
• changes in the dimension are most correlated with width when reproducing classical music, width and spaciousness (boxy) with pop, and reverberance with speech;
• absorption has a similar relationship with the perceptual dimension for all excerpts;
• the greatest difference in perception as a result of position occurs when comparing narrow directivity reproductions with pop music, though ANOVA results suggest it has no significant effect;
• the greatest difference in perception as a result of wall separation occurs when comparing omni and cardioid directivities with classical and pop music and narrow directivity reproductions with speech.
### Table 6.16: Number of resulting dimensions based on the RSQ, and the associated RSQ value in brackets, for each of the MDS analysis types, for the four-boundary experiment.

<table>
<thead>
<tr>
<th>Method</th>
<th>Analysis</th>
<th>Pop</th>
<th>Classical</th>
<th>Speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mapping</td>
<td>Metric WMDS</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>(0.79426)</td>
<td></td>
<td>(0.66192)</td>
<td>(0.85394)</td>
</tr>
<tr>
<td></td>
<td>Non-Metric WMDS</td>
<td>2</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>(0.83032)</td>
<td></td>
<td>(0.94006)</td>
<td>(0.88336)</td>
</tr>
<tr>
<td></td>
<td>Aggregated Metric CMDS</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(0.99337)</td>
<td></td>
<td>(0.961)</td>
<td>(0.98593)</td>
</tr>
<tr>
<td></td>
<td>Aggregated Non-Metric CMDS</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(0.95004)</td>
<td></td>
<td>(0.92255)</td>
<td>(0.99043)</td>
</tr>
<tr>
<td>Sorting</td>
<td>Metric CMDS</td>
<td>5</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>(0.84274)</td>
<td></td>
<td>(0.75503)</td>
<td>(0.83911)</td>
</tr>
<tr>
<td></td>
<td>Non-Metric CMDS</td>
<td>4</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>(0.84638)</td>
<td></td>
<td>(0.7784)</td>
<td>(0.76076)</td>
</tr>
</tbody>
</table>

Table 6.16: Number of resulting dimensions based on the RSQ, and the associated RSQ value in brackets, for each of the MDS analysis types, for the four-boundary experiment.

### 6.6 Four-Boundary Test (Condition 2)

#### 6.6.1 Description

The second condition tests (four-boundary model) were carried out with 7 expert listeners (all the same as for the first condition, except one) in the listening room at the University of Surrey, following the same procedure as above.

#### 6.6.2 Analysis

As for the two-boundary test, a range of MDS analyses was carried out in order to determine the number of dimensions in which the data can best be represented spatially. The best solution in terms of number of dimensions for each analysis type (based on when negligible improvement occurs), is presented in Table 6.16 for the second condition (four-boundary).

Results summarised in this table suggest that:

- a good-fit, one dimensional solution exists for all excerpts with a non-metric Classic MDS analysis of listeners' aggregated data;
\begin{table}
\centering
\begin{tabular}{|l|c|}
\hline
\textbf{Group} & \textbf{Total Frequency of Words} \\
\hline
\textit{Frequency Response} (Ordered) & 270 \\
\textit{Artifacts} (Ordered) & 250 \\
\textit{Proximity} (Ordered) & 193 \\
\textit{Reverberation} (Ordered) & 185 \\
\textit{Width} (Ordered) & 168 \\
\textit{Spaciousness} (Ordered) & 161 \\
\textit{Timbral Descriptors} (Order Attempted) & 128 \\
\textit{Flatness of Frequency Response} (Ordered) & 77 \\
\textit{Loudness} (Ordered) & 64 \\
\textit{Filtering Terms} (Unordered) & 63 \\
\textit{Room Description} (Ordered) & 51 \\
\textit{Clarity} (Ordered) & 42 \\
\textit{Prominence} (Unordered) & 7 \\
\textit{Pleasantness} (Ordered) & 0 \\
\textit{Envelopment} (Unordered) & 0 \\
\hline
\end{tabular}
\end{table}

\textsuperscript{a} closely tied to Spaciousness

Table 6.17: Total frequency of words in groups/scales defined by panel members in verbal analysis session (four-boundary).

- generally, a reduction in solution dimensionality is seen with a non-metric analysis compared to the related metric analysis;

- solution dimensionality is generally highest for the sorted data analysis;

- Both metric and non-metric Weighted MDS, which take into account individual listener scores, generally show higher dimensionality solutions than the corresponding CMDS of aggregated matrices and also lower goodness of fit.

As in the two-boundary test, the data can be represented well in one dimension (for pop and classical, results for aggregated metric CMDS are slightly higher, but the benefits of having a single dimension more than make up for the small improvement in RSQ) and the antonyms most frequently used are the same (see Tables 6.17 and 6.18, and Figures D.3 to D.8 in the Appendix). As a result, the same analysis of the qualitative (verbal) data was used so that the correlation of the attributes and the dimension could be calculated and plotted. For the four-boundary test, the correlation values are shown in Table 6.21 and plotted in Figures 6.17, 6.18 and 6.19.

Analysis of timbral attributes, frequency response and bass increase/decrease, using the
Table 6.18: High-frequency antonyms and the relative percentage of each term within its group (four-boundary)

<table>
<thead>
<tr>
<th>Antonym 1</th>
<th>Freq</th>
<th>%</th>
<th>Antonym 2</th>
<th>Freq</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wide</td>
<td>65</td>
<td>39</td>
<td>Narrow</td>
<td>45</td>
<td>27</td>
</tr>
<tr>
<td>Natural</td>
<td>48</td>
<td>19</td>
<td>Unnatural</td>
<td>46</td>
<td>18</td>
</tr>
<tr>
<td>Distant</td>
<td>40</td>
<td>21</td>
<td>Close</td>
<td>70</td>
<td>36</td>
</tr>
<tr>
<td>Loud</td>
<td>48</td>
<td>75</td>
<td>Quiet</td>
<td>11</td>
<td>17</td>
</tr>
<tr>
<td>Spacious</td>
<td>51</td>
<td>32</td>
<td>Small</td>
<td>58</td>
<td>36</td>
</tr>
<tr>
<td>Spacious</td>
<td>51</td>
<td>32</td>
<td>Boxy</td>
<td>22</td>
<td>14</td>
</tr>
<tr>
<td>Reverberant</td>
<td>39</td>
<td>21</td>
<td>Dry</td>
<td>25</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 6.19: $R^2$ Correlation between attribute values (frequency of use in the case of 'Frequency') and the single dimension output from CMDS of the aggregated dissimilarity matrices for the four-boundary experiment.

<table>
<thead>
<tr>
<th>Excerpt</th>
<th>Bright/Muffled</th>
<th>Bright/Boomy</th>
<th>Bright/Dull</th>
<th>Bright/Muffled, Boomy and Dull</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical</td>
<td>0.28</td>
<td>0.05</td>
<td>0.05</td>
<td>0.26</td>
<td>0.09</td>
</tr>
<tr>
<td>Pop</td>
<td>0.15</td>
<td>0.08</td>
<td>0.16</td>
<td>0.22</td>
<td>0.04</td>
</tr>
<tr>
<td>Speech</td>
<td>0.10</td>
<td>0.02</td>
<td>0.10</td>
<td>0.03</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Table 6.20: $R^2$ Correlation between bass increase/decrease and the single dimension output from CMDS of the aggregated dissimilarity matrices for the four-boundary condition.

<table>
<thead>
<tr>
<th>Excerpt</th>
<th>Bass increase/decrease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical</td>
<td>0.1375</td>
</tr>
<tr>
<td>Pop</td>
<td>0.1429</td>
</tr>
<tr>
<td>Speech</td>
<td>0.15708</td>
</tr>
</tbody>
</table>

Table 6.21: $R^2$ Correlation between attribute values and the single dimension output from CMDS of the aggregated dissimilarity matrices for the four-boundary condition.

<table>
<thead>
<tr>
<th>Excerpt</th>
<th>Wide/Narrow</th>
<th>Natural/Unnatural</th>
<th>Distant/Close</th>
<th>Loud/quiet</th>
<th>Spacious/Small</th>
<th>Spacious/Boxy</th>
<th>Reverb./Dry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical</td>
<td>0.84</td>
<td>0.22</td>
<td>0.35</td>
<td>0.72</td>
<td>0.05</td>
<td>0.13</td>
<td>0.06</td>
</tr>
<tr>
<td>Pop</td>
<td>0.72</td>
<td>0</td>
<td>0</td>
<td>0.72</td>
<td>0.58</td>
<td>0.22</td>
<td>0</td>
</tr>
<tr>
<td>Speech</td>
<td>0.79</td>
<td>0.53</td>
<td>0.09</td>
<td>0.71</td>
<td>0.56</td>
<td>0.13</td>
<td>0.82</td>
</tr>
</tbody>
</table>
same antonyms as earlier which were again used the most, showed little correlation with
the perceptual dimension in this case (see Tables 6.19 and 6.20).
CHAPTER 6. MAIN EXPERIMENTS

Attribute Correlation - Classical Four Walls

(a) $R^2$ Correlation of attributes with perceptual dimension

(b) Absorption - classical, four-boundary

(c) Position - classical, four-boundary

(d) Wall-separation - classical, four-boundary

Figure 6.17: Boxplots showing the effect of different experimental parameters upon the relationship between directivity and the perceptual dimension for the classical excerpt reproduced in condition S2 (four-boundary): (b) absorption, (c) listener position and (d) boundary separation. Subfigure (a) shows the correlation of various attributes with the perceptual dimension, the label indicating the direction of a positive change in that dimension (i.e. increase in dimension is well correlated to the stimuli being perceived as wider).
CHAPTER 6. MAIN EXPERIMENTS

190

Figure 6.18: Boxplots showing the effect of different experimental parameters upon the relationship between directivity and the perceptual dimension for the pop excerpt reproduced in condition S2 (four-boundary): (b) absorption, (c) listener position and (d) boundary separation. Subfigure (a) shows the correlation of various attributes with the perceptual dimension, the label indicating the direction of a positive change in that dimension (i.e. increase in dimension is well correlated to the stimuli being perceived as wider).
(a) $R^2$ Correlation of attributes with perceptual dimension

