



**University of Salford**  
A Greater Manchester University

# Perception of Noise from Large Wind Turbines (EFP-06 Project)

Sabine von Hünenbein<sup>1</sup>, Andrew King<sup>1</sup>, Jonathan Hargreaves<sup>1</sup>, Andrew Moorhouse<sup>1</sup>, Chris Plack<sup>2</sup>



University of Salford, Acoustics Research Centre UK  
University of Manchester, Division of Human Communication and Deafness, UK



# **Perception of Noise from Large Wind Turbines**

**(EFP-06)**

## **Sabine von Hünenbein**

University of Salford, Acoustics Research Centre  
s.vonhunerbein@salford.ac.uk

## **Andrew King**

University of Salford, Acoustics Research Centre  
A.King@pgr.salford.ac.uk

## **Jonathan Hargreaves**

University of Salford, Acoustics Research Centre  
J.A.Hargreaves@salford.ac.uk

## **Andrew Moorhouse**

University of Salford, Acoustics Research Centre  
a.t.moorhouse@salford.ac.uk

## **Chris Plack**

The University of Manchester, Division of Human Communication and Deafness  
chris.plack@manchester.ac.uk

**The University of Salford, Greater Manchester, UK**  
**October 2010**

# Contents

Figures.....	iv
Tables.....	vii
List of Terms, Symbols, and Abbreviations.....	viii
1 Preface .....	9
2 Executive summary.....	9
3 Definitions .....	11
3.1 Types of noise.....	11
3.2 Low frequency sound .....	11
3.3 Audibility, masking thresholds and equal annoyance contours.....	11
3.4 Relative annoyance .....	11
4 Context and aims .....	12
4.1 Why study noise perception from large wind turbines? .....	12
4.2 Wind turbine noise characteristics and propagation .....	12
4.3 General methods for assessing noise annoyance.....	13
Survey studies .....	13
Laboratory experiments .....	14
Low frequency noise.....	14
Dose-response relations.....	15
4.4 Perception of wind turbine noise – questionnaire based studies .....	15
Survey based studies.....	15
Laboratory based studies .....	16
4.5 Aims and objectives .....	16
5 Listening Test Design .....	17
5.1 The listening environment .....	17
Acoustics .....	17
Room design.....	18
5.2 Participant recruitment and screening .....	19
Participant protection .....	20
5.3 Listening test software .....	21
5.4 Test procedure Part A.....	22
Aims and objectives.....	22
Description.....	22
Stimuli Part A.....	23
Stimuli overview .....	25
5.5 Test procedure Part B .....	26
Aims and objectives.....	26
Description.....	26
Stimuli Part B .....	26
Stimuli overview .....	27
6 Part A results: Audibility and annoyance from low frequency tones .....	29
6.1 Audibility threshold dependence on frequency and level .....	29
6.2 Validity of ISO1996-2 masking thresholds and broadband audibility threshold.....	31
Case studies on audibility of broadband noise near the hearing threshold .....	33
6.3 Equal Annoyance .....	35
6.4 Masking comparisons.....	37
7 Part B: The annoyance of large wind turbines compared to small wind turbines.....	39
8 Conclusions .....	42
8.1 Audibility and masking thresholds.....	42
8.2 Annoyance from idealised stimuli.....	42
8.3 Comparison of different masking sounds .....	43
8.4 Annoyance from recorded wind turbines in comparison to a reference noise source .....	43
8.5 Reproduction of realistic, calibrated sound at low frequencies .....	44
9 References .....	45
9.1 Web-Sources .....	48
10 Appendix I: Advertisement for participant recruitment and consent form .....	49

10.1	Advertisement for testing in April 2010 .....	49
10.2	Advertisement for testing in August 2010 .....	50
10.3	Consent form .....	51
11	Appendix II: Participant Screening .....	52
12	Appendix III – Sound Reproduction System .....	56
12.1	Hardware .....	56
12.2	Subwoofer System .....	56
12.3	Room Measurement Routine .....	57
12.4	Room Inversion Routine.....	59
12.5	Modification of Stimuli .....	60
12.6	Run Routines .....	60
12.7	Daily Checks .....	60
12.8	Loudspeaker Calibration Data .....	61
13	Appendix IV – Wind Turbine Signals.....	67
13.1	Part A – artificial stimuli.....	67
	Derivation of Part A stimuli specification from measured data .....	67
	Generation of auralisable stimuli from third-octave specification .....	67
13.2	Part B – real stimuli .....	69
	Derivation of Wind Turbine propagation attenuation spectra .....	69
	Selection of recordings .....	69
	Looping of recordings.....	70
	Filtering of recordings.....	71
13.3	Measurements of Stimuli.....	71
14	Appendix V: Pilot Tests.....	76
14.1	Audibility threshold .....	76
	Audibility threshold for Tones without masking.....	76
	Audibility threshold in indoor and outdoor scenarios .....	77
14.2	Equal annoyance of tones.....	78
	Equal annoyance with and without garden noise .....	78
	Conclusions of pilot tests .....	79
15	Appendix VI: Participant instructions and Graphical User Interfaces.....	80
15.1	Written instructions to Users.....	80
15.2	Graphical User Interfaces in Pilot Test.....	81
15.3	Graphical User Interfaces in Main Test.....	84
16	Appendix VII: Statistical validation of audibility threshold.....	87
17	Appendix VIII Statistical evaluation of equal annoyance results.....	89
17.1	Outdoors, omission versus inclusion of garden noise .....	89
17.2	Outdoors in comparison to Indoors .....	89
18	Appendix IX: Statistical evaluation of masking comparison .....	92
18.1	Difference in tone SPL .....	92
18.2	Difference in $L_{Aeq}$ .....	93
19	Appendix X: Statistical evaluation Part B results .....	94
	Appendix XI: Traffic recordings.....	95
20	Appendix XII: Results from first test series: April 2010 .....	98
20.1	Part A.....	98

# Figures

Figure 5.1a) Plan of listening position (white rectangle in centre), screen (transparent square), ambisonic loudspeaker ring (black squares), subwoofers (grey squares), suspended frame and curtain (octagon denoted position of frame, rounded octagon denotes approximate position of curtain at floor level). b) Photo of listening environment. ....	18
Figure 5.2 Approximate outdoor lighting arrangement.....	19
Figure 5.3 Top-level function data dependencies.....	21
Figure 5.4 1/3 octave masking Spectra: The indoor spectrum with garden noise was virtually identical to the one without garden noise below 400 Hz and was therefore not included separately in the listening test.....	24
Figure 5.5 Part B Recorded spectra a) outside b) inside, SPL inside lower due to façade attenuation .....	28
Figure 6.1: Outdoor scenario – Masking thresholds (blue), equal annoyance to 180 Hz tone at 5 dB audibility (green) and 10 dB audibility (red) of low frequency tones within masking noises as labelled. Additional green and red dots at 180 Hz denote the results from the masking comparisons discussed in Section 6.4 (for 5 dB and 10 dB above $L_{pn}$ respectively). Black solid line: Hearing threshold according to ISO 389-7). Dashed black line: Masking threshold according to ISO 1996-2. Error bars denote 95% confidence intervals.....	30
Figure 6.2 Indoor Scenario – Masking thresholds (blue), equal annoyance to 180 Hz tone at 5 dB audibility (green) and 10 dB audibility (red) of low frequency tones within façade attenuated masking noises as labelled. Black solid line: Hearing threshold according to ISO 389-7. Dashed black line: Masking threshold according to ISO 1996-2. Error bars denote 95% confidence intervals.....	30
Figure 6.3 Room background without masking noise - Audibility threshold (blue), equal annoyance to 180 Hz tone at 5 dB audibility (green) and 10 dB audibility (red) of low frequency tones. Black solid line: Hearing threshold according to ISO 389-7 (2005). Error bars denote 95% confidence intervals.....	31
Figure 6.4 a) Difference between tonal audibility levels in masking noise and tonal audibility levels in background as seen in Figure 6.3. All outdoor scenarios and some frequencies in the indoor scenario exceed tonal audibility threshold in quiet by more than 10 dB. This indicates that AT in masking noise are governed by the masking threshold. The indoor scenario has a number of frequencies within 5 dB of AT. b) Difference between tonal audibility levels in masking noise and masking threshold in dB tone level according to ISO 1996-2 Error bars denote 95% confidence intervals. Note that most of the error bars in b) include 0.....	32
Figure 6.5 Correlation of measured audibility thresholds in masking noise with masking threshold calculated in accordance with ISO1996-2. Linear regression shows a slope of 1. $R^2 = 0.96$ . Note that there is a high correlation between level, frequency and residuals.....	33
Figure 6.6 Audibility of background noise in listening room calculated according to Pedersen (2008).....	34
Figure 6.7 Audibility of masking noise in indoor scenario calculated according to Pedersen (2008).....	34
Figure 6.8 Relative sensation level for equal annoyance averaged over all masking noise types and scenarios at reference tone prominence levels of 5 dB (green) and 10 dB (red). Error bars show 95% confidence intervals.....	36
Figure 6.9 Alternatively derived relative sensation level for equal annoyance averaged over all masking noise levels a) outdoors, b) indoors at reference tone prominence levels of 5 dB (green) and 10 dB (red). Error bars show 95% confidence intervals. The alternative relative sensation level was derived by using $L_{pn}$ instead of the measured values of the audibility threshold.....	37
Figure 6.10 Equal annoyance level of a tone in quiet (room background noise) when matched to a tone in masking noise. The tone frequency was 180 Hz and the levels within masking were played at 5 dB (blue) and 10 dB (red) above $L_{pn}$ . Error bars denote 95% confidence intervals. ....	37
Figure 7.1 Equal annoyance contours for recorded turbines small wind turbine (blue lines) and a large wind turbine (red lines) matched to a neutral noise source (TN) a) Outdoor scenario without garden noise, b) Outdoor scenario with garden noise. Error bars are confidence intervals ( $\alpha = 0.05$ ). ....	39
Figure 7.2 Indoor scenario: Equal annoyance contours for recorded turbines: Small wind turbine (blue line) and large wind turbine (red line) matched to a neutral noise source (TN). Error bars show 95% confidence interval .....	40
Figure 7.3 Outdoor scenario: Critical band levels computed from hearing threshold weighted frequency analysis of the noise for the noise from the large turbine (blue) and the low turbine (magenta) with and	

without garden noise, dotted and solid lines respectively. The A-weighted level was 44 dB(A) for all curves.....	41
Figure 11.1 Zimmer-Ellermeier screening results a) Average score for volunteers living in the country compared to living in the city with red standard error bars and blue standard deviation bars, b) Distribution of scores in overall sample. ....	55
Figure 11.2 Test sample statistics, a) Age distribution of participants; b) Zimmer and Ellermeier NSS score distribution for participants. ....	55
Figure 12.1: Room Measurement GUI .....	58
Figure 12.2: Measurement Microphones.....	59
Figure 12.3: Loudspeaker Level Check GUI.....	61
Figure 12.4: Front Loudspeaker Response and Correction.....	61
Figure 12.5: Front Left Loudspeaker Response and Correction .....	62
Figure 12.6: Left Loudspeaker Response and Correction .....	62
Figure 12.7: Back Left Loudspeaker Response and Correction .....	63
Figure 12.8: Back Loudspeaker Response and Correction.....	63
Figure 12.9: Back Right Loudspeaker Response and Correction.....	64
Figure 12.10: Right Loudspeaker Response and Correction.....	64
Figure 12.11: Front Right Loudspeaker Response and Correction.....	65
Figure 12.12: Subwoofer Response and Correction .....	65
Figure 12.13: Measurement of total corrected loudspeaker system response.....	66
Figure 13.1: large turbine repetition period = 1.25 seconds.....	70
Figure 13.2: small turbine repetition period = 0.74 seconds .....	71
Figure 13.3 Measured and intended Part A Stimuli spectra in comparison – Indoors, Green Intended wind turbine broadband spectrum attenuated from 39 dB(A), Measured stimuli: Black Background noise in room, Blue including 180 Hz tone at 10 dB above wind turbine $L_{pn}$ , Magenta including 180 Hz tone at 5 dB above wind turbine $L_{pn}$ . The stimuli are dominated by the intended wind turbine spectrum below 600 Hz and by the background noise above 600 Hz. The domination of the local background would be even stronger in a typical living room environment because that would not be as sound-proof as the laboratory environment. The effect of the reference stimulus can be clearly seen in the respective octave band. The rest of the measured tone spectra are identical to within 3 dB.....	72
Figure 13.4 Measured and intended Part A Stimuli spectra in comparison – Indoors, Green Intended wind turbine broadband spectrum at 49 dB(A), Measured stimuli: Black Background noise in room, Blue including 32 Hz tone, Magenta including 44 Hz. The stimuli are dominated by the intended wind turbine below 6 kHz and by the room characteristics above 6 kHz. The effect of the reference stimuli can be clearly seen in the respective octave bands. The rest of the measured audible tone spectra are identical to within 1 dB. ....	72
Figure 13.5 Measured and intended Part A Stimuli spectra in comparison – Indoors, Green Intended wind turbine broadband spectrum at 49 dB(A), Measured stimuli: Black Background noise in room, Blue including 72 Hz tone, Magenta including 115 Hz. The stimuli are dominated by the intended wind turbine below 6 kHz and by the room characteristics above 6 kHz. The effect of the reference stimuli can be clearly seen in the respective octave bands. The rest of the measured audible tone spectra are identical to within 1 dB. ....	73
Figure 13.6 Measured and intended Part A Stimuli spectra in comparison – Indoors, Green Intended wind turbine broadband spectrum at 49 dB(A), Measured stimuli: Black Background noise in room, Blue including 180 Hz tone, Magenta including 400 Hz. The stimuli are dominated by the intended wind turbine below 6 kHz and by the room characteristics above 6 kHz. The effect of the reference stimuli can be clearly seen in the respective octave bands. The rest of the measured audible tone spectra are identical to within 1 dB. ....	73
Figure 13.7 Measured and intended Part B garden noise spectra in comparison – Green – Intended Garden Noise broadband spectrum at 44 dB(A), Black – Measured background noise in room, Blue – Measured Garden Noise spectrum. At frequencies above 600 Hz differences between the specified spectrum and the recorded spectrum exceed the intended accuracy of 5 dB due to destructive interference effects in the ambisonic system. This does not significantly affect the masking of Part A tones and it has also been confirmed that both turbine spectra in Part B are affected in a similar way. When listening to the garden noise, the stimulus was perceived as sounding like natural garden noise in spite of the reduced high frequency levels. ....	74

Figure 13.8 Examples of measured Part A outdoor stimuli FFT spectra at broadband level of 44 dB(A), including a) 72Hz and b) 400 Hz tones.....	75
Figure 14.1 Audibility threshold of tones without masking noise.....	76
Figure 14.2 Audibility thresholds of tones in wind turbine noise masking, no garden noise. a) indoor, b) outdoor.....	77
Figure 14.3 Equal annoyance, tones only for two different reference tone levels 5 dB apart and two participants.....	78
Figure 14.4 Equal annoyance: Effect of using garden noise.....	78
Figure 15.1 Audibility threshold Instructions.....	81
Figure 15.2 Audibility threshold Control .....	81
Figure 15.3 Audibility threshold Scientific Monitoring Panel.....	82
Figure 15.4 Equal annoyance Instructions .....	83
Figure 15.5 Equal annoyance Control .....	83
Figure 15.6 Audibility threshold Instructions.....	84
Figure 15.7 Audibility threshold Control .....	84
Figure 15.8: Audibility threshold scientific monitoring panel.....	85
Figure 15.9 Equal annoyance instructions .....	85
Figure 15.10 Equal annoyance control.....	86
Figure 17.1 Outdoor scenario: Relative sensation level for equal annoyance with various masking noises at tone prominence levels of 5 dB (green) and 10 dB (red). a) $L_{Aeq,WT}$ 39 dB(A), no garden noise, b) $L_{Aeq,WT}$ 44 dB(A), no garden noise, c) $L_{Aeq,WT}$ 49 dB(A), no garden noise, d) $L_{Aeq,WT}$ 39 dB(A), with garden noise, e) $L_{Aeq,WT}$ 44 dB(A), with garden noise and f) $L_{Aeq,WT}$ 49 dB(A), with garden noise. Error bars show 95% confidence intervals.....	90
Figure 17.2 Indoor scenario: Relative sensation level for equal annoyance in the outdoor scenarios with various masking noises at tone prominence levels of 5 dB (green) and 10 dB (red). a) $L_{Aeq,WT}$ 39 dB(A), b) $L_{Aeq,WT}$ 44 dB(A), c) $L_{Aeq,WT}$ 49 dB(A), Error bars show 95% confidence intervals. ....	90
Figure 17.3 Relative sensation level for equal annoyance without masking noise at tone prominence levels of 5 dB (green) and 10 dB (red). Error bars show 95% confidence intervals. ....	91
Figure 20.1 Map of measurement side including distances from motorway. ....	95
Figure 20.2 a) Traffic recording position 75m away from the motorway. b) Dense motorway traffic consisting of a combination of cars, vans and lorries. ....	97
Figure 21.1 Room background without masking noise - Audibility threshold (blue), equal annoyance to 180 Hz tone at 5 dB audibility (green) and 10 dB audibility (red) of low frequency tones. Error bars denote standard errors.....	98
Figure 21.2 Outdoor scenario – Masking thresholds (blue), equal annoyance to 180 Hz tone at 5 dB audibility (green) and 10 dB audibility (red) of low frequency tones within masking noise a) 39 dB(A), b) 44 dB(A), c) 49 dB(A), d) 39 dB(A) and garden noise, e) 44 dB(A) and garden noise and f) 49 dB(A) and garden noise. Additional green and red dots at 180 Hz denote the results from the masking comparisons discussed in Section 6.4 (for 5 dB and 10 dB above $L_{pn}$ respectively). Error bars denote 95% standard error. ....	99
Figure 21.3 Indoor Scenario – Masking thresholds (blue), equal annoyance to 180 Hz tone at 5 dB audibility (green) and 10 dB audibility (red) of low frequency tones within façade attenuated masking noise comprising of $L_{Aeq,WT}$ a) 39 dB(A), b) 44 dB(A), c) 49 dB(A), d) 39 dB(A) and garden noise, e) 44 dB(A) and garden noise and f) 49 dB(A).. Error bars denote standard error. ....	100
Figure 21.4 Equal annoyance level difference between a tone in quiet (room background noise) and a tone in masking noise. The tone frequency was 180 Hz and the levels within masking were played at 5 dB (green) and 10 dB (red) above $L_{pn}$ . a) indoor, b) outdoor Error bars denote standard error. ....	100
Figure 21.5 Outdoor scenario: Relative sensation level for equal annoyance with various masking noises at tone prominence levels of 5 dB (green) and 10 dB (red). a) $L_{Aeq,WT}$ 39 dB(A), no garden noise, b) $L_{Aeq,WT}$ 44 dB(A), no garden noise, c) $L_{Aeq,WT}$ 49 dB(A), no garden noise, d) $L_{Aeq,WT}$ 39 dB(A), with garden noise, e) $L_{Aeq,WT}$ 44 dB(A), with garden noise and f) $L_{Aeq,WT}$ 49 dB(A), with garden noise. Error bars show standard error. ....	101
Figure 21.6 Indoor scenario: Relative sensation level for equal annoyance with various masking noises at tone prominence levels of 5 dB (green) and 10 dB (red). a) $L_{Aeq,WT}$ 39 dB(A), no garden noise, b) $L_{Aeq,WT}$ 44 dB(A), no garden noise, c) $L_{Aeq,WT}$ 49 dB(A), no garden noise, d) $L_{Aeq,WT}$ 39 dB(A),	



with garden noise, e) LAeq,WT 44 dB(A), with garden noise and f) LAeq,WT 49 dB(A), with garden noise. Error bars show standard error. .... 102

Figure 21.7 Relative sensation level for equal annoyance without masking noise at tone prominence levels of 5 dB (green) and 10 dB (red). Error bars show standard error. .... 102

## Tables

Table 5.1: Part A1: Audibility Threshold: LF Tones at six frequencies, three different background noise levels (WTS) of 39 dB, 44 dB and 49 dB, FA = Façade Attenuation, GN is Garden Noise	25
Table 5.2: Part A2 Equal Annoyance: Reference Tone at 180 Hz at levels 5dB and 10 dB above audibility threshold. Tones at five frequencies to be adjusted to equal annoyance to reference tone, three different background noise levels (WTS) of 39 dB, 44 dB and 49 dB, FA = Façade Attenuation, GN is Garden Noise	25
Table 5.3: Part A3 Equal Annoyance of reference tone in added background spectra in comparison to the pure reference tone.	26
Table 5.4: Part B Equal Annoyance distance: Levels of Small Wind Turbine (SWT) and Large Wind Turbine (LWT) to be adjusted to equal annoyance to traffic reference stimulus, FA = Façade Attenuation, GN = Garden Noise	27
Table 13.1 Tonal analysis output from noiseLAB software.	74
Table 14.1 Number of pilot test participants for Part A Tone Audibility Threshold (AT), Equal Annoyance (EA) and Equal Loudness (EL) in the three different scenarios as specified in Section 5.5.	76
Table 16.1 Details of the 2 ANOVAs on AT data	87
Table 16.2 ANOVAGN Tests of Within-Subjects Effects	88
Table 16.3 ANOVAGN Tests of Within-Subjects Contrasts	88
Table 16.4 ANOVAScen Tests of Within-Subjects Effects	88
Table 16.5 ANOVAScen Tests of Within-Subjects Contrasts	88
Table 17.1 Details of the 2 ANOVAs on relative sensation levels	89
Table 17.2 ANOVA <sub>GN</sub>	91
Table 17.3 ANOVA <sub>Scen</sub>	91
Table 18.1 ANOVA <sub>GN</sub>	92
Table 18.2 ANOVA <sub>Scen</sub>	92
Table 19.1 Details of the 2 ANOVAs on TN changes for EA to recorded WTN.	94
Table 19.2 ANOVA <sub>GN</sub>	94
Table 19.4 ANOVA <sub>Scen</sub>	94
Table 18.1 Overview of traffic noise recordings.	96

# List of Terms, Symbols, and Abbreviations

Symbol or abbreviation	Meaning
AT	Audibility threshold of tones either in quiet or above masking noise
BN	Environmental or ambient background noise. In the listening environment this is the measured noise in the room in quiet consisting of room and sound reproduction system background noise.
EA	Equal Annoyance
FA	Façade attenuation
GN	Garden noise: Sum of vegetation and background noise
HH	Distance in units of turbine hub height
$L_{EA}$	Measured Equal Annoyance Sound Pressure Level re 20 $\mu$ Pa
$L_{EA-T}$	Equal Annoyance sound pressure level of test tones above audibility threshold or masking threshold
$L_{pn}$	Total sound pressure level of masking noise in critical band according to ISO 1996-2
$L_{pt}$	Sound pressure level of tone according to ISO 1996-2
$L_R$	Sound pressure level $L_{pt}$ of the reference tone at 180 Hz above $L_{pn}$
$L_{R-T}$	Sound pressure level of reference tone at 180 Hz above measured audibility/masking threshold
$L_T$	Sound pressure level of measured audibility threshold for tone re 20 $\mu$ Pa.
RSL	Relative Sensation Level: $L_{EA-T} - L_{R-T}$
$L_{Aeq,WT}$	A-weighted equivalent-continuous sound level of broadband, wind turbine noise present in background noise.
LF(N)	Low Frequency (Noise).
LWT	Large Wind Turbine.
MN	Masking noise: Combination of all or some of wind turbine noise, Vegetation Noise and background noise This is broadband noise with no audible tonal components.
MT	Masking threshold: the threshold of tonal audibility determined by masking noise
RT	Reverberation Time
SPL	Sound Pressure Level
SWT	Small Wind Turbine.
TN	(Road) Traffic Noise.
VN	Vegetation noise: Tests Part, A artificial vegetation noise, Part B, recording of coniferous trees rustling in the wind as detailed in Appendix IV, Section 13.3
WT	Wind turbine
WTN	Wind turbine (broadband) noise

# 1 Preface

The work presented in this report is part of the EFP-06 project called “Low Frequency Noise from Large Wind Turbines – Perception of Noise from Large Wind Turbines”. The project is funded by the Danish Energy Agency under contract number 033001/33032-008. Supplementary funding to the project is given by Vestas Wind Systems A/S, Siemens Wind Power A/S, Vattenfall AB Vindkraft, E.ON Vind Sverige AB.

The project has been carried out in cooperation between DELTA and Salford University.

## 2 Executive summary

Is noise from large wind turbines more annoying than noise from small wind turbines? This is a question that is discussed widely in the context of a new generation of large wind turbines replacing the traditional smaller ones. To date legislation takes noise levels and the tonality of noise sources into account. However, many more influencing factors are known from the psychoacoustic literature. Examples are the nature of the listening environment, spectral and temporal characteristics of the sound, and the influence of masking noise. An earlier part of the EFP-06 project established the measurable differences between large and small wind turbines. It concluded that spectral characteristics are generally very similar apart from a slight increase in the low frequency content of large turbines.

In this study on the perception of wind turbine noise, audibility thresholds and equal annoyance contours have been established for idealised wind turbine sounds containing low frequency tones. The listening test simulated an indoor scenario and an outdoor scenario with and without masking garden noise.

The focus has been on the question whether annoyance changes with the frequency of a tone. The test sounds consisted of a broadband spectrum with a specific tone at one of the frequencies 32, 44, 72, 115, 180 and 400 Hz. Idealised sounds with features broadly representative of wind turbine sounds were used. The participant were asked to imagine being in different scenarios. The outdoor scenario presented sounds broadly representative of a wind turbine at three A-weighted sound pressure levels, each with and without garden noise, whereas the indoor scenario omitted the garden noise since the facade attenuation rendered it inaudible. A comparative adaptive method was used to establish relative equal annoyance levels in the form of equal annoyance contours. The results enable comparisons between different scenarios, broadband levels, tone frequencies, masked and unmasked ‘wind turbine’ sound, and two different prominence levels for the reference tone. Temporal variation like “swishing” was avoided to keep the research questions well focused.

In a second part of the study wind turbine recordings from a large and a small wind turbine were compared in annoyance with steady traffic noise. The recordings were manipulated to include the effect of sound propagation and façade attenuation. They were also normalised to equal A-weighted levels.

The study concludes:

Tones in quiet were heard at levels that agree well with hearing thresholds published elsewhere. As the broadband noise level increases the tones were heard at levels that were determined by the masking level. Masking thresholds predicted by the ISO 1996-2 standard have been shown to agree well with the measured tonal audibility thresholds as long as the masking noise clearly exceeds the hearing threshold of the tones. As low levels can frequently occur indoors in the neighbourhood of wind turbines when the Danish noise regulations are observed it would be useful to extend the standard to include a method to evaluate the hearing threshold.

One possible method published by Pedersen (2008) to establish the audibility of broadband spectra has been successfully tested for two examples: a broadband spectrum of room background noise and the broadband spectra of wind turbine noise at levels close to the hearing threshold. The calculated critical band levels agree to within 2 dB with perceived audibility. Theoretical considerations support the conclusion that the method should be adequate for use in standard applications.

Low frequency tones had to be adjusted to higher tone levels above the masking threshold to be equally annoying as higher frequency tones. Garden noise was not shown to reduce annoyance because different scenarios could not be compared easily.

It was shown that increasing the tone level by 5 dB increases the equal annoyance level by a smaller value both for tone frequencies lower than 180 Hz and at 400 Hz. This casts doubt on the appropriateness of the adjustment used in the ISO 1996-2 standard which adds penalty adjustments which are increasing linearly with sound pressure level above masking.

Relative sensation levels were calculated from equal annoyance contours to determine whether low frequency tones are relatively more annoying than high frequency tones. The frequency dependence was not shown to be significant. The main influence on these levels is the tone level above masking level: Tones at higher levels are more annoying than tones at lower levels above masking. Both findings are common for the indoor and outdoor scenarios.

To compare real recordings of a large and a small wind turbine a test protocol was developed. This method was successfully trialled.

The comparison between normalised recordings showed the spectral characteristics of the small turbine to be more annoying outdoors than those of the large turbine recording. This has been attributed to the different spectral characteristics of the two turbines. These differences are effectively masked by garden noise and the equal annoyance ratings change accordingly. The indoor scenario does also not find the turbines to be differently annoying. If these results can be reproduced in other listening experiments then it follows that the specific differences in spectral content will determine the annoyance levels from a wind turbine more than whether it is a small or a large turbine. It would also mean that the differences in annoyance between wind turbines get smaller when sufficient masking noise is present. Presently, the finding that the small turbine is more annoying cannot be generalised to large and small wind turbines or to a wider range of wind and terrain conditions than were used in the test. The listener responses were however consistent and therefore demonstrate the potential of the comparison method.

Another significant achievement of this project was of technical nature: It was the design of an immersive sound reproduction system that is calibrated to high precision over the largest part of the audible frequency range including low frequencies down to 30 Hz. It has been shown that this design is possible and that the stimuli sound realistic. Future listening test with similar requirements will therefore be easy to design and fast to perform.

In answer to the initial question whether large turbines are more annoying than small wind turbines, the results of this study find no evidence for a significant difference in annoyance between small and large wind turbines as long as total noise levels and tonal characteristics are taken into account in the assessment. Temporal variations of wind turbine noise such as the level of swishing might also have to be evaluated in the future.

## 3 Definitions

Before motivating the test design a number of terms need to be introduced.

### 3.1 Types of noise

Different types of noise are present in a listening environment. **Background noise** (BN) is environmental or ambient background noise. In the listening room environment this is the measured noise in the room in quiet consisting of room and sound reproduction system background noise.

**Wind turbine** (broadband) **noise** (WTN) is the noise from a wind turbine only. Its equivalent A-weighted level is used to characterise the noise and denoted by  $L_{Aeq,WT}$ . In general wind turbine noise can also contain tones which might be heard above the broadband noise.

The wind turbine noise at the listener position can be masked by ambient noise. One such type of noise is **vegetation noise** (VN). It is also commonly called **garden noise** (GN) which in the context of this study is the sum of vegetation and background noise. The term **masking noise** (MN) is used for a combination of all or some of wind turbine noise, vegetation noise and background noise.

### 3.2 Low frequency sound

All of the above noise types except the tones have broadband spectral characteristics. In the debate about wind turbine noise, the role of low frequency sound has been discussed intensely. Various definitions of the frequency range of low frequency sound and infrasound have been published. In this report we follow Pedersen (2008) and S ndergaard & Madsen (2008) who define infrasound to occur below 20 Hz and low frequency sound between 20 Hz and 200 Hz. The focus of this study is on low frequency sound as turbines are not thought to emit audible infrasound (Pedersen, 2008).

### 3.3 Audibility, masking thresholds and equal annoyance contours

The **hearing threshold** (HT) is defined as the minimum sound level of a pure tone that a person of normal hearing can detect when there is no other sound present. Measured hearing thresholds have been published by researchers such as Fastl & Zwicker (2007) and Pedersen (2008) and specifically for low frequencies by M ller & Pedersen (2004). In the current study the hearing threshold in room background has also been measured for low frequencies. The threshold that was measured in the current study will be referred to as **audibility threshold** (AT) for easy distinction.

To assess the audibility of tones within broadband noise the standard ISO 1996-2 is commonly used. It defines the widths of so called **critical bands**. The total sound pressure level (SPL) of a sound in a critical band determines whether a sound can be heard and how loud it is. Within the critical band the total sound pressure level of the tone  $L_{pt}$  and the total sound pressure level of the masking noise  $L_{pn}$  are calculated separately. The **masking threshold** (MT) is then defined by  $L_{pn}$  corrected by 2 dB and a frequency dependent term as defined in the ISO 1996-2 standard. The tone is thought to be audible when  $L_{pt}$  exceeds the masking threshold.

### 3.4 Relative annoyance

As a reaction to noise annoyance is influenced by the type and character of noise and by the listener's attitudes, personality and context (Pedersen, 2007).

Annoyance is frequently measured on scales that are subdivided into a number of categories from "highly annoyed" to "not annoyed". The disadvantage of using absolute scales in a laboratory study is that the test participants will be outside their usual context, they might be ignorant about the type of noise they are listening to, and noise exposure timescales are different from those heard

from the natural noise source. The participants will therefore react differently which makes the absolute annoyance scale meaningless.

In this study we have consequently designed the listening tests on annoyance comparisons between sounds so that test participants are adapting the sound pressure level of a particular stimulus to match the annoyance of another. The results are then seen as **equal annoyance** (EA) contours.

## 4 Context and aims

The following literature review starts off answering the question why there is a need for a study on noise from large wind turbines separately from general wind turbine noise studies. The chapter also serves to introduce the concepts of the nature of wind turbine noise and the effects of sound propagation on the characteristics of the noise, both of which have an impact on perception. It then moves on to review a number of methods that have been used to study the perception of general wind turbine noise and low frequency noise in the context of other noise sources. The context chapter does not attempt to be an exhaustive literature review but only cites work necessary to motivate the study design.

### 4.1 Why study noise perception from large wind turbines?

The modern generation of industrial wind turbines with typical hub heights of 80 m and more is a lot taller than the first generation. Noise legislation in many countries was tailored towards this first generation of wind turbines and concerns have been raised as to whether the legislation is still adequate. Therefore the project on Low Frequency Noise from Large Wind Turbines (EFP-06) was specifically devised to address these concerns.

### 4.2 Wind turbine noise characteristics and propagation

Concern about wind turbine noise is one of the major obstacles to more widespread use of wind energy (Oerlemans et al., 2007). Previous research of wind turbine emission has identified the main noise sources in a wind turbine. There are two types of noise. Mechanical based emissions are mainly created by gears and generators, whilst aerodynamic based emissions arise from the interaction between blades and air flow (Jakobsen, 2005; Oerlemans et al. 2007). Unlike aerodynamic noise, mechanical noise tends to be more tonal in character and can also be modulated in amplitude and frequency (Jakobsen, 2005; McKenzie, 1997).

Aerodynamic noise tends to be broadband in spectrum and its amplitude is modulated with the blade passing frequency of typically ~1 Hz resulting in a 'swishing' sound (Jakobsen, 2005). This swishing sound has been previously studied (van den Berg, 2005) and a metric for it was quantified from a laboratory study (Legarth, 2007). A number of authors such as Oerlemans et al. (2009) and Leloudas et al. (2007) have published methods for reducing aerodynamic noise by modifying the wind turbine response to airflow. However, some new, larger turbine emissions can have a minor increase in low frequency noise (LFN) due to mechanical noise.

In the first stage of EFP-06, DELTA, Riso DTU, and DONG Energy collaborated to characterise the physical properties of wind turbine noise from the source along the propagation path all the way to the recipient. The summary report (Søndergaard and Madsen, 2008) specifies the main outcomes of this part of the project: It confirmed that modern upwind turbines do not emit audible infrasound.

Public concern about the health effects of low frequency sound from wind turbines has been reported elsewhere (Jung et al. 2008). Jakobsen (2005) have shown that older turbines where the blades pass downwind of the tower create more LFN and infrasound than upwind turbines.