(b) Absorption - speech, four-boundary

(c) Position - speech, four-boundary

(d) Wall-separation - speech, four-boundary

Figure 6.19: Boxplots showing the effect of different experimental parameters upon the relationship between directivity and the perceptual dimension for the speech excerpt reproduced in condition S2 (four-boundary): (b) absorption, (c) listener position and (d) boundary separation. Subfigure (a) shows the correlation of various attributes with the perceptual dimension, the label indicating the direction of a positive change in that dimension (i.e. increase in dimension is well correlated to the stimuli being perceived as wider).
CHAPTER 6. MAIN EXPERIMENTS

Table 6.22: Significant Factors - Main Experiment (four-boundary)

<table>
<thead>
<tr>
<th>Excerpt</th>
<th>Sig.Factors/Interactions</th>
<th>df</th>
<th>F</th>
<th>Sig.</th>
<th>$\eta^2$</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Classical</strong></td>
<td>Directivity</td>
<td>2</td>
<td>85.511</td>
<td>.000</td>
<td>.905</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Absorption</td>
<td>1</td>
<td>109.014</td>
<td>.000</td>
<td>.858</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Wall Separation</td>
<td>1</td>
<td>25.956</td>
<td>.000</td>
<td>.590</td>
<td>.998</td>
</tr>
<tr>
<td><strong>Pop</strong></td>
<td>Directivity</td>
<td>2</td>
<td>69.097</td>
<td>.000</td>
<td>.885</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Absorption</td>
<td>1</td>
<td>154.720</td>
<td>.000</td>
<td>.896</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Wall Separation</td>
<td>1</td>
<td>12.068</td>
<td>.003</td>
<td>.401</td>
<td>.907</td>
</tr>
<tr>
<td></td>
<td>Position</td>
<td>1</td>
<td>4.898</td>
<td>.040</td>
<td>.214</td>
<td>.553</td>
</tr>
<tr>
<td><strong>Speech</strong></td>
<td>Directivity</td>
<td>2</td>
<td>115.445</td>
<td>.000</td>
<td>.928</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Absorption</td>
<td>1</td>
<td>891.574</td>
<td>.000</td>
<td>.980</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Table 6.23: Further Effects - Main Experiment (four-boundary)

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Factor</th>
<th>Post-Hoc Bonferroni</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Classical</strong></td>
<td>Directivity</td>
<td>Omni.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Omni.</td>
</tr>
<tr>
<td><strong>Pop</strong></td>
<td>Directivity</td>
<td>Omni.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Omni.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cardioid</td>
</tr>
<tr>
<td><strong>Speech</strong></td>
<td>Directivity</td>
<td>Omni.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Omni.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cardioid</td>
</tr>
</tbody>
</table>

ANOVA

A UniANOVA was carried on the data used for the boxplots. Significant factors and interactions for each audio excerpt are presented in Table 6.22 and further effects in 6.23.
Observations of the boxplots and results from the ANOVA suggest that:

for the classical excerpt (four-boundary condition),

- the perceptual dimension was found to be most highly correlated with the following attributes: width (correlation of 0.84) and loudness (0.72);
- the directivity had a large effect on the perceptual dimension, and hence is correlated with the terms above (a narrowing in directivity may be related to a perceived narrowing in width and decrease in loudness);
- higher absorption leads to the reproduction being perceived as lower on the perceptual dimension (which may relate to it being narrower and quieter) compared to lower absorption;
- position has no significant effect;
- greater wall separation causes reproduction to be perceived as lower on the perceptual dimension (which may related to it being narrower and quieter), but this effect is small.

for the pop excerpt (four-boundary condition),

- the perceptual dimension was found to be most highly correlated with the following attributes: width (correlation of 0.72), loudness (0.72) and spaciousness (small) (0.58);
- the directivity had a large effect on the perceptual dimension, and hence is correlated with the terms above (a narrowing in directivity may be related to a perceived narrowing in width, a decrease in loudness and a reduction in spaciousness);
- higher absorption leads to the reproduction being perceived as much lower on the perceptual dimension (which may relate to it being narrower, quieter and less spacious) compared to lower absorption;
- whilst position appears to have little effect, ANOVA results suggest that it is significant. The F-ratio and Observed Power values are much lower than the other factors;
- wall separation has a small yet significant effect.
CHAPTER 6. MAIN EXPERIMENTS

for the speech excerpt (four-boundary condition),

- the perceptual dimension was found to be most highly correlated with the following attributes: width (correlation of 0.79), loudness (0.71), reverberance (0.82), naturalness (0.53) and spaciousness (small) (0.56);

- the directivity had a large effect on the perceptual dimension, and hence is correlated with the terms above (a narrowing in directivity may be related to a perceived narrowing in width, a reduction in loudness, reverberance and spaciousness, and an increase in naturalness);

- higher absorption leads to the reproduction being perceived as very much lower on the perceptual dimension (which may relate to it being narrower, quieter, less reverberant, less spacious and more natural) compared to lower absorption;

- position does not have a significant effect;

- wall separation does not have a significant effect.

across all excerpts (four-boundary condition),

- a narrowing in directivity (increased directivity index) corresponds to a change in the perceptual dimension;

- changes in the dimension are most correlated with width and loudness for all excerpts, and also reverberance when reproducing speech;

- the greatest difference in perception as a result of absorption occurs when comparing reproductions with speech;

- position has little effect;

- wall position has little effect.
6.7 Six-Boundary Test (Condition 3)

6.7.1 Description

The third condition tests (six-boundary model) also took place in the listening room at the University of Surrey, using the same listeners as the first condition. Again, the procedure was the same, though there appeared to be some delay in the updating of BRIRs as the listener moved across large angles, and so the crossfade time between adjacent BRIRs was increased to 100ms. This was found to have no perceivable effect, other than to improve the spatial quality.

6.7.2 Analysis

As for the two-boundary test and four-boundary test, a range of MDS analyses was carried out in order to determine the number of dimensions in which the data can best be represented spatially. The best solution in terms of number of dimensions for each analysis type (based on when negligible improvement occurs), is presented in Table 6.24 for the third condition (six-boundary).

<table>
<thead>
<tr>
<th>Method</th>
<th>Analysis</th>
<th>Pop</th>
<th>Classical</th>
<th>Speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mapping Metric WMDS</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>(0.84141)</td>
</tr>
<tr>
<td>Non-Metric WMDS</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>(0.85887)</td>
</tr>
<tr>
<td>Aggregated Metric CMDS</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>(0.97336)</td>
</tr>
<tr>
<td>Aggregated Non-Metric CMDS</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>(0.99019)</td>
</tr>
<tr>
<td>Sorting Metric CMDS</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>(0.90757)</td>
</tr>
<tr>
<td>Non-Metric CMDS</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>(0.97853)</td>
</tr>
</tbody>
</table>

Table 6.24: Number of resulting dimensions based on the RSQ, and the associated RSQ value in brackets, for each of the MDS analysis types, for the six-boundary experiment.
Results summarised in this table suggest that:

- a good-fit, one dimensional solution exists for all excerpts with a non-metric Classic MDS analysis of listeners' aggregated data;
- a reduction in solution dimensionality is seen with a non-metric analysis of sorted data compared to the related metric analysis;
- solution dimensionality is highest for the sorted data analysis;
- Both metric and non-metric Weighted MDS, which take into account individual listener scores, generally show higher dimensionality solutions than the corresponding CMDS of aggregated matrices and also lower goodness of fit.

As in the two-boundary and four-boundary test, the data can be represented well in one dimension and the antonyms most frequently used are the same (see Tables 6.25 and 6.26, and Figures D.9 to D.14 in the Appendix). As a result, the same analysis of the qualitative (verbal) data was used so that the correlation of the attributes and the dimension could be calculated and plotted. For the six-boundary test, the correlation values are shown in Table 6.29 and plotted in Figures 6.20, 6.21 and 6.22.

Analysis of timbral attributes, frequency response and bass increase/decrease, using the same antonyms as earlier which were again used the most, showed little correlation with the perceptual dimension in this case (see Tables 6.27 and 6.28).
 CHAPTER 6. MAIN EXPERIMENTS

<table>
<thead>
<tr>
<th>Group</th>
<th>Total Frequency of Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reverberation (Ordered)</td>
<td>330</td>
</tr>
<tr>
<td>Frequency Response (Ordered)</td>
<td>297</td>
</tr>
<tr>
<td>Timbral Descriptors (Order Attempted)</td>
<td>231</td>
</tr>
<tr>
<td>Proximity (Ordered)</td>
<td>170</td>
</tr>
<tr>
<td>Spaciousness (Ordered)</td>
<td>169</td>
</tr>
<tr>
<td>Artifacts (Ordered)</td>
<td>167</td>
</tr>
<tr>
<td>Width (Ordered)</td>
<td>140</td>
</tr>
<tr>
<td>Loudness (Ordered)</td>
<td>130</td>
</tr>
<tr>
<td>Room Description (Ordered)</td>
<td>99</td>
</tr>
<tr>
<td>Filtering Terms (Unordered)</td>
<td>95</td>
</tr>
<tr>
<td>Flatness of Frequency Response (Ordered)</td>
<td>37</td>
</tr>
<tr>
<td>Prominence (Unordered)</td>
<td>35</td>
</tr>
<tr>
<td>Clarity (Ordered)</td>
<td>30</td>
</tr>
<tr>
<td>Pleasantness (Ordered)</td>
<td>19</td>
</tr>
<tr>
<td>Envelopment (Unordered)</td>
<td>9</td>
</tr>
</tbody>
</table>

a closely tied to Spaciousness

Table 6.25: Total frequency of words in groups/scales defined by panel members in verbal analysis session (six-boundary).