Frequency spectra from the large wind turbines included in EFP-06 were found to be very similar to typical spectra from small wind turbines (Søndergaard & Madsen, 2008). The only observed difference in spectral characteristics of large turbines was a relative increase in LFN of less than 2-3 dB which was mainly due to mechanical noise such as gear-noise. Søndergaard & Madsen concluded that low frequency indoors levels are not expected to increase for large turbines. This finding was opposed by Pedersen & Møller (2010) who found a small enhancement of the low frequency levels of 1.5 – 3.2 dB from the same measurements where 2 large wind turbine recordings are compared to 37 small wind turbine recordings. The authors claim that these small level differences could potentially make a significant difference in noise perception. Recent results by Madsen that included a larger number of large wind turbines, suggest that differences in sound emission between the individual wind turbine makes and models are generally much larger than the general difference between small and large turbines (Madsen, 2010).

In general noise propagation from wind turbines is determined by source directivity, geometric spreading and atmospheric absorption, ground reflection and absorption, meteorological effects and terrain complexity. The audibility of noise from wind turbines is then determined by, amongst other factors, the ratio between turbine and masking noise, the masking effect. Geometric spreading of sound decreases sound levels with increasing distance from any source as does atmospheric absorption. The latter is more effective at high frequencies with the consequence that low frequency sound travels further. The strength of ground reflections is determined by the surface properties where soft, porous ground reflects a lot less efficiently than hard surfaces such as tarmac and water. The amount of reflected sound also depends on the source height and the reception of reflected sound on the listener height. The main influence of meteorology is the change in sound speed profile due to wind shear and temperature profiles which is well described in Wagner et al. (1996). Most pronounced is the “shadow zone” upwind of the noise source and the increased noise levels in the downwind direction. In situations with large wind shear – low wind speed near the ground and high wind speed at hub height – the masking noise created for example by vegetation noise close to the ground is small. The turbine blades experience a higher wind speed resulting in higher aerodynamic noise levels than might be expected based on the wind speed near the ground. The lack of masking noise then leaves the turbine noise more audible.

### **4.3 General methods for assessing noise annoyance**

The mentioned concerns about wind turbine noise have given rise to a number of studies on noise annoyance. However, this is a relatively recent area of research whereas the general subject of noise annoyance is much older. The following two sections outline general methods of perception research before the details of recent wind turbine related work is reviewed in Section 4.4.

The perception of sound has been studied using two main methods. The first uses scaling magnitude estimation: A participant assigns a numerical value to a stimulus that quantifies the property (loudness, annoyance, etc). The method was developed by Stevens (1955) and has been applied to the assessment of noise annoyance in laboratory studies (Hellman, 1982, 1984; Berglund et al 1975, 1976). Another method is paired comparison, whereby two stimuli containing examples of the property are presented and a two-way rating scale is adjusted to indicate the relative rating of the two stimuli. Alternatively, one of the two stimuli can be actively adjusted by the participant until the two are equally representative for the property being studied. While the former method can be used in survey studies and laboratory experiments the latter one is naturally restricted to laboratory environments.

### **Survey studies**

At present, the majority of work focuses on measuring the environmental noise levels, either at the source and using propagation algorithms or at nearby residences, and acquiring annoyance ratings

via surveys. These two measures are combined to create dose-response relations for noise levels (or any other characteristic) and community annoyance directed at source. Schultz (1978) provides a good synthesis of 11 such examples for various forms of transport noise.

Survey studies have the advantage that they measure in the listener's natural environment. Therefore context and attitude can be taken into account. The disadvantages are that these studies are retrospective studies on an emission that already exists. Apart from the well known problems with this method it is not applicable in a situation where there are few large wind turbines. This limits the available data base where a large number of participants would be necessary because of the source variability and other factors.

## **Laboratory experiments**

Another way of studying environmental noise annoyance is to present either recordings of noise or similar, synthesised sounds to participants in the controlled environment of a laboratory such as an anechoic chamber or listening room/booth. Many of the physical properties of sound and environment and some personal variables can be controlled, thereby allowing accurate estimates of how acoustical parameters affect noise annoyance. To an extent, researchers can choose to study noise annoyance with respect to its non-acoustical factors, although never quite as realistically as in field studies, by including non-acoustic sensory stimuli typically associated with the noise source.

Additionally, an experimental design can either allow or restrict the influence of context by, for example, asking participants to imagine a particular scenario during exposure. For attitude, the participant can be explicitly informed of the source thereby allowing their expectations and previous experiences of the source to influence their ratings of annoyance. Alternatively, researchers can study noise annoyance purely from an acoustical perspective by limiting other sensory stimuli or keeping them constant, removing contextual cues and keeping participants naïve to the source.

However well designed, a laboratory experiment will never give the same absolute ratings as a survey study because of the laboratory environment is incompatible with the natural environment where the noise occurs. Therefore relative annoyance measures will give a better impression when comparative results are useful. Because this current project studies the perception of large turbine noise as compared to small turbine noise, it needs to be explored whether laboratory measurements using relative annoyance measures can be an appropriate method to use.

## **Low frequency noise**

One example of such a study in a different context was published by Bradley (1994). The author studied the effect of LFN level on simulated heating, ventilation and air conditioning systems by having participants change the level of a 'test' sound until they felt it was equally annoying as a reference sound. The test sound was the reference plus varying amounts of extra LFN. Using this simple paradigm, Bradley (1994) was able to accurately predict the attenuation needed to counteract additional annoyance from increased and/or modulated LF content. Bradley (1994) recognized that any noise rating procedure derived from the study would need evaluating in actual office environments, but this relative annoyance rating design provides a good way to quantify annoyance from certain acoustical properties in the laboratory. Any absolute rating of an environmental noise's annoyance in a laboratory will be out of context and therefore unrepresentative of the noise's annoyance within context. However, relative ratings provide their own reference and do not necessarily require any context.

Niedzweicki and Ribner (1979) studied low frequency and infrasound equal loudness and annoyance levels using a similar method. They found that filtering out the low frequencies produced a barely perceptible reduction in subjective loudness, but *increased* annoyance. This is supported by Key (1979) and Turner and Burns (1977). The latter showed that high frequency, 1/3 octave band white noise could be made less annoying by including low frequency 1/3 octave band white noise, despite this resulting in an increase in sound energy. However, more recent research



by Huang et al. (2008), Pawlaczyk-Łuszczynska et al. (2003, 2009), Persson and Björkman (1988) and Nilsson (2007) has suggested that noise with higher low frequency content is more annoying than neutral or high frequency noise. Although other work on low frequency noise has been published by authors such as Møller and Leventhall, the literature cited above shows already that there are different findings over whether LFN has a positive or negative effect on annoyance. This suggests that either techniques have improved and the latter is true, or that annoyance from LFN is a complex phenomenon whose effect is dependent on definition of LF being used and the specific paradigm or scenario being studied.

The published work using relative annoyance measures shows promising results. This fact led to the decision to design the current study using a laboratory environment and relative annoyance measures. The work on low frequency noise shows that there is still a need to understand the relation between the noise emissions and the resulting annoyance.

## **Dose-response relations**

This relation is important because many countries have introduced legislation that uses, often time- and frequency-weighted, sound level limits to restrict source emissions to acceptable levels. To determine this level limit, source emissions are measured and then the annoyance caused to local residents by that source either estimated or determined through survey studies. If sufficient data is available then it might be possible to derive a dose-response relationship which the annoyance, averaged over the population, can be predicted for a source type whose noise emissions are known. Once an acceptable level of annoyance is decided, legislation can prevent new noise sources from being too close or too noisy to local residents.

For traffic noise an example are the dose-response relations by Schultz (1978). He found a single dose-response relation for all types of traffic noise. However, work by Kryter (2007) and Pedersen (2009) amongst others suggest dose-response relations vary for different sources. This shows that while reliable dose-response relations are desirable to find legal limits on noise emissions they might not be easily found. If a dose-response relation between low frequency wind turbine emissions and annoyance was found this would be useful to guide legislation.

## **4.4 Perception of wind turbine noise – questionnaire based studies**

In the previous section general methods for studying the annoyance from noise emissions have been introduced. Wind turbine noise research is much more recent than the studies mentioned before. The following paragraph summarises results from recent wind turbine related studies most of which were survey based. The results from previous laboratory based studies shows that they have not focussed on the low frequency aspect of large wind turbines.

### **Survey based studies**

Pedersen and Persson Waye (2007) and a follow up study by Pedersen, van den Berg, Bakker and Bouma (2009) came up with a dose-response relationship for wind turbine noise in Sweden and in the Netherlands respectively. The results of the two studies were very similar suggesting that the relation can be generalised.

Pedersen and Persson Waye (2004) and Pedersen et al. (2009) also found that wind turbine noise is more annoying than transport or other industrial noises, which Pedersen et al. state may be due to its persistence throughout the night. In addition to this, it may be more pronounced at night due to atmospheric conditions (van den Berg, 2008) and lower sound levels from other sources. In this context the role of atmospheric stability has been widely discussed. Atmospheric stability can lead to the wind velocity at wind turbine hub height being larger than at ground level. This results in less wind-induced masking noise from vegetation as compared to wind turbine noise (van den Berg, 2004, 2005, 2008), which could make the wind turbine noise more prominent within the general

soundscape. According to dose-response relations, this would increase the risk of residents becoming annoyed and augment pre-existing annoyance. Supporting evidence for this has been found for aircraft noise in high or low masking noise regions (Lim, Kee Hong & Lee, 2008), with annoyance being higher in low masking noise regions, despite aircraft noise levels being the same.

Pedersen and Larsman (2008) suggest that visibility increases and economic benefit decrease risk of noise related annoyance significantly.

## **Laboratory based studies**

Although the majority of wind turbine noise annoyance is studied using methods similar to Schultz (1978), a few papers and reports have been produced from laboratory studies of subjective annoyance of wind turbine noise. For example, Persson Waye and Öhrström (2002) used samples from recordings of five wind turbines to study the variability in annoyance between different turbine models and how well 14 psychoacoustic descriptors accounted for annoyance. Persson Waye and Öhrström (2002) found significant variability between the different turbine recordings in overall annoyance despite calculated loudness, sharpness, roughness, fluctuation strength and modulation being invariant across the five recordings. Of the 14 descriptors, "lapping", "swishing", "whistling", "uneven", "low frequency" and "grinding" were most strongly related to annoyance. The wind turbine recordings were only considered for one distance.

Similarly, Legarth (2007) studied the annoyance of five turbines from distances at 6 and 12 hub heights using the Nord2000 propagation model (Plovsing & Kragh, 2006) for those who live (or wish to live) in the countryside. Using an 11-point magnitude estimation assessment, participants rated how annoying the sounds would be in their gardens. Next, they judged annoyance, loudness, swishing sound, tonality and pace for short excerpts of the turbine recordings.

Legarth (2007) found that annoyance decreased with distance for 4 of 5 of the turbines and that garden noise decreased annoyance. Subjective ratings of loudness and swishing sound agreed well with respective calculated metrics, whilst subjective tonality and pace did not. Finally, Legarth (2007) found noise sensitivity did not correlate with annoyance.

For the test environment, Persson Waye and Öhrström (2002) used a semi-reverberant, sound insulated room with two loudspeakers at the opposite end to the entrance. These loudspeakers were hidden by thin curtains. Additionally a garden chair, sun umbrella and recorded bird song were used to simulate an outdoors scenario as much as possible. During the longer exposures, participants were allowed to read a book of their choice. Such procedures are designed to help participants imagine the scenario better. Legarth (2007) used an acoustically treated listening room, a large screen displaying a wind turbine from a distance equal to the auralised distance, and presented the sounds through headphones. They asked participants to imagine sitting in their garden drinking a cup of tea or coffee and reading a newspaper or book.

Whilst some ideas for helping facilitate imagination of the scenario used by Legarth (2007) and Persson Waye and Öhrström were adopted in the current project, the 11-point magnitude estimation rating scale was not adopted following the discussion in Section 4.3. Instead, a paired comparison was used where one sound is adjusted to be equally annoying as a reference. Also, unlike Legarth (2007) and Persson, Waye and Öhrström (2002), the current project concerns low frequency acoustics and therefore extra attention had to be given to accurately reproduce the low frequency spectrum.

## **4.5 Aims and objectives**

Based on the reviewed previous research the project aims have been formulated in two parts A and B. These parts correspond to two different test setups which will be described in the following chapter.

Part A aims

- to establish relations for the audibility of low frequency (LF) tones within masking noise from wind turbines and for annoyance compared with a similar noise source, say, a single frequency test tone;
- to compare audibility and annoyance of LF tones within broadband masking noise of varying spectral content broadly representative of wind turbine noise;
- to use the results to test the ISO 1996-2 and other audibility models at low frequencies;

Part A specific objectives are:

- To establish audibility and relative annoyance thresholds for LF tones in the presence of broadband masking noise;

Part B aims:

- to develop a method to compare annoyance of the sound from small wind turbines and large wind turbines with respect to distance of the listener, with and without the masking noise typical of a residential environment.

Part B specific objectives are:

- To establish relative wind turbine levels that produce equal annoyance for two sizes of turbines taking into account the effect of masking noise on these estimates;
- To determine if turbine equal annoyance can be predicted from the LF tone equal-annoyance contours measured in part A.

## 5 Listening Test Design

In this chapter we describe all aspects of the tests designed to achieve the aims and objectives described in Section 4.5. Detailed aspects of the listening environment, participant recruitment and screening are first presented, which were common to both Part A and B tests. Choice of test stimuli and other aspects of the Part A and B test procedure are then described.

### 5.1 The listening environment

#### Acoustics

The multi-faceted nature of the test design placed stringent requirements on the listening environment. On the one hand, precise sound reproduction was required to ensure accurate reproduction of each scenario, particularly at low frequencies, but subjective aspects such as spatial cues were also of primary concern. For the majority of listening tests either one or the other of these requirements is dominant, but this test's combination of presenting tonal stimuli and measuring emotional response required that both be satisfied. A planar ambisonic reproduction system was chosen for its well-defined source-direction-rendering properties, supplemented by a low-frequency reinforcement system. The system was calibrated using a custom measurement system and modification of the stimuli signals sent to each loudspeaker. The performance of this approach was later evaluated using a 01 dB Symphony PC measurement system and DELTA's NoiseLab software. The superior spatial performance of this reproduction technique was judged to be more important than the disadvantage of restricted calibration accuracy at frequencies above 600 Hz due to the interference of the ambisonic ring. When listening to the stimuli the characteristics of the broadband sounds were not audibly changed compared to other reproduction systems. For full details of the sound reproduction setup and calibration see Appendix III.

Selection of a test space was the first issue that had to be resolved. Critical requirements included extremely low background noise, low reverberance, good low-frequency modal control, adequate space for the ambisonic loudspeaker ring, and minimal adverse non-acoustical cues. The Acoustics

Research Centre at the University of Salford has three laboratories which could have satisfied these requirements, a full-anechoic chamber, a hemi-anechoic chamber, and a listening room conforming to ITU-R BS 1116-1. After evaluation of the advantages and disadvantages of each space the listening room was chosen, primarily because its similarity to a domestic environment was thought to be the least unsettling to the participant. It is of room-within-a-room construction, the inner room floating on a bed of compressed mineral wool, and achieves a background noise level of below 6 dB(A). At 6.6m by 5.5m by 3m it is large enough to permit a 4m diameter space within the ambisonic loudspeaker ring, which avoids the feeling of claustrophobia which can sometimes occur with such setups. The room contains extensive passive absorption, resonant absorption and diffusion, giving it a low reverberation time (RT) and eliminating false spatial cues.

## Room design

As the experiment involved the assessment of annoyance, attention needed to be paid to the non-acoustical properties of the listening environment which could affect subjective response.

The room design followed ideas by Persson, Waye and Öhrström (2002): An acoustically transparent white taffeta curtain was installed, suspended from a 4m diameter octagonal frame just inside the loudspeakers, to create a neutral visual background, and to screen adverse visual stimuli. A folding garden chair was selected for its comfort and absence of low-frequency resonances. Interaction with the test software was via a 15" touch-screen located for easy operation whilst sitting in the specified listening position.

Figure 5.1a) shows a floor plan of the positions of the ambisonic loudspeaker ring, subwoofers, suspended octagon frame, curtain within the listening room, chair and screen positions.

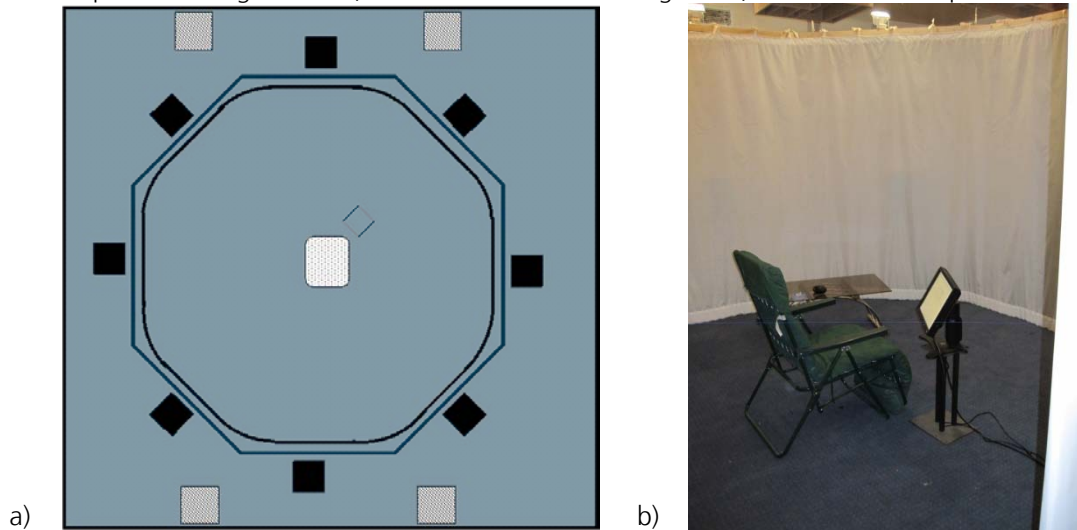


Figure 5.1a) Plan of listening position (white rectangle in centre), screen (transparent square), ambisonic loudspeaker ring (black squares), subwoofers (grey squares), suspended frame and curtain (octagon denoted position of frame, rounded octagon denotes approximate position of curtain at floor level). b) Photo of listening environment.

The listening test is designed to encourage contextual influence on annoyance ratings by asking participants to imagine themselves at home in the countryside and to rate how annoying these sounds would be if they were a constant presence. To avoid interference from the participant's attitudes they were not informed about the type of noise sources they would be exposed to. Although some participants made guesses as to what noise sources the stimuli represented, none mentioned wind turbine noise. However, most did identify the road traffic noise. It should be noted that some of the participants who did not comment on the noise sources, may still have believed them to be wind turbines.

Two typical listening scenarios, an evening time living room and a summer garden scenario were chosen, to study how the annoyance from wind turbine noise varies with the change in characteristic level and spectrum. This is necessary because building façades attenuate high frequencies more significantly, leaving the low frequencies relatively more prominent, therefore potentially leading to very different frequency content for the outdoors and indoors environment. The colour temperature of the lighting was varied so as to suggest indoor and outdoor scenarios. Indoor lighting colour temperature was tungsten at about 3000 K and daylight colour was produced by HMI at about 5400 K. The daylight illumination was achieved by one lamp that angled down over the back left side of the octagon toward the listener and four 150W halogen lamps with 201 Lee filters opposite the metal-halide lights, facing the curtain (Figure 5.2). The latter lights were designed to remove some of the shadowing effect. The indoor lights were conventional ceiling mounted lamps. All other non-acoustical cues remained constant.

The furniture and décor of the listening environment was kept minimalistic and detailed instructions about the scenario (given in full in Chapter 15) were given to aid the imagination. It is important that participants imagined being at home rather than just 'in the countryside' as appraisal of ownership and personal space affect annoyance by an intruding noise (Devine-Wright, 2009). Because every home is different, visual cues of turbines, gardens or living rooms were not used in contrast to Persson, Waye and Öhrström (2002). The instructions therefore specified:

Indoors/Evening scenario

"Imagine you are at home in the countryside in your living room, trying to relax..."

Outdoors/Day time scenario:

"Imagine you are at home in the countryside in your garden trying to relax..."

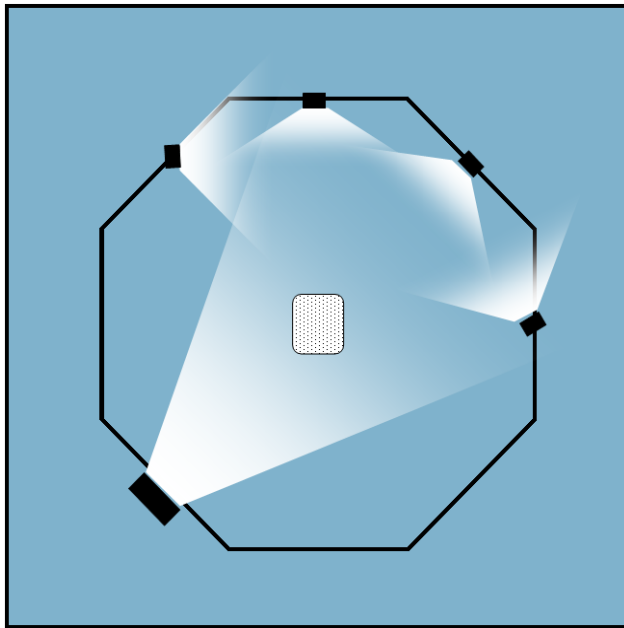


Figure 5.2 Approximate outdoor lighting arrangement

Full participant instructions can be found in Appendix VI.

## 5.2 Participant recruitment and screening

For the a first set of tests, the sampled population was initially contacted via an article placed on the main University of Salford website and the separate websites for the staff and students at the university in March 2010. The details of the article are in Appendix I. The article simply referred to the project as a study of the annoyance caused by LFN in environmental sounds. The sources such as wind turbines were intentionally omitted from the article as we wanted to test noise annoyance

without attitude towards the source affecting judgements (see Section 5.1). The article mentioned that volunteers for participation were requested and that they would be paid £10 an hour. It also mentioned that participation was subject to a screening procedure. This procedure entailed the volunteer providing their names, age, nationality, occupation, sex, and previous listening test experience; then volunteers completed several multiple-choice questions about the type of area they live in (see Appendix II). Non-leading questions were used to prevent responders to the advert from falsely claiming to belong to the population of interest.

To find out whether residents in cities have different noise sensitivity to volunteers living in the countryside or wanting to live in the country side, the screening procedure also included the Zimmer and Ellermeier short noise sensitivity measure (Appendix II). This is a 9-item self-reported questionnaire that asks the participant to either strongly agree, slightly agree, slightly disagree or strongly disagree with statements about disruptions caused by everyday noises. This was deemed a useful measure as an individual's sensitivity to noise may influence how annoying they perceive sounds to be (Weinstein, 1978; van Kamp et al., 2004; Zimmer & Ellermeier, 1998) and would therefore inform the choice of participants. Figure 11.1a) shows the difference in sensitivity. As there is a clear difference in rating between the two different groups of volunteers it was decided to restrict the study to the group of volunteers who either live in the country side or want to live in the countryside.

For a second set of tests, all participants who completed the testing from the first set in April were contacted again in July to ask if they would be willing to participate a second time. They had not been formally debriefed at this point and therefore had not been informed of the nature of the stimuli. This time participants were paid £6 an hour. 9 of the original 21 participants returned to redo the tests in August. To achieve a sample size of about 20 participants and to widen the age range, the researchers re-advertised (see Appendix II) within the University of Salford staff community. Figure 11.2 b) shows the Zimmer-Ellermeier scores for this group of participants. The average score is about 50. That is 5 points lower than the score of the first group of participants (Figure 11.1a, city dwellers) and very similar to the overall average of the full sample of volunteers for the first set of trials (Figure 11.1 b).

Another criterion for participant selection was that their hearing was not impaired. This aimed to recruit participants with 'normal' hearing for their age, rather than a sample of particularly sensitive or impaired hearing participants and was necessary because some stimuli were close to the hearing threshold. If not heard, the results for these stimuli would have been meaningless. Additionally to choosing participants on their assertion that they were of normal hearing, the audibility threshold for each participant was also tested as described in Section 5.5. This allowed excluding participants with significant hearing loss.

20 participants (10 male, 10 female) were recruited using the advertisement in Appendix I, Section 10.2 . The mean age was 40.3 years, ranging from 19 to 65 years with a standard deviation of 14.6. All data from one male participant was omitted due to hearing loss. Participant statistics and screening results can be found in Appendix II.

## **Participant protection**

The data set from the screening process contained potentially sensitive personal information. Therefore it was stored in password protected spreadsheets on a secure server that was only accessible by project staff. No other copies were kept. Outside these files participant information was made anonymous by the use of ID numbers; therefore the data could only be traced back to the participant via the protected spreadsheet. Informed consent forms as specified in Appendix I were signed by each participant in accordance to standard University procedures. Specific noise sources were not mentioned in the project explanation in order to avoid possible bias.

### 5.3 Listening test software

The test software was implemented in MATLAB<sup>†</sup>, which supports high-precision data manipulation and analysis, GUIs, and multi-channel synchronous audio I/O via PortAudio<sup>‡</sup> and the Playrec<sup>§</sup> wrapper. The latter accesses the soundcard using Steinberg's ASIO drivers, and the complete system has the desirable property that it contains no software audio mixers or similar, so the data that is sent from MATLAB arrives at the DA converter unmodified – an important concern given the system was to be calibrated.

The functional organisation of the software is depicted in Fig. 5.3. The details of the part-specific stimuli-defining routines are discussed in Appendix IV, while the common room equalisation routines are discussed in Appendix III. Each function saved a unique .mat data file upon completion, which was subsequently loaded by other routines requiring that data. The structure was devised such that the stimuli for each part could be specified or modified independently of each other or the room equalisation, culminating in a single “run-file” being compiled for each part. This contained all the data necessary to execute the test. So just ensuring the stimuli data files and the run script remained unchanged was adequate to guarantee all participants heard the same stimuli.

The only data not contained in the run-file was the participant information, since it was quite possible that more participants might be recruited after testing had begun. In the interests of anonymity the participant data was never loaded into the runtime workspace. Instead a list of participant names and IDs was saved in a data file, and this was then loaded by a function which displayed a participant name selection dialog box and then returned the associated participant ID. Selecting participants by name was thought to give lower chance of human error than simply entering participant ID manually.

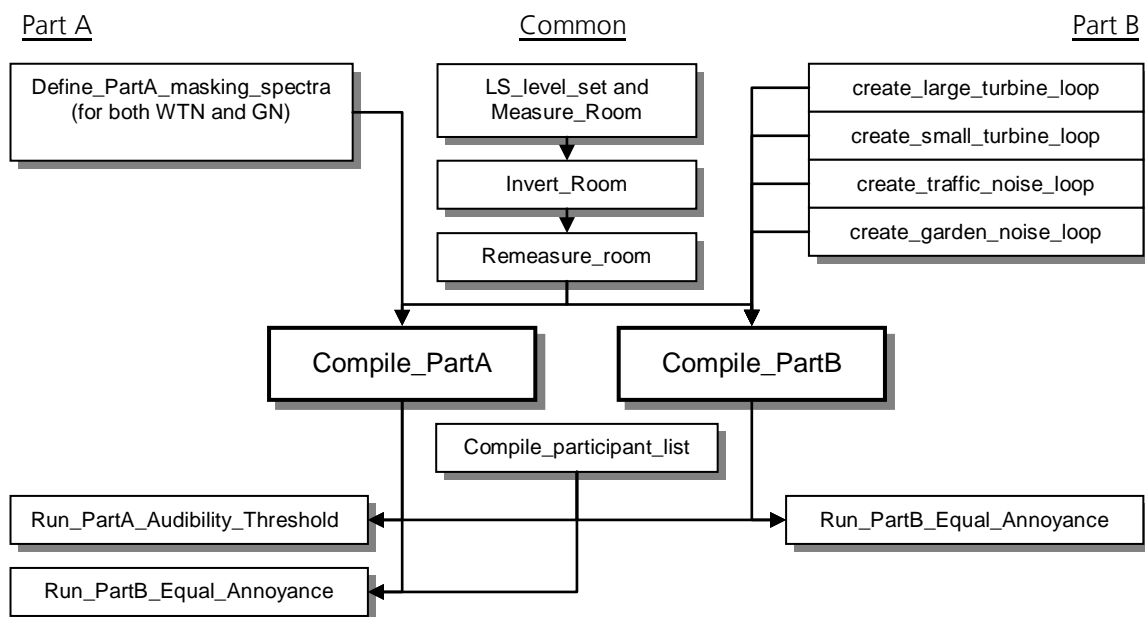


Figure 5.3 Top-level function data dependencies

<sup>†</sup> [www.mathworks.com](http://www.mathworks.com), release 14

<sup>‡</sup> [www.portaudio.com](http://www.portaudio.com), version 19

<sup>§</sup> [www.playrec.co.uk](http://www.playrec.co.uk), version 2.1.1

## 5.4 Test procedure Part A

### Aims and objectives

The aims of Part A, recalled from section 4.5 were:

- to establish relations for the audibility of low frequency (LF) tones within masking noise from wind turbines and for annoyance compared with a similar noise source, say, a single frequency test tone;
- to compare audibility and annoyance of LF tones within broadband masking noise of varying spectral content broadly representative of wind turbine noise;
- to use the results to test the ISO 1996-2 and other audibility models at low frequencies;

The objectives were:

- To establish audibility and relative annoyance thresholds for LF tones in the presence of broadband masking noise;

### Description

Participants were asked first to perform the audibility threshold tests for the tones to find out the individual's threshold of hearing for pure tones generated in the range  $30 \leq f \leq 150$  Hz in approximately logarithmic steps and one of 400 Hz. This is because individual thresholds are thought to vary widely especially in the low frequency range between 20 – 200 Hz. Two methods, Békésy tracking and Two-Alternative-Forced Choice (2AFC), were discussed and tested. 2AFC was used in the main tests because of concerns over masking noise exciting room modes, potentially creating tonal artefacts that were difficult to distinguish from the target stimuli. With Békésy tracking, occasionally participants would believe they could still hear the stimuli when it was well below their threshold. In these cases the mistakenly identified 'tone' was probably an artefact of the listening environment. The room calibration routine was refined to address this issue for the second phase of tests and no further evidence of tonal artefacts in the masking sound were observed.

For the 2AFC procedure two buttons on the participant's screen (see Figure 15.2) briefly changed colour, to alert the participant that the stimulus was playing. The left button lit when the first stimulus was played and the right button lit with the second stimulus. These were the visual cues of the two 'intervals'; one 'interval' included just the masking noise (or no stimulus in the no masking noise conditions), whilst the other included both the masking noise and the test tone mentioned above. The sounds played for the full 550 ms of the interval. The participants' task was to press the button which lit when the target stimulus was present. The '2 down, 1 up' procedure was used (Figure 15.3), such that the participant had to correctly identify the target stimulus in two consecutive iterations before the level was reduced. If the participant failed to identify the interval with the target stimulus once, the stimulus became louder in the next iteration. This procedure was performed for each stimulus until it reversed the direction of level changes 8 times. For the first two reversals, the increments and decrements are by 8 dB, then 4 dB for the next two, then 2 dB for the final four reversals. The threshold is calculated for that stimulus by the mean of the levels at the final four reversals.

Instructions were given on screen before the practice trials (see Appendix VI, Figures 15.6-8). For at least one practice trial the experimenter explained the procedure and demonstrated the task. The context instructions were displayed again before the experimental trials started.

The masking noise was created for three scenarios:

- no masking noise
- typical indoor wind turbine sound without garden noise
- typical outdoor turbine sound with and without garden noise



Apart from the visual cues the scenarios were created by adding the extra sound components described in Section 5.3. Scenarios were presented in counterbalanced order to avoid fatigue bias. The tone and masking sound orders were randomised.

The same test tones combined with the masking noises were then played at higher levels and compared to a reference tone at 180 Hz at a fixed tone level  $L_{pt}$  above masking  $L_{pn}$  (ISO 1996-2, 2007). The participant adjusted the test tone level until they deemed it as annoying as the reference stimulus (Figure 15.5). The frequency of the reference stimulus was chosen to be at the upper end of the low frequency tone range. The broadband wind turbine noise was played at three different levels in both the indoor and outdoor scenarios. This procedure resulted in equal annoyance contours for each combination of wind turbine level, scenario, and reference tone prominence.

For the choice of reference frequency a procedure following Robinson & Dadson (1956) had been considered using reference tones at adaptive frequencies. However, due to the large number of frequencies one fixed reference was more appropriate in this case to avoid a fixed presentation order of stimuli and potential fatigue bias.

It has been suggested that increasingly different sound characteristics between reference and test sounds make comparison harder, thus increasing the standard deviation of participant responses (Niedzweicki & Ribner, 1979). The reference tone frequency in the current study was therefore not chosen to be the highest frequency in the test, because it was considered that 400 Hz may be too different in pitch from the low frequency tone for participants to confidently compare. The 400 Hz tone was included as a test tone to reflect the fact that wind turbine spectra can have tonal components in the 400-800 Hz frequency range.

Wind turbine noise levels were chosen to be 39 dB(A), 44 dB(A) and 49 dB(A). These levels correspond to Danish outdoor noise limits where 44 dB(A) represents the noise limit for a residence in open country at 8m/s wind speed referenced to 10 m height, 39 dB(A) represents the noise limit for noise sensitive areas at 8m/s and the 49 dB(A) represents a level exceeding all limits. The reference tone levels were chosen such that they had prominences of 5 and 10 dB above  $L_{pn}$  using ISO1996-2 (2007). Lower broadband noise levels were not considered as they were not audible.

Participants were first given written instructions regarding rating noise annoyance (see Appendix VI, Figure 15.6) before being given on-screen instructions to the equal annoyance task (see Appendix VI, Figures 15.9-10) before the practice trials. Again, the experimenter demonstrated the task for the first practice and the context instructions were displayed again on-screen before the experimental trials began.

To be able to compare the equal annoyance contours arising from the different masking noise conditions, a further equal annoyance procedure was devised where masking noise containing the reference tone was compared to the pure reference tone without masking. First the tone in masking noise was played then the participant matched the pure tone level such that it was equally annoying to the tone in masking. This masking comparison allowed judgments on the relative annoyance of different scenarios.

Equal annoyance tests, masking comparisons and the wind turbine recording comparison (Part B, Section 5.6) were presented to participants in counterbalanced order.

## Stimuli Part A

In order to be realistic, the test stimuli had to include both a test sound, representing a noise source at some distance, and local masking noise. In the outdoor scenario the wind turbine sound was presented as a plane wave arriving from in front of the participant and the garden noise from all around. In the indoor scenario sources were presented from behind the participant so as to simulate sound coming from the facade. Garden noise was not added in this scenario because the

spectra with and without garden noise were virtually identical in the frequency range at and below 400 Hz. These considerations applied to both Part A and Part B of the test.

In Part A, the test sound comprised a tone and a broadband masking sound. The tone was played at one of six pre-determined frequencies. The masking sound was either a sound broadly representative of wind turbine sound at one of three pre-determined levels or this broadband wind turbine sound combined with surrounding garden noise. The wind turbine masking spectra were derived from third-octave sound power measurements of 45 wind turbines, each attenuated to simulate propagation to the minimum distance permitted to a dwelling by Danish regulations (4 total heights), scaled to the target A-weighted SPL, and finally averaged. Propagation attenuation was predicted in third octave bands by the Nord2000 model for a receiver 1.5 m above flat grassy terrain with a wind speed of 8m/s in the downwind direction.

The local garden noise was created to match a noise spectrum of 8 m/s wind through deciduous foliage. A similar process was followed for the stimuli for the indoor scenario, except that a third octave façade attenuation spectrum was applied to all spectra, and the propagation attenuation for the wind turbine was averaged over three heights (0.5 m, 1.5 m, 2.5 m) to account for excitation over the full height of the façade. The façade attenuation was chosen from the values presented in the paper from Hoffmeyer and Jacobsen (2010). Stimuli had to be looped. For Part A, a stimulus loop length of  $2^{19}$  samples (approximately 10 seconds at 48 kHz) was found to be the best compromise between obvious repetition and computational resource usage. Masking spectra are shown in Figure 5.4 while examples of third octave spectra of the stimulus recordings and more details on stimuli creation can be found in Appendix IV.

It should be noted that the stimuli were steady sounds. Temporal variation or “swishing” was explicitly excluded from the stimuli both, because its variability is still not well understood and the focus of the study is on the effect of low frequency tones.

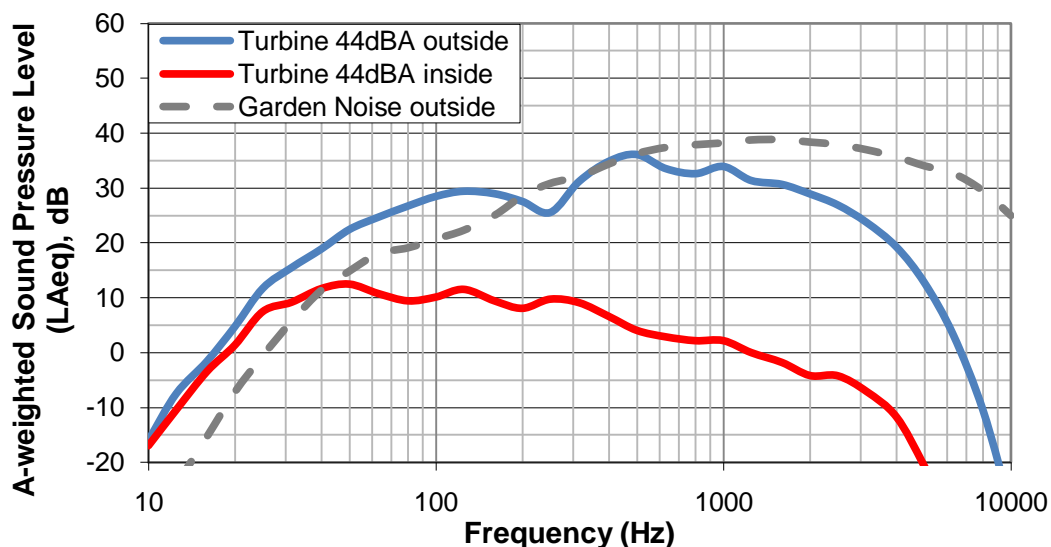


Figure 5.4 1/3 octave masking Spectra: The indoor spectrum with garden noise was virtually identical to the one without garden noise below 400 Hz and was therefore not included separately in the listening test.

## Stimuli overview

Table 5.1 and Table 5.2 summarise the stimuli used for the final audibility threshold and equal annoyance tests, respectively. Location stimuli were presented in manual counterbalanced order. Masking noise was presented in random order groups with randomised tone order within the group.

Test/ scenario	Masking	Test sound	No of Trials
Practice (always first)	Random selection	Prac Tones (2)	1-3
Audibility threshold	None	32, 44, 72, 115, 180 and 400 Hz	6
Indoor	FA[L <sub>Aeq,WT</sub> (3)]	32, 44, 72, 115, 180 and 400 Hz	18
Outdoor	L <sub>Aeq,WT</sub> (3)	32, 44, 72, 115, 180 and 400 Hz	18
Outdoor	L <sub>Aeq,WT</sub> (3)+GN	32, 44, 72, 115, 180 and 400 Hz	18
<b>Total stimuli</b>			<b>60</b>

Table 5.1 Part A1: audibility threshold: Tones at six frequencies, three different levels L<sub>Aeq,WT</sub> of 39 dB(A), 44 dB(A) and 49 dB(A), FA = Façade Attenuation

Test/ scenario	Masking	Test sound	L <sub>R</sub>	No of Trials
Manual counterbalance	Group, randomised	Randomise	Randomly either 5 or 10 dB	
Practice (always first)	Random selection	Random selection	5 or 10 dB	1-3
Audibility threshold	None	32, 44, 72, 115 and 400 Hz	5 and 10 dB	10
Indoor	FA[L <sub>Aeq,WT</sub> (3)]	32, 44, 72, 115 and 400 Hz	5 and 10 dB	30
Outdoor	L <sub>Aeq,WT</sub> (3)	32, 44, 72, 115 and 400 Hz	5 and 10 dB	30
Outdoor	L <sub>Aeq,WT</sub> (3)+GN	32, 44, 72, 115 and 400 Hz	5 and 10 dB	30
<b>Total stimuli</b>				<b>100</b>

Table 5.2 Part A2 equal annoyance: Reference Tone at 180 Hz at levels 5 dB and 10 dB above L<sub>pn</sub>. Tones at five frequencies to be adjusted in level to equal annoyance to reference tone, three different levels L<sub>Aeq,WT</sub> of 39 dB(A), 44 dB(A) and 49 dB(A), FA = Façade Attenuation

Table 5.3 gives an overview of the comparison of a pure reference tone to the reference tone in different scenarios. The first stimulus, the participant heard was the reference tone in masking noise. They then adjusted the level of a reference pure tone without added masking noise to match the first stimulus in annoyance. This comparison was designed to link the test results in Table 5.2 to give a relative annoyance relation between all stimuli. Location was manually counterbalanced. Masking was presented in random order groups with randomised L<sub>R</sub>.

Location	Masking	$L_R$	No of Trials
Practice (always first)	Random selection	Randomly either 5 or 10 dB	1-3
Indoor	$FA[L_{Aeq,WT}(3)]$	5 and 10 dB	6
Outdoor	$L_{Aeq,WT}(3)$	5 and 10 dB	6
Outdoor	$L_{Aeq,WT}(3)+GN$	5 and 10 dB	6
<b>Total stimuli</b>			<b>18</b>

Table 5.3 Part A3 equal annoyance of reference tone in added masking spectra in comparison to the pure reference tone.

## 5.5 Test procedure Part B

### Aims and objectives

The aims of Part B, recalled from section 4.5 were:

- to develop a method to compare annoyance of the sound from small wind turbines and large wind turbines with respect to distance of the listener, with and without the masking noise typical of a residential environment.

The objectives were:

- To establish relative wind turbine levels that produce equal annoyance for two sizes of turbines taking into account the effect of masking noise on these estimates;
- To determine if turbine equal annoyance can be predicted from the LF tone equal-annoyance contours measured in part A.

### Description

The participant played a wind turbine stimulus at one of three levels  $L_{p,A} = 39, 44$  and  $49$  dB(A) and compared this with a reference traffic signal using the equal annoyance control from Part A (Figure 15.10). The participant then adjusted the A-weighted level of the traffic noise signal until they perceived it as equally annoying to the wind turbine signal. The participant could switch between hearing the reference stimulus and the adjustable test stimulus as many times as they liked. The scenarios of the wind turbine stimuli were small wind turbine/large wind turbine, garden noise/no garden noise presented in group randomised order and indoors/outdoors presented in manually counterbalanced order.

### Stimuli Part B

In Part B, field recordings of real wind turbines were modified to simulate the same scenarios as Part A. As garden noise was directly derived from a measurement, that recording could be used directly with only a slight volume adjustment to match the A-weighted SPL of the selected cut to the long-term average. Modification of the wind turbine recordings involved reversing the propagation attenuation from the hub to the plate microphone and applying propagation attenuation to a distance of 4 turbine heights, normalising that level to 44 dB(A). Normalisation of wind turbine levels was used to focus listener attention on spectral characteristics, to avoid unrealistic small distances and therefore near-field directional effects as well as to reduce the effects of differences in wind speed and source strength. This is reasonable because regulations ensure a certain maximum noise level that cannot be exceeded. The result of normalisation is a less direct comparison, as relative required distances between turbines cannot be calculated easily.

After the normalisation step additional propagation attenuation was applied such that the other target A-weighted SPLs were achieved. The propagation attenuation was calculated using the same model and parameters as Part A, and once again façade attenuation was applied to all indoor stimuli. More details on Part B stimuli can be found in Appendix IV, Section 13.2.

Part B also required Traffic Noise (TN) to be presented as a reference sound. This was recorded specifically for the project so distance and microphone height were chosen to match that described above and no further propagation attenuation was applied. TN recordings had to be as free from

wind noise as possible and preferably continuous in nature. The sample of TN used was therefore recorded from 100m approximately South-South-West of the M6 motorway 12:53 GMT. Weather conditions were overcast and calm, with occasional drizzle (see Appendix XI). Road surfaces started out dry and became slightly damp during the periods of drizzle.

Selection of sound samples for Part B was a critical task. All recordings contained variations due to the natural random fluctuations in wind speed, plus sound from additional sources (e.g. birdsong). The latter must not be present in the selected cuts, and the former is beyond the scope of this study and needed to be minimised.

#### **Adaptive Signal: Traffic recordings**

The reference stimulus was selected from the traffic recordings that are documented in Appendix IV. It was a 10.92 sec recording taken from 2 min into the 12:53 recording interval and chosen for its steady character for the same reasons given in Section 5.4 and for the lack of other noise sources. This recording was then looped to produce the final stimulus using a 1.32 second equal-power cross-fade profile to ensure that the looping could not be identified while listening.

#### **Wind turbine signals**

The small turbine stimulus was derived from a free-field recording. The turbine was rated at 600 kW with a hub height of 35 m. The 10 minute wind speed average at recording time was  $5 \text{ ms}^{-1}$ . The large 2.3 MW turbine with an 80m hub height was recorded using the plate method at a distance of 128 m in reference orientation to wind direction at a wind speed of  $8.4 \text{ ms}^{-1}$ . The stimuli were looped and attenuation was applied to the spectra (shown in Figure 5.5) using the Nord 2000 model. Details on stimuli production are given in Appendix IV.

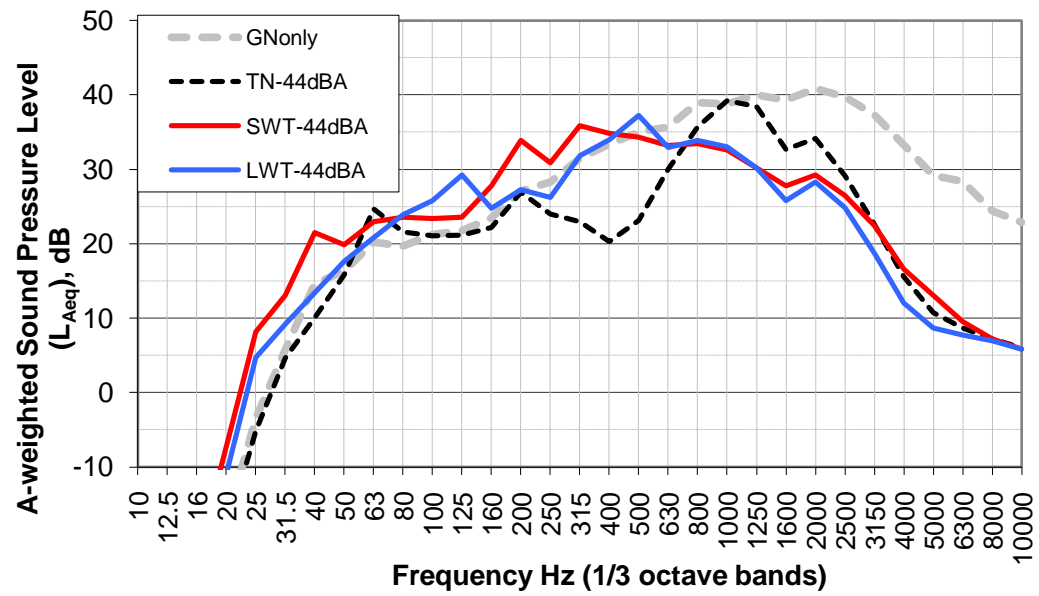
Looping for the wind turbine sounds was determined by their rhythmic blade swish sounds caused by the blades passing. The loop length had to be fixed to a multiple of this characteristic interval, otherwise the turbine would appear to stutter. The loop length was found by inspecting the peaks in the auto-correlation of the signal.

The small wind turbine swishing sounds from the rotor blades were heard as amplitude and frequency modulation of mid frequency noise bands. The modulation frequency was approximately 1.4 Hz. The large wind turbine sound sounded more constant with weak swishing sounds probably due to amplitude and frequency modulation of mid frequency noise bands. The modulation frequency was approximately 0.8 Hz. A weak tone at 530 Hz was detected by FFT analysis just below the average masking threshold and was just audible. In both turbine sounds low frequency rumbling was heard. This was weaker for the large wind turbine than for the small wind turbine which is also evident from the 1/3 octave band analysis where the large wind turbine SPL at 40 Hz is 7.6 dB lower than the small wind turbine SPL. Weak irregular hissing sounds which could be due to turbulence at the blade tips were also noticed for both turbines.

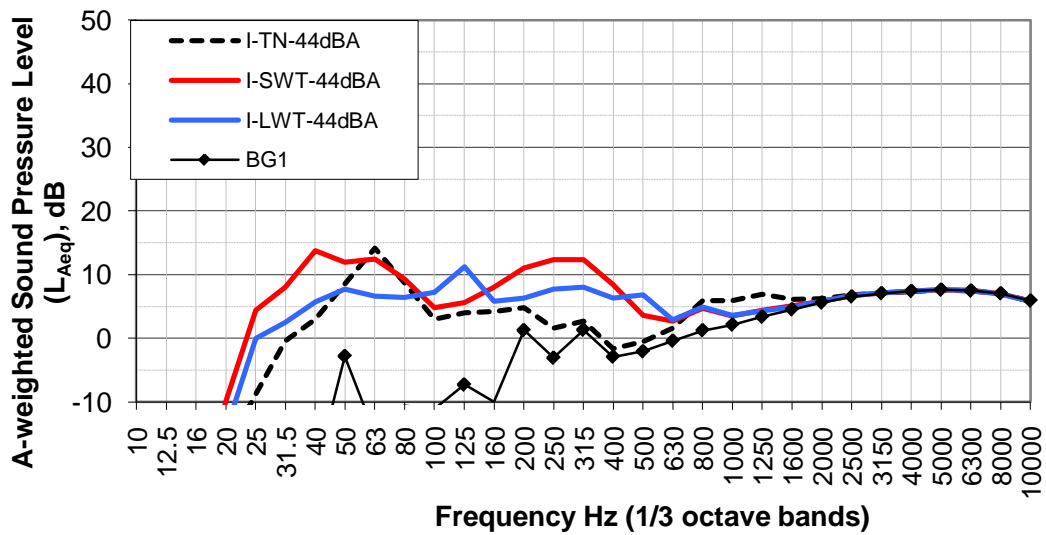
### **Stimuli overview**

<b>Location/Scenario</b>	<b>Masking</b>	<b>Adaptive stimulus</b>	<b>Reference sound</b>	<b>No. of trials</b>
Practice	Random	TN	1-3	2
Indoor	None	TN	SWT (3)	3
Indoor	None	TN	LWT (3)	3
Outdoor	None	TN	SWT(3)	3
Outdoor	None	TN	LWT (3)	3
Outdoor	GN	TN	SWT (3)	3
Outdoor	GN	TN	LWT (3)	3
<b>Total stimuli</b>				<b>18</b>

Table 5.4 Part B equal annoyance level:  $L_{Aeq,WT}$  of small wind turbine and large wind turbine to be adjusted to equal annoyance to traffic reference stimulus.



a)



b)

Figure 5.5 Part B Recorded spectra a) outside b) inside, SPL inside lower due to façade attenuation

## 6 Part A results: Audibility and annoyance from low frequency tones

The present chapter contains the results of Part A of the study. It focuses on the audibility and annoyance of tones in masking noise and in quiet. Both tones and masking noise are idealised sounds with features broadly representative of wind turbine sounds. To distinguish clearly between the tones and the simulated masking noise caused by a wind turbine, the term “wind turbine noise” in this chapter specifically refers to broadband noise that does not contain tones.

Section 6.1 gives details on the trends of the measured audibility thresholds and compares the results to published thresholds. Section 6.2 looks in more detail into the low frequency applicability of the masking threshold calculations according to ISO 1996-2. Calculations of the audibility of broadband noise near the hearing threshold according to Pedersen (2008) are also presented in this section and compared to the observations. Equal annoyance trends are specified in Section 6.3 independently for each scenario while Section 6.4 finds relations between the scenarios for equal annoyance.

### 6.1 Audibility threshold dependence on frequency and level

Audibility thresholds ( $L_T$ ) of six tones in different types of masking noise were established according to the procedure in Section 5.5. The audibility threshold contours are shown in blue in Figure 6.1 and Figure 6.2, for the outdoor and indoor scenarios respectively. These figures also contain the masking thresholds calculated using ISO 1996-2. Figure 6.3 shows the data for the background noise level in the room without masking. All three figures also contain recently published hearing threshold data (ISO 389-7). Generally, audibility threshold decreases with increasing frequency regardless of the masking noise spectrum or level for the types of masking noise used in this experiment. The same trend can be observed for the hearing threshold.

In the outdoor scenario the audibility thresholds exceed the hearing threshold by at least 5 dB for all frequencies. The slope of the hearing threshold is steeper than the audibility threshold. This is due to the nature of the masking spectra where SPL/bandwidth decreases significantly below 30 Hz. The masking threshold calculated by the ISO 1996-2 standard is close to the measured audibility thresholds at all frequencies. The largest differences of up to 5 dB are observed for the frequencies 180 and 400 Hz whereas the low frequencies are within 2 dB. The measured audibility threshold is therefore clearly governed by the masking and will be referred to as the measured masking threshold. This finding is supported by Figure 6.4 a) that shows the outdoor masking threshold to be in excess of 10 dB above the measured audibility threshold in quiet.

In the indoor scenario the situation is less clear. The measured audibility threshold values are all above or equal to the published hearing threshold in Figure 6.2 and the calculated masking threshold at low frequencies does not match the audibility threshold as well as in the outdoor scenario. The comparison in Figure 6.4 makes that situation even clearer for low masking levels: When the A-weighted wind turbine level  $L_{Aeq,WT}$  equals 39 dB(A), the confidence intervals in Figure 6.4a) all suggest that the measured audibility thresholds in masking could be identical to the audibility thresholds in quiet. The higher the masking level  $L_{Aeq,WT}$  the more reliably is the measured audibility threshold in masking noise above the measured audibility threshold in quiet.

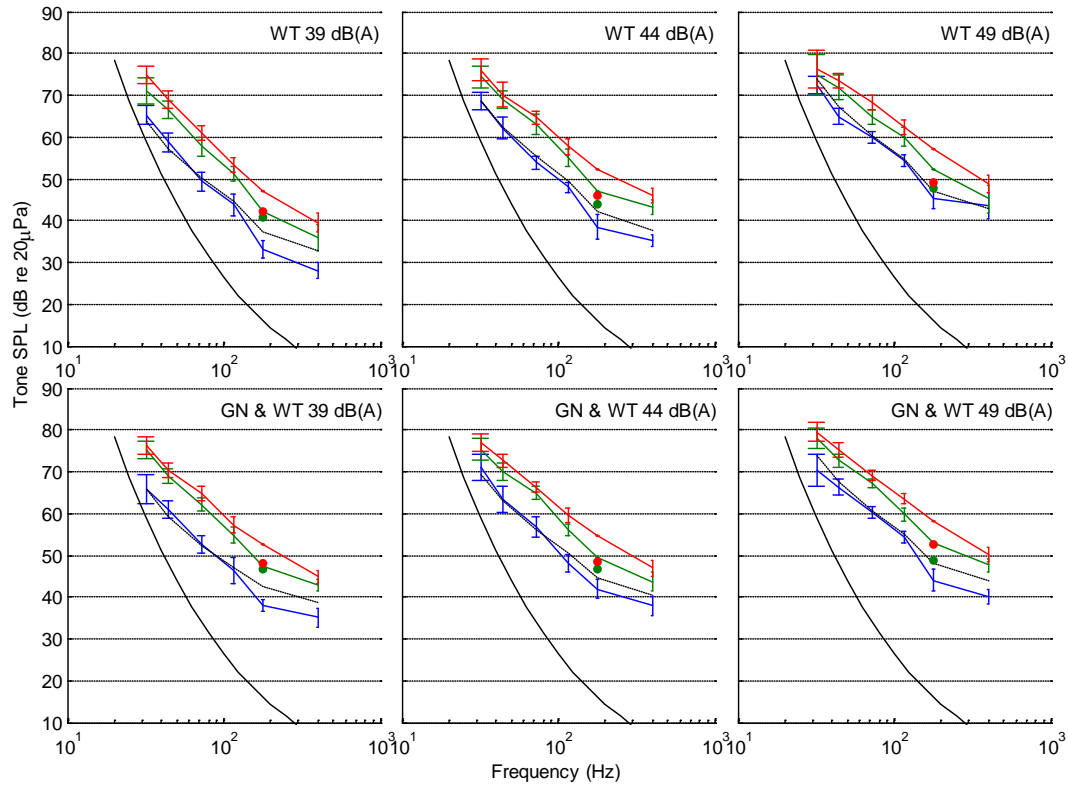


Figure 6.1: Outdoor scenario – Masking thresholds (blue), equal annoyance to 180 Hz tone at 5 dB audibility (green) and 10 dB audibility (red) of low frequency tones within masking noises as labelled. Additional green and red dots at 180 Hz denote the results from the masking comparisons discussed in Section 6.4 (for 5 dB and 10 dB above  $L_{pn}$  respectively). Black solid line: Hearing threshold according to ISO 389-7). Dashed black line: Masking threshold according to ISO 1996-2. Error bars denote 95% confidence intervals.

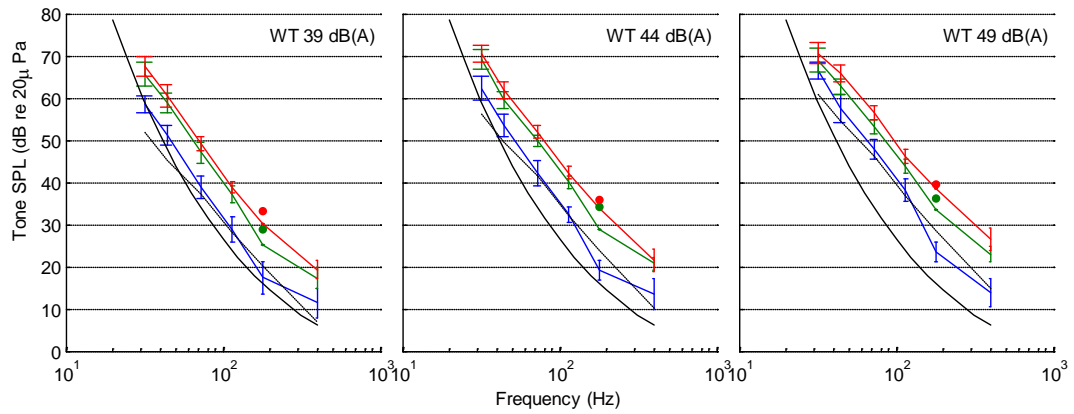


Figure 6.2 Indoor Scenario – Masking thresholds (blue), equal annoyance to 180 Hz tone at 5 dB audibility (green) and 10 dB audibility (red) of low frequency tones within façade attenuated masking noises as labelled. Black solid line: Hearing threshold according to ISO 389-7. Dashed black line: Masking threshold according to ISO 1996-2. Error bars denote 95% confidence intervals.



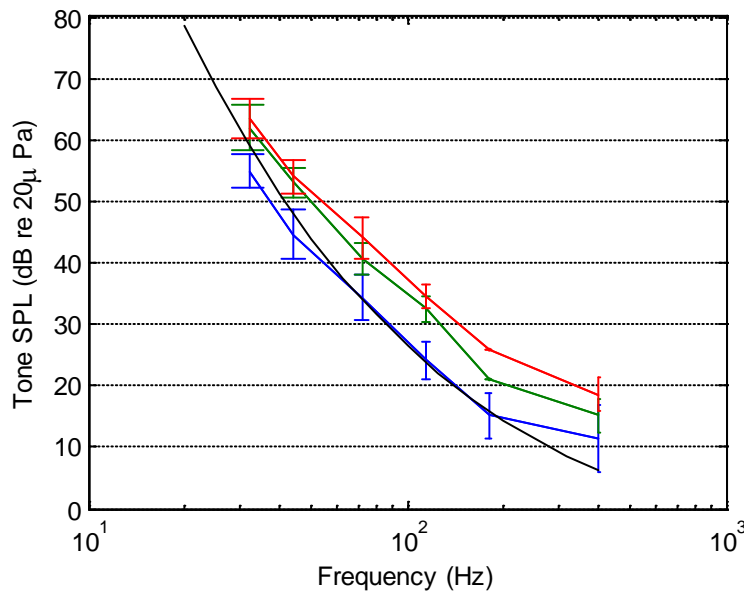


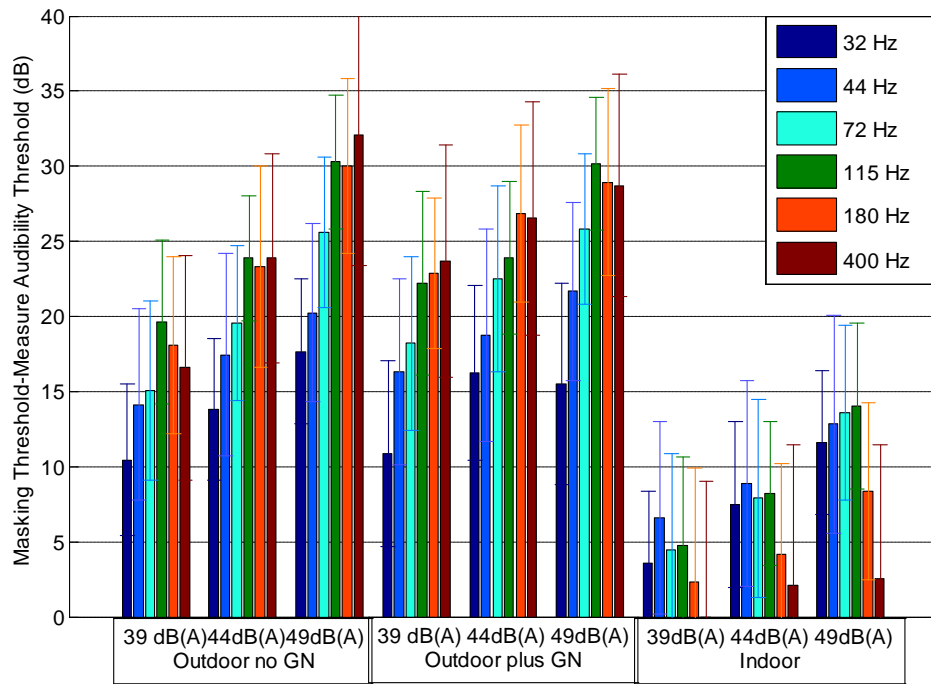
Figure 6.3 Room background without masking noise - Audibility threshold (blue), equal annoyance to 180 Hz tone at 5 dB audibility (green) and 10 dB audibility (red) of low frequency tones. Black solid line: Hearing threshold according to ISO 389-7 (2005). Error bars denote 95% confidence intervals

While the audibility threshold in quiet agrees very well with the published hearing threshold ISO 389-7 (2005) between 72 Hz and 180 Hz, the hearing threshold exceeds the audibility threshold at lower frequencies and is lower at 400 Hz. The higher audibility threshold at 400 Hz might be an effect of the room. Figure 6.6 shows that the loudspeakers, when not playing, contribute some sound power in that particular band. This could not be heard as the figure indicates. But being so close to the masking threshold it might have masked the 400 Hz tone to some degree and might have caused the audibility of the 400 Hz tone to be about 5 dB higher than it would have been in the room with the loudspeakers turned off. This needs to be kept in mind when interpreting the audibility thresholds in Figure 6.2 and Figure 6.3. The audibility thresholds toward lower frequencies deviate from the published hearing threshold by about 5 dB. It is interesting to note that this deviation is larger than that reported by Fastl & Zwicker (2007, Fig. 2.4). The presented results have been found to be statistically significant. The method and results as well as non significant effects can be found in Appendix VII.

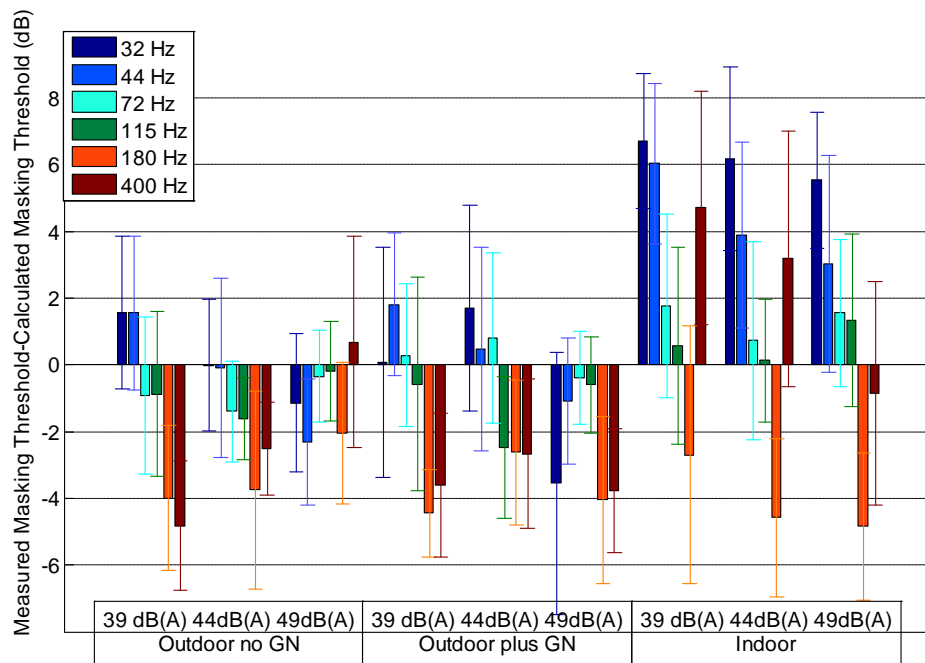
In the indoor scenario where both the lowest and highest frequencies were observed to be very close to the audibility threshold in quiet the deviation between audibility thresholds and masking threshold can be in excess of 6 dB. In these cases it would be more appropriate to evaluate the hearing threshold than the masking threshold.

## 6.2 Validity of ISO1996-2 masking thresholds and broadband audibility threshold

In Section 6.1 it was discussed how Figure 6.4a) shows which tonal audibility was determined by the masking threshold and which by the hearing threshold. Figure 6.4b) also shows the obvious difference between the indoor and the outdoor scenario due to the issue of audibility by subtracting the calculated masking threshold from the measured one. However, it is surprising that in the outdoor scenario at frequencies between 32 and 115 Hz the audibility threshold values agree generally better (not significantly different from 0 when taking the error bars into account) with the masking thresholds than at the two higher frequencies, 180 and 400 Hz. This could be due to the fact that the subwoofer set-up is similar to standard LF test chamber equipment for free-field low frequency tone audibility threshold testing, whereas the ambisonic system is not. Therefore it is expected that low frequency results agree better with published values than the high frequency results.



a)



b)

Figure 6.4 a) Difference between tonal audibility levels in masking noise and tonal audibility levels in background as seen in Figure 6.3. All outdoor scenarios and some frequencies in the indoor scenario exceed tonal audibility threshold in quiet by more than 10 dB. This indicates that AT in masking noise are governed by the masking threshold. The indoor scenario has a number of frequencies within 5 dB of AT. b) Difference between tonal audibility levels in masking noise and masking threshold in dB tone level according to ISO 1996-2 Error bars denote 95% confidence intervals. Note that most of the error bars in b) include 0.

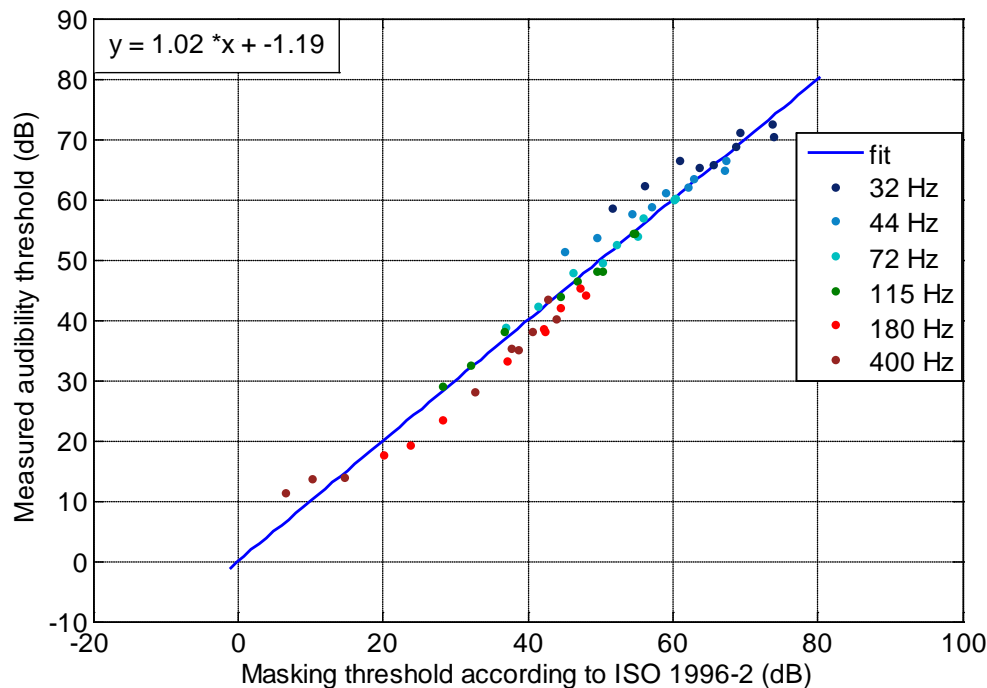


Figure 6.5 Correlation of measured audibility thresholds in masking noise with masking threshold calculated in accordance with ISO1996-2. Linear regression shows a slope of 1.  $R^2 = 0.96$ . Note that there is a high correlation between level, frequency and residuals.