<table>
<thead>
<tr>
<th>Antonym 1</th>
<th>Freq</th>
<th>%</th>
<th>Antonym 2</th>
<th>Freq</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wide</td>
<td>41</td>
<td>29</td>
<td>Narrow</td>
<td>55</td>
<td>39</td>
</tr>
<tr>
<td>Natural</td>
<td>48</td>
<td>29</td>
<td>Unnatural</td>
<td>46</td>
<td>28</td>
</tr>
<tr>
<td>Distant</td>
<td>92</td>
<td>54</td>
<td>Close</td>
<td>56</td>
<td>33</td>
</tr>
<tr>
<td>Loud</td>
<td>69</td>
<td>53</td>
<td>Quiet</td>
<td>47</td>
<td>36</td>
</tr>
<tr>
<td>Spacious</td>
<td>52</td>
<td>31</td>
<td>Small</td>
<td>53</td>
<td>31</td>
</tr>
<tr>
<td>Spacious</td>
<td>52</td>
<td>31</td>
<td>Boxy</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td>Reverberant</td>
<td>145</td>
<td>44</td>
<td>Dry</td>
<td>46</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 6.26: High-frequency antonyms and the relative percentage of each term within its group (six-boundary)
CHAPTER 6. MAIN EXPERIMENTS

Excerpt | Bright/ Muffled | Bright/ Boomy | Bright/ Dull | Bright/ Muffled, Boomy and Dull | Frequency
---|---|---|---|---|---
Classical | 0.08 | 0.00 | 0.18 | 0.02 | 0.00
Pop | 0.10 | 0.01 | 0.02 | 0.01 | 0.05
Speech | 0.05 | 0.00 | 0.00 | 0.01 | 0.39

Table 6.27: $R^2$ Correlation between attribute values (frequency of use in the case of 'Frequency') and the single dimension output from CMDS of the aggregated dissimilarity matrices for the six-boundary experiment.

Excerpt | Bass increase/decrease
---|---
Classical (six-boundary) | 0.0006
Pop (six-boundary) | 0.052
Speech (six-boundary) | 0.00394

Table 6.28: $R^2$ Correlation between bass increase/decrease and the single dimension output from CMDS of the aggregated dissimilarity matrices for the six-boundary condition.

Excerpt | Wide / Narrow | Natural / Unnatural | Distant / Close | Loud / Quiet | Spacious / Small / Boxy | Reverb. / Dry
---|---|---|---|---|---|---
Classical | 0.55 | 0.11 | 0.48 | 0.88 | 0.39 | 0.1 | 0.9
Pop | 0.34 | 0.36 | 0.09 | 0.67 | 0.05 | 0.06 | 0.85
Speech | 0 | 0.63 | 0.55 | 0.71 | 0.79 | 0.67 | 0.91

Table 6.29: $R^2$ Correlation between attribute values and the single dimension output from CMDS of the aggregated dissimilarity matrices for six-boundary
CHAPTER 6. MAIN EXPERIMENTS

199

Attribute Correlation - Classical Six Walls

(a) $R^2$ Correlation of attributes with perceptual dimension

(b) Absorption - classical, six-boundary

(c) Position - classical, six-boundary

(d) Wall-separation - classical, six-boundary

Figure 6.20: Boxplots showing the effect of different experimental parameters upon the relationship between directivity and the perceptual dimension for the classical excerpt reproduced in condition S3 (six-boundary): (b) absorption, (c) listener position and (d) boundary separation. Subfigure (a) shows the correlation of various attributes with the perceptual dimension, the label indicating the direction of a positive change in that dimension (i.e. increase in dimension is well correlated to the stimuli being perceived as more narrow).
Figure 6.21: Boxplots showing the effect of different experimental parameters upon the relationship between directivity and the perceptual dimension for the pop excerpt reproduced in condition S3 (six-boundary): (b) absorption, (c) listener position and (d) boundary separation. Subfigure (a) shows the correlation of various attributes with the perceptual dimension, the label indicating the direction of a positive change in that dimension (i.e. increase in dimension is well correlated to the stimuli being perceived as more dry).
CHAPTER 6. MAIN EXPERIMENTS

Attribute Correlation - Speech Six Walls
Narrow Unnatural Distant Loud spacious Reverb.

(a) $R^2$ Correlation of attributes with perceptual dimension

(b) Absorption - speech, six-boundary

(c) Position - speech, six-boundary

(d) Wall-separation - speech, six-boundary

Figure 6.22: Boxplots showing the effect of different experimental parameters upon the relationship between directivity and the perceptual dimension for the speech excerpt reproduced in condition S3 (six-boundary): (b) absorption, (c) listener position and (d) boundary separation. Subfigure (a) shows the correlation of various attributes with the perceptual dimension, the label indicating the direction of a positive change in that dimension (i.e. increase in dimension is well correlated to the stimuli being perceived as more reverberant).
### Table 6.30: Significant Factors - Main Experiment (six-boundary)

<table>
<thead>
<tr>
<th>Excerpt</th>
<th>Sig. Factors/Interactions</th>
<th>df</th>
<th>F</th>
<th>Sig.</th>
<th>(\eta^2)</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical</td>
<td>Directivity</td>
<td>2</td>
<td>122.365</td>
<td>.000</td>
<td>.931</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Absorption</td>
<td>1</td>
<td>783.603</td>
<td>.000</td>
<td>.978</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Wall Separation</td>
<td>1</td>
<td>15.983</td>
<td>.001</td>
<td>.470</td>
<td>.965</td>
</tr>
<tr>
<td>Pop</td>
<td>Directivity</td>
<td>2</td>
<td>215.439</td>
<td>.000</td>
<td>.960</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Absorption</td>
<td>1</td>
<td>2146.683</td>
<td>.000</td>
<td>.992</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Wall Separation</td>
<td>1</td>
<td>30.446</td>
<td>.000</td>
<td>.628</td>
<td>.999</td>
</tr>
<tr>
<td>Speech</td>
<td>Directivity</td>
<td>2</td>
<td>68.400</td>
<td>.000</td>
<td>.884</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Absorption</td>
<td>1</td>
<td>1172.683</td>
<td>.000</td>
<td>.985</td>
<td>1.000</td>
</tr>
</tbody>
</table>

### Table 6.31: Further Effects - Main Experiment (six-boundary)

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Factor</th>
<th>Post-Hoc Bonferroni</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical</td>
<td>Directivity</td>
<td>Omni. Cardioid Beam</td>
</tr>
<tr>
<td></td>
<td>Omni.</td>
<td>-.001 -.000</td>
</tr>
<tr>
<td></td>
<td>Cardioid</td>
<td>-.         -.000</td>
</tr>
<tr>
<td>Pop</td>
<td>Directivity</td>
<td>Omni. Cardioid Beam</td>
</tr>
<tr>
<td></td>
<td>Omni.</td>
<td>-.000 -.000</td>
</tr>
<tr>
<td></td>
<td>Cardioid</td>
<td>-.         -.000</td>
</tr>
<tr>
<td>Speech</td>
<td>Directivity</td>
<td>Omni. Cardioid Beam</td>
</tr>
<tr>
<td></td>
<td>Omni.</td>
<td>-.000 -.000</td>
</tr>
<tr>
<td></td>
<td>Cardioid</td>
<td>-.         -.000</td>
</tr>
</tbody>
</table>

ANOVA

A UniANOVA was carried on the data used for the boxplots. Significant factors and interactions for each audio excerpt are presented in Table 6.30 and further effects in 6.31.
CHAPTER 6. MAIN EXPERIMENTS

Observations of the boxplots and results from the ANOVA suggest that:

for the classical excerpt (six-boundary condition),

- the perceptual dimension was found to be most highly correlated with the following attributes: loudness (correlation of 0.88), reverberance (0.9) and width (0.55);
- the directivity had a large effect on the perceptual dimension, and hence is correlated with the terms above (a narrowing in directivity may be related to a perceived reduction in loudness, reverberance and width);
- higher absorption leads to the reproduction being perceived as much higher on the perceptual dimension (which may relate to it being quieter, less reverberant and narrower) compared to lower absorption;
- position has no significant effect;
- wall separation has a small yet significant effect.

for the pop excerpt (six-boundary condition),

- the perceptual dimension was found to be most highly correlated with the following attributes: reverberance (correlation of 0.85) and loudness (0.67);
- the directivity had a large effect on the perceptual dimension, and hence is correlated with the terms above (a narrowing in directivity may be related to a perceived reduction in reverberance and loudness);
- higher absorption leads to the reproduction being perceived as much higher on the perceptual dimension (which may relate to it being less reverberant and quieter) compared to lower absorption;
- position has no significant effect;
- wall separation has a small yet significant effect.

for the speech excerpt (six-boundary condition),

- the perceptual dimension was found to be most highly correlated with the following attributes: reverberance (correlation of 0.91), spaciousness (small and boxy) (0.79 and 0.67), loudness (0.71), distance (0.55) and naturalness (0.63);
• the directivity had a large effect on the perceptual dimension, and hence is correlated with the terms above (a narrowing in directivity may be related to a perceived reduction in reverberance, spaciousness, loudness and distance, and an increase in naturalness);

• higher absorption leads to the reproduction being perceived as very much lower on the perceptual dimension (which may relate to it being less reverberant, less spacious, quieter, closer and more natural) compared to lower absorption;

• position has no significant effect;

• wall separation has no significant effect.

across all excerpts (six-boundary condition),

• a narrowing in directivity (increased directivity index) corresponds to a change in the perceptual dimension;

• changes in the dimension are most correlated with loudness and reverberance for classical music, reverberance for pop music and loudness, spaciousness (small) and reverberance when reproducing speech;

• the greatest difference in perception as a result of absorption occurs when comparing reproductions with speech;

• position has little effect;

• wall position has little effect.

6.8 Summary of Experimental Observations

6.8.1 Dimensionality of Solution

Observing the general results of the multidimensional analyses, the following outcomes are apparent:

• a good-fit, one dimensional solution exists for all excerpts and conditions with a non-metric Classic MDS analysis of listeners' aggregated data. The RSQ values
for these solutions are the highest of all analysis methods in 6/9 cases and the benefits of having a single dimension outweigh the small improvement in the cases where the metric RSQs were higher;

- generally, a reduction in solution dimensionality is seen with a non-metric analysis compared to the related metric analysis;

- solution dimensionality is highest for the sorted data analysis;

- Both metric and non-metric Weighted MDS, which take into account individual listener scores, generally show higher dimensionality solutions than the corresponding CMDS of aggregated matrices and also lower goodness of fit. For example, in the non-metric case, Pop and Speech have 2 dimensional solutions for all conditions, and Classical has 3, 6 and 2 dimensional solutions for the two-boundary, four-boundary and six-boundary condition respectively, compared to the one dimensional solution for CMDS.