Another way of telling how well measured masking thresholds correlate with calculated masking thresholds is to look at a linear regression as shown in Figure 6.5 which contains data from both the indoor and the outdoor scenario. Although single values deviate by more than 6 dB from the regression line the slope has a value of 1.02 which indicates excellent agreement between measured and calculated thresholds. If the values of the indoors scenario were left out of the regression because of concerns over the applicability of the masking threshold then the correlation equation would change to  $y = 1.09 * x - 6.38$  with an  $R^2$  value of 0.98. In conclusion, the calculations defined by ISO 1996-2 are in good agreement with the low frequency measurements.

## Case studies on audibility of broadband noise near the hearing threshold

So far the focus of the results has been on the audibility of tones in masking. Another question is how audible are stimuli with broadband spectra. In the literature wind turbine spectra are often found measured at different frequency resolutions which when compared to the hearing threshold then appear to be more or less audible when in fact the total sound pressure level in the critical bands is the same for all of them. Pedersen (2008) therefore proposed to attenuate the measured spectra by the inverse hearing threshold and subsequently evaluate the total sound pressure level within each critical band to be able to compare spectra of different frequency resolution. Although the results in Figure 6.3 are for tones it demonstrates that the method should work because it is based on the hearing threshold which agrees well with the audibility threshold results. No systematic threshold measurements were conducted for testing this method but the observations described below have been made.

Figure 6.6 shows an example of broadband spectra. When the listening room was characterised for the testing, the background noise in the room was measured in three different operational states: with all electronic systems off; with the loudspeakers on; and with the ventilation system on. Nothing could be heard by the researchers when everything was switched off and when the loudspeakers were on. But when the ventilation system was on this could be clearly heard. The

results in Figure 6.6 are calculated according to the Pedersen method. It can be seen that when all systems were off the level did not exceed -5 dB SPL per critical band. 0 dB denotes the hearing threshold. When the speakers were on all critical band values were well below the hearing threshold except the band between 300 and 400 Hz that had a value very close to the hearing threshold. And still the loudspeakers could not be heard. The critical band levels of the inverse hearing threshold weighted spectra from the ventilation system exceeded the hearing threshold only by 2 dB in one critical band and by less than 1 dB in a second band and yet the vent could be heard clearly. The vent was therefore always switched off during the listening tests. These results show that the Pedersen method worked well in this particular case.

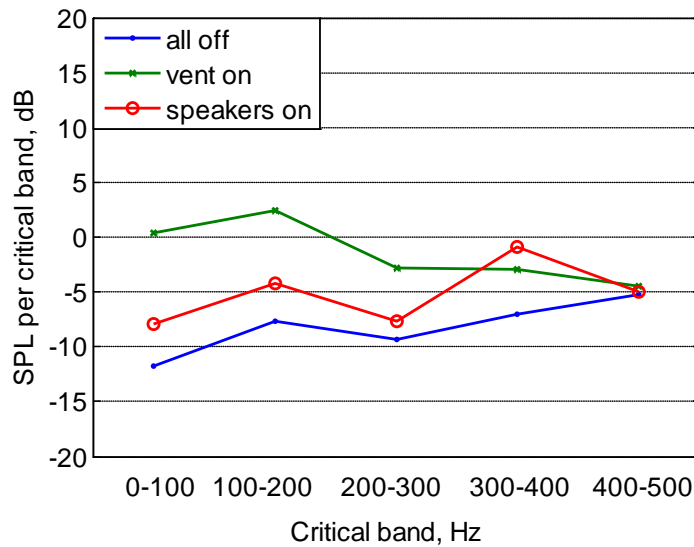


Figure 6.6 Audibility of background noise in listening room calculated according to Pedersen (2008)

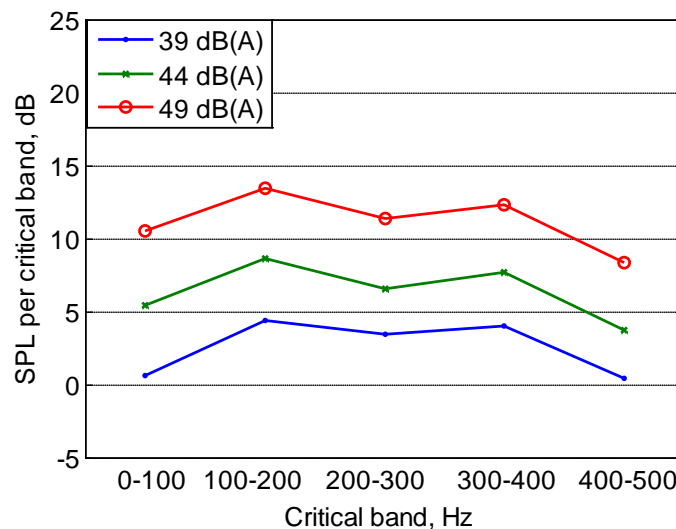


Figure 6.7 Audibility of masking noise in indoor scenario calculated according to Pedersen (2008)

Another example is the audibility of masking noise in the indoor scenario. Participants reported that the masking noise was difficult to hear for the low masking level of 39 dB(A). Figure 6.7 makes clear why. All critical band levels are within 5 dB of the hearing threshold for this level whereas the other level exceeded the hearing threshold by more than 5 dB. This agreement is further strengthened by the audibility threshold in Figure 6.2: The tonal audibility at 32 Hz and 400 Hz is identical to the hearing threshold in quiet whereas the values for the other frequencies are slightly higher which suggest the audibility was restricted by masking rather than the hearing threshold. In conclusion two examples have been shown that demonstrate that the Pedersen method for determining the audibility of broadband noise near the hearing threshold works well in the frequency bands between 0 and 500 Hz.

### 6.3 Equal Annoyance

After the analysis of audibility thresholds in Sections 6.1 and 6.2 the focus of the following is on equal annoyance. The green and red lines in Figures 6.1 - 6.3 are equal annoyance contours as a function of tone frequency. They were established by measuring the equal annoyance levels ( $L_{EA}$ ) for the five test tones in masking noise each matched in annoyance to a 180 Hz reference tone at either 5 or 10 dB ( $L_R$ ) above  $L_{pn}$  as described in Section 5.4. It is worth remembering that the equal annoyance contours were derived by focussing the participant's attention specifically on tone levels and not on the masking noise. The additional single red and green markers in these figures were however established by explicitly asking the participant to rate the whole reference sound (tone + masking) by adjusting the level of a reference tone in quiet to match the annoyance of that whole stimulus. The stimulus for the green marker was played at  $L_R = 5$  dB above  $L_{pn}$  and the one for the red marker at  $L_R = 10$  dB. These values were measured to provide a link between the equal annoyance contours in Figure 6.1 and Figure 6.2 with those in Figure 6.3.

It can be seen that the equal annoyance contours in Figs. 6.1-6.3 follow similar contours to the audibility thresholds, indicating the physiological constraints depicted in the audibility threshold have the greatest impact on equal annoyance level. However, the figures show consistently different contour levels for the two reference tone levels. In the outdoor scenario the contours are consistently higher and the slopes less steep than in the indoor scenario. This suggests that both reference tone levels and scenario also affect the equal annoyance contours significantly. Higher tone level above background noise is perceived as more annoying. The influence of scenario and frequency on relative annoyance cannot be judged immediately from Fig. 6.1 and 6.2. The addition of garden noise in the outdoor scenario did not seem to change the response strongly possibly because the tones are not masked effectively by the garden noise at low frequencies as seen in Figures 5.4 and 5.5. For each tone there are slight variations in the differences between equal annoyance contours and audibility thresholds and also in how close the equal annoyance contours are together.

To examine the equal annoyance data more closely with respect to frequency dependence and to find fine details in the different level distance between contours a new parameter was introduced. For this, audibility threshold was first subtracted from equal annoyance level for each tone. Next the tone  $L_{EA-T}$  was subtracted from the  $L_{EA-T}$  of the reference ( $L_{R-T}$ ) giving a parameter that we call the relative sensation level. If the same difference between equal annoyance level and audibility threshold occurs at the reference tone and at another tone frequency, the tone will have a relative sensation level value of 0. This showed any effect of tone frequency and masking noise type more clearly in the variations in relative sensation level values relative to the reference. These values are presented in Figure 17.1, Figure 17.2 and Figure 17.3 in Appendix VIII and all show very similar behaviour. Therefore, Figure 6.8 shows the relative sensation levels averaged over all scenarios and masking noises.

If annoyance scaled linearly with tone levels and was independent of frequency then all relative sensation levels would be equal to zero because all the tone levels would be the same 5 dB and 10 dB above audibility threshold as the reference. Negative values would signify that tones have to be further away from the audibility threshold than 5/10 dB to be equally annoying: the sound would be relatively less annoying than the reference.

Positive values of relative sensation levels indicate that the equal annoyance contour and audibility threshold are closer together at the test tone compared to the reference tone. Or in other words, the tone starts to be relatively more annoying closer to the audibility threshold. If we assume that low frequency tones become more annoying closer to the audibility threshold than high frequency tones then it is expected that relative sensation level steadily decreases from low to high frequencies. Then the relative sensation level at 400 Hz should be negative.

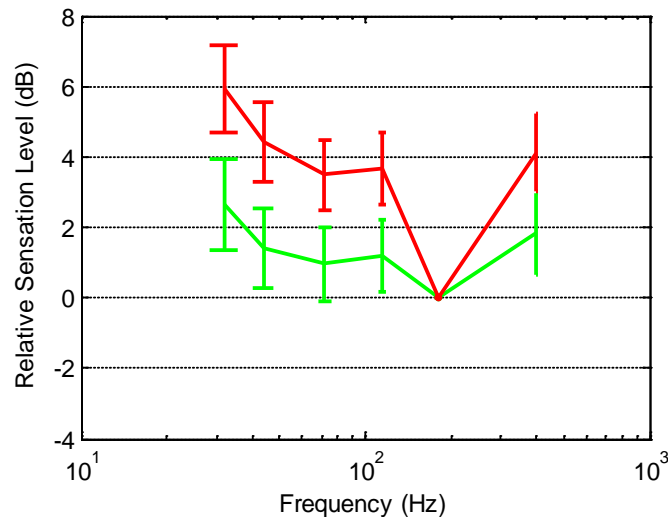


Figure 6.8 Relative sensation level for equal annoyance averaged over all masking noise types and scenarios at reference tone prominence levels of 5 dB (green) and 10 dB (red). Error bars show 95% confidence intervals.

Clearly the situation is not as simple as that. We find that ALL test tones have higher relative sensation levels than the reference tone. A possible reason could be that the reference tone becomes less annoying during the course of the experiment due to a habituation effect. If that is the case it would be reasonable to interpolate between the two relative sensation levels at 115 Hz and 400 Hz and bring the curves down so that the minimum is zero. This would in effect cancel any habituation effect. The resulting maximum relative sensation levels would then be about 1.5 dB for the green curve and 2.5 dB for the red curve. In that case all confidence intervals would include the zero value. The presence of a habituation effect is further supported by anecdotal remarks of participants.

Even if habituation is the reason for the pattern of results an obvious conclusion cannot be drawn: It is tempting to say that the more a subject listens to a tone the less annoying it becomes. While this might have been true for the reference tone for the duration of the experiment it would most certainly not be true over longer timescales and in a different context.

Another possible explanation for the low relative sensation levels at 180 Hz is the experimental design and specifically the choice of frequencies: The physical step size between 180 Hz and 400 Hz was larger than the approximated 1/3 octave between the low frequency tones. Additionally the reference tone was not the median frequency of the tone frequencies. The frequency choice of the reference tone may therefore have introduced a bias. There is however no obvious explanation why this bias would have caused all relative sensation levels to be positive.

To investigate this further, relative sensation levels have been alternatively computed by using the masking threshold calculated from the ISO 1996-2 instead of the measured audibility thresholds. The results are shown in Figure 6.7. The indoor and outdoor behaviour in that case is different because the hearing threshold is higher than the masking threshold for a number of tones in the indoor scenario. The alternative relative sensation levels in the outdoor scenario range between 3 and -2. In this case the relative sensation levels are higher for the higher frequencies than for the lower frequencies in contrast to Figure 6.6. The lower frequencies of the green curve are relatively less annoying than the reference while the values in the red curve are above zero. Most confidence intervals include the zero level. The frequency dependence is therefore not strong. This is a very similar result as Figure 6.8 when the habituation effect is discounted. It can be concluded that while the frequency dependence is probably very weak there might be a tendency for low frequency tones to be slightly more annoying than higher frequency tones. In this study we have not found that trend to be statistically significant.

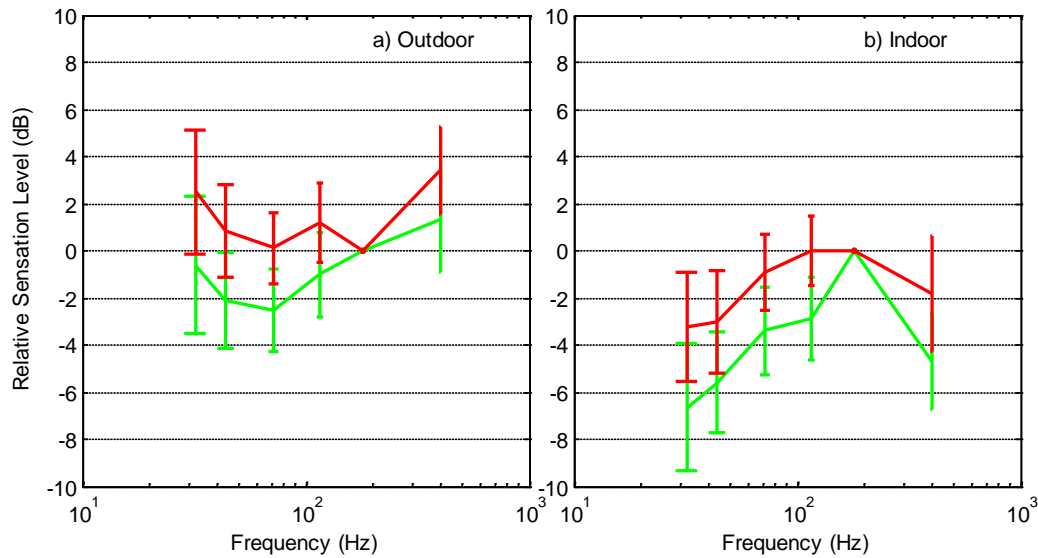


Figure 6.9 Alternatively derived relative sensation level for equal annoyance averaged over all masking noise levels a) outdoors, b) indoors at reference tone prominence levels of 5 dB (green) and 10 dB (red). Error bars show 95% confidence intervals. The alternative relative sensation level was derived by using  $L_{pn}$  instead of the measured values of the audibility threshold.

In addition to the last finding all other observations above were also statistically validated. The variations in relative sensation level were tested for significance. Effects of the five independent variables were analysed using two four-factor, repeated measures GLM ANOVAs. Details on methods and results can be found in Appendix VIII.

Having derived equal annoyance contours for a number of independent scenarios, the participants' attention was drawn to the masking sound to find annoyance relations between the scenarios.

## 6.4 Masking comparisons

Figure 6.10 shows the mean adjustments that participants made to the tone level of a 180 Hz pure tone when they matched its annoyance level with a test stimulus that consisted of a 180 Hz tone in masking noise. The higher the values the more annoying the stimulus including masking is. Test procedure and types of masking noise were the same as for the equal annoyance results in Section 6.3.

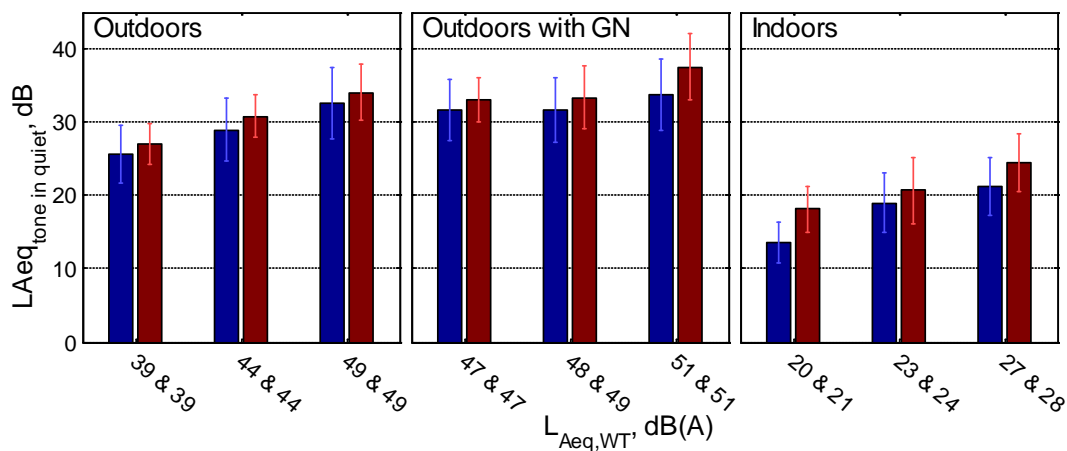


Figure 6.10 Equal annoyance level of a tone in quiet (room background noise) when matched to a tone in masking noise. The tone frequency was 180 Hz and the levels within masking were played at 5 dB (blue) and 10 dB (red) above  $L_{pn}$ . Error bars denote 95% confidence intervals.

The graph shows that the higher the level  $L_{Aeq}$  of the stimulus with masking the higher is the annoyance. The A-weighted levels of the stimuli containing the masking noise were louder than the ones without masking noise and the annoyance increases accordingly. The indoor levels are a lot less annoying than the outdoor levels. However the levels of the pure tone are not halved as the level  $L_{Aeq}$  is halved. This is due to the fact that the indoor tone levels would have been below hearing threshold if halved. The fact that the tone levels scale well with the A-weighted levels of the stimulus including the masking possibly reflects a tendency to match level or loudness rather than annoyance. This seems likely as the perception of a pure tone is very different from the perception of a tone in masking. We therefore conclude that we cannot compare the relative annoyance levels between different scenarios from those results.

The equal annoyance values in the blue columns ( $L_r = 5$  dB) are always at higher values than the red columns. This is in accordance to expectations as the stimuli with lower prominence above masking thresholds are expected to be less annoying as seen in Figures 6.1 and 6.2. It is worth noting though that the difference between red and blue columns is always smaller than the 5 dB difference between the  $L_r = 5$  and 10 dB levels. This has also been seen before when the equal annoyance contours in Section 6.1 were closer together for all tones compared to the reference tone. In effect the annoyance does not scale with tone level above masking threshold and tone levels closer to the masking threshold are relatively more annoying than tones exceeding the masking threshold by higher values. This finding is relevant for the application of correction factors  $K$  in the ISO 1996-2. If annoyance does not vary linearly with tone level above masking threshold this needs to be taken into account by corrections that decrease with level. It would take listening tests at a number of levels to determine relevant corrections.

The presented results have been found to be statistically significant. Both ANOVAs showed the effect of  $L_r$  to be significant. In comparing the indoors and outdoors scenarios  $L_r$ , the presence of garden noise, and scenario show a significant effect on the results. The method and results as well as non significant effects can be found in Appendix IX.



## 7 Part B: The annoyance of large wind turbines compared to small wind turbines

This chapter contains the results of Part B of the study. Equal annoyance contours are shown of the road traffic noise level compared to recordings of a small 600 kW turbine (small wind turbine) and a large 2.3 MW turbine (large wind turbine). In contrast to the previous chapter the term “wind turbine noise” is now used for the recorded sounds of the two wind turbines. Stimuli were created according to the procedure described in Section 5.5. Analysis showed that the spectra had broadband characteristics at low frequencies but no evident tonal content.

Figure 7.1 and Figure 7.2 show the equal annoyance contours for the outdoor and the indoor scenarios respectively. With increasing wind turbine level, the equal annoyance level of traffic noise increases. This is consistent with Part A results and with other published work. The conclusion is strengthened by the fact that participant response is very consistent.

The large wind turbine showed consistently lower equal annoyance levels than the small wind turbine in the outdoors scenario when no garden noise was present. Therefore the large turbine was perceived as being less annoying than the small turbine at the same A-weighted levels. The finding can be explained from the turbine sound characteristics detailed in Section 5.5: Both turbines show mainly broadband characteristics with no significant tonal content. The small turbine showed higher sound pressure levels at low frequencies and the temporal variation or “swishing” sound was more pronounced. Figure 7.1b) shows how in the presence of garden noise the two curves are not significantly different anymore. The reason is evident from Figure 7.3 where the total critical band levels for both turbines have been computed with and without garden noise. The figure shows that the garden noise effectively decreases the spectral differences at low frequencies: They become indistinguishable as shown in the equal annoyance results. While spectral characteristics might have been the main factor in the masking it is impossible from the current results to tell how effectively the swishing was masked by the garden noise. This will be an important area of future work.

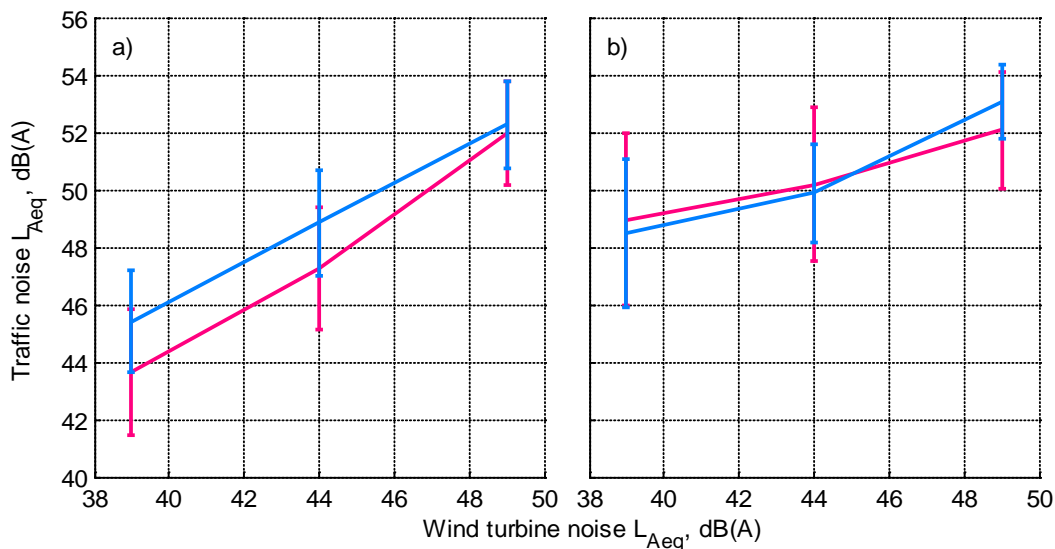


Figure 7.1 Equal annoyance contours for recorded turbines small wind turbine (blue lines) and a large wind turbine (red lines) matched to a neutral noise source (TN) a) Outdoor scenario without garden noise, b) Outdoor scenario with garden noise. Error bars are confidence intervals ( $\alpha = 0.05$ ).

In surveys wind turbine noise at the same A-weighted sound pressure levels as traffic noise are perceived as more annoying. Therefore the traffic noise levels that match the annoyance of the wind turbine noise would be expected to be much lower than the levels of the wind turbine noise. Compared to survey studies (Pedersen, 2007 and Pedersen & van den Berg, 2009) the traffic noise levels shown in Figure 7.1 are surprisingly large. Pedersen (2007) find A-weighted traffic noise

levels  $L_{DEN,Aeq}$  to be between 5 and 15 dB(A) louder to be equally annoying whereas the difference in Figure 7.1a) ranges much lower between 3 and 6 dB(A). There are two possible explanations for this. Firstly the study by Pedersen & van den Berg contained stall regulated wind turbines which are possibly louder sources than the modern pitch regulated turbines used in this study. Tonal content was also not assessed by Pedersen & van den Berg. Therefore the degree to which wind turbines are more annoying than traffic noise will vary depending on the types of wind turbines included in the studies. Or the difference in traffic noise level could be due to the fact that the outdoor measurements were perceived to be out of context in the listening room: When listening to outdoor noise levels in a listening room the outdoors levels are often perceived excessively loud. This would apply to traffic noise in the same way as it applies to wind turbine noise. But because the participants were not informed about the source of the noise they would not have introduced a bias due to attitude. Another explanation would be that the two sounds were matched in loudness rather than in annoyance. In that case we would expect the slope to be about 1. This is not the case: The slope is 0.6 in Figure 7.1a and less in 7.1b. However there might be an influence of that type of rating behaviour on the results.

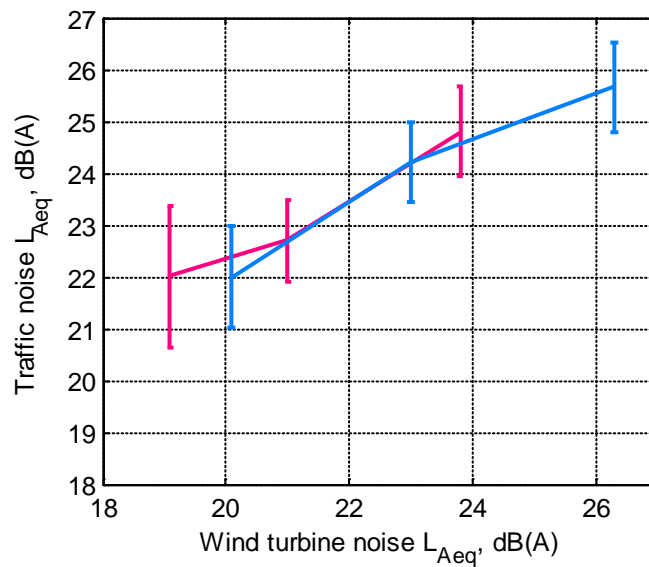


Figure 7.2 Indoor scenario: Equal annoyance contours for recorded turbines: Small wind turbine (blue line) and large wind turbine (red line) matched to a neutral noise source (TN). Error bars show 95% confidence interval

Although it is important to keep in mind that to achieve the same A-weighted levels the large wind turbine needs to be further away from the receiver, this result also shows more generally that annoyance from small wind turbines is not necessarily larger than from large wind turbines at the same A-weighted levels. Careful analysis of the spectral content of the turbine sound will be necessary. The spectral characteristics of the sound will depend on source characteristics which will be different at different wind speeds, on propagation distance, direction of the receiver in relation to wind direction and many other factors. The fact that so many variations are possible will have an impact on the logistics and value of listening tests.

Figure 7.1b) shows that the equal annoyance contours with garden noise start at higher levels and increase to similar maximum values as in the case without garden noise. Therefore annoyance with garden noise is perceived as higher than without garden noise for low wind turbine levels whereas the difference at higher levels is negligible. This result is similar to the results in Part A and can be explained in the same way: When listening to outdoor noise levels in a listening room the outdoors levels are often perceived excessively loud because they are out of context. The reason that this effect decreases for higher turbine levels is then due to the fact that the turbine dominates the sound at higher levels. In this context it should be mentioned that because the experimental design draws the participant attention specifically to level variations other annoyance related factors than context such as affective state might also be excluded from the response.

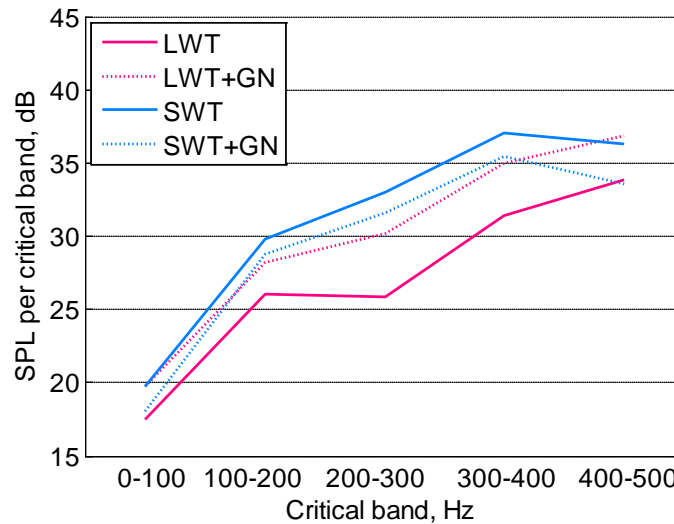


Figure 7.3 Outdoor scenario: Critical band levels computed from hearing threshold weighted frequency analysis of the noise for the noise from the large turbine (blue) and the low turbine (magenta) with and without garden noise, dotted and solid lines respectively. The A-weighted level was 44 dB(A) for all curves.

The indoors results are shown in Figure 7.2. There is no difference in annoyance between the two turbines. When comparing the spectra for the two wind turbines in Figure 5.5 b) this result is surprising because the small turbine has significantly more low frequency content than the large turbine. However, the A-weighted indoor level was very low at 19-26 dB(A). This might explain why the annoyance perception is not very different for the two turbines. The slope of the indoor equal annoyance levels is 0.57 which again is indicative that the results are not mainly an artefact of loudness rating.

The fact that Part B levels are quieter indoors than outdoors agrees well with the Part A findings. However, levels cannot be compared directly between the two parts of the study because the Part A levels refer to the sound pressure levels of tones while the Part B values are the A-weighted average spectrum levels.

Fig. 7.1b shows that the inclusion of garden noise removes any significant differentiation between equal annoyance levels for the two turbine types. This is possibly due to the masking of spectral characteristics of the specific turbines which would also explain higher TN levels for the quietest  $L_{Aeq,WT}$ . The presented results have been found to be statistically significant. The method and results as well as non significant effects can be found in Appendix X.

## 8 Conclusions

In this study on the perception of wind turbine noise, audibility thresholds and equal annoyance contours have been established for idealised wind turbine sounds containing low frequency tones at frequencies between 32 Hz and 400 Hz. The listening test simulated an indoor scenario and an outdoor scenario without and without masking garden noise.

### 8.1 Audibility and masking thresholds

Audibility thresholds for tones were first established in quiet. The thresholds agreed well with published values (Fastl & Zwicker, 2007 and ISO 389-7, 2005), between 72 and 180 Hz. An overestimation of the audibility threshold of about 5 dB at 400 Hz was attributed to a possible masking effect caused by the loudspeaker system. At the lowest frequencies the audibility threshold agreed better with hearing thresholds published by Fastl & Zwicker (2007) than the more commonly used thresholds in ISO 389-7.

Masking noise was then added to the tones, consisting of sound broadly representing broadband wind turbine sound. For the outdoor scenario, sound was also added that simulated garden noise modelled on deciduous foliage. The measured audibility thresholds in the outdoor scenario with and without garden noise were found to be at least 10 dB higher than the threshold in silence which suggests that the tonal audibility was determined by masking. In the indoor scenario the audibility of the tones between 44 and 180 Hz was also determined by the masking threshold whereas the tones at 32 Hz and 400 Hz were audible at the level of the hearing threshold.

For the outdoor scenario masking thresholds found in the listening test were therefore in good agreement with the masking thresholds predicted using the ISO 1996-2 Annex C standard. Tone frequencies below 180 Hz agreed to within  $\pm 2$  dB while the maximum deviation at 400 Hz was just under 5 dB. Some indoor thresholds were so close to the hearing threshold that using the masking threshold to describe audibility became meaningless.

A method to assess the audibility of broadband stimuli (Pedersen 2008) was tested for two different examples which were a) the background noise in the listening room with vents and loudspeakers switched on and off and b) the indoor scenario masking noise at three different levels. The method has been found to give reliable audibility estimates in a number of critical bands in the frequency range between 0 and 500 Hz.

### 8.2 Annoyance from idealised stimuli

Equal annoyance contours were established by comparing tones in different types and levels of masking noise with a reference tone at 180 Hz. The equal annoyance tone levels follow very similar contours to the audibility thresholds at higher levels. The two different tone prominence levels are consistently different and follow almost parallel contours.

To answer the question as to whether low frequency tones in the wind turbine sound are more annoying than higher frequencies with the same audibility, relative sensation levels have been defined. They are calculated by subtracting the equal annoyance tone levels from the audibility threshold values and the reference tone levels. The relative sensation level was not found to be significantly dependent on frequency. It can therefore be concluded that low frequency tones with the same prominence as tones of higher frequencies are either not more annoying or that the increase in annoyance is very weak.

When investigating the effect of garden noise the equal annoyance contours with and without garden noise were not found to be significantly different for the sound pressure levels tested in this project.

### **8.3 Comparison of different masking sounds**

It was intended to compare the annoyance of different scenarios by matching the level of a pure tone with a tone at the same frequency in different types of masking noise. The results show that the equal annoyance level increased at a similar rate as the level of the A-weighted stimulus including masking. The equal annoyance levels were lower in the indoor scenario which means that the stimuli were less annoying than the outdoor stimuli. The equal annoyance levels were larger for stimuli including garden noise because their A-weighted total level was higher. In summary, it seems rather than matching annoyance this test has in effect matched levels. That means that relative annoyance levels between different scenarios cannot be compared directly.

Generally it is seen that annoyance increases the more tonal energy content there is in the spectrum. This agrees with published results. However, the relative annoyance level does not increase as fast as the tone level in masking. This is probably due to the fact that the A-weighted level of the tone in masking does not increase as fast as the tone level.

### **8.4 Annoyance from recorded wind turbines in comparison to a reference noise source**

Part B of the listening tests developed an objective method of comparing recordings of pairs of turbines. Normalisation of wind turbine levels was used to focus listener attention on spectral characteristics, to avoid unrealistic small distances and therefore near-field directional effects as well as to reduce the effects of differences in wind speed and source strength. This is reasonable because regulations ensure a certain maximum noise level that cannot be exceeded. The result of normalisation is a less direct comparison, as relative required distances between turbines cannot be calculated easily.

The presented comparison of two specific turbine samples shows consistent participant response and therefore serves to illustrate the usefulness of the comparison method. The results are not meant to be and cannot be representative of general large/small wind turbine noise behaviour given the vast number of possible different turbine models, terrain types and sound propagation conditions.

The results show that the small wind turbine is perceived as more annoying than the large wind turbine in the outdoor scenario when the stimuli are not masked by garden noise. As soon as garden noise is added the differences in annoyance disappear, because the spectral differences between turbines are effectively masked. The indoor comparison shows that both turbines are perceived as identically annoying. This is at much lower A-weighted sound pressure levels but for sounds with clear spectral differences in the low frequency part of the spectrum. It follows that spectrum analysis might be useful to predict relative annoyance at sufficiently high sound pressure levels. The results from Part A of the study will be useful for this purpose too.

When comparing the results from idealised stimuli and the comparison of the recordings, the increase in relative annoyance levels with increasing wind turbine noise levels is common to both. Indoors equal annoyance levels are consistently lower in the indoor scenario compared to the outdoor scenario in both parts of the study. The difficulty arises when relative annoyance is to be forecast for different scenarios. The high equal annoyance levels for garden noise indicated excessive equal annoyance levels. This has been attributed to the problems with context for the outdoor scenario and leads to the conclusion that the outdoor levels without garden noise might also be exaggerated. This is a general problem with laboratory studies and can best be clarified using a survey study.

## **8.5 Reproduction of realistic, calibrated sound at low frequencies**

To achieve immersive realistic wind turbine sound in a garden environment a fully calibrated spatial audio reproduction system has been designed. The setup consisting of an ambisonic high frequency loudspeaker ring and 8 subwoofers in a so-called CABS configuration provides a high degree of room mode suppression and calibration accuracy of  $\pm 5\text{dB}$  at frequencies between 30 and 600 Hz. The superior spatial performance of this reproduction technique was judged to be more important than the disadvantage of restricted calibration accuracy at frequencies above 600 Hz due to the interference of the ambisonic ring. When listening to the stimuli the characteristics of the broadband sounds were not audibly changed compared to other reproduction systems. A Matlab software package was developed that allows largely automatic calibration and fully reproducible listening test conditions. Spatial variability of sound was kept to a minimum. This listening test setup allows highly reproducible, calibrated listening tests.

In summary the study has shown that listening tests can be successfully used to find answers to the perception of low frequency tonal wind turbine noise and to compare recordings of wind turbine sounds. Further work is needed to investigate the role of temporal variation on annoyance and to relate the annoyance between different scenarios.

## 9 References

- Berglund, B., Berglund, U. & Lindvall, T. (1975) Scaling Loudness, Noisiness and Annoyance of Aircraft Noise. *Journal of the Acoustical Society of America*, 57, 930-934.
- Berglund, B., Berglund, U. & Lindvall, T. (1976) Scaling Loudness, Noisiness and Annoyance of Community Noise. *Journal of the Acoustical Society of America*, 60, 1119-1125.
- Bradley, J. S. (1994) Annoyance Caused by Constant-Amplitude and Amplitude-Modulated Sounds Containing Rumble. *Noise Control Engineering Journal*, 42, 203-208.
- Celestinos, A. & Nielsen, S. B. (2008) Controlled Acoustic Bass System (CABS): A Method to Achieve Uniform Sound Field Distribution at Low Frequencies in Rectangular Rooms., *J. Audio Eng. Soc.*, 56(11), 915 – 931
- Devine-Wright, P. (2009) Rethinking NIMBYism: The role of place attachment and place identity in explaining place-protective action. *Journal of Community & Applied Social Psychology* 19, 426-441
- Fastl, H. & Zwicker, E. (2007) *Psychoacoustics, Facts and Models*. Springer, Berlin, Germany.
- Gerzon, M. A. (1985) Ambisonics in Multichannel Broadcasting and Video, *J. Audio Eng. Soc.*, 33(11), pp. 859-871
- Hellman, R. P. (1982) Loudness, annoyance, and noisiness produced by single-tone-noise complexes. *J Acoust Soc Am*, 72, 62-73.
- Hellman, R. P. (1984) Growth rate of loudness, annoyance, and noisiness as a function of tone location within the noise spectrum. *J Acoust Soc Am*, 75, 209-18.
- Hoffmeyer, D. and Jakobsen, J. (2010) Sound insulation of dwellings at low frequencies. *J. Low Frequency Noise, Vibration and Active Control*, 29(1), 15-23.
- Huang, Y. F., Di, G. Q., Zhu, Y. T., Hong, Y. P. & Zhang, B. J. (2008) Pair-wise comparison experiment on subjective annoyance rating of noise samples with different frequency spectrums but same A-weighted level. *Applied Acoustics*, 69, 1205-1211.
- ISO 389-7:2005 (E) Acoustics – Reference zero for the calibration of audiometric equipment. Part 7: Reference threshold of hearing under free-field and diffuse-field listening conditions. Second edition 2005-11-01.
- ISO 1996-2 (2007) Acoustics – Description, assessment and measurement of environmental noise – Part 2: Determination of environmental noise levels. Annex C: Objective method for assessing the audibility of tones in noise. Second edition 15-3-2007.
- Jakobsen, J. (2005) Infrasound Emission from Wind Turbines. *Journal of Low Frequency Noise, Vibration and Active Control*, 24(3), 145-155.
- Jung, S. S., Cheung, W. S., Cheong, C. & Shin, S. H. (2008) Experimental Identification of Acoustic Emission Characteristics of Large Wind Turbines with Emphasis on Infrasound and Low-Frequency Noise. *Journal of the Korean Physical Society*, 53, 1897-1905.
- Key, K. F. (1979) Effects of Low-Frequency Noise on Performance and Annoyance. *Journal of the Acoustical Society of America*, 65, S119-S119.
- Kryter, K. D. (2007) Acoustical, sensory, and psychological research data and procedures for their use in predicting effects of environmental noises. *Journal of the Acoustical Society of America*, 122, 2601-14.

- Legarth, S. V. (2007) Auralization and Assessments of Annoyance from Wind Turbines. *Second International Meeting on Wind Turbine Noise*. Lyon, France, 20-21 September 2007, pp 14.
- Leloudas, G., Zhu, W. J., Sorensen, J. N., Shen, W. Z. & Hjort, S. (2007) Prediction and reduction of noise from a 2.3 MW wind turbine - art. no. 012083. *Science of Making Torque from Wind*, 75, 12083-12083.
- Lim, C., Kim, J., Hong, J. & Lee, S. (2008) Effect of background noise levels on community annoyance from aircraft noise. *Journal of the Acoustical Society of America*, 123, 766-71.
- McKenzie, A. R. (1997) Objective and subjective assessment of tonal components in noise from UK wind farm sites. *Wind Energy Conversion* 1996, 467-472.
- Madsen, K. D. (2010) Windturbine spectra for LF project, Technical Note, DELTA, DK, pp. 9
- Madsen, K.D. (2010a) Personal communication.
- Moorhouse, A. T. (2005) Virtual Acoustic Prototypes: listening to machines that don't exist, *Acoustics Australia*, 33(3) 97-105
- Møller, H. and Pedersen, C.S. (2004) Hearing at low and infrasonic frequencies. *Noise & Health*, 6(23), Apr-Jun 2004.
- Niedzwieki, A. & Ribner, H. S. (1979) Subjective loudness and annoyance of filtered N-wave sonic booms. *Journal of the Acoustical Society of America*, 65, 705-7.
- Nilsson, M. E. (2007) A-weighted sound pressure level as an indicator of short-term loudness or annoyance of road-traffic sound. *Journal of Sound and Vibration*, 302, 197-207.
- Oerlemans, S., Sijtsma, P. & Lopez, B. M. (2007) Location and quantification of noise sources on a wind turbine. *Journal of Sound and Vibration*, 299, 869-883.
- Pawlaczyk-Luszczynska, M., Dudarewicz, A., Waszkowska, M. & Sliwinska-Kowalska, M. (2003) Assessment of annoyance from low frequency and broadband noises. *International Journal of Occupational Medicine and Environmental Health*, 16, 337-43.
- Pawlaczyk-Luszczynska, M., Dudarewicz, A., Waszkowska, M. & Sliwinska-Kowalska, M. (2009) Annoyance Related to Low Frequency Noise in Subjective Assessment of Workers. *Journal of Low Frequency Noise Vibration and Active Control*, 28, 1-17.
- Pedersen, C. S. & Møller, H. (2010) An analysis of low frequency noise from large wind turbines. *14<sup>th</sup> International Meeting on Low Frequency Noise and Vibration and its Control*, 339-360.
- Pedersen, E. & Larsman, P. (2008) The impact of visual factors on noise annoyance among people living in the vicinity of wind turbines. *Journal of Environmental Psychology*, 28, 379-389.
- Pedersen, E. & Waye, K. P. (2004) Perception and annoyance due to wind turbine noise - a dose-response relationship. *Journal of the Acoustical Society of America*, 116, 3460-3470.
- Pedersen, E. & Waye, K. P. (2007) Wind turbine noise, annoyance and self-reported health and well-being in different living environments. *Occupational and Environmental Medicine*, 64, 480-486.
- Pedersen, T.H. (2007) The "Genlyd\*" Noise Annoyance Model, Dose-Response Relationships Modelled by Logistic Functions, Technical Report, DELTA, DK, pp. 121.



Pedersen, E., Van Den Berg, F., Bakker, R. & Bouma, J. (2009) Response to noise from modern wind farms in The Netherlands. *Journal of the Acoustical Society of America*, 126, 634-43.

Pedersen, J. A. (2006) Loudspeaker-Room Adaptation for a specific Listening Position using Information about the Complete Sound Field, AES 121st Convention, San Francisco, US, Oct 5, Paper number 6908.

Pedersen, T. H. (2007) The "Genlyd" Noise Annoyance Model – Dose-Response Relationships Modelled by Logistic Functions. Technical Report, DELTA, DK, pp. 121.

Pedersen, T. H. (2008) EFP-06 project Low frequency noise from large wind turbines – A procedure for evaluation of the audibility for low frequency sound and a literature study. Technical Report, DELTA, DK, pp 78.

Persson, K. & Bjorkman, M. (1988) Annoyance Due to Low-Frequency Noise and the Use of the Db(a) Scale. *Journal of Sound and Vibration*, 127, 491-497.

Persson Waye, K. P. & Ohrstrom, E. (2002) Psycho-acoustic characters of relevance for annoyance of wind turbine noise. *Journal of Sound and Vibration*, 250, 65-73.

Plovsing, B. & Kragh, J. (2006) Nord2000. Comprehensive Outdoor Sound Propagation Model. Part 1: Propagation in an Atmosphere without Significant Refraction. Revised Report AV 1849/00, DELTA Acoustic & Vibration.

Robinson, D. W. & Dadson, R. S. (1956) A re-determination of the equal-loudness relations for pure tones. *British Journal of Applied Physics*, 7, 166-181.

Schultz, T. J. (1978) Synthesis of social surveys on noise annoyance. *Journal of the Acoustical Society of America*, 64 (2), 377-405.

Søndergaard, B. & Madsen, K. D. (2008) Low Frequency Noise from Large Wind Turbines. Summary and Conclusions on Measurements and Methods. Technical Report, DELTA, DK, pp. 43.

Stevens, S. S. (1955) THE MEASUREMENT OF LOUDNESS. *Journal of the Acoustical Society of America*, 27, 815-829.

Turner, C. W. & Burns, E. M. (1977) Reduction of Annoyance as a Result of Addition of Low-Frequency Noise. *Journal of the Acoustical Society of America*, 61, S78-S78.

Van Den Berg, G. P. (2004) Effects of the wind profile at night on wind turbine sound. *Journal of Sound and Vibration*, 277, 955-970.

Van Den Berg, G. P. (2005) The beat is getting stronger: The effect of atmospheric stability on low frequency modulated sound of wind turbines. *Journal of Low Frequency Noise Vibration and Active Control*, 24, 1-23.

Van Den Berg, G. P. (2008) Wind turbine power and sound in relation to atmospheric stability. *Wind Energy*, 11, 151-169.

Van Kamp, I., Job, R. F., Hatfield, J., Haines, M., Stellato, R. K. & Stansfeld, S. A. (2004) The role of noise sensitivity in the noise-response relation: a comparison of three international airport studies. *Journal of the Acoustical Society of America*, 116, 3471-9.

Wagner, S., Bareiss, R., and Guidati, G. (1996) Wind Turbine Noise: Springer

Weinstein, N. D. (1978) Individual Differences in Reactions to Noise: A Longitudinal Study in a College Dormitory. *Journal of Applied Psychology*, 63 (4), 458-466.

Zimmer, K. & Ellermeier, W. (1998) Ein Kurzfragebogen zur Erfassung der Lärmempfindlichkeit. *Umweltpsychologie*, 2 (2), 54-63.

## **9.1 Web-Sources**

[www.mathworks.com](http://www.mathworks.com), release 14 (10/9/2010)

[www.portaudio.com](http://www.portaudio.com), version 19 (10/9/2010)

[www.playrec.co.uk](http://www.playrec.co.uk), version 2.1.1 (10/9/2010)

# 10Appendix I: Advertisement for participant recruitment and consent form

## 10.1 Advertisement for testing in April 2010



**University of Salford**  
A Greater Manchester University

### **UNIVERSITY OF SALFORD NEWS RELEASE February 2010**

#### **Salford seeks paid volunteers for sound study**

The University of Salford's Acoustics Research Centre is seeking volunteers for a study into the effect of low frequency outdoor noises.

The study will involve testing each volunteer's hearing sensitivity at low frequencies, and comparing how different types of outdoor noises would affect them if they were heard at home.

Volunteers will take part in three one-hour sessions and the sounds will not be played at harmful levels.

**Eligible volunteers will be paid £10 an hour for each session.**

People who are over 18 years old and have normal or corrected hearing are asked to apply. Applicants will be asked some questions for screening purposes.

For more information or to apply contact Andrew King by emailing xxx or call xxx

## 10.2 Advertisement for testing in August 2010

### Want to volunteer in a sound study?

Wednesday, 21 July 2010

The University's [Acoustics Research Centre \(ARC\)](#) is calling all countryside-loving colleagues to volunteer for a study on the impact of environmental noises being carried out.

A team within the Centre is currently looking at how annoying sounds are depending on certain characteristics.

Characteristics associated with annoyance include low frequency content, which can introduce a deeper or 'rumbling' quality to a sound while monotonous tones, which have a clear and distinct pitch, are also commonly cited as irritating.

The team are studying the changes in annoyance of environmental sounds with different tones and deep, 'rumbling' content.

Firstly, the study will involve testing each volunteer's hearing sensitivity at low frequencies, which indicates how loud a sound needs to be, to be noticeable. Secondly, each volunteer will compare how loud one sound needs to be, to be as annoying as another. Volunteers are tested separately.

Volunteers will take part in two or three sessions lasting around an hour. Taking place in the acoustics laboratories, the sounds will be played at levels naturally encountered.

In particular, the team is looking for participants who are middle aged and above with relatively normal hearing for their age range.

If you would like to register to participate or you are interested in more information, please contact Andrew King on [a.king@pgr.salford.ac.uk](mailto:a.king@pgr.salford.ac.uk) or ext 54669.

## 10.3 Consent form

# Consent Form

Project : Comparative Annoyance from Noise  
Researcher : Andrew King  
Contact Details: a.king@pgr.salford.ac.uk  
Supervisor : Sabine von Hünenbein  
Contact Details: s.vonhunerbein@salford.ac.uk

Thank you for agreeing to participate in this study, taking place on .....

This form outlines the objectives of the study and your involvement.

The objectives are:

- To measure your hearing threshold in an audiometric test
- To compare how annoying different types of noise are

First we will carry out an audiometric test in which you will be asked to identify which, out of two intervals, contains a certain sound. Sometimes the background noise will be presented too. In the second and third test you will be asked to imagine you are at home, in the garden or living room whilst you play pairs of sounds and adjust one until it is equally as annoying as the other sound. The second test uses artificial sounds whilst the third test uses recorded sounds from the environment.

The levels of sound are quite low, typically what you might hear in the countryside, and so there is no risk to your hearing or your health. There will be regular breaks approximately every 30 to 40 minutes. However, if you feel tired or uncomfortable at any time or would like a break please press the 'help' button to pause the test and alert the researcher.

The information gathered from this study will be used for no other purpose except the completion of this study and the publication of its results. The results of this test will be stored anonymously. **Your participation is voluntary – you have the right to withdraw at any time without giving any reason.**

Please feel free to ask any questions at any time about the nature of the study and methods being used – the contact details are listed above.

☐ Please tick this box if you would like to be de-briefed after the current study.

☐ Please tick this box if you are happy to be contacted about participating in the future.

Participant : I agree to the terms

Name ..... Signature ..... Date .....

Researcher : I agree to the terms

Name ..... Signature ..... Date .....

## 11 Appendix II: Participant Screening

This section details the exact wording and layout of the screening form sent to prospective participants. The Noise Sensitivity Scale at the end is the short version of the Zimmer and Ellermeier Noise sensitivity scale (1998):

All information provided here will be kept confidential (only available in its raw form to project members) and shall not be published in any way that identifies the participant. Information shall only be kept if applicant participates.			
Forename:		Surname:	
Age:		Sex:	Male / Female (Delete as appropriate)
Occupation:		Nationality:	
Previous listening test experience	<b>A lot</b> (participated in more than 5 tests) <b>Some</b> (participated in between 2 and 5 tests before) <b>A Little</b> (participated in 1 test before) <b>None</b> (never participated in a test before) (Delete as appropriate)		
<p>For the following questions, please give one answer by deleting the answers that do not apply to you.</p> <p><b>Q1. What best describes the area surrounding your home?</b></p> <p>Inner city          Suburb (eg. City outskirts)          In the countryside          Other (please specify) .....</p> <p><b>Q2. How content are you with the area surrounding your home?</b></p> <p>Very unhappy          Unhappy          Neither unhappy or happy          Happy          Very Happy</p> <p>If you wish you lived in a different area type, please answer Qs 3 and 4. If not, please go to Q5.</p> <p><b>Q3. Which of the following area types do you wish you lived in?</b></p> <p>Inner city          Suburb (eg. City outskirts)          In the countryside          Other (please specify) .....</p>			

**Q4. How strong is your desire to live in the area selected in q.3?**

Strong  
Moderately Strong  
Moderately Mild  
Mild

**Q5. How good is your hearing, in general?**

Very good  
Good  
Moderate  
Poor  
Very Poor

**Q6. Do you have any specific problems with your hearing?**

Yes  
No

If yes, please provide details in the space below

.....

.....

.....

.....

Now please complete the Noise Sensitivity Scale below. Show whether you agree fully, rather agree, rather disagree or fully disagree with each statement by putting a tick in the relevant box.

Noise Sensitivity Scale	Agree fully	Rather agree	Rather disagree	Disa fully
1. It is no fun keeping up a conversation while the radio is on.				
2. I tend to notice disturbing sounds later than do other persons.				
3. I avoid noisy pastimes such as going to soccer matches or fairs.				
4. I wake up at the slightest sound.				
5. Even in noisy surroundings. I am able to work quickly and with concentration.				
6. On doing my shopping in the city. I hardly hear the street noise.				
7. After having passed an evening in a noisy pub I feel drained.				
8. When I want to fall asleep, hardly any sound can disturb me.				
9. On weekends I like to be in quiet places.				
<p><b>Thank you very much. We will contact you soon about participation, which is scheduled to take place in April.</b></p>				

The items within the scale are scored: 1: 3-0, 2: 0-3, 3: 3-0, 4: 3-0, 5: 0-3, 6: 0-3, 7: 3-0, 8: 0-3, 9: 3-0 (scored left to right across response boxes).



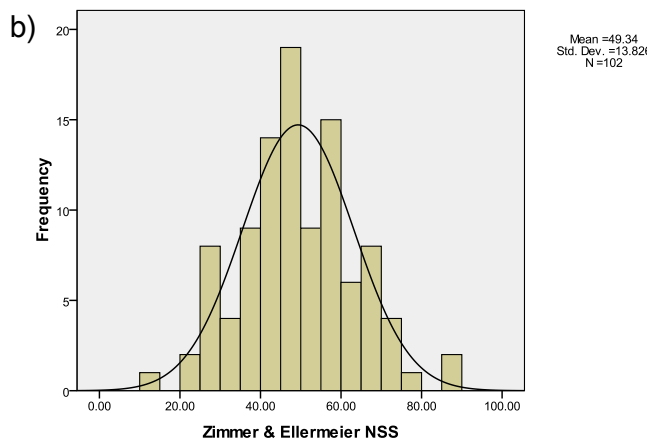
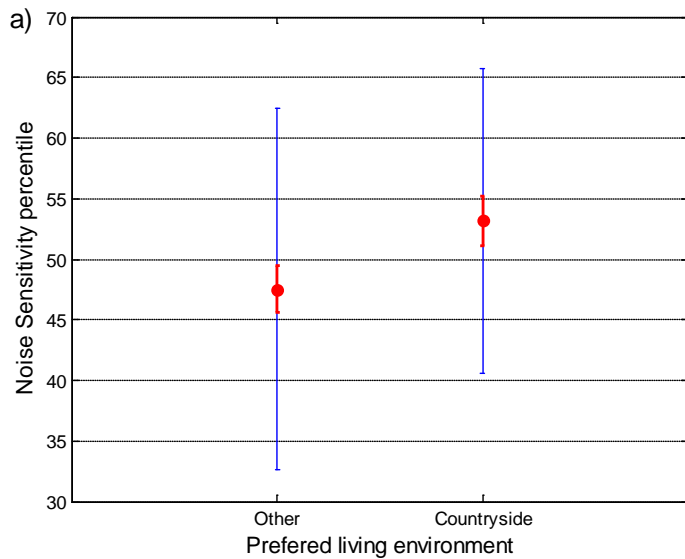


Figure 11.1 Zimmer-Ellermeier screening results a) Average score for volunteers living in the country compared to living in the city with red standard error bars and blue standard deviation bars, b) Distribution of scores in overall sample.

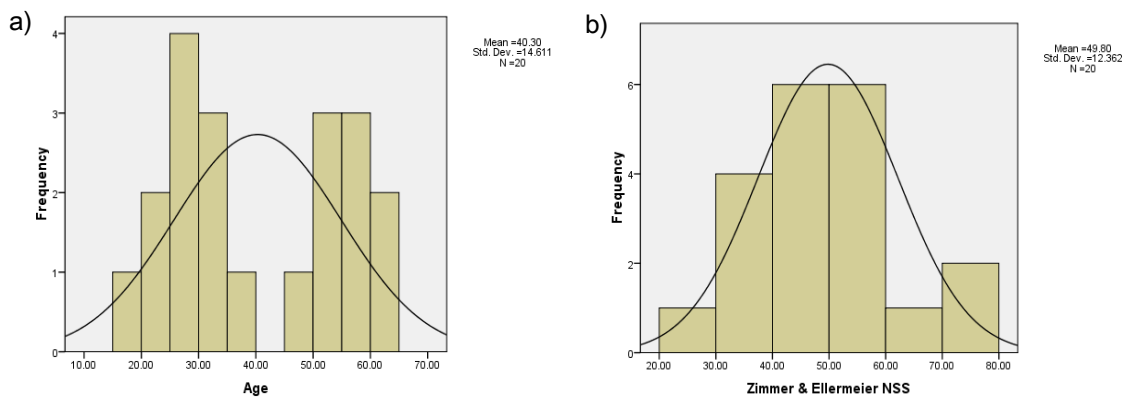


Figure 11.2 Test sample statistics, a) Age distribution of participants; b) Zimmer and Ellermeier NSS score distribution for participants.

## 12 Appendix III – Sound Reproduction System

The requirements placed by the test design on the listening environment and sound reproduction system were outlined in Section 5.1, along with the key decisions on selection of listening room and sound reproduction format. In this section details of the sound reproduction system and its calibration procedure are discussed.

### 12.1 Hardware

The PC that hosted the listening test software was located in the lobby area between the listening room's outer and inner walls. It contained an RME HDSP 9652 PCI card which provides up to 24 synchronous channels of 24 bit input and output on ADAT optical I/O. Optical data and coaxial word-clock cables linked this to two RME ADI-8 DS 8-channel DA/AD converters located inside the listening room. The analogue gain of the loudspeakers was set so as make the best use of soundcard bit-depth and all cabling was balanced to ensure low noise. The PC was also equipped with a dual head graphics card, one output of which was split and connected to a 15" touch-screen located within easy reach of the specified listener position. This allowed GUI windows to be displayed to the participant, while other windows showing tracking data, test progress, system status and live web-cam video of the participant were visible only to the researcher.

Eight Genelec 8030a were located in a 4 metre diameter ring centred on the participant, who in turn was at the centre of the room (see Section 5.1), and eight Genelec 7050B subwoofers were arranged in a CABS configuration (described later). Input sensitivity was set such that peak soundcard output level matched the maximum required SPL so as to make best use of the available soundcard bit-depth. Harmonic distortion was investigated for all loudspeakers but not found to be significant due to the low output levels used.

For the purposes of system calibration it was desirable to measure the entire reproduction chain including the DA converter, so a bespoke measurement system was implemented in the test software (described below). For this a measurement microphone system was connected to the soundcard, comprising B&K 4165 ½" capsules (or equivalent) with type 2269 preamps and Nordsonic type 336 Front Ends (gain adjustable in 10dB steps). The microphone was calibrated with a pistophone calibrator and factoring in all gains found the capsule sensitivity to typically be within 0.1dB of the value stated in its documentation.

### 12.2 Subwoofer System

Room modes are an unavoidable problem in small listening rooms and given the importance of accurate control of low frequency tone SPLs in this project, commissioning the low frequency reproduction system was of primary concern. After investigating various configurations a down-scaled version of what is used in low-frequency test chambers was selected. This is called the Controlled Acoustic Bass System (CABS) (Celestinos and Nielsen 2008) and comprises symmetrical sets of subwoofers on the front and back walls, with the rear delayed and in anti-phase (with a small attenuation) to cancel out the reflection from the rear of the room and attempt to mimic free-field plane wave propagation. Control of lateral and vertical modes is achieved by locating the subwoofers such that low-order modes are not excited either due to the subs being on a node or driving it in anti-phase. Eight subwoofers are installed in the Listening Room – four at each end located at floor and ceiling  $\frac{1}{4}$  and  $\frac{3}{4}$  of the way across the room.

CABS has two tuneable parameters for the rear subwoofers, being the delay relative to the front, typically the length of the room divided by the speed of sound, and an attenuation due to damping, which is harder to predict. To optimally choose these the transfer functions from each subwoofer to the listening position were measured separately. These were then combined numerically, varying both parameters and calculating a cost function (standard deviation of the total transfer function in dB, designed to penalise rapidly variation with frequency) in the range

from 25 Hz to 140 Hz. The values used for the parameters were chosen from the minima in the cost function, being a delay of 20.1ms and an attenuation of 1.8dB.

The effect on the low frequency response was measured and the range and rate of variation was significantly improved. Of particular significance was the removal of deep troughs in the response, due to the microphone and/or subwoofer being on a room mode node, since these cannot be robustly corrected by equalisation. This spatial variation in the sound field was assessed by playing pure tones while a trained listener deliberately moved their head in a volume close to the defined listening position; very little variation could be perceived. This quality meant the system could have compensation equalisation applied with confidence and that the resulting stimuli would be heard as intended by all participants.

## 12.3 Room Measurement Routine

As shown in Figure 5.3 the calibration routine was a three stage process; first the transfer functions were measured from each loudspeaker to the listener position, then correction functions were designed for each of these, and finally the correction functions were used to modify the stimuli such that the correct levels arrived at the listener position. Such room inversion (sometimes called "de-reverberation") is in general a non-trivial task because of the non-minimum-phase nature of loudspeakers and rooms, however it was possible to use inversion by Discrete Fourier Transform for this test because all the stimuli were approximately equal length loops so satisfied the periodicity requirement.

A necessary task prior to measurement was to set the analogue gain control on each loudspeaker to make best use of the soundcard bit depth. This required an approximate measurement of the broadband SPL at the listener position due to a given loudspeaker and this was performed before the measurement proper was commenced. White noise was used as the test signal and the transfer function from the loudspeaker was calculated by the Cross Power Spectral Density method with 16 averages of  $2^{15}$  samples ( $\approx 0.68$  second) measurements. This was compared to the transfer function required to match the maximum SPL required by the test to peak soundcard output, and the mean difference in the loudspeaker's pass-band was displayed to the operator to indicate how the gain control should be adjusted. Settings achieving a broadband level 0dB to 5dB above that required were deemed acceptable.

The interface is shown in Figure 12.1, where a listbox allows the operator to select which loudspeaker to measure, radio buttons input the Front End Gain, and meters display the peak microphone level, loudspeaker level relative to that required, and measurement coherence. Tick boxes prompt the operator to check all criteria are satisfied, upon which the "Measure" button becomes active and the measurement proper can be initiated. Upon completion the transfer function and coherence are calculated and displayed to the operator along with a dialog box asking if the measurement should be saved or repeated.

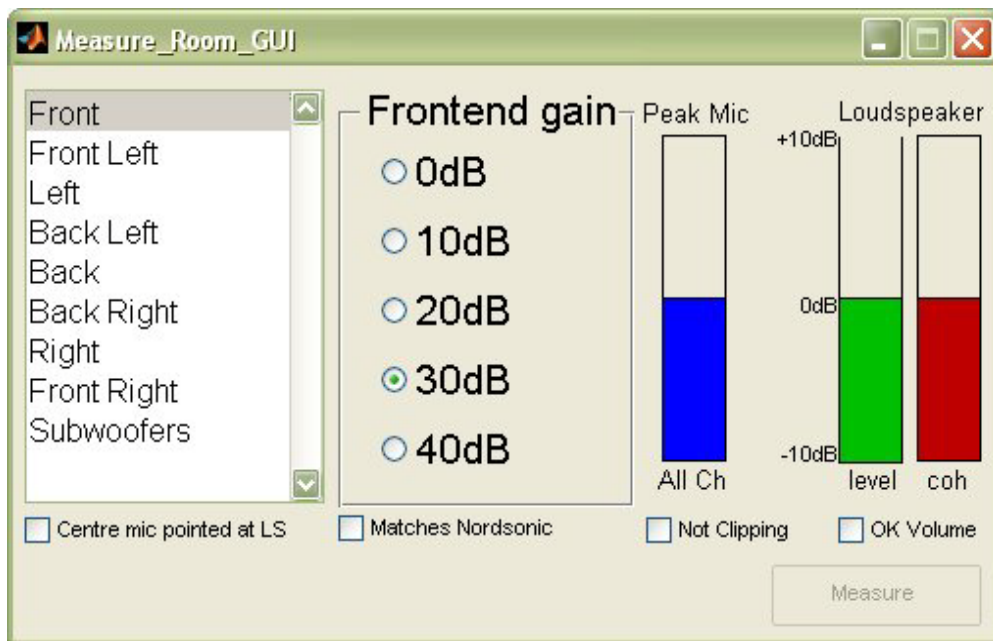


Figure 12.1: Room Measurement GUI

In order to be consistent with the stimuli periodicity assumption used for the compensation equalisation, a Cross Power Spectral Density measurement method was used, with white noise as the test signal and 20 averages. Preliminary measurements (by swept sine) showed the room mode decay to be quite long at some frequencies so it was decided to use a measurement period that matched the loop length ( $2^{19}$  samples  $\approx$  10.9 seconds); this meant the measurements were quite time consuming but immune to error from impulse response wrapping.

When measuring a room response with a single microphone, dips (nulls) in the frequency response will often be observed because destructive interference occurs at the microphone position at those particular frequencies. This is not a characteristic of the overall sound field and should not be compensated for by equalisation as such a boost would be heard as a tone at other locations in the room. Because this project is concerned with prominence of tones in masking noise it is crucial that other spurious tonal artefacts are avoided, hence it is preferable to have a null in the frequency response of the masking noise rather than risk creating a tonal artefact.

In order to assess what features of the measured response were genuine artefacts in the sound reproduced by the loudspeaker and which were local effects at the measurement mic position, multiple microphones were used. One microphone was located at the listener position (the focal microphone) and another three were scattered randomly up to 0.5 m away, located to avoid symmetry and blocking of any direct path from a loudspeaker to the focal microphone (Figure 12.2). The focal microphone transfer function was regarded as the primary measurement, and a power sum of all four microphones provided as estimation of the sound energy in the global sound field (Pedersen, 2006).



Figure 12.2: Measurement Microphones

## 12.4 Room Inversion Routine

In order to achieve a controlled crossover between the two loudspeaker systems and avoid excessive boosting of out-of-band frequency content, target responses were designed for both the subwoofers and full-range loudspeakers. Because of the nature of the equalisation these could be arbitrary zero-phase responses, and because the calibration routine would phase-compensate all loudspeakers the crossover curve could be constant-pressure rather than constant-power, hence the target responses were all defined by simple piecewise-linear profiles in transfer-function magnitude (scalar / not dB). Based on the low and high pass filtering already present in the Genelec 7050Bs, a high-pass profile was set between 18 Hz and 30 Hz (the low-frequency reproduction limit in the test specification) and a low-pass crossover profile between 100 Hz and 140 Hz. The ambisonic ring had a complementary high-pass profile between 100 Hz and 140 Hz, and a low-pass profile between 20kHz and 24kHz to match a decaying trend in response measured there (stimuli spectra were specified up to the 10kHz third octave band).

It was noted in the previous section that local nulls in the measured response should not be compensated for as this would give rise to tonal artefacts elsewhere. To satisfy this condition the nulls were removed by taking the maximum value from the focal and global magnitude responses at each frequency before inversion took place (see Figure 12.12a for example). This was then re-sampled into a log-frequency base, smoothed, and re-sampled back to linear-frequency base, a process which preserved the low-frequency variations exactly and then gradually introduced smoothing so that only the general shape of the high-frequency response was correct and the detail left unchanged. This was consistent with the test design, which considered that low-frequency accuracy was paramount, whereas high-frequency variations could for Part A be regarded as aiding envelopment due to the plethora of sound diffusing treatment installed in the listening room. For Part B it was confirmed that masking of both wind turbine spectra were affected similarly even at the higher frequencies. To ensure the phase was not corrupted by this process the focal response was first decomposed into its minimum-phase and all-pass parts. The operations above were applied to the magnitude of the transfer function and then a new matching minimum-phase phase-response was calculated. This was finally combined with the original all-pass part to create a response with consistent magnitude and phase and all the delays of the original measurement. Inversion was then the simple process of dividing the target response by the processed measurement. These correction spectra were calculated for each loudspeaker in the ambisonic ring and the CABS system, and defined what digital data must be sent to the sound-card to achieve a desired pressure (in Pascals) at the listener position. Graphs showing the calibration results for each loudspeaker are shown in Section 12.8.

## 12.5 Modification of Stimuli

The stimuli were represented by a variation of ambisonic B-format scaled such that the mono  $w$  channel exactly equalled the desired pressure in Pascals at the listener position and the remaining  $x$  and  $y$  channels encoded the surround information, effectively the particle velocity vector normalised by  $\rho c$ . For each loudspeaker this was decoded according to

$$p_i(t) = \frac{1}{N} \left[ w(t) + \sqrt{2} \cos(\theta_i) x(t) + \sqrt{2} \sin(\theta_i) y(t) \right], \quad \text{Equation 1}$$

where  $p_i$  is the pressure contributed by loudspeaker  $i$  located at angle  $\theta_i$  (anti-clockwise relative to front) and  $N$  is the total number of loudspeakers. The factors of  $\sqrt{2}$  arise because the first order spherical harmonics are multiplied by two relative to the zero<sup>th</sup> order, and this gain is split between the encoding and decoding equations (Gerzon, 1985). The subwoofer system was fed by only the  $w$  channel. The task of combining this with the correction equalisation to obtain the signal that would be sent to the soundcard was performed in the compile routines, since that was where the stimuli specification and sound reproduction data were brought together (see Figure 5.3).

In Part A the length of all masking sound loops exactly matched the periodicity of the room measurement, hence the frequency array arising from an FFT operation matched too. Application of the room correction could therefore be simply achieved by taking the FFT of the stimuli loop, multiplying by the room correction spectra, and taking an inverse FFT. The tone loops had much shorter periodicity, the optimum length in samples being found by expressing the ratio of the tone frequency to the sample frequency as rational fraction of the smallest possible two integers, and were corrected by the appropriate magnitude scaling. Ambisonic decoding was performed for each loudspeaker as above, and the magnitude and phase modified as required by the correction equalisation.