It is concluded that the data can be well-fitted in one dimension for all cases, based on the CMDS analysis of aggregated listener matrices. WMDS of individual matrices shows that, typically, 2 dimensions are affected (though the solutions are at a lower goodness-of-fit than that CMDS solutions).

6.8.2 Relationship Between Dimension and Loudspeaker Directivity

A trend showing that a narrowing in directivity corresponds to a change in the perceptual dimension is apparent in all conditions and excerpts.

6.8.3 Relationship Between Dimension and Attributes

Results show that different attributes correspond with the dimension, depending on the excerpt and condition. Multiple attributes have been shown to be affected ‘in parallel’, suggesting that changes in the test variables are associated with similar magnitudes of changes in multiple attributes at the same time (for example, in the four-boundary condition, loudness and width were perceived to change with an increase in directivity). Table 6.32 shows which attributes are most highly correlated with the dimension for each condition and excerpt and Table 6.33 describes the perceptual effects associated
## Table 6.32: Attributes that are highest correlated with the single perceptual dimension

with a change in the dimension occurring in the same direction as a change in perception resulting from a narrowing in directivity.

### 6.8.4 Effect of Boundary Conditions and Listener Position

As well as directivity, three additional factors; absorption, wall-separation and listener position, were included in these experiments. According to the results earlier, absorption has a similar magnitude of effect for all excerpts with respect to the dimension with the two-boundary condition, and has greatest effect when using speech as an excerpt in both the four-boundary and six-boundary conditions. Position causes little effect with respect to the dimension for most conditions, with the greatest notable effect caused when comparing narrow directivity reproductions with pop excerpts (though, no significant
<table>
<thead>
<tr>
<th>Condition</th>
<th>Excerpt</th>
<th>Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-Boundary</td>
<td>Classical</td>
<td>narrower, farther, more boxy</td>
</tr>
<tr>
<td></td>
<td>Pop</td>
<td>narrower, quieter, less spacious</td>
</tr>
<tr>
<td></td>
<td>Speech</td>
<td>drier, quieter</td>
</tr>
<tr>
<td>Four-Boundary</td>
<td>Classical</td>
<td>narrower, quieter</td>
</tr>
<tr>
<td></td>
<td>Pop</td>
<td>narrower, quieter, less spacious</td>
</tr>
<tr>
<td></td>
<td>Speech</td>
<td>narrower, quieter, drier, more natural, less spacious</td>
</tr>
<tr>
<td>Six-Boundary</td>
<td>Classical</td>
<td>quieter, less reverberant, narrower</td>
</tr>
<tr>
<td></td>
<td>Pop</td>
<td>less reverberant, quieter</td>
</tr>
<tr>
<td></td>
<td>Speech</td>
<td>less spacious, less reverberant, quieter, closer, more natural</td>
</tr>
</tbody>
</table>

Table 6.33: Associated change in perception of sound reproduction according to change in dimension in same direction as a change in perception caused by reduction in modelled loudspeaker directivity. Changes are presented in order of magnitude.
CHAPTER 6. MAIN EXPERIMENTS

<table>
<thead>
<tr>
<th>Factor</th>
<th>Directivity</th>
<th>Absorption</th>
<th>Wall Separation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td>Directivity Index, dB (Ratio)</td>
<td>Absorption Coefficient</td>
<td>Meters, m</td>
</tr>
<tr>
<td>1</td>
<td>0dB (1) Omni</td>
<td>0.05</td>
<td>2m</td>
</tr>
<tr>
<td>2</td>
<td>5.3dB (3.39) Cardioid</td>
<td>0.5</td>
<td>6m</td>
</tr>
<tr>
<td>3</td>
<td>15.3dB (33.9) Narrow beam</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Table 6.34: Levels of factors that vary according to a scale

effect was reported following ANOVA). Finally, wall separation shows most effect with respect to the dimension when comparing omni and cardioid directivities with classical and pop excerpts, and narrow directivity reproductions with speech excerpts.

It is also possible to plot the correlation between the different levels of some factors (those which vary along a scale - see Table 6.34) and the perceived change in dimension. Results can be used to suggest how they might compare for each condition (simulated environment type), yet it should be made clear that these factors only have 2 or three levels with which to correlate - to draw strong conclusions, more levels would be desirable. The outcome is shown in Figure 6.23 and suggests that, whilst different wall distances show little correlation with the perceived changes in the dimension for all conditions, directivity index is most correlated with perceived dimensional changes in the first condition (two-boundary) and absorption is most correlated with the third condition (six-boundary). It appears that, as more boundaries are included in the model, a change in directivity has less of an effect on the perceived changes, relative to those caused by other factors, and a change in absorption becomes more important. It should be noted that the six-boundary condition was perceived by listeners to be very reverberant in comparison to the other conditions and this may have exaggerated the trend. Although, the differences between the other conditions were less noticeable and the trend is still shown to exist.

An infographic to demonstrate this trend, including the main attributes perceived, is included below (see Figure 6.24).
6.8.5 Summary

- the effects that changing the experimental factors have upon the perception of sound reproduction are best explained in one dimension;

- change in loudspeaker directivity is seen to correspond to a change in this dimension for every experimental condition - observing Tables 6.15, 6.23 and 6.31, there is a significant difference between all directivity types with the exception of cardioid and omnidirectional for the two-boundary condition;

- different attributes are associated with this dimension - they depend on the excerpt and condition and may be affected together, in parallel;

- the attributes that most highly correlate to the dimension for each excerpt and condition have been identified (see Table 6.32);
• the perceptual effects that correspond to a change in loudspeaker directivity have been identified (see Table 6.33);
• the magnitude of effect of absorption, wall-separation and listener condition according to these results have been measured;
• the relationship between the different levels of some experimental factors and the effect upon perception of reproduction has been identified (Figure 6.23).

6.9 Discussion

6.9.1 Dimensionality

Results of the MDS analysis suggested that the perceptual data acquired from listeners was best fit to a one dimensional solution, via non-metric CMDS analysis of aggregated listener matrices. RSQ values that resulted from WMDS were also high and generally suggested 2 dimensional solutions. It was noted earlier that the decision to base the main analysis on the one dimensional solution was justified on the basis that a consensual
outlook was desired, as opposed to an individualised, more listener-specific viewpoint. If a more detailed listener-specific viewpoint was of interest, then the experiments carried out here could be extended to include the analysis of multiple dimensions. The outcomes, however, would have to be noted as being based on a multidimensional solution which is not as good compared to the aggregated CMDS solution.

6.9.2 Relationship Between Perceptual Dimension and Loudspeaker Directivity

The results indicate that the changes in directivity caused a perceivable change. A generally noticeable trend occurred across all conditions and excerpts whereby with a change in directivity index, a corresponding change in the dimension occurs.

The correlation between the change in directivity and the change in the perceptual dimension, relative to changes in the other tested parameters, is seen to be highest for the first condition (two-boundary) and lowest for the third condition (six-boundary). This may be linked to the earlier mentioned study by Olive and Toole (1989), who found that the effects of single lateral reflections (typically image shift and spaciousness) were reduced in the presence of more reflections.

Also, the finding in Figure 6.23 that changes in directivity when reproducing the classical excerpt are typically more highly correlated with the dimensional changes than when producing pop or speech is perhaps related to the earlier idea that long, sustained musical notes lead to lower detection thresholds of single reflection effects (Barron 1971; Toole 2008)

6.9.3 Relationship Between Perceptual Dimension and Attributes

It has been established that the types of attributes associated with the dimension are dependent on the excerpt and condition. Observing Tables 6.32 and 6.33, it appears that the most correlated attributes are:

\textit{width, spaciousness (small), reverberance, loudness,}

with width and brightness (muffled) being most highly correlated with the perceptual dimension in two-boundary conditions, and loudness and reverberance, becoming most highly correlated as the number of reflective walls increase.

These attributes correspond well with those found in the literature and in pre-
liminary experiments (which, to recap, were: ASW, timbre, localisation, loudness, envelopment/spaciousness and source distance) and, thus, the objective to provide experimentally derived quantitative and qualitative data regarding direction and boundary characteristics, which is in-line with theory and previous literature, has been met.

6.9.4 Effect of Boundary Conditions (Absorption and Wall-Separation)

For the two-boundary condition, a variation in absorption causes a small, yet statistically significant, change in the perception of reproduction. From the graphs, this appears to be for the omni and cardioid directivities, but not for the narrow beam - this would support earlier discussion which suggested that boundary absorption has less effect when the sound is focused towards a listener and does not interact with the room. According to these results, the fact that the low frequencies were still omnidirectional with the narrow beam directivity, does not affect this expected outcome. It should be noted that for this condition the change in directivity to the beam has more effect upon the perceptual dimension than the changes in absorption. For the four-boundary condition, the absorptive case causes a distinct difference from the reflective case with respect to the dimension, for all directivities - now, the beam is reflecting off the rear wall behind the listener and so the absorptive properties become more important than in the previous condition. For the six-boundary condition, different absorption cases cause the largest differences in dimension to occur, suggesting that, whereas directivity is the key influence on the perception of reproduction in a two-boundary environment, absorption is the key influence in the perception of reproduction in the six-boundary environment, i.e. with an increase in the number of reflective boundaries, the absorption characteristics of those boundaries become dominant compared the the directivity characteristics of the loudspeakers. Listeners reported that the task was easier with the six-boundary condition and the discussion of the likely subtlety of directivity effects was introduced earlier, in Chapter 3.

In all conditions, the variation in the perceptual dimension for the two absorption values is shown to be highest when listening to speech excerpts. Speech perhaps serves as the most revealing signal where the absorption differences begin to dominate the effects of directivity, specifically due to its non-consistent, time-variant nature which is likely to be more revealing of differences relating to loudness and reverberance - both of which would be expected to be less noticeable with a prolonged musical signal.
The correlation of boundary separation with the perceptual changes is greatest in the two-boundary conditions. In this case, differences are greatest between omni and cardioid directivities reproducing classical and pop excerpts, and narrow directivities reproducing speech. Despite these findings, it is shown to have much less influence than directivity and absorption in all conditions.

6.9.5 Effect of Listener Position

For the two-boundary condition, listener position was only observed to affect the reproduction when comparing the narrow directivity simulation (yet was not shown to have significant effect according to ANOVA). For the four-boundary condition, only narrow directivity appeared to be affected (and for the pop excerpt a statistically significant effect was measured). For the six-boundary condition, position seemed to have no effect. Compared to the effect of directivity and absorption in all conditions and for all excerpts, listener position is less significant.

6.9.6 Overall Outcome

In light of the discussion here, it is possible to draw a general outcome: differences in directivity, which are more perceptible in the two-boundary case and with classical music excerpts (relative to differences in the other experimental factors), correspond to perceived differences mostly in width (whereby the reproduction becomes narrower with a narrowing in directivity). As the absorption differences become more dominant than differences in directivity (highlighted most by speech excerpts), loudness and reverberance are terms most associated with the changes perceived (with a narrowing in directivity linked to a reduction in loudness and reverberance). Spaciousness is also a key affected attribute for all conditions (seen to reduce with a narrowing in directivity and dependent on signal type) and the effects of wall-separation and listener position, for the levels tested here, are less significant in comparison to the effects of directivity and absorption.
6.9.7 Evaluation of Methodology

The experimental technique employed here was designed with reference to several previous studies and methods. Outcomes from the experiments were in-line with existing literature and contributed numerical data to this area of research. Despite this, a few improvements to the method would be seen as beneficial in future:

- fewer stimuli for projective mapping - some listeners found the task a little taxing at first and this was attributed to the high number of stimuli. Fewer stimuli would reduce the mental load for listeners and perhaps improve accuracy in judgement;

- additional instructions for labelling - the variation in language, even when describing the same perceptual effect, led to increased processing being necessary after the tests. If additional instructions were to be included, listeners could be directed to present attributes in a more systematic way, making collection and processing easier. This may also help to improve accuracy in analysis.

Some overall considerations regarding the test design:

- limited number of absorption conditions - only two absorption conditions meant that the linearity of the relationship between directivity and absorption could not be investigated;

- limited number of wall combinations - the four-boundary condition in these experiments was chosen to use four boundaries in the horizontal plane - it would be of interest to compare test results using four boundaries in the vertical plane (perhaps timbre would become more prominent);

- limited number of listening positions - despite the two listening positions being chosen so as to highlight differences, perceived effects were minimal - to be absolutely sure of the relative influence of listener position compared to the other factors, it would be of interest to evaluate more positions.

6.10 Chapter Summary

This chapter documented the main experimental work of this thesis. Following on from the outcomes of the previous chapter, a set of experiments was designed to acquire
qualitative and quantitative data regarding the effects of different directivities, boundary characteristics and listener positions upon the perception of sound reproduction. The experiments were designed to be ‘first look’ in nature and to indicate the effects likely to occur in a real listening situation, and used an auralisation system to allow instant comparison between reproductions.

To reduce the impact of experimenter bias and to avoid fatigue-inducing tasks, a novel method was employed. This began with a projective mapping (or Napping®) experiment in conjunction with MDS analysis to determine the dimensionality of the perceived effects. Verbal data elicited from the test stage was then processed with a combination of qualitative assessment techniques to determine, for the dimensions revealed, attribute names and weightings.

Results from the experiment showed that the perceptual changes caused by changes in the test variables could be represented by a single dimension and it was confirmed that a change in loudspeaker directivity always corresponded to a change in this dimension for several modelled environments/audio excerpts. The effects of changing loudspeaker directivity upon the perception of sound reproduction were found to be much greater than those caused by changing boundary conditions and listener position in a two-boundary condition. However, in more reverberant environments, like the four and six-boundary conditions tested, effects caused by changes in absorption became dominant. This reinforces the notion that the effects of directivity upon reproduction are subtle and are most audible when fewer reflections are present. Classical music was shown to highlight differences in directivity most, whereas speech was shown to be more revealing of differences in absorption - both of these results concur with earlier theoretical discussion.

The attribute(s) associated with the observed one-dimensional perceptual changes depended, to some extent, on the programme material auditioned. In some cases multiple attributes varied in parallel. The attributes most highly correlated to the revealed dimension were identified as width, spaciousness (small), reverberance and loudness; brightness (muffled), distance, spaciousness (boxy) and naturalness also showed notable correlation. These findings concur with those of the literature and preliminary experiments. In conditions with few/low-level reflections, variable changes had the largest effect on width, spaciousness and timbre; in conditions with many/high-level reflections, loudness and reverberance were most affected. Although changes to each independent variable were perceived, overall, changes to
directivity and absorption had larger effects than those to boundary separation and listener position.
Chapter 7

Conclusions and Further Work

The principal question to which this thesis refers is: how is loudspeaker directivity likely to affect the perception of reproduced sound in domestic listening rooms?

The answer to this question is dependent on the boundary conditions of the room. The effects of directivity upon perception are likely to be most evident in a room with few acoustically reflective surfaces, with width and brightness being the key affected attributes. With more reflective surfaces, the effects of differences in surface absorption are likely to become dominant, with loudness and reverberance being the key affected attributes. The effects of boundary separation and listener position are likely to be small. Classical music is likely to be a good choice of source material to highlight the effects of differences in directivity, and speech a good choice for highlighting the effects of differences in absorption.

This answer has been arrived at via a series of component questions. Each of these questions will now be answered according to the work undertaken, with specific reference to each contribution to the existing field of knowledge and identification of work required to advance it still further. Real-world considerations as a result of research outcomes will also be presented, as well as any general learned outcomes that may be useful to readers in this area.
CHAPTER 7. CONCLUSIONS AND FURTHER WORK

7.1 R1: To what extent can loudspeaker directivity be controlled?

Loudspeaker directivity is a multivariate function that describes the radiation of a loudspeaker in all directions. In Chapter 2 an overview of acoustic principles relating to source directivity and the technological/pragmatic limitations regarding the design of loudspeakers intended for domestic use indicated that the directivity of a loudspeaker is affected by the transduction type, arrangement, mounting and size of the drive mechanisms used, and that control of directivity is not likely below ~200Hz. This confines research in this area to be able to consider only changes in directivity above ~200Hz and for traditional loudspeakers to be generally thought of as omnidirectional below this frequency.

Recent technological advances have meant that the control of directivity is relatively straightforward to achieve above this frequency, thus validating the motivation behind this research; loudspeakers can be designed to exhibit variable directivity in most of their operational frequency range.

As the purpose of this thesis is to focus on the perceptual effects caused by differences in loudspeaker directivity, a highly detailed investigation into the technological possibilities was not carried out. Although it is felt that the level of detail was sufficient to establish the limitations of experimental work undertaken here, it would be of interest to refer to a study which focused primarily on the practical aspects of loudspeaker directivity design, and the performance of such systems. After all, in order to generate real-life reproductions in-line with the simulations used here, a full understanding of the level of control possible is required.

Conclusions

- Low-frequency directivity control of loudspeakers intended for domestic use is unlikely below ~200Hz and therefore research in this area should only consider perceptual effects of different directivities above this frequency.
Further Work

- A detailed study regarding the design of multi-directivity loudspeakers would be necessary to predict the performance of a real-life system based on simulations used here.

7.2 R2: In what ways can loudspeaker directivity and room acoustics affect the sound field at the listener?

In the knowledge that directivity is a parameter worth considering, it was necessary to consider its role within the acoustic system that causes a sound field to exist within a domestic room. Chapter 2 explained that loudspeaker directivity determines the angle and relative level of sound radiated into a listening space and that the geometry of the room, the absorption characteristics of the room and the listener position subsequently affect the reflections that arrive at the listener (in terms of relative level, spectrum and arrival time). Differences in the combination of direct and reflected sound at the listener position cause phenomena such as comb filtering and signal repetition/delay, changing the spectral and temporal properties of the sound field. Discussion with reference to theory in Chapter 2 highlighted that any changes that may be caused to the sound field at the listener as a result of changing the directivity of the loudspeaker are wholly dependent on the room boundaries and the loudspeaker/listener position. It was at this point that the idea that considering only the effect of directivity upon the perception of reproduction in domestic rooms would be of no value if the outcomes were going to be applicable to a general situation, where room geometry and relative placement of receiver/source are additional parameters. Thus, it was accepted that the study of directivity effects must be carried out in association with room boundary and loudspeaker/listener position effects, which subsequently increases the complexity and breadth of the study - what originally may have been intended as a study based on measuring the change of one parameter now became more complicated. This contributed somewhat to the generality of the investigation and the decision to later use idealistic, basic room/source simulations in order to obtain information regarding the 'bigger picture', as opposed to a single facet relating to specific rooms and directivities that may have been of no practical use to readers.

Chapter 2 presented theory which could be used to infer that, with a narrowing in
directivity, the ‘imprint’ of the room is reduced. The idea that increased amounts of absorption in a room would serve to lessen the effects of a change in directivity was also stated on the grounds that it was becoming more like an anechoic chamber, where the reflections that should contribute to the physical changes at the listener are reduced. A realistic consideration regarding domestic rooms, however, would be that absorption at low frequencies is costly and impractical, and that coupled with the likely omnidirectionality of loudspeakers at low frequencies, some room imprint will probably always be apparent - the idea that a very narrow directivity could eliminate the room imprint completely is denounced as a result of the pragmatic limits of loudspeaker size.

Conclusions

- The direct sound at the listener position is affected by loudspeaker directivity and listener position. The reflected sound is also affected by these factors, as well as by the geometry and absorption characteristics of the room. Changes to any of these factors can lead to spectral and temporal changes in the sound field at the listener position, since this is determined by the combination of direct and reflected sound.

- For data regarding the perception of reproduction in domestic rooms to be applicable to a general range of real-world situations, it is necessary to consider the effects of listener position, room boundaries and loudspeaker directivity, which are all influential on the arrival of sound at the listener.

Further Work

- The breadth of this study is large, with a high number of experimental parameters, and so a wider range of parametric levels would be necessary to extend the level of detail achieved.
7.3 R3: What changes in loudspeaker directivity and boundary characteristics are perceivable, and what are the relative magnitudes of these changes?