Part B involved no tones, but the turbine loop lengths were fixed by the turbine's natural periodicity so deviated slightly from the measurement periodicity. To compensate for this the correction equalisation spectrum was resampled for each loudspeaker using spline interpolation up to the Nyquist frequency, and then the negative frequency data was copied and conjugated to ensure the resulting audio data was real. Other than this, the process for modifying the Part B stimuli was identical to that for the Part A masking.

## 12.6 Run Routines

When writing the run routines, one code management objective was to ensure that the test structure was clearly evident and not masked by details of writing to the soundcard. Accordingly all the audio playback functionality was written in a sub-function called `loop_player.m`, which kept all the stimuli loops and their current playback position in local persistent memory as well as keeping track of the playback buffering, and accepted playback gains (scalar or dB) for each stimuli loop as its arguments. Unfortunately Matlab does not support multi-threading, so this function had to be repeatedly called from the top level code to avoid glitches in the audio reproduction, but it has still led to a tidy, high level code with a low risk of errors occurring if the test procedure code was updated. The modified stimuli loops were sent directly to the soundcard outputs with only a uniform scalar gain applied. Typical gains were 0dB or -infdB for the garden noise, or -5dB, 0dB and +5dB for the three wind turbine levels (the wind turbine stimuli was generated at 44dB(A)). "Gain fades" were also supported to allow gradual introduction of stimuli and "click-less" switching.

## 12.7 Daily Checks

A program was available that played pink noise through any soundcard channel and this was used to ensure all loudspeakers were functional before tests began. Then a program similar to the room measurement loudspeaker level setting routine was used to ensure loudspeaker levels had not been accidentally altered (Figure 12.3). Finally a third program played corrected white noise out of all loudspeakers so the system could be checked with a sound level meter.

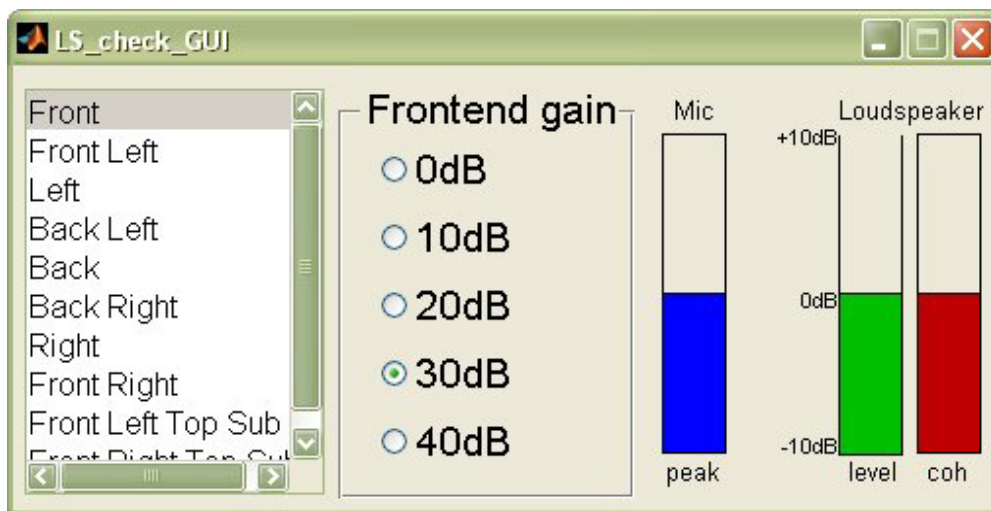


Figure 12.3: Loudspeaker Level Check GUI

## 12.8 Loudspeaker Calibration Data

Graphs follow showing the calibration results for each loudspeaker follow, where  $H(f)$  is the measured response and  $G(f)$  is the correction applied. "Focal", "Global", and "Processed" are defined in Section 12.4.

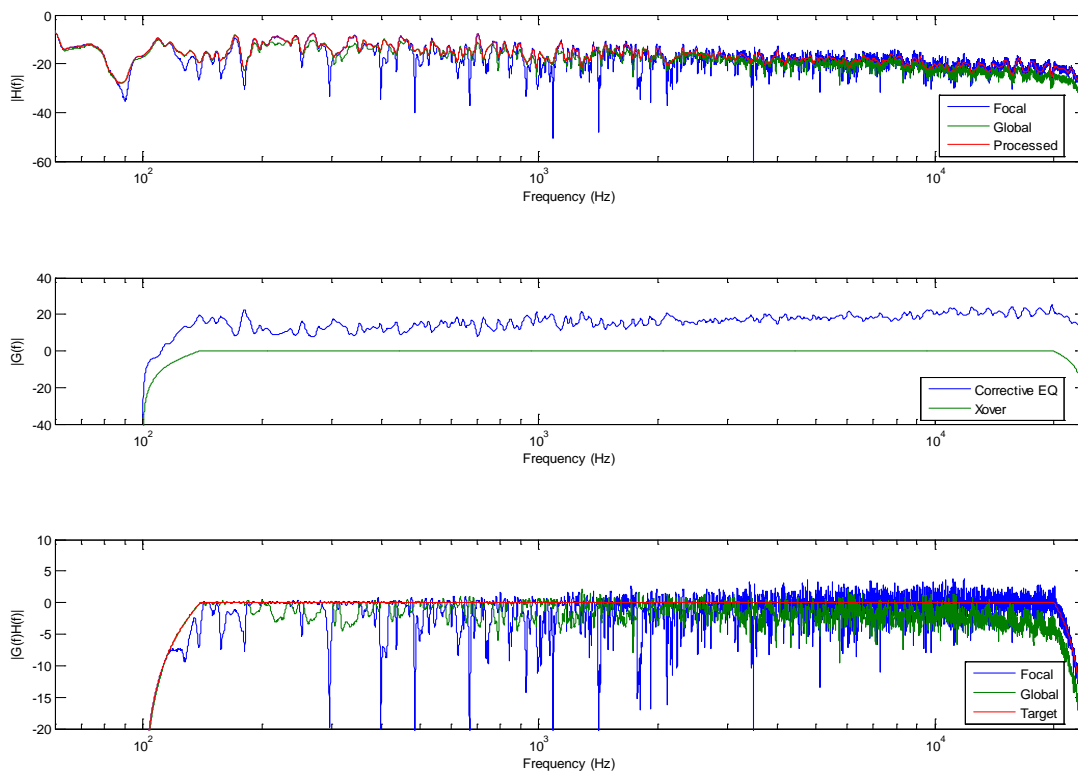


Figure 12.4: Front Loudspeaker Response and Correction

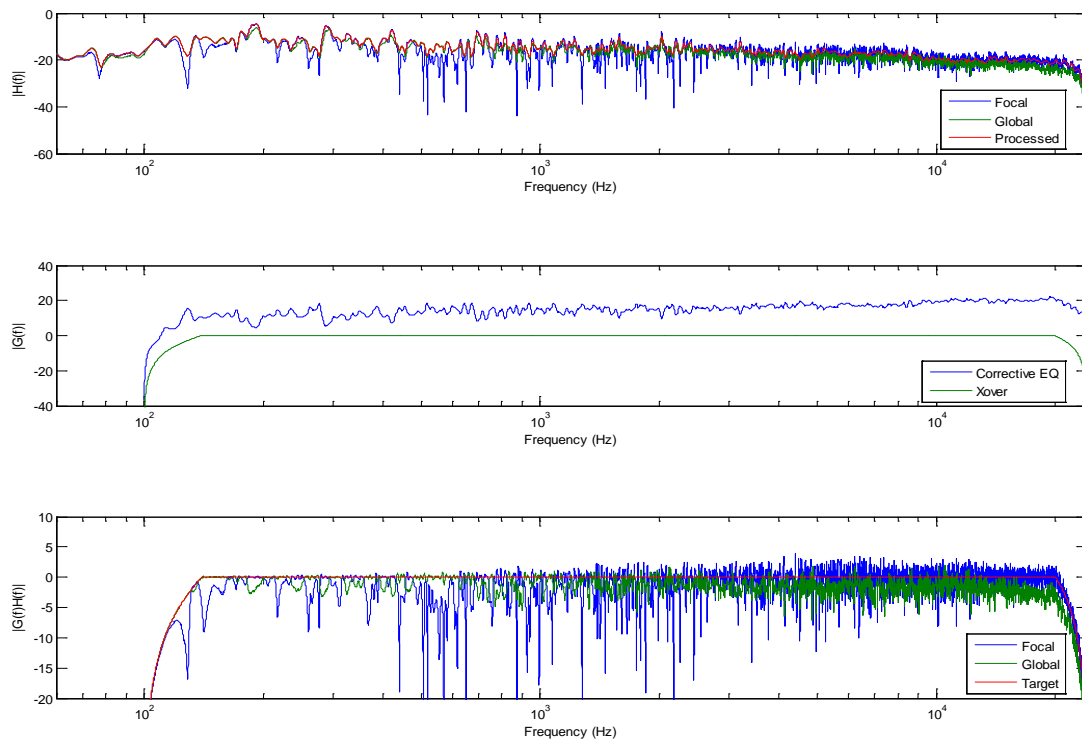


Figure 12.5: Front Left Loudspeaker Response and Correction

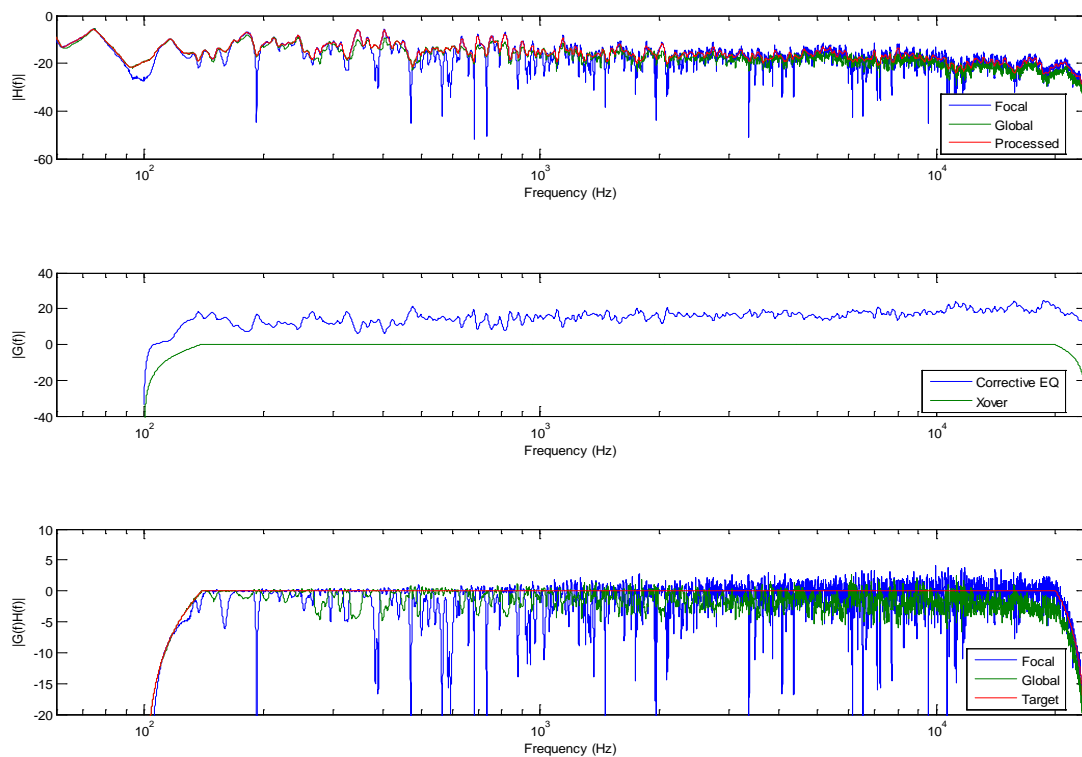


Figure 12.6: Left Loudspeaker Response and Correction



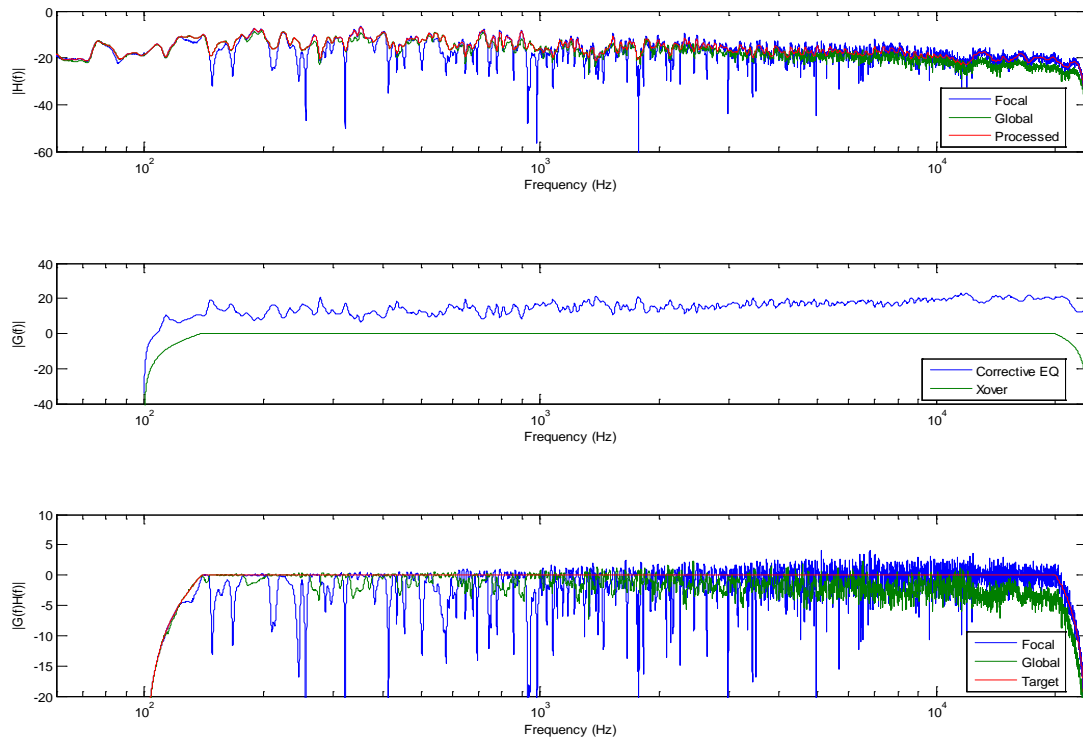


Figure 12.7: Back Left Loudspeaker Response and Correction

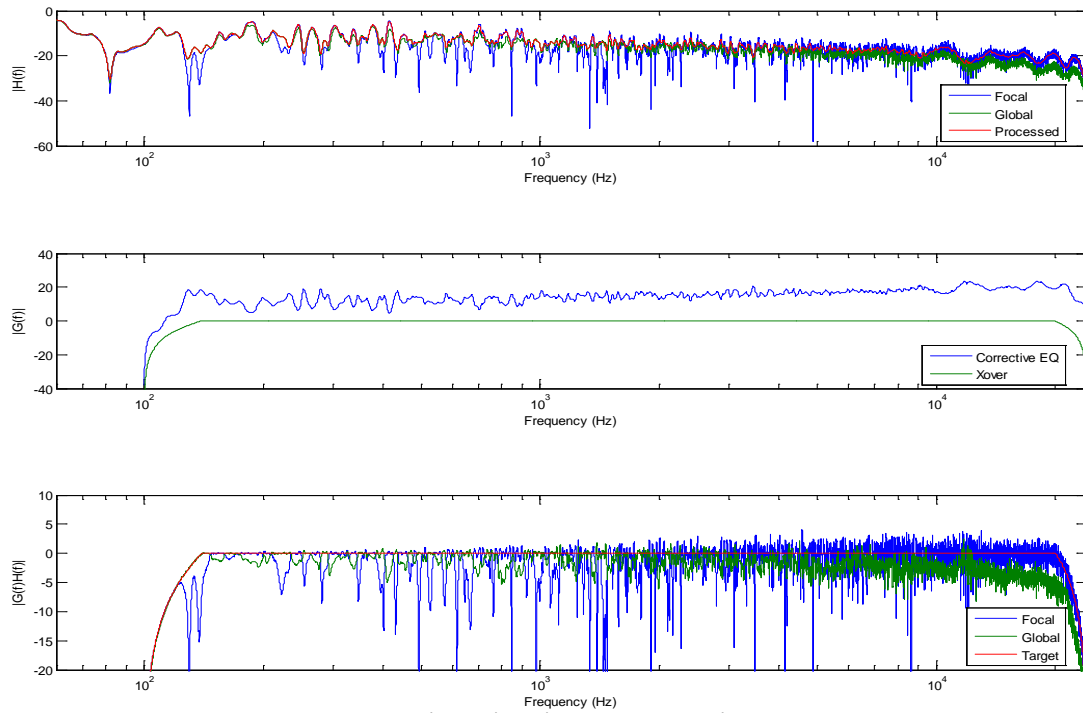


Figure 12.8: Back Loudspeaker Response and Correction

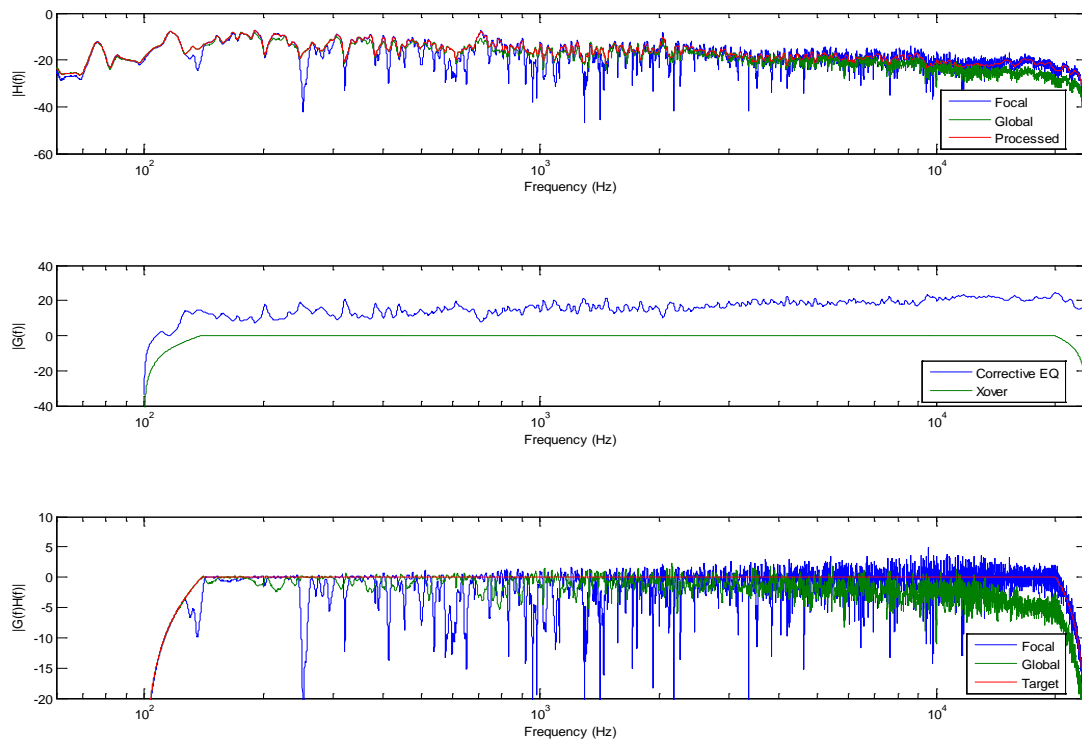


Figure 12.9: Back Right Loudspeaker Response and Correction

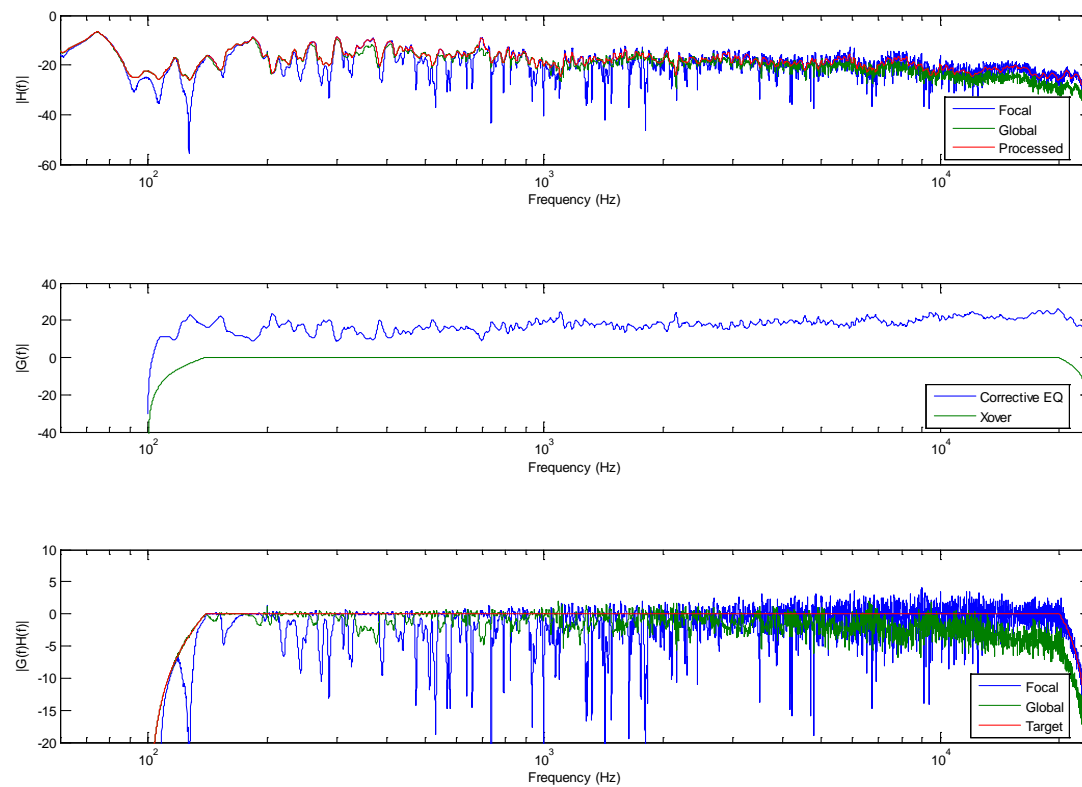


Figure 12.10: Right Loudspeaker Response and Correction

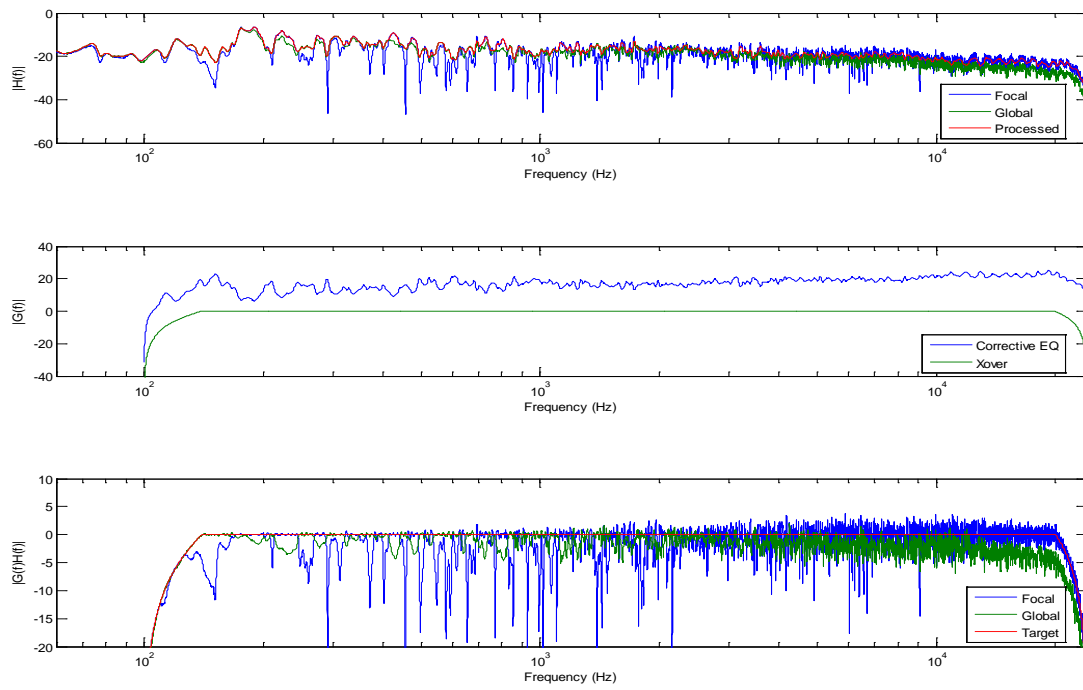


Figure 12.11: Front Right Loudspeaker Response and Correction

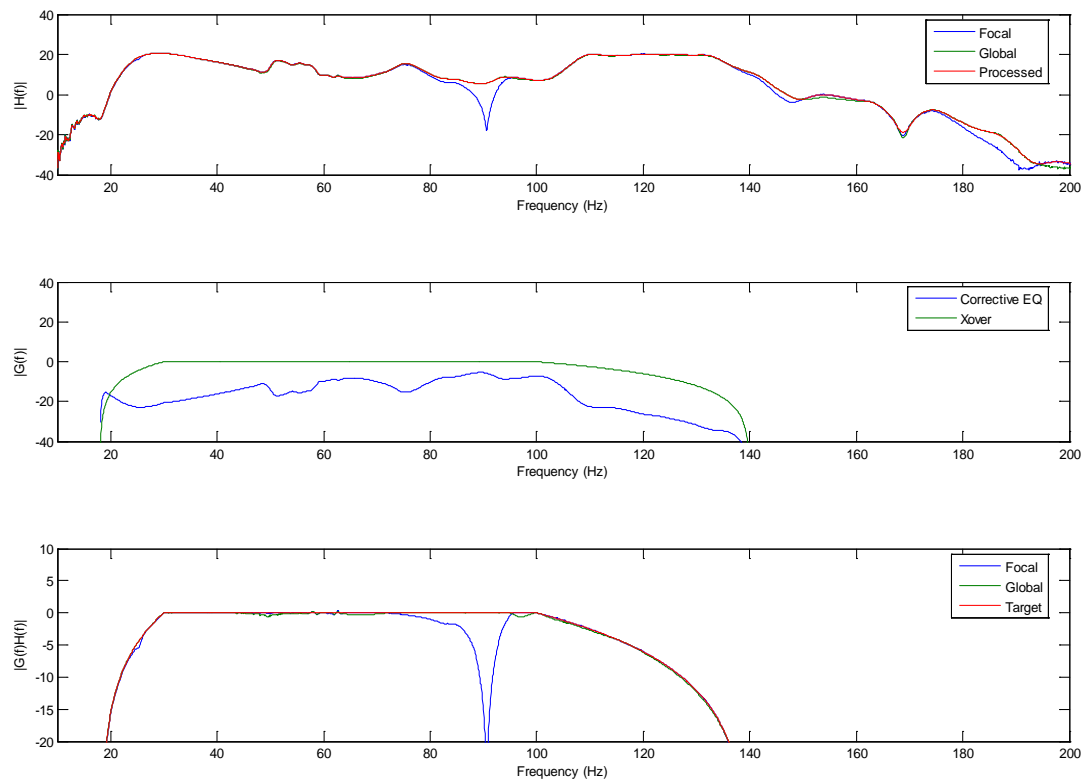


Figure 12.12: Subwoofer Response and Correction

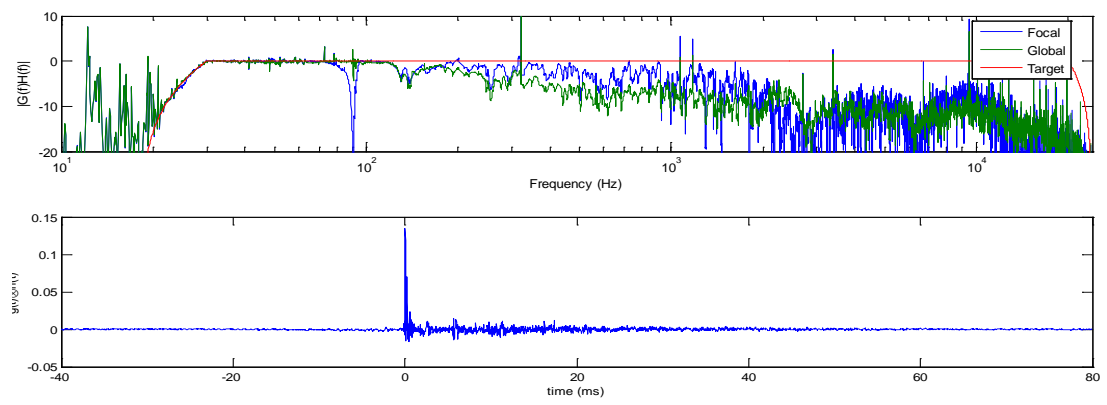


Figure 12.13: Measurement of total corrected loudspeaker system response

## 13 Appendix IV – Wind Turbine Signals

This appendix contains information on how the third octave specifications for the stimuli were generated, how they were made into auralisable sound loops, and finally what the participant actually heard. This structure is consistent with the overview in Section 6.3. It is divided into two because the process differed slightly.

### 13.1 Part A – artificial stimuli

#### Derivation of Part A stimuli specification from measured data

Typical broadband wind turbine spectra are designed by Madsen (2010): Measurements from 45 wind turbines are used. Any distinct tonal peaks are removed from the spectra by smoothing thus achieving the broadband character. This allows later superposing of controlled discrete tones for the listening tests.

Sound propagation effects due to terrain, air absorption, meteorology and distance are included using the Nord2000 model for each turbine. A receiver position of 1.5 m above ground level is assumed at the distance corresponding to 4 times the total turbine height; a distance that is chosen because of the minimum requirements of 4 turbine heights separation between turbines and dwellings in Danish regulations. The terrain type is assumed to be grass, the receiver position to be downwind to include the worst case scenario and the wind speed to be 8 m/s which is a typical value used in Danish regulations. A comparison of the propagated spectra reveals that differences in spectra shape are mostly due to spherical spreading and air absorption which predominantly affect the higher frequencies.

Using the processing described above allows calculating the sound pressure level representing the “free field” (over ground) for each turbine.

The procedure of synthesizing “typical” spectra consists of averaging the 45 free-field turbine spectra and normalising the resulting spectra to 44 dB(A). This approach is straight forward and seems reasonable after attempts of using “typical extreme” spectra to represent particularly noisy and particularly quiet wind turbines have been abandoned when it turned out that the spectra were virtually parallel for the whole frequency range of interest (Madsen, 2010). The variation of the LAeq to the values of 39 dB(A), 44 dB(A) and 49 dB(A) of this spectrum can then either be interpreted as the effect of distance or source strength. The final spectra are displayed in Figure 5.4 a).

#### Generation of auralisable stimuli from third-octave specification

A third octave sound pressure spectrum does not contain adequate information for auralisation so a process is required to interpolate and extrapolate it in a realistic way. A methodology (Moorhouse, 2005) has been developed at Salford for auralising narrowband spectra so this was extended to support third-octave spectra. The process must perform four main operations:

- Compensate for the varying bandwidths of the third-octave bands
- Extrapolate the spectrum to the frequency reproduction limits of the audio hardware
- Interpolate the spectrum so it is free from sharp transitions, check the result against the third-octave band specification and refine if necessary
- Add phase information and perform ambisonic encoding

The A-weighted sound power\*\* in each third-octave band is defined as the integral of the A-weighted power spectral density, so that is the appropriate quantity to design the above algorithm

---

\*\* Because the stimuli were specified in terms of sound pressure level but we want to evaluate the power spectral density the quantities of sound pressure level and sound power will be interchanged

around. However the required auralisation data set is not A-weighted and working in linear scale causes the interpolation scheme to ignore small magnitude parts of the spectra, so instead the unweighted power spectral density in dB was used, and is referred to herein as the “sound power density”. For a third-octave band with bandwidth  $b$  containing sound power  $L_w$ , this can be calculated as:

$$L_{w,D} = L_w - 10 \log_{10}(b) \quad \text{Equation 2}$$

Because third-octave bandwidth is proportional to frequency, the sound power density spectra of all stimuli rolled off above 10 kHz, so a suitable level for the Nyquist frequency was found by fitting a straight line through the 8 kHz and 10 kHz points and interpolating. The sound power below 10 Hz was linearly interpolated down to zero at 0 Hz – this was only necessary to eliminate low-frequency artefacts in the stimuli loops as the reproduction system would not reproduce them anyway.

The loop length determines the number of lines in the FFT spectrum, so these specified and extrapolated points were interpolated at those frequencies by spline fitting. Then the power in each FFT line  $L_w$  was found from the sound power density  $L_{w,D}$  by the inverse of Equation 2:

$$L_w = L_{w,D} + 10 \log_{10}(F_s / N_{fft}) \quad \text{Equation 3}$$

Finally the third-octave sound power in this narrowband spectrum was evaluated and compared to the specification, and if necessary the target values for sound power density were refined and the process repeated until a maximum deviation of 0.1 dB from the specification was achieved.

The algorithm above produced a narrowband spectrum that covered the entire reproducible frequency range with enough resolution to produce a loop of the desired length; the final step was to produce auralisable audio loops in B-format. This process differed slightly between the wind turbines, which were intended to be rendered as a distant point source, and the garden noise, which was intended to arrive from all around. To create a mono audio loop the procedure in Moorhouse (2005) was followed: first the phase of the narrowband spectrum was randomised, then copied to the negative frequencies and scaled to the reference pressure, and finally the inverse FFT was calculated to give a time domain signal in Pascals. For the wind turbine this loop was encoded into the B-format channels according to the following equations:

$$\begin{aligned} w(t) &= p(t) \\ x(t) &= p(t)\sqrt{2} \cos(\theta) \\ y(t) &= p(t)\sqrt{2} \sin(\theta) \end{aligned} \quad \text{Equation 4}$$

where  $\theta$  is the angle the wind turbine was located at anti-clockwise relative to front. These differ slightly from the standard B-format encoding equations because the double-precision floating point storage and arithmetic used in Matlab make the usual scaling of the  $w$ -channel to achieve similar amplitude an unnecessary complication. The factors of  $\sqrt{2}$  arise because the first order spherical harmonics are multiplied by two relative to the zero<sup>th</sup> order, and this gain is split between the encoding and decoding equations (Gerzon, 1985). The garden noise required the  $x$  and  $y$  channels to be decorrelated from the  $w$  channel so additional mono audio loops were created for each of these, and attenuated by approximately 3 dB in line with what is typical for B-format recordings of immersive soundscapes. One interesting feature of the ambisonic decoding process is that it is only the  $w$ -channel (here specified in Pascals) that contributes to pressure at the listener position, and therefore to the microphone measurements in Section 13.3, so in terms of stimuli specification the

---

in this section. Ultimately the reproduction system was calculated in sound pressure level so no error is introduced, and the dB references are approximately equal so this should not introduce undue ambiguity.

levels of the x and y channels for the garden noise are arbitrary. However they clearly affect the extent of envelopment and lateral energy so were empirically set to a suitable level; the optimum level for this is an open research question.