According to theory presented in Chapter 3, a number of perceptual attributes were expected to be affected as a result of the physical consequences of changing directivity, boundary characteristics and loudspeaker/listener position (which all affect the nature of direct and indirect sound arriving at a listener). The magnitude of the effect upon these attributes was expected to be small with respect to the overall listening experience, based on the fact that reflections in typical rooms are low in level with short delays (and also according to experimental research based upon single-reflections of reproduced sound). Despite this, preliminary experiments in Chapter 4 using real loudspeakers with differing directivity caused similar attributes to be elicited, and so it can be concluded that the effects are not so small that they go unnoticed.

In Chapter 4, it was found that two speakers with markedly different directivity properties did not show significant difference with respect to chosen attributes - despite there being acknowledged limitations with regard to this type of real-room/multiple loudspeaker testing, it still highlights the subtlety of the parameters in question with respect to four supposedly key attributes.

The concept of adaptation was also introduced in Chapter 3, which may mean that perceivable differences in source/room acoustics are unnoticeable if the listener has time to adapt. This raises an important point of consideration - changes in directivity, boundary characteristics and listener/loudspeaker position may only be discernible when compared in close succession, thus potentially causing this research to only be of interest to systems where instantaneous changes in these parameters are available. It is unlikely that absorption and loudspeaker position can be changed in an instant, and thus outcomes are restricted to reference for: variable-directivity loudspeakers, multiple position listening, or simulated environment listening.

The necessity to improve the control/resolution of the experiments, as well as the pragmatic consideration that this research is only informative for the situations mentioned above, meant that an experimental system with full parameter control (loudspeaker position, directivity, room characteristics and listener position) was designed. Having variable-directivity loudspeakers and multiple position listener would
provide useful information, yet it would not allow consideration of the characteristics of the room and so a reproduction system based on acoustic simulations was chosen.

The auralisation system was used for experiments in Chapter 6, and it allows the source and environment to be designed and simulated. It was shown that differences in the directivity of the source and the absorption characteristics of the room do cause perceivable differences in the reproduction of audio for three typical audio excerpts. Differences in directivity were found to be more notable than absorption changes with just two reflective boundaries, and, with six reflective boundaries, absorption changes were found to be more notable than directivity changes (Figure 6.23 shows that correlation between parameter and dimension change is around 0.7/0.1 for directivity/absorption in the two-boundary condition and 0.1/0.8 for the six-boundary condition). This further proves the subtlety of the effects of directivity upon perception, which appear to be subdued in the presence of typical domestic reflections. With more time, it would be of interest to measure the effects of an additional, intermediate boundary absorption to find out whether directivity effects are linearly related to changes in absorption in a typical six-boundary environment. Also, as the magnitudes of changes measured here have only been described with reference to each other, it would be beneficial in future to find a more universally descriptive way of quantifying the magnitude of effects caused.

Classical music was found to highlight differences between different loudspeaker directivities (dominant in two-boundary environment) most effectively, and speech was found to be most effective at highlighting differences between different boundary absorptions (dominant in six-boundary environment). It was only with speech as a signal that wall-separation was found to have an effect (on narrow directivity sources in the two-boundary case). Wall-separation, in general, had little effect on perception of reproduction in comparison to directivity and wall absorption.

Conclusions

- **An auralisation system is suitable to use for controlled studies of directivity effects (where it is necessary for multiple loudspeaker directivities, environments and listening positions to be compared directly)**

- Theory suggests that the effects of different directivities, combined with different boundary absorption, boundary separation and listener positions, upon the
perception of reproduction should be subtle. Original research carried out here using an auralisation system to present simulations of different combinations of these parameters shows that all parameters have an audible and distinguishable effect: changes in loudspeaker directivity, boundary absorption, boundary separation and listener position, can each cause changes in the perception of sound reproduction.

- Using this system, directivity is found to be the most influential factor when considering just two reflective boundaries. It becomes less influential in the presence of more reflections (shown to occur experimentally with four/six reflective boundaries), where absorption is the most influential factor. This decrease in the influence of directivity with increased reflections is consistent with previous studies based on the effects of single-reflections. The magnitude of effects caused by different loudspeaker directivities is reduced with the increased presence of reflections.

- Using this system, although speech is found to highlight the effects of different absorption types best, classical music is found to highlight the effects of different directivities most effectively. This is also consistent with previous literature, which suggest that continuous sounds reduced the perceptual thresholds for single-reflection effects. Classical music is an effective signal to highlight the perceptual differences between different loudspeaker directivities.

- Boundary separation has less effect on the perception of sound reproduction in comparison to directivity and wall absorption.

Further Work

- Since only two absorption coefficients were tested, it would be of interest to investigate the effect of an intermediate absorption value. This could provide more detail about the relationship between the effects of directivity and the effects of absorption.

- The way that the magnitude of effects are measured and reported in these experiments could be improved to be more meaningful with regards to typical listening situations.
7.4 R4: Which perceptual attributes are affected by the combination of loudspeaker directivity and boundary characteristics?

Attributes most affected in the main experiments include: width, timbre, distance, spaciousness (associated with differences in directivity), loudness and reverberence (associated with changes in absorption). These are well-aligned with existing literature on the topic and with the attributes elicited in the preliminary experiments. It was found that all of the attributes differed in parallel along one MDS dimension and that the key attributes affected in all simulations are width, loudness and reverberence. With narrower directivity (or more absorption), each was found to decrease.

Changes in timbre were noted in the two-boundary condition (it becomes less bright with a narrowing in directivity) and distance is affected for classical excerpts in both two-boundary and six-boundary cases.

It should be noted that the preliminary tests in Chapter 4 provided evidence that liking can also be affected, indicating that the factors investigated can impact listener preference.

Conclusions

- Changes in loudspeaker directivity, in combination with changes in absorption, mostly affect perceived width, loudness and reverberence.

- A narrowing in on-axis directivity is associated with a perceived reduction in width, brightness, closeness and spaciousness and an increase in absorption associated with reduced loudness and reverberence.

- The perceptual changes caused by different loudspeaker directivity and absorption combinations are affected in parallel, along one dimension.

- These changes can also affect listener preference.
Chapter 7. Conclusions and Further Work

7.5 R5: What part does listener position play in the relationship between loudspeaker directivity and perception?

The outcomes from preliminary tests using two positions in Chapter 4 suggested that there was minimal difference between reproduction at one listening position and reproduction at another with respect to the presented attributes. However, when listeners were invited to evaluate loudspeakers at several positions around the room (with the freedom to move between positions) a directivity which causes a very diffuse, non-direct sound field was favoured and reported to be more consistent.

Whilst this suggests that the loudspeaker/listener position has some influence upon the perception of reproduction, results from the main experiments using the auralisation system with two adjacent, but expectedly different with regards to perception of reproduction, positions suggest that listener position has little effect, particularly when compared to the effect of changing directivity/absorption.

Conclusions

• With the auralisation system, the influence of listener position (based on two adjacent positions at equal distance away from the loudspeaker) is small in comparison to the other parameters. However, if the listener position is continuously varied over a wide area then it can have a significant effect - more work is needed in order to determine the exact size and nature of this effect.

Further Work

• In order to obtain more information regarding the role of listener position, a greater range of positions should be tested, perhaps allowing the listener to move freely within the acoustic environment.
7.6 R6: What is the best experimental method for this type of investigation?

It was established early on that experiments which identified and measured the attributes affected by changes in directivity were necessary for this investigation. Preliminary tests, however, showed that methods using real loudspeakers in rooms were limited by uncontrollable confounding factors which would make it difficult to identify clearly the effect of the variables of interest. It was also expected, at this stage, that the perceptual effects of changes in these variables would be small and therefore a system able to control and measure small differences was required (the variables of interest would be difficult to control in real rooms with real loudspeakers). It was established that an auralisation system would be most suitable, where acoustic simulations could be used to create any desired source/environment - it was shown that results from tests in virtual environments were representative of real-world data at a level sufficient to draw useful conclusions about what is likely to happen in real domestic listening rooms, with real loudspeakers.

After establishing that the auralisation system could be used, the final test method and analysis had to be decided. The rating of elicited or pre-selected attributes was considered, but concerns regarding the confinements of such methods arose - subjects would be unable to freely express their perceptual experience via their own lexicon and, although this may be acceptable in cases where a specific attribute is under investigation, this investigation was exploratory and so a method allowing more listener freedom was deemed most suitable. A number of options were available, all of which involved subjects comparing the stimuli to each other. Projective mapping in conjunction with MDS analysis was found to be most suitable, providing better quality results than sorting, and taking less time than pairwise comparison.

This method allowed the listeners to have immediate access to all stimuli and to quickly indicate the similarities by organising them within a two-dimensional space. Data relating to the similarities were extracted and used to find perceptual dimensions which were then linked to attributes which had been elicited from the listeners.

This method therefore: reduced confounding factors found in real rooms with real loudspeakers; allowed parameter control so that small perceptual differences could be measured; allowed listeners to label and measure stimuli according to their own lexicon and without fixed linear scales; and could be completed quickly and easily with a simple
test setup.

Conclusions

- **Listening tests using real loudspeakers and rooms are not ideal for the detailed investigation of directivity and its effect upon perception of reproduced sound in domestic listening rooms.**

- **Auralisation can be used as part of a listening experiment to provide useful information regarding perceptual effects of changes in loudspeaker directivity.**

- **Projective mapping in conjunction with MDS analysis is a suitable method for exploring the effects of directivity upon the perception of reproduced sound, where the requirement is to allow listener freedom with regard to attributes and scales, and to have short duration, simple tests.**

7.7 Research Contribution

A number of contributions to this area of research have been made as a result of the work in this thesis:

1. A detailed assessment of current knowledge regarding loudspeaker directivity and its likely affect upon the perception of sound reproduction in domestic listening rooms.

2. A strategy to measure the relationship between loudspeaker directivity and the perception of sound reproduction.

3. Details of a novel test system and analysis method which has been used in an experiment to show that previously assumed theoretical relationships between loudspeaker directivity and certain perceptual attributes do exist.

4. Numerical data and conclusions regarding loudspeaker directivity and its likely affect upon the perception of sound reproduction in domestic listening rooms.
All of these contributions may serve as a basis for further research in this area. As far as the author is aware, all are original and the first of their kind.