## 13.2 Part B – real stimuli

### Derivation of Wind Turbine propagation attenuation spectra

The attenuation spectra for both the large and the small turbine were derived by Madsen (2010a) in a 4 step procedure:

- 1) The sound propagation based on the chosen sound sample is calculated for the different distances using the NORD2000 model.
- 2) For the distance corresponding to the min. allowed distance (4 x total height) large wind turbine 512 m, small wind turbine 212 m the scaling for the calculated sound levels are calculated so that we will get 44 dB(A) at this distance.
- 3) This linear scaling is used for all other distances as well
- 4) The distances corresponding to resulting levels 39 - 44 - 49 dB is chosen and the corresponding attenuations are calculated.

This approach avoids the problem of a wind turbine being more or less noisy in different wind conditions and that a really quiet wind turbine ends up closer to the receiver than 4 turbine height. However, the different levels cannot be directly related to distance.

### Selection of recordings

The criteria for the selection of sound samples to form the Part B stimuli were outlined in Section 5.3. These focussed on finding cuts where there was little fluctuation in wind speed and no spurious artefacts. The following tables specify when and from which recordings in DELTA's sound library the stimuli were taken.

#### Small Turbine

Calibration tone file	D_15HH ff caltone.wav
Recording file	D09-11_15HH_2min_ff.wav
Offset	00:42
Duration	00:10

#### Large Turbine

Calibration tone file	SWT-2.3 B2_Ch1_RAW_KALIB_020909_155422_01.wav
Recording file	CUT_SWT-2.3 B2_Ch1_RAW_020909_145108_05.wav
Offset	01:00
Duration	00:10

#### Vegetation Noise (coniff)

Calibration tone file	Not available. Gain applied such that $L_{pA} = 48.6\text{dB}$ (to match 8m/s spectra used in Part A)
Recording file	outside260807_160613_40dB cut1.wav

Offset	00:00
Duration	00:12

### Traffic Noise

Recording file	Measurement_02.txt
Offset	02:00
Duration	00:12

## Looping of recordings

The wind turbine sounds contained a rhythmic blade swish sound caused by the blades passing the mast (roughly 1Hz intervals), so the loop length had to be set to a multiple of this interval else the turbine would appear to stutter. This periodicity was identified by inspecting the peaks in the auto-correlation of the signal (Figure 13.1 and Figure 13.2), so the loop was made the maximum multiple of this available from the selected cuts and equal-power cross-fading was performed to ensure seamless looping.

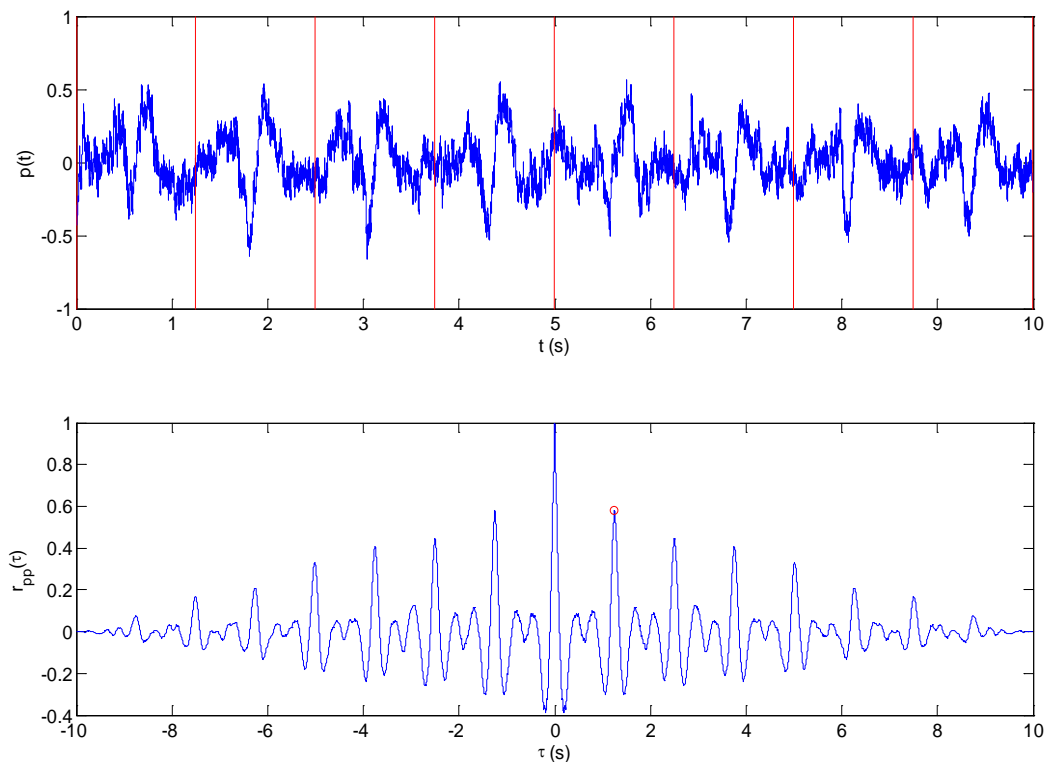


Figure 13.1: large turbine repetition period = 1.25 seconds



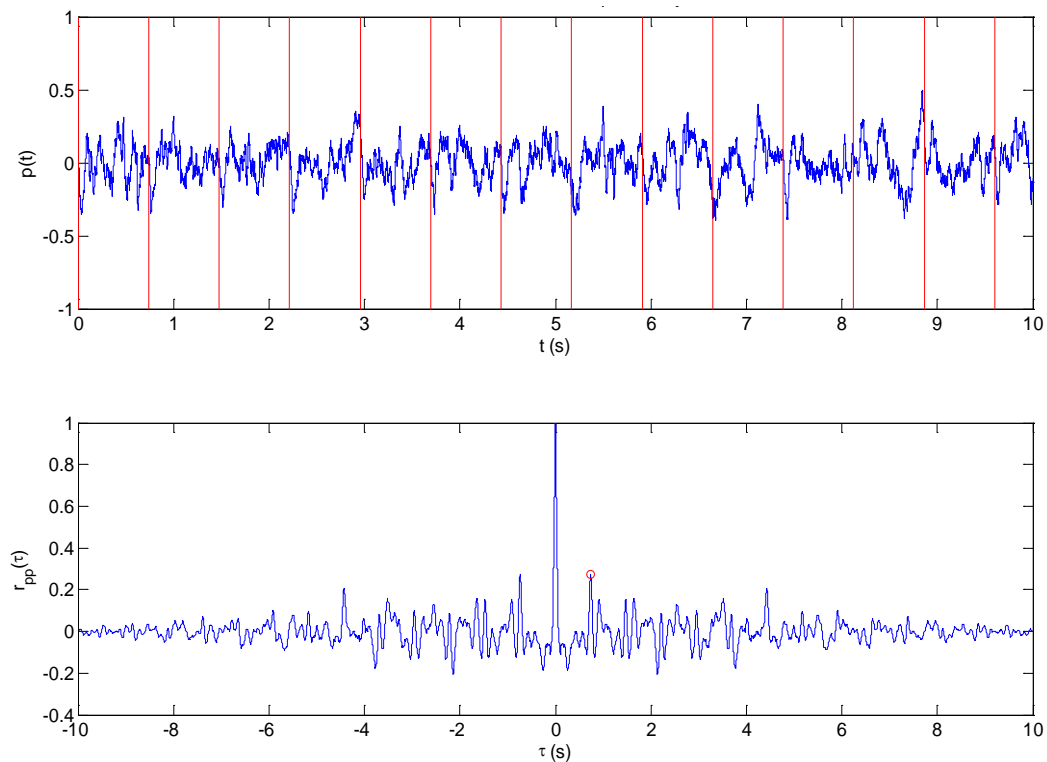


Figure 13.2: small turbine repetition period = 0.74 seconds

## Filtering of recordings

In a process akin to that described in Section 13.1 the recordings had to be modified according to third-octave attenuation spectra, defined by the spreadsheet described in Section 13.1, and then encoded to B-format. Similar challenges were present; particularly that the attenuation data does not have the range and resolution required for auralisation and must be extrapolated and interpolated. However whereas the power summation inherently included in the third-octave sound power spectra of Part A meant third-octave bandwidth need be carefully considered, the attenuation spectra used here are frequency-by-frequency transfer functions and third-octave data is simply samples of that continuous function. This means weighting or bandwidth concerns do not need to be taken into account and they were extrapolated and interpolated directly using the process described for Part A. This created a narrowband transfer function so the recorded loops were easily modified by a process of FFT, multiplication, and inverse FFT.

Ambisonic panning of the mono wind turbine and traffic noise loops was achieved in the same manner as Part A. The garden noise still required the  $x$  and  $y$  channels, but the temporal decorrelation present in the recording meant these could just be the  $w$  channel delayed by small amounts ( $<0.4s$ , necessary to ensure realistic wind variation) and attenuated by approximately 3 dB in line with what is typical for B-format recordings of immersive soundscapes. The result sounded convincing and only the  $w$  channel should contribute significantly to the measured pressure.

## 13.3 Measurements of Stimuli

Each stimulus was recorded using a standard B&K measurement microphone at the listener position and 01dB's Symphonie hardware. The following graphs show a 1/3 octave band analysis, performed in 01dB's dBFA suite, of the masking recordings including the specified and intended spectra.

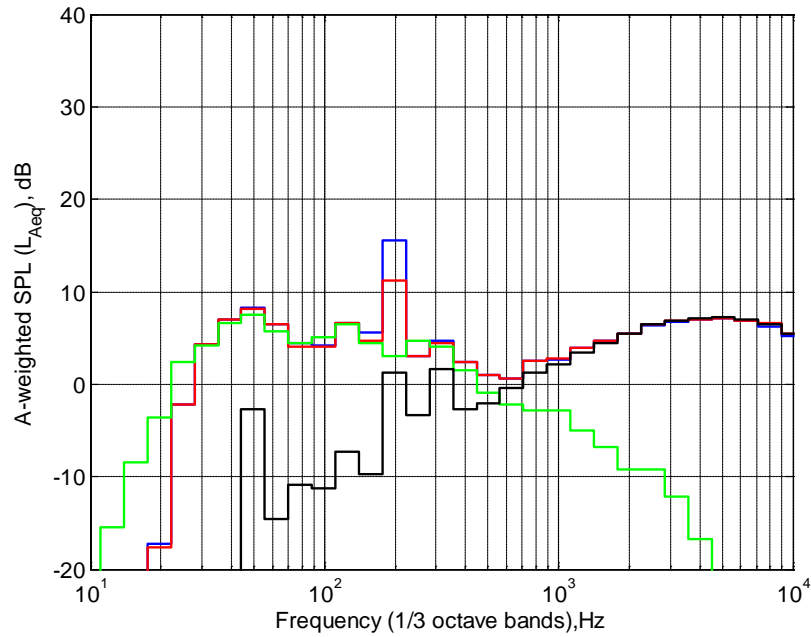


Figure 13.3 Measured and intended Part A Stimuli spectra in comparison – Indoors, Green Intended wind turbine broadband spectrum attenuated from 39 dB(A), Measured stimuli: Black Background noise in room, Blue including 180 Hz tone at 10 dB above wind turbine  $L_{pn}$ , Magenta including 180 Hz tone at 5 dB above wind turbine  $L_{pn}$ . The stimuli are dominated by the intended wind turbine spectrum below 600 Hz and by the background noise above 600 Hz. The domination of the local background would be even stronger in a typical living room environment because that would not be as sound-proof as the laboratory environment. The effect of the reference stimulus can be clearly seen in the respective octave band. The rest of the measured tone spectra are identical to within 3 dB.

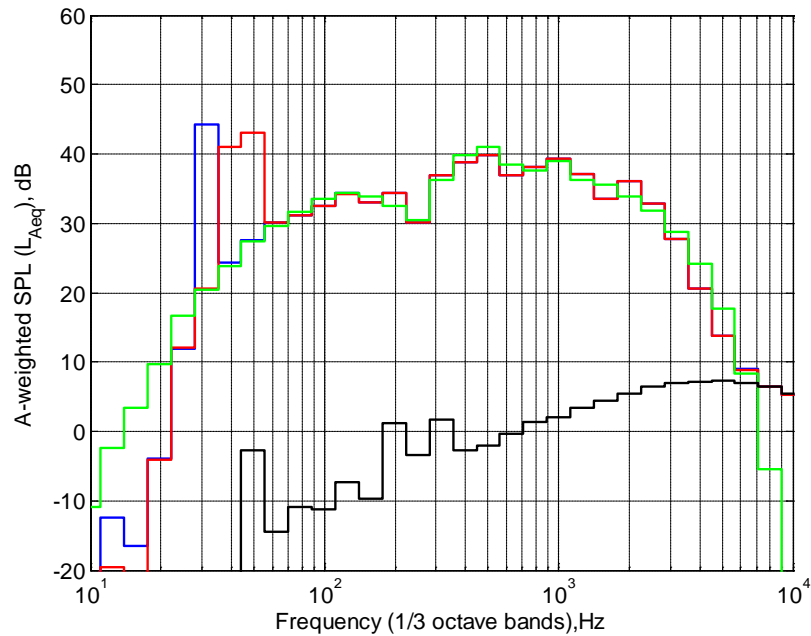


Figure 13.4 Measured and intended Part A Stimuli spectra in comparison – Indoors, Green Intended wind turbine broadband spectrum at 49 dB(A), Measured stimuli: Black Background noise in room, Blue including 32 Hz tone, Magenta including 44 Hz. The stimuli are dominated by the intended wind turbine below 6 kHz and by the room characteristics above 6 kHz. The effect of the reference stimuli can be clearly seen in the respective octave bands. The rest of the measured audible tone spectra are identical to within 1 dB.

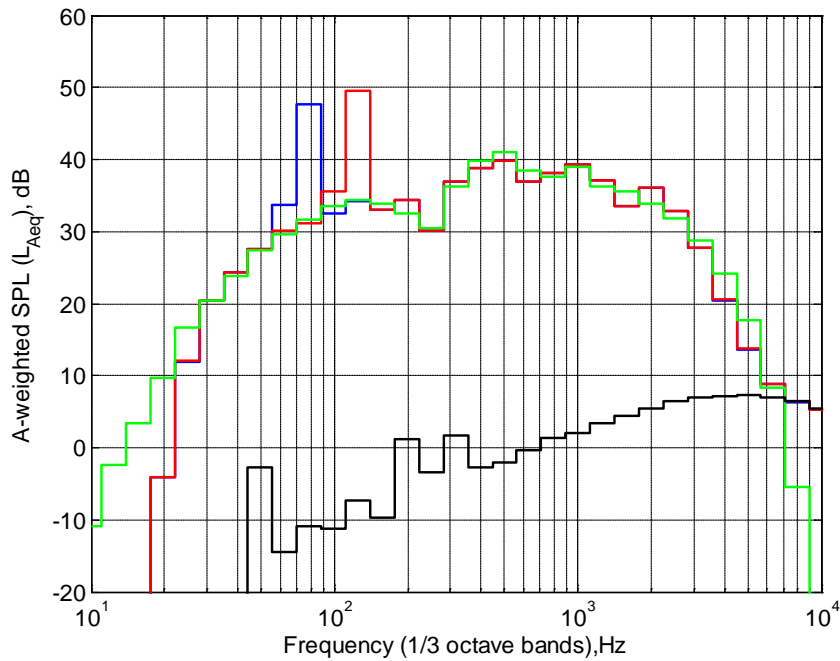


Figure 13.5 Measured and intended Part A Stimuli spectra in comparison – Indoors, Green Intended wind turbine broadband spectrum at 49 dB(A), Measured stimuli: Black Background noise in room, Blue including 72 Hz tone, Magenta including 115 Hz. The stimuli are dominated by the intended wind turbine below 6 kHz and by the room characteristics above 6 kHz. The effect of the reference stimuli can be clearly seen in the respective octave bands. The rest of the measured audible tone spectra are identical to within 1 dB.

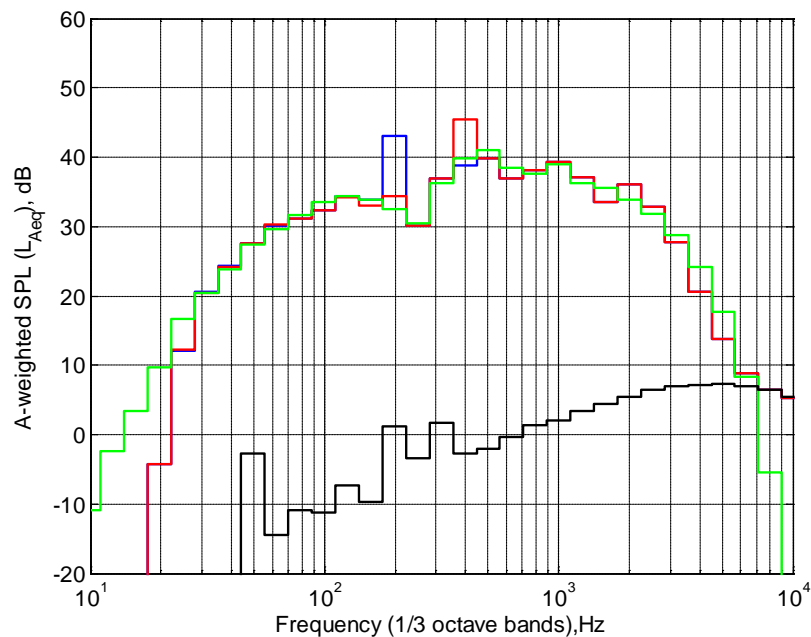


Figure 13.6 Measured and intended Part A Stimuli spectra in comparison – Indoors, Green Intended wind turbine broadband spectrum at 49 dB(A), Measured stimuli: Black Background noise in room, Blue including 180 Hz tone, Magenta including 400 Hz. The stimuli are dominated by the intended wind turbine below 6 kHz and by the room characteristics above 6 kHz. The effect of the reference stimuli can be clearly seen in the respective octave bands. The rest of the measured audible tone spectra are identical to within 1 dB.

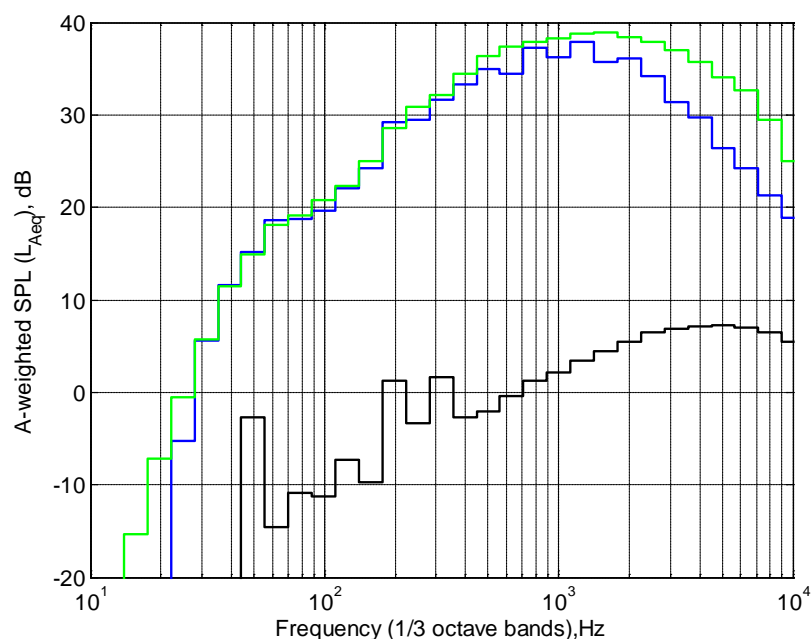


Figure 13.7 Measured and intended Part B garden noise spectra in comparison – Green – Intended Garden Noise broadband spectrum at 44 dB(A), Black – Measured background noise in room, Blue – Measured Garden Noise spectrum. At frequencies above 600 Hz differences between the specified spectrum and the recorded spectrum exceed the intended accuracy of 5 dB due to destructive interference effects in the ambisonic system. This does not significantly affect the masking of Part A tones and it has also been confirmed that both turbine spectra in Part B are affected in a similar way. When listening to the garden noise, the stimulus was perceived as sounding like natural garden noise in spite of the reduced high frequency levels.

Stimulus				Linear		A-weighted					
$L_{Aeq,WTr}$ dB(A)	Tone Freq.Hz	Scenario	Intended lin nom Tone	Lin nom Tone	$L_{eq}$ (lin)	$\Delta L_{ta}$	$L_{pt}$	$L_{pn}$	$L_{pt}-L_{pn}$	$L_{Aeq}$	A (LF)
None	N/A	Room BN		-	40.7	-	-	-	-	17.5	0.3
None	N/A	GN		-	54.1	-	-	-	-	46.2	28.5
39	180	Outdoor	25.1	22.2	48.7	6.7	10.8	6.1	4.7	20.1	15.2
39	180	Outdoor	30.1	27.3	48.8	11.4	15.5	6.1	9.4	21.0	15.6
49	32	Outdoor	83.7	83.8	83.9	11.0	44.6	35.7	8.9	49.9	45.9
49	44	Outdoor	77.2	77.3	77.7	12.0	44.7	34.7	10.0	49.9	46.0
49	72	Outdoor	70.3	70.3	71.6	11.7	46.2	36.5	9.7	50.4	48.0
49	115	Outdoor	64.6	66.0	69.0	13.0	48.7	37.7	11.0	51.6	50.1
49	180	Outdoor	57.2	54.6	66.4	7.9	42.6	36.7	5.9	49.4	40.8
49	400	Outdoor	52.8	49.2	66.2	7.1	44.4	39.5	4.9	49.9	39.9

Table 13.1 Tonal analysis output from noiseLAB software.

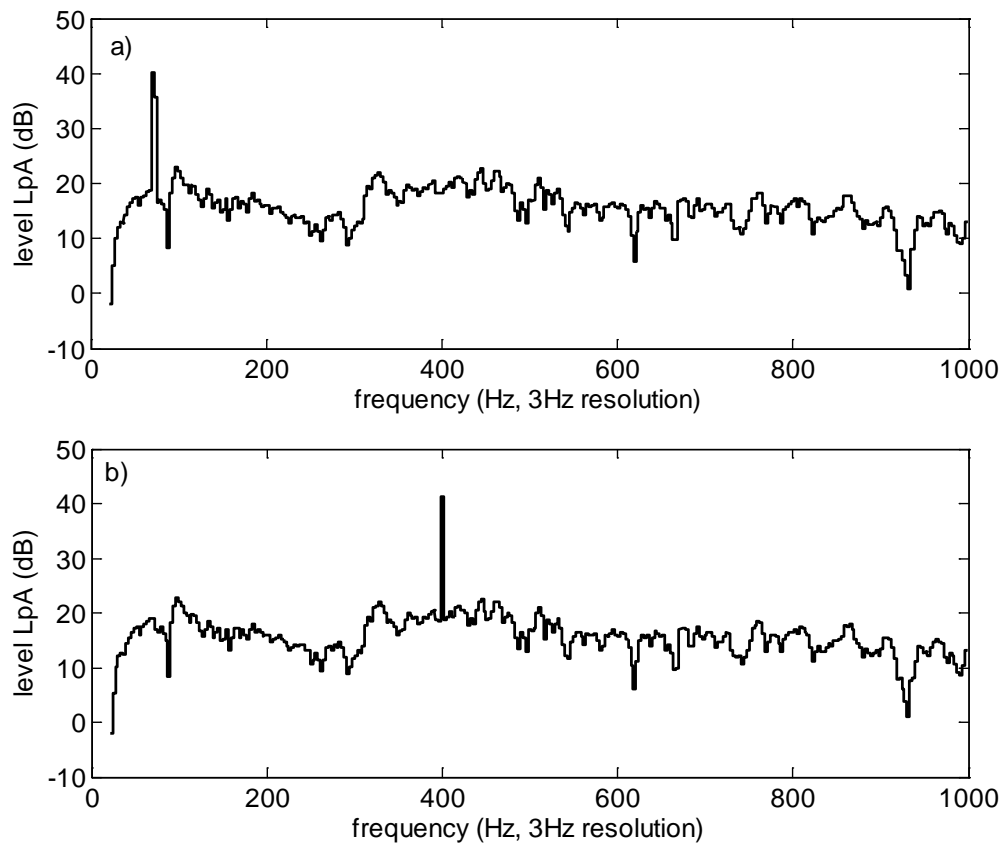


Figure 13.8 Examples of measured Part A outdoor stimuli FFT spectra at broadband level of 44 dB(A), including a) 72Hz and b) 400 Hz tones.

## 14Appendix V: Pilot Tests

Pilot tests were conducted with a limited number of participants to establish the suitability of the audibility threshold and equal annoyance procedures. Additionally the equal annoyance procedure was adapted to ask specifically for Equal Loudness (EL) to give an indication whether the equal annoyance results would be different from EL results.

The number of pilot test participants for each test is summarised in Table 14.1

	Audibility threshold	Equal annoyance	Equal loudness
No masking noise	5	4	3
Indoor	7	6	3
Outdoor	4	6	3

Table 14.1 Number of pilot test participants for Part A Tone audibility threshold, equal annoyance and equal loudness in the three different scenarios as specified in Section 5.5.

### 14.1 Audibility threshold

The pilot tests used a tracking Békésy method to assess 94 detection thresholds: First the participant heard the garden noise played and a user interface on screen (Figure 15.1, Appendix VI) explained the procedure. When the participant pressed start an audible tone would be played in addition to the masking and a user control GUI (Figure 15.2) appeared. This required the participant to press the “Audible” button until the Tone level is reduced to be inaudible. The participant then pressed the “Inaudible” button which caused the Tone level to increase until audible. The process was terminated when steps converged as can be seen in the trial operator control window in Fig. V.3.

### Audibility threshold for Tones without masking

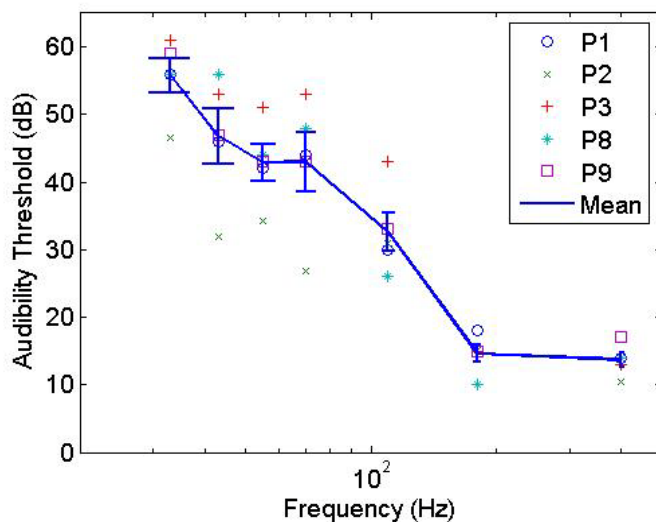


Figure 14.1 Audibility threshold of tones without masking noise

Fig. 6.1 shows the results of the nominal audibility thresholds for five pilot test participants and the average values in the absence of masking noise. It can be seen that the audibility threshold

decreases towards higher frequencies in agreement with standard literature (Fastl & Zwicker, 2007). The exception occurs at 70 Hz where the audibility threshold seemed to be consistently higher than the 53 Hz threshold. This is a test artefact and was traced back to the effect of a room mode.

It is also evident that the variation between participants expressed through the standard error bars are larger from frequencies at and below 100 Hz compared to the 200 Hz and 400 Hz test tones. This could be due to individual differences of perception or to spatial variability of the low frequency tones with head movement. While spatial variability was found to be a problem at 70 Hz the tone levels did not change perceptibly at other frequencies.

## Audibility threshold in indoor and outdoor scenarios

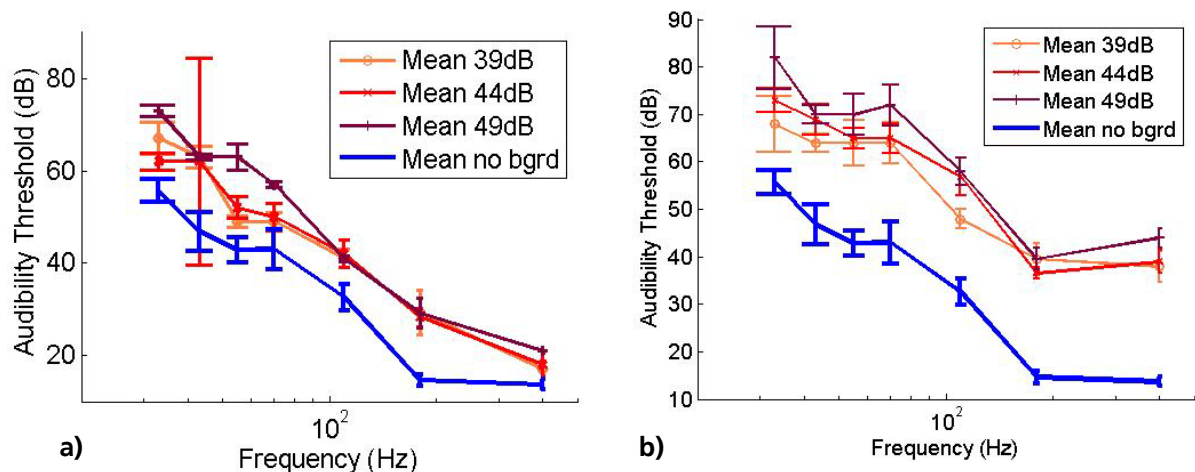


Figure 14.2 Audibility thresholds of tones in wind turbine noise masking, no garden noise. a) indoor, b) outdoor

When comparing audibility thresholds in the indoor and outdoor scenarios in relation to the no masking scenario, Fig. 6.2 shows the same general shape of the audibility curves with consistently higher audibility threshold levels in the outdoor scenario. This is entirely expected as the dB(A) level values in the figure refer to the outdoor levels and the façade attenuation therefore causes the indoor levels to be considerably quieter.

The figure also shows that the three different chosen turbine levels show very similar tone audibility with many of the data points lying within the error bars of the other data points and some audibility threshold levels of the 39 dB(A) wind turbine noise higher than the levels of the 44 dB(A) wind turbine noise. At the same time participants reported to be able to distinguish clearly between the different scenarios. This suggests that at least part of the masking spectrum was audible if not necessarily the low frequency masking bands.

The high standard deviation at 43 Hz (Figure 6.2 a) was due to problem with tone identification when a participant erroneously identified part of the masking sound to be the tone which caused an unrealistically low audibility threshold. Two measures were taken to avoid this type of problem for the final test. Firstly, the room calibration procedure was enhanced (see Appendix IV for details of the final algorithm). Secondly the audibility threshold methodology was changed to Two-Alternative Forced Choice (2AFC) to enable detection of erroneous tone identification.

When comparing these results with those that contain additional masking garden noise audibility threshold does not change significantly.

## 14.2 Equal annoyance of tones

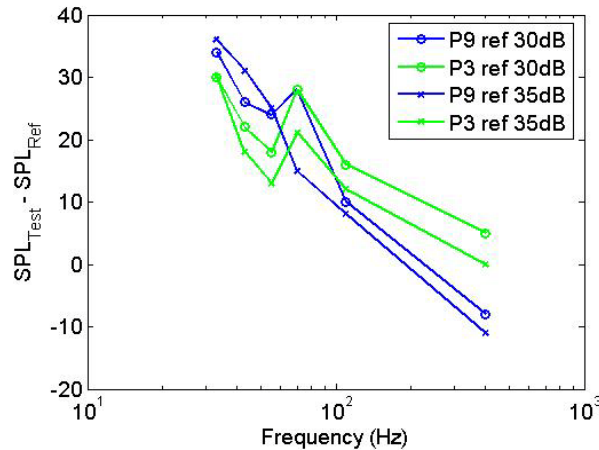


Figure 14.3 Equal annoyance, tones only for two different reference tone levels 5 dB apart and two participants.

When two participants were asked to compare the annoyance of a reference tone to the annoyance of the test tones defined in Section 5.5 P9 adjusted the tone levels of equal annoyance to be higher towards the lower end of the spectrum and lower towards the higher end of the spectrum which is the expected behaviour given the audibility threshold results presented above. P3 chose a higher equal annoyance level for the 400 Hz tone which could suggest a certain habituation effect with respect to the reference tone. The 70 Hz irregularity is the same as in the audibility threshold results. Fig. 6.3 shows the difference between the SPL of the test tone and the SPL of the reference tone. If annoyance scaled linearly with loudness the green and blue lines should therefore fall together. Instead the figure shows parallel lines with the lower reference tone levels at a higher than expected tone level. The curves show a more arbitrary gradient at the lowest 3 frequencies

## Equal annoyance with and without garden noise

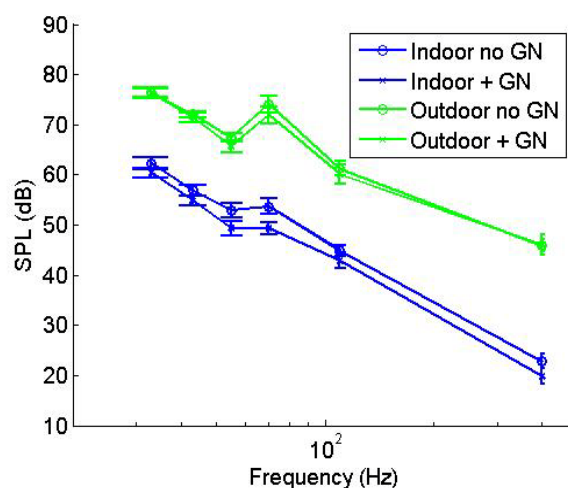


Figure 14.4 Equal annoyance: Effect of using garden noise



Tone levels of equal annoyance at garden noise and no garden noise annoyance (Fig. 6.4) for the indoor and outdoor scenarios are mostly parallel and different especially for the indoors scenario. Statistical significance is to be explored in main listening tests.

## Conclusions of pilot tests

The pilot tests were conducted to explore the appropriateness of the methodology with respect to participant response and expected outcomes. The effect of using a variety of  $L_{Aeq\_WT}$  reference tone levels and types of masking noise was explored.

In general the participants found the tests to be intuitive and workable in spite of prior concerns about the length of the procedure. Results on audibility threshold and equal annoyance look promising and reasonable. The audibility threshold method was changed from Békésy tracking to Two-Alternative Forced Choice (2AFC) to generally avoid false identification of tones in noise. Stimuli loop length was also increased to avoid periodicity being perceptible; this solved the identification problems with the 35 Hz tone.

In response to technical problems during the pilot tests and first round of listening tests, the loudspeaker calibration procedure was enhanced to use multiple microphones to avoid measurement nulls, as compensating for these can create tonal artefacts, and a technical issue with the subwoofers (which particularly affected the 70Hz tone) was corrected.

# 15 Appendix VI: Participant instructions and Graphical User Interfaces

## 15.1 Written instructions to Users

### Instructions

Thank you for participating in this study. It is designed to investigate, firstly, the effect of varying the frequency and intensity of tonal components on annoyance when the component is presented in broadband noise. Secondly, we want to see if the annoyance scales from this match with the annoyance ratings of recorded outdoor noise sources.