7.8 Limitations of Research

The primary limitations of this research lie in its use of simulated loudspeakers and acoustic environments. These exhibited some non-realistic characteristics (e.g. non-frequency-dependent directivity, non-frequency-dependent absorption characteristics, no low-frequency modelling of wave behaviour) and were rendered via non-individualised HRTFs. It can therefore not be guaranteed that the results would be identical if the research were to be repeated in real acoustic spaces. However, the use of real rooms would introduce many more variables and it is likely that, consequently, results would differ from room to room. The findings reported in this thesis indicate general trends that are likely to be observed across a majority of domestic listening environments.
Appendix A

Preliminary Studies

This appendix relates to Chapter 4. Here, selected screenshots of test interfaces, scans of information sheets given to listeners and information about the statistical analysis procedure relevant to both experiments in Surrey and Denmark are presented.

A.1 Attribute Elicitation

Items relating the initial attribute elicitation experiments, which took place in both Surrey and Denmark, are included here. Figure A.1 shows the interface used by listeners to control the playback via different loudspeakers and Figure A.2 shows the accompanying form used to describe the differences that they heard.

A.2 Attribute Rating

Screenshots of the GUIs used to control playback and gather rating data based on different attributes are shown here: Figure A.3 is the interface for liking, Figure A.4 the interface for timbral fidelity, Figure A.5 the interface for ensemble width and Figure A.6 the interface for loudness.

Also, scans of the instructions presented to listeners for the attribute rating part of the experiment are included (Figures A.7 and A.8).
"A sounds -------------- than B."

Describe up to 3 differences between A and B, then rank them in order of prominence.

Figure A.1: Interface of Max/MSP patch used for elicitation tests
Elicitation Experiment – William Evans, IoSR, University of Surrey

In this test you are asked to describe the differences between two randomly selected audio samples, A and B. You may switch freely between sample A and B, and should consider the most striking perceptual differences.

Once you have noted down three differences in the table below, please rank them in order of prominence, with 1 being the most prominent difference, and 3 the least prominent difference.

When you have completed a comparison between A and B, please press NEXT and repeat the above process. You will hear nine pairs of audio samples in total.

On completion of the test, please save the file as “Your Name”.

Thank you for your time. If you have any additional questions, please ask.

** To Start Audio, please press SPACE before choosing A or B. If at any time you should wish to stop Audio, Simply press SPACE.

<table>
<thead>
<tr>
<th>Page Number</th>
<th>Difference (Eg. “A sounds ______ than B”)</th>
<th>Rank (1, 2, or 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comments</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Page Number</th>
<th>Difference (Eg. “A sounds ______ than B”)</th>
<th>Rank (1, 2, or 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure A.2: Assessment form for Elicitation Testing
How much do you like these sounds?:

- Extremely Like
- Extremely Dislike

What is the timbral fidelity of these sounds?:

- Highest Imaginable
- Lowest Imaginable
What is the width of the ensemble?:

- A: 180°
- B: 135°
- C: 90°
- D: 45°
- E: 0°

Figure A.5: Max/MSP GUI for Ensemble Width Test

How loud are these sounds?:

- A: Extremely Loud
- B: Extremely Quiet

Figure A.6: Max/MSP GUI for Loudness Test
Listening Experiment – William Evans, IoSR, University of Surrey / Bang&Olufsen

For the following experiment you will be asked to complete a series of rating tests.

The format of all tests are the same – You must listen to the three audio samples - A, B and C, and provide a score/rating for each.

Once you have provided a rating for each sample you can move on to the next page. There are 6 pages in total for each test. When you have finished the test, please let me know and you will be moved on to the next test.

For the duration of these tests, imagine you are seated comfortably in the room below;

1. Loudness

For this test you must indicate the loudness of each sound based upon your own idea of what "loud" is.

2. Timbral Fidelity

Please indicate the timbral fidelity of these sounds as if they were being reproduced by the system in room pictured above. The system is a 2-channel stereo hi-fi.

3. Liking

In this test, please indicate how much you like each sound.
4. Ensemble Width

In this test, please indicate the width of the ensemble, as indicated in the diagram below.

Figure A.8: Listener instructions (Page 2)
A.3 Statistical Analysis Procedure

This section details the analysis procedure used for the results of the preliminary experiments at Surrey and Denmark.

• **Procedure A - Customised UniANOVA** - A 'custom' univariate analysis, with main effects and 2nd order interactions between the experimental factors is carried out in SPSS. Studentized residuals are saved for further inspection.
  
  - The factors 'repetition' and 'listener' are included as random variables (Bech and Zacharov 2006).
  
  - If 'repetition', or its interaction with another factor is significant, then its effect size must be evaluated. If it is negligible, then Procedure B should be used. If it is large, the listener scores should be observed more closely and screening considered. If 'repetition' is not significant, the next procedure should be used.
  
  - ANOVA requires that the data be from a normally distributed population and so kurtosis and skew of studentized residuals are assessed. Studentized residuals are a variation on standardized residuals and provide a more precise estimate of the error variance of a specific case.
  
  - For example:
    * Main effects - Listener, Recording, Loudspeaker, Repetition
    * 2nd order interactions -
      - Listener vs recording
      - Listener vs loudspeaker
      - Listener vs repetition
      - Recording vs loudspeaker
      - Recording vs repetition
      - Loudspeaker vs repetition

• **Procedure B - Full factorial UniANOVA** - A 'full factorial' univariate analysis is carried out in SPSS. Studentized residuals are saved for further inspection.
‘Listener’ is included as a random variable. ‘Repetition’ is excluded from the model (Bech and Zacharov 2006).

- The normality, kurtosis and skew of studentized residuals are assessed.
- The significance of factors and interaction between factors is assessed.

**Procedure C - Reduced UniANOVA** - A 'reduced' univariate analysis is carried out in SPSS (with any non-significant factors from Procedure B removed. Studentized residuals are saved for further inspection.

- ‘Listener’ is included as a random variable. ‘Repetition’ is excluded from the model (Bech and Zacharov 2006).
- The normality, kurtosis and skew of studentized residuals are assessed.
- The significance of factors and interaction between factors is assessed.
- Significant factors and interactions are plotted with confidence intervals.
Appendix B

Preliminary Studies - Surrey, UK

This appendix includes additional graphs and result details for the preliminary experiment carried out in Surrey.

B.1 Vertical Directivity Plots

Vertical directivity plots for each of the loudspeakers used in the preliminary experiment in Surrey are shown here: Loudspeaker 1 in Figure B.1, Loudspeaker 2 in Figure B.2 and Loudspeaker 3 in Figure B.3.
Figure B.1: Loudspeaker 1 (Surrey) Vertical Directivity Contour Plot
Figure B.2: Loudspeaker 2 (Surrey) Vertical Directivity Contour Plot
Figure B.3: Loudspeaker 3 (Surrey) Vertical Directivity Contour Plot
B.2 Results

More detailed information regarding the attribute rating results from the preliminary study in Surrey is presented here.

- All normality histograms have been checked and appear normal.
- Interactions in **bold** are plotted in Section B.3.
- If **Procedure C** is not included, it follows from the model not being able to be reduced further from **Procedure B**.

Liking (Surrey)

Significant factors (*Kolmogorov-Smirnov, Shapiro-Wilk*)

**Procedure A - Custom UniANOVA**

- Loudspeakers (.172, .024)
- Loudspeakers vs Listener
- Recording vs Loudspeakers

**Procedure B - Full factorial UniANOVA**

- Loudspeakers (.010, .002)
- Loudspeakers vs Listener

**Procedure C - Reduced UniANOVA**

- Loudspeakers (.034, .087)
- Loudspeakers vs Listener
Loudness (Surrey)

Significant factors (*Kolmogorov-Smirnov, Shapiro-Wilk*)

**Procedure A - Custom UniANOVA**

- Loudspeakers (.157, .027)
- Listener
- Loudspeakers vs Listener
- Recording vs Listener

**Procedure B - Full factorial UniANOVA**

- Loudspeakers (.040, .006)
- Listener
- Loudspeakers vs Listener
- Recording vs Listener

Timbral Fidelity (Surrey)

Significant factors (*Kolmogorov-Smirnov, Shapiro-Wilk*)

**Procedure A - Custom UniANOVA**

- Loudspeakers (.200, .385)
- Listener
- Loudspeakers vs Listener
- Recording vs Listener
- Loudspeakers vs Recording
- Recording vs Repetition *(\(\eta^2 = 0.001\))
Procedure B - Full factorial UniANOVA

- Loudspeakers (.000, .000)
- Listener
- **Loudspeakers vs Recording**
- **Loudspeakers vs Listener**
- **Recording vs Listener**

Ensemble Width (Surrey)

Significant factors (*Kolmogorov-Smirnov, Shapiro-Wilk*)

Procedure A - Custom UniANOVA

- Loudspeaker (.010, .000)
- Listener
- **Loudspeaker vs Listener**

Procedure B - Full factorial UniANOVA

- Loudspeaker (.200, .003)
- Recording
- Listener
- **Loudspeaker vs Recording**
- **Loudspeaker vs Listener**
B.3 Significant Interactions

Additional graphs showing the interactions not included in Chapter 4 (where loudspeaker and/or recording are significant factors) are presented in Figures B.4 to B.11 (for the results from preliminary experiments in Surrey).

![Figure B.4: Liking (Surrey): Loudspeakers vs Listeners](image)

Figure B.4: Liking (Surrey): Loudspeakers vs Listeners
Figure B.5: Loudness(Surrey): Loudspeakers vs Listeners

Figure B.6: Loudness(Surrey): Listeners vs Recording
Figure B.7: Timbral Fidelity (Surrey): Listeners vs Recording

Figure B.8: Timbral Fidelity (Surrey): Loudspeakers vs Listeners
APPENDIX B. PRELIMINARY STUDIES - SURREY, UK

Figure B.9: Timbral Fidelity (Surrey): Loudspeakers vs Recording

Figure B.10: Ensemble Width (Surrey): Loudspeakers vs Listeners
Figure B.11: Ensemble Width (Surrey): Loudspeakers vs Recording
Appendix C

Preliminary Studies - Struer, Denmark

This appendix includes additional graphs and result details for the preliminary experiment carried out in Denmark.

C.1 Horizontal Directivity Plots

The additional horizontal directivity plot for the array (Loudspeaker 3), based on the right channel output, is included here in Figure C.1.
Figure C.1: Loudspeaker 3 - Right Channel (Denmark) Horizontal Directivity Contour Plot
C.2 Vertical Directivity Plots

Vertical directivity plots for each of the loudspeakers used in the preliminary experiment in Denmark are shown here: Loudspeaker 1 in Figure C.2, Loudspeaker 2 in Figure C.3 and Loudspeaker 3 in Figure C.4.

![Figure C.2: Loudspeaker 1 (Denmark) Vertical Directivity Contour Plot](image)
Figure C.3: Loudspeaker 2 (Denmark) Vertical Directivity Contour Plot
Figure C.4: Loudspeaker 3 - Left Channel (Denmark) Vertical Directivity Contour Plot
Figure C.5: Loudspeaker 3 - Right Channel (Denmark) Vertical Directivity Contour Plot
C.3 Results

More detailed information regarding the attribute rating results from the preliminary study in Denmark is presented here.

- All normality histograms have been checked and appear normal.
- Interactions in bold are plotted in Section C.4.
- If Procedure C is not included, it follows from the model not being able to be reduced further from Procedure B.

Liking (Denmark)

Significant factors (Kolmogorov-Smirnov, Shapiro-Wilk)

Procedure A - Custom UniANOVA

- Recording vs Listener (.200, .410)
- Position vs Listener
- Loudspeaker vs Listener

Procedure B - Full factorial UniANOVA

- Recording vs Loudspeaker (.000, .001)
- Loudspeaker vs Listener

Procedure C - Reduced UniANOVA

- Recording vs Loudspeaker (.200, .222)
- Loudspeaker vs Listener
Loudness (Denmark)

Significant factors (*Kolmogorov-Smirnov, Shapiro-Wilk*)

Procedure A - Custom UniANOVA

- Recording (.000, .000)
- Loudspeaker vs Listener
- Position vs Listener
- Listener vs Repetition *(η² = 0.078)*
- Loudspeaker vs Position

Procedure B - Full factorial UniANOVA

- Recording (.000, .000)
- Loudspeaker vs Position
- Loudspeaker vs Listener
- Other 3rd and 4th order interactions...

Timbral Fidelity (Denmark)

Significant factors (*Kolmogorov-Smirnov, Shapiro-Wilk*)

Procedure A - Custom UniANOVA

- Recording vs Listener (.200, .768)
- Recording vs Loudspeaker
- Loudspeaker vs Listener
- Position vs Listener

Procedure B - Full factorial UniANOVA
• Recording vs Loudspeaker (.000, .000)
• Loudspeaker vs Listener
• Loudspeaker vs Position vs Listener

Procedure C - Reduced UniANOVA

• Recording vs Loudspeaker (.200, .706)
• Loudspeaker vs Listener

Ensemble Width (Denmark)

Significant factors (Kolmogorov-Smirnov, Shapiro-Wilk)

Procedure A - Custom UniANOVA

• Loudspeaker (.064, .012)
• Listener
• Recording vs Listener
• Recording vs Position
• Loudspeaker vs Listener
• Position vs Listener
• Loudspeaker vs Position

Procedure B - Full factorial UniANOVA

• Recording (.000, .000)
• Loudspeaker
• Listener
• Recording vs Loudspeaker
- Loudspeaker vs Position vs Listener

Procedure C - Reduced UniANOVA

- Recording (.006, .004)
- Loudspeaker
- Listener
- Recording vs Loudspeaker
- Loudspeaker vs Listener
C.4 Significant Interactions

Additional graphs showing the interactions not included in Chapter 4 (where loudspeaker and/or recording are significant factors) are presented in Figures C.6 to ?? (for the results from preliminary experiments in Denmark).

Figure C.6: Liking(Denmark): Loudspeakers vs Listeners
Figure C.7: Liking (Denmark): Loudspeakers vs Recording

Figure C.8: Loudness (Denmark): Loudspeakers vs Listener
Figure C.9: Loudness (Denmark): Loudspeakers vs Position

Figure C.10: Loudness (Denmark): Loudspeakers vs Recording
Figure C.11: Timbral Fidelity (Denmark): Loudspeakers vs Listener

Figure C.12: Timbral Fidelity (Denmark): Loudspeakers vs Recording
Figure C.13: Ensemble Width (Denmark): Loudspeakers vs Listeners
C.5 ‘Roaming’ Preference Test

The questionnaire given to listeners in the ‘Roaming’ preference test in Denmark is presented here in Figure C.14.
William Evans – PhD Research Student, University of Surrey / Bang & Olufsen

Listening Questionnaire

Please answer the following questions.

Listen to each of the sounds A, B and C around the room;

1. Which sound, A, B or C do you like the most?

2. Give the main reason for your answer in Question 1.

3. Where do you think they sound best?

   A ..........................................................................

   B ..........................................................................

   C ..........................................................................

4. Describe how you normally listen to music / audio at home.

   ...........................................................................

   ...........................................................................

   ...........................................................................

   ...........................................................................

Thank you for completing the questionnaire. Have a good day!

Figure C.14: Questionnaire given to listeners in ‘Roaming’ preference test
Appendix D

Main Experiment - Surrey, UK

This appendix includes additional instruction scans and verbal analysis results from the main experiment (Chapter 6).

D.1 Listener Instructions

The instructions presented to listeners for the main experiment are included here in Figures D.1 and D.2.

D.2 Occurrence of Descriptive Terms

The occurrence of the top antonyms compared to the next highest term and other terms used are presented here for the four-boundary (Figures D.3 to D.8) and six-boundary (Figures D.9 and D.14) experimental conditions.

D.3 Normality Tests

Results from normality tests of the main experimental data are presented in D.1.
<table>
<thead>
<tr>
<th>Condition</th>
<th>Excerpt</th>
<th>Kolmogorov-Smirnov</th>
<th>Shapiro-Wilk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-boundary</td>
<td>Classical</td>
<td>.200</td>
<td>.148</td>
</tr>
<tr>
<td></td>
<td>Pop</td>
<td>.200</td>
<td>.320</td>
</tr>
<tr>
<td></td>
<td>Speech</td>
<td>.200</td>
<td>.243</td>
</tr>
<tr>
<td>Four-boundary</td>
<td>Classical</td>
<td>.200</td>
<td>.381</td>
</tr>
<tr>
<td></td>
<td>Pop</td>
<td>.200</td>
<td>.540</td>
</tr>
<tr>
<td></td>
<td>Speech</td>
<td>.200</td>
<td>.081</td>
</tr>
<tr>
<td>Six-boundary</td>
<td>Classical</td>
<td>.116</td>
<td>.069</td>
</tr>
<tr>
<td></td>
<td>Pop</td>
<td>.168</td>
<td>.102</td>
</tr>
<tr>
<td></td>
<td>Speech</td>
<td>.065</td>
<td>.009</td>
</tr>
</tbody>
</table>

Table D.1: Normality Tests for Main Experiment Data
Listener Instructions

- You will be presented with 24 audio stimuli.
- Double-click any stimulus and it will play (after a short delay).
- Arrange the stimuli within the browser based on how similar you think they sound to each other, with similar stimuli close together and dissimilar stimuli further apart. You can move a stimulus by clicking, holding and dragging it.
- Once you feel the arrangement of all stimuli is representative of their similarities and differences, press ‘End’.
- Next, label the stimuli, either individually or in groups. Click ‘Start Class Definition’. Select the stimuli that you wish to label and then enter your description in the comments box. If you have selected more than one stimulus to be defined by the same term, you should choose which of those stimuli is the most typical of that class before you confirm your definition.
- Repeat this process until all stimuli are labelled.
- Once you have finished labelling all stimuli, you may reduce the volume of the audio playback (on the soundcard) and notify the experimenter that you have finished.
- The experimenter will save the data and close the experiment.
William Evans – *The Influence of Loudspeaker Directivity upon the Perception of Reproduced Sound in Domestic Listening Rooms*

*Napping* advice (adapted from Pages (2005))

**Principle.** You are asked to evaluate the similarities (or dissimilarities) between several audio stimuli. You have to do this according to your own criteria, those that are significant for you.

**Procedure.** You have to position the stimuli in the browser in such a way that two stimuli are very near if they seem identical to you and that two stimuli are distant from one another if they seem different to you. This must be done according to your own criteria. Do not hesitate to express strongly the differences you perceive by using the most part of the browser space.
Figure D.3: Number of occurrence of top antonyms and next highest term in the *Width* group (four-boundary)

Figure D.4: Number of occurrence of top antonyms and next highest term in the *Artifacts* group (four-boundary)
Figure D.5: Number of occurrence of top antonyms and next highest term in the Proximity group (four-boundary)

Figure D.6: Number of occurrence of top antonyms and next highest term in the Loudness group (four-boundary)
Figure D.7: Number of occurrence of top antonyms (spacious vs boxy/small) and next highest term in the *Spaciousness* group (four-boundary)

Figure D.8: Number of occurrence of top antonyms and next highest term in the *Reverberation* group (four-boundary)
Figure D.9: Number of occurrence of top antonyms and next highest term in the *Width* group (six-boundary)

Figure D.10: Number of occurrence of top antonyms and next highest term in the *Artifacts* group (six-boundary)
Figure D.11: Number of occurrence of top antonyms and next highest term in the Proximity group (six-boundary)

Figure D.12: Number of occurrence of top antonyms and next highest term in the Loudness group (six-boundary)
Figure D.13: Number of occurrence of top antonyms (spacious vs boxy/small) and next highest term in the Spaciousness group (six-boundary).

Figure D.14: Number of occurrence of top antonyms and next highest term in the Reverberation group (six-boundary).
Bibliography


BIBLIOGRAPHY


Bibliography


——— (2005), *High Performance Loudspeakers (Sixth Edition)*, Chichester, UK: John Wiley & Sons Ltd.


Gilbert, J. and Heymann, H. (1995), Comparison of four sensory methodologies as alternatives to descriptive analysis for the evaluation of apple essence aroma, The Food Technologist (NZIFST) 24, 28-32.

BIBLIOGRAPHY


Kruskal, J. (1964), Multidimensional scaling by optimising goodness of fit to a non metric hypothesis, Psychometrika 29 1-27.


Wickelmaier, F. (2003), An introduction to MDS, *Internal Document - Sound Quality Research Unit, Aalborg University, Denmark*.