Unlike loudness, annoyance is an attribute that is dependent on context. Therefore, we will be asking you to imagine you are hearing the sounds in specific situations. On screen, you will see message asking you to imagine a particular scenario. When we ask you to imagine you are relaxing at your home in the garden or in the living room, try to imagine you are doing the thing you most enjoy doing to relax or wind down.

Please find a comfortable seating position facing the front of the room. Try not to move around too much.

Now, whilst you maintain this 'frame of mind', press 'Begin' and a control panel will appear. The 'Ref tone' button will be initially selected, this plays the reference sound. Listen to it for a while and think about how annoying this sound would be to hear in the imagined scenario.

Now press the 'Test tone' play button, this stops the reference sound and begins the test sound. Again, try to imagine how annoying this sound would be to hear in your garden or living room. Press the plus or minus buttons to change the level of part of the sound until it is of equal annoyance to the reference sound. You can toggle between the two sounds indefinitely until you are satisfied with your response. Press 'Next' when you are ready to move onto the next trial.

Although it essential that you are satisfied that you have changed the 'test' sound to be equally annoying as the 'ref' sound, please do not spend too long or think too hard about your answer as an initial answer is often the most natural.

Please remember that this is not a loudness test. Whilst changing the loudness of sounds can change their annoyance, do not try to make the sounds equally loud. Your goal should always be to compare annoyance, if the sounds were heard at home, in your garden.

## 15.2 Graphical User Interfaces in Pilot Test

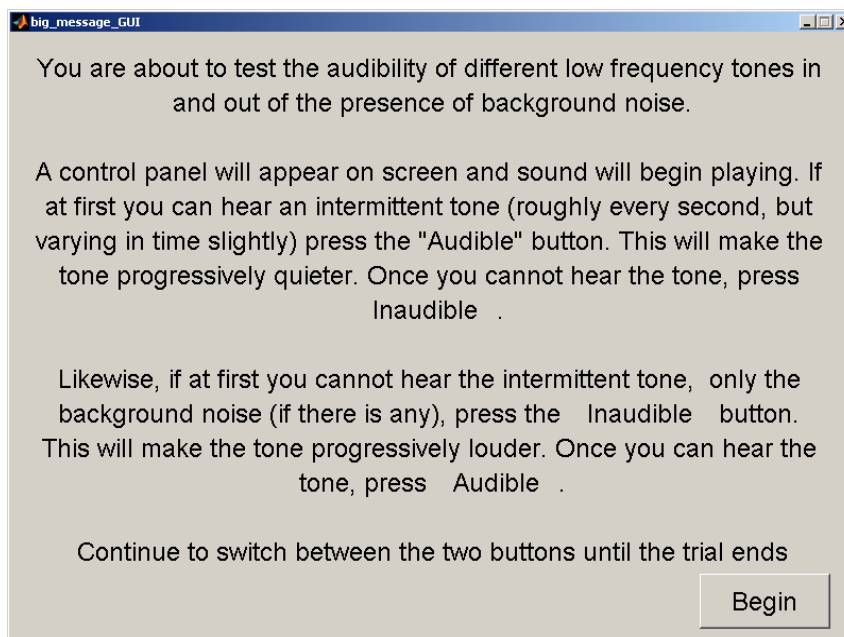


Figure 15.1 Audibility threshold Instructions

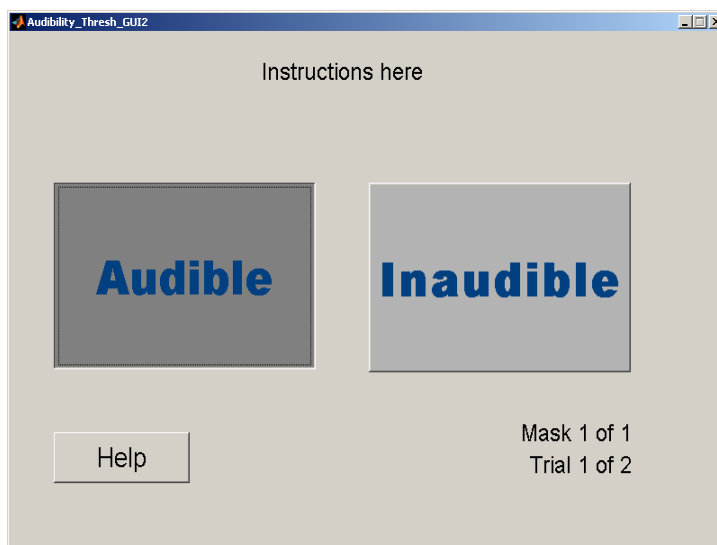


Figure 15.2 Audibility threshold Control

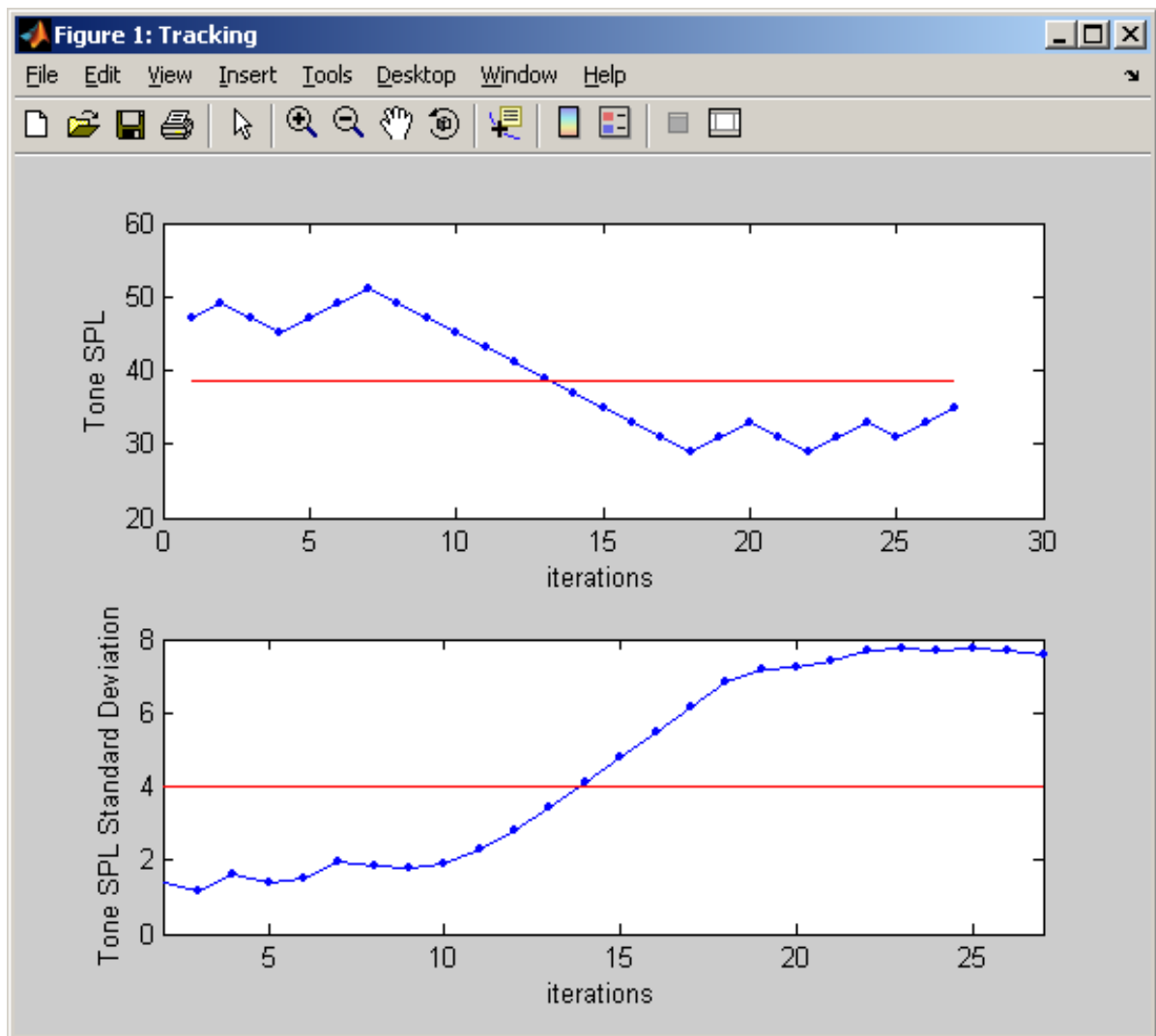


Figure 15.3 Audibility threshold Scientific Monitoring Panel

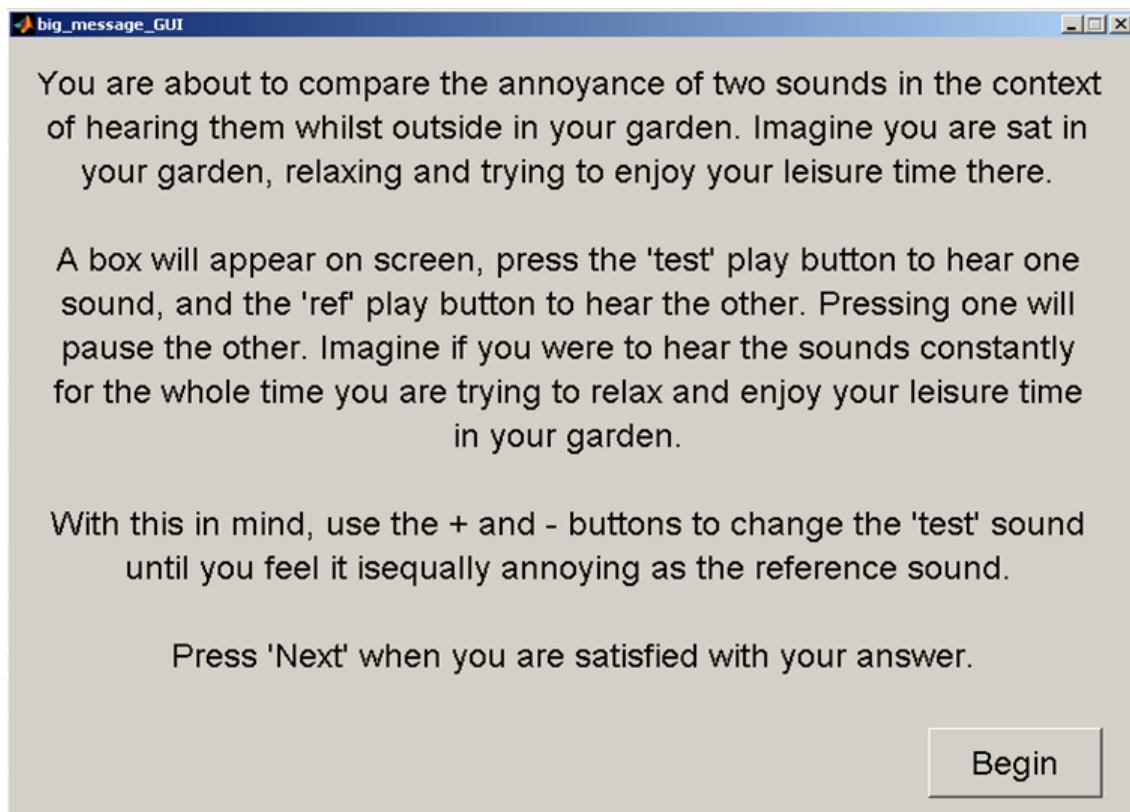


Figure 15.4 Equal annoyance Instructions

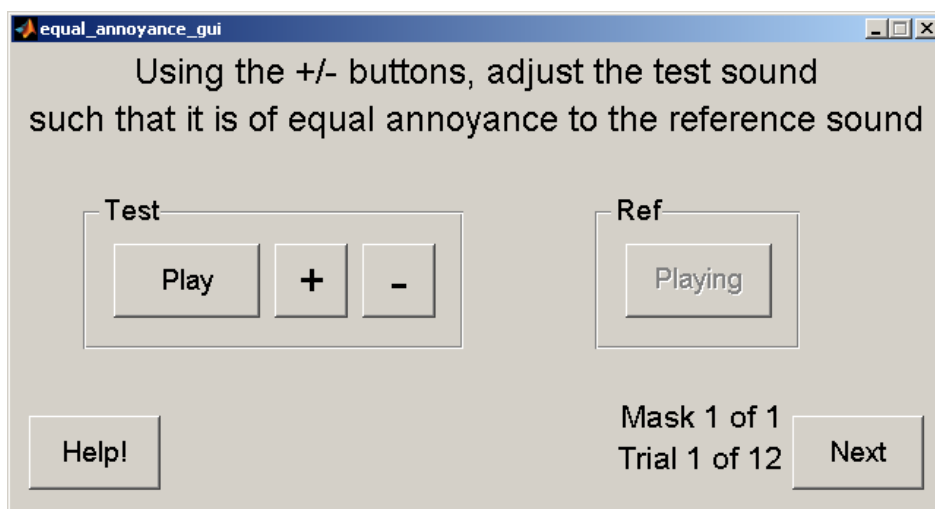


Figure 15.5 Equal annoyance Control

## 15.3 Graphical User Interfaces in Main Test

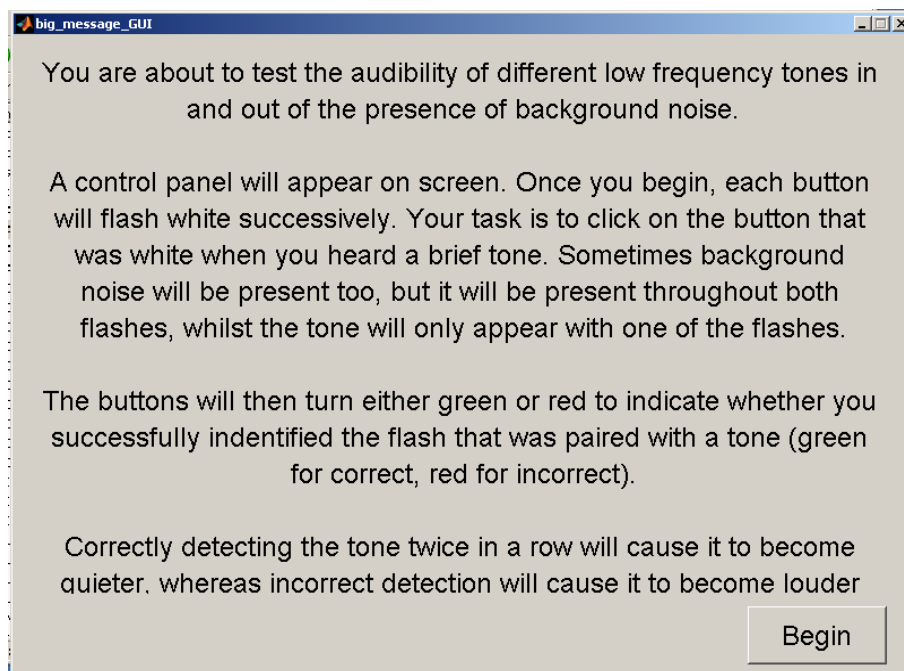


Figure 15.6 Audibility threshold Instructions

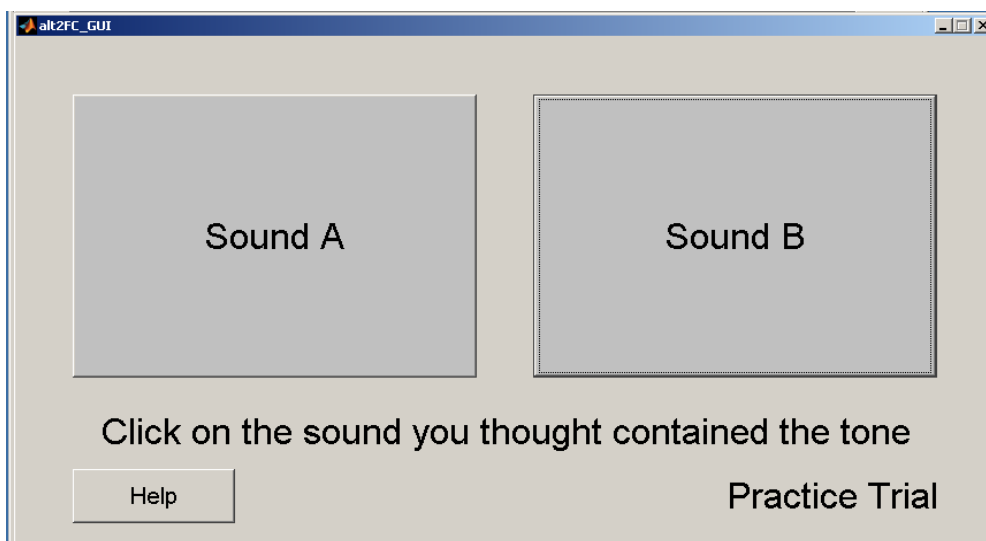


Figure 15.7 Audibility threshold Control

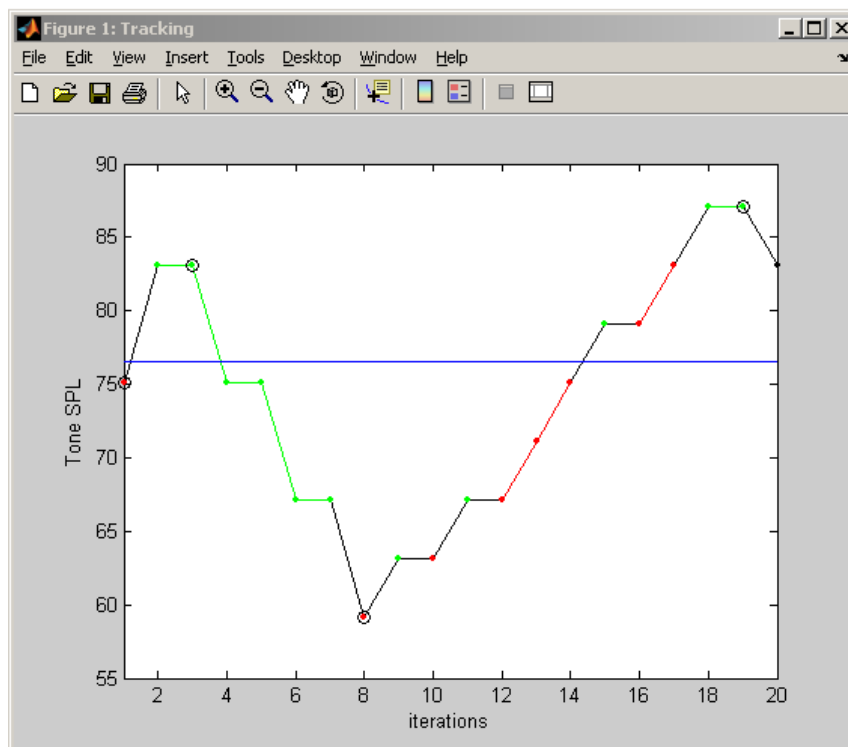


Figure 15.8: Audibility threshold scientific monitoring panel

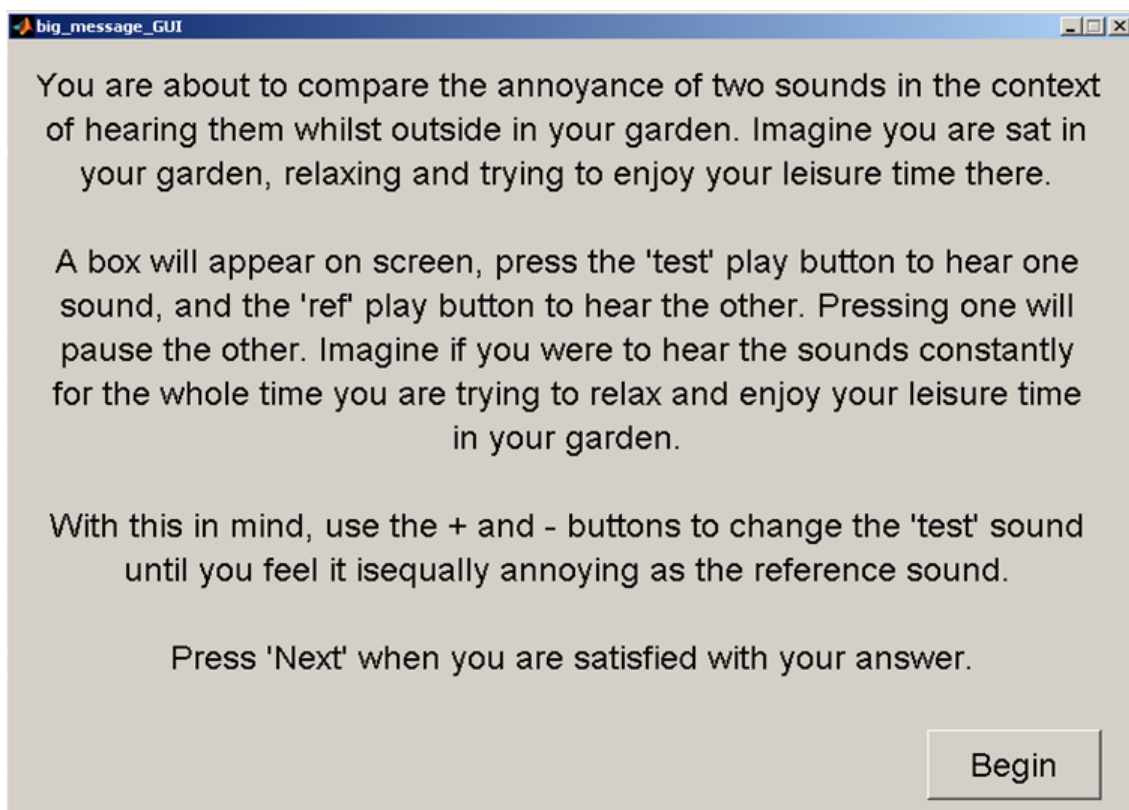


Figure 15.9 Equal annoyance instructions

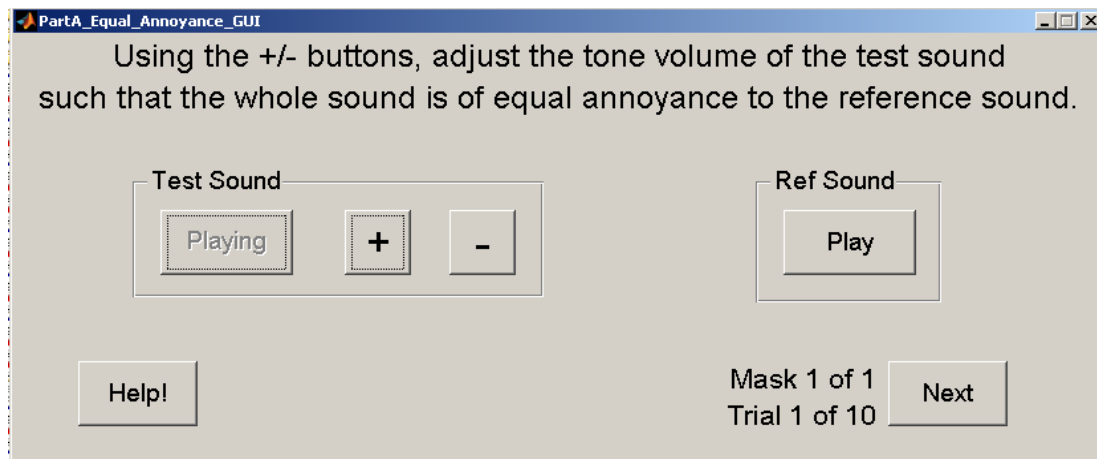


Figure 15.10 Equal annoyance control



## 16 Appendix VII: Statistical validation of audibility threshold

The General Linear Model (GLM) was used with a Type III sum of squares model, as the data was balanced to conduct analysis-of-variance (ANOVA) as calculated by the statistics software package SPSS™. A GLM ANOVA allows testing the significance of any relationship between a number of IVs (with discrete levels of treatment) and a dependent variable (DV) for which it is possible to calculate means and variances. It indicates which IVs cause variation in DV scores, but does not show which levels of treatment of IV are responsible for the effect. For this reason, any significant effects caused by IVs with more than two levels need to be examined further with comparisons between the levels themselves.

Two three-factor repeated measures analyses-of-variance (ANOVA) were used to test if any of the independent variables (IVs) had a significant effect on the audibility threshold.

The IVs were the presence or absence of GN, the A-weighted level of broadband artificial wind turbine noise ( $L_{Aeq,WT}$ ) and the frequency of the target tones. The dependent variable was the level (dB re 20  $\mu$ Pa) of the target tone required to be audible. Table 16.1 shows the factors and their number of levels for the two ANOVAs. It was decided that treating the scenarios: outdoors without GN, outdoors with GN and indoors without GN as 3 levels of the same factor in a single ANOVA would not be appropriate, because the presence or absence of GN tailored to a different research question to changing the simulated listening environment.

Label	IV Factors			DV
ANOVA <sub>gn</sub>	GN	$L_{Aeq,WT}$ , dB(A)	Frequency Hz	Threshold level (dB re 20 $\mu$ Pa)
Levels	No GN With GN	39 44 49	32 44 72 115 180 400	
ANOVA <sub>scen</sub>	Scenario	$L_{Aeq,WT}$ , dB(A)	Frequency	Threshold level (dB re 20 $\mu$ Pa)
Levels	Outdoors Indoors	39 44 49	32 44 72 115 180 400	

Table 16.1 Details of the 2 ANOVAs on AT data

Both ANOVAs found highly significant effects for frequency and  $L_{Aeq,WT}$  as well as GN and Scenario for ANOVA<sub>GN</sub> and ANOVA<sub>scen</sub> respectively (see below for details). For GN, the estimated means suggest that adding GN increases the level needed for a tone to be audible ( $M = 53$  dB) compared to the same masking without GN ( $M = 51$  dB). The Indoors scenario (with façade-attenuated masking noise) lowered the thresholds levels ( $M = 38$  dB) compared to the outdoors without GN scenario.

The estimated means for  $L_{Aeq,WT}$  show similar linear increases in threshold level as  $L_{Aeq,WT}$  increases for both ANOVAs (see below for details). Pairwise comparisons show the differences in masking threshold level to be highly significant ( $p < 0.001$ ).

The ANOVAs suggest the effect of frequency has a strongly linear trend. Figures 6.1 and 6.2 and the estimated means suggest that as the frequency increases, the threshold level diminishes. Pairwise comparisons show that threshold levels for all frequencies tested are strongly significantly different from each with  $p < 0.001$ , with the exception of 180 and 400 Hz in ANOVA<sub>GN</sub>, which differ by  $p < 0.01$ .

Inspection of Figure 6.1 suggests that without GN, the increase in  $L_{AeqWT}$  has a larger effect on threshold levels and that increased  $L_{AeqWT}$  decreases the slope of the frequency trend.

Source	df	F	Sig.
<b>GN</b>	<b>1</b>	<b>14.057</b>	<b>.001</b>
<b><math>L_{AeqWT}</math></b>	<b>2</b>	<b>143.238</b>	<b>.000</b>
<b>Freq</b>	<b>5</b>	<b>498.835</b>	<b>.000</b>
<b>GN * <math>L_{AeqWT}</math></b>	<b>2</b>	<b>10.122</b>	<b>.000</b>

Table 16.2 ANOVA<sub>GN</sub> Tests of Within-Subjects Effects (significant effects are emboldened).

Source	GN	Freq	df	F	Sig.
<b>GN</b>	<b>Linear</b>	<b><math>L_{AeqWT}</math> * Freq</b>	<b>1</b>	<b>14.057</b>	<b>.001</b>
<b><math>L_{AeqWT}</math></b>	<b>GN * <math>L_{AeqWT}</math></b>	<b>Freq Linear</b>	<b>1</b>	<b>281.539</b>	<b>.000</b>
<b>Freq</b>	<b>GN * <math>L_{AeqWT}</math></b>	<b>Freq <math>L_{AeqWT}</math></b>	<b>1</b>	<b>2190.536</b>	<b>.000</b>
<b>GN * <math>L_{AeqWT}</math></b>	<b>Linear</b>	<b>Linear</b>	<b>1</b>	<b>14.497</b>	<b>.001</b>

Table 16.3 ANOVA<sub>GN</sub> Tests of Within-Subjects Contrasts (significant effects are emboldened).

### 3.3 ANOVA<sub>Scen</sub>

Source	df	F	Sig.
<b>Scen</b>	<b>1</b>	<b>263.902</b>	<b>.000</b>
<b><math>L_{AeqWT}</math></b>	<b>2</b>	<b>203.518</b>	<b>.000</b>
<b>Freq</b>	<b>5</b>	<b>607.840</b>	<b>.000</b>

Table 16.4 ANOVA<sub>Scen</sub> Tests of Within-Subjects Effects (significant effects are emboldened).

Source	Scen	$L_{AeqWT}$	Freq	df	F	Sig.
<b>Scen</b>	<b>Linear</b>	<b><math>L_{AeqWT}</math> * Scen</b>	<b>Freq</b>	<b>1</b>	<b>263.90</b>	<b>.000</b>
<b><math>L_{AeqWT}</math></b>	<b>Scen * Linear</b>	<b>Scen Freq</b>	<b>Scen Freq</b>	<b>1</b>	<b>319.40</b>	<b>.000</b>
<b>Freq</b>	<b>Scen * <math>L_{AeqWT}</math></b>	<b>Freq</b>	<b>Linear</b>	<b>1</b>	<b>2023.0</b>	<b>.000</b>

Table 16.5 ANOVA<sub>Scen</sub> Tests of Within-Subjects Contrasts (significant effects are emboldened).

# 17 Appendix VIII Statistical evaluation of equal annoyance results

Again, the ANOVAs were split into ANOVA<sub>GN</sub> and ANOVA<sub>Scen</sub>, the breakdown of IV factors and the DVs are given in Table 17.1.

It is worth noting that the 'No scenario' conditions were omitted from the ANOVAs, as they would not occur in reality or fit into the matrix for the above ANOVAs. Furthermore, the reference tone was omitted from the frequency factor as its relative sensation level would always equal 0.

Label	IV Factors				DV
ANOVA <sub>GN</sub>	GN	$L_{AeqWT}$ , dB(A)	$L_R$ , dB	Frequency, Hz	RSL value
Levels	No GN With GN	39 44 49	5 10	32 44 72 115 400	
ANOVA <sub>Scen</sub>	Scenario	$L_{AeqWT}$ , dB(A)	$L_R$ , dB	Frequency, Hz	RSL value
Levels	Outdoors Indoors	39 44 49	5 10	32 44 72 115 400	

Table 17.1 Details of the 2 ANOVAs on relative sensation levels

## 17.1 Outdoors, omission versus inclusion of garden noise

The ANOVA<sub>GN</sub> demonstrated a strong, significant effect of  $L_R$  and no other significant effects (see below for more detail). It is notable that the equal annoyance levels for the high reference levels (+10 dB) did not need to be 5 dB higher above audibility threshold ( $M = 4.2$  dB) than the low reference levels (for +5 dB,  $M = 1.6$  dB) to be equally annoying. In other words the two reference contours are closer together than the lower reference contour and the audibility threshold contour in Figure 6.1. This might be due to the fact that audibility threshold is according to ISO 1996-2 2 - 6 dB below the 0dB level of the tone compared to  $L_{pn}$ . It is also likely that annoyance does not scale linearly with level.

## 17.2 Outdoors in comparison to Indoors

The ANOVA<sub>Scen</sub> showed the same significant main effect of  $L_R$  as ANOVA<sub>GN</sub>. With increasing reference tone levels the test tones had to be higher in level to be equally annoying. However, when  $L_R$  was 10 dB the mean tone level for EA and not 0 dB but 4.8 dB below the reference tone level, and only 2.1 dB below when  $L_R$  was 5 dB based on estimated means. This was the case regardless of tone frequency or masking noise type or level (Figs. 6.1, 6.2 and 6.3) and is similar to the effect seen in equal loudness contours with different level 1 kHz reference tones (Fastl & Zwicker, 2007).

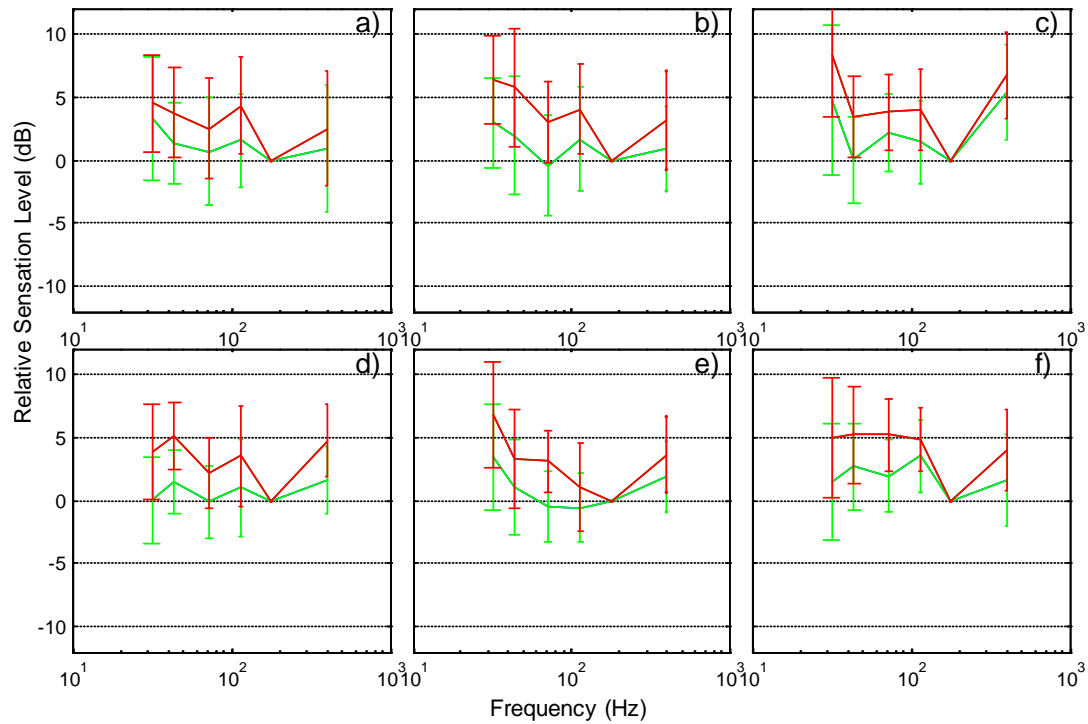


Figure 17.1 Outdoor scenario: Relative sensation level for equal annoyance with various masking noises at tone prominence levels of 5 dB (green) and 10 dB (red). a)  $L_{Aeq,WT}$  39 dB(A), no garden noise, b)  $L_{Aeq,WT}$  44 dB(A), no garden noise, c)  $L_{Aeq,WT}$  49 dB(A), no garden noise, d)  $L_{Aeq,WT}$  39 dB(A), with garden noise, e)  $L_{Aeq,WT}$  44 dB(A), with garden noise and f)  $L_{Aeq,WT}$  49 dB(A), with garden noise. Error bars show 95% confidence intervals.

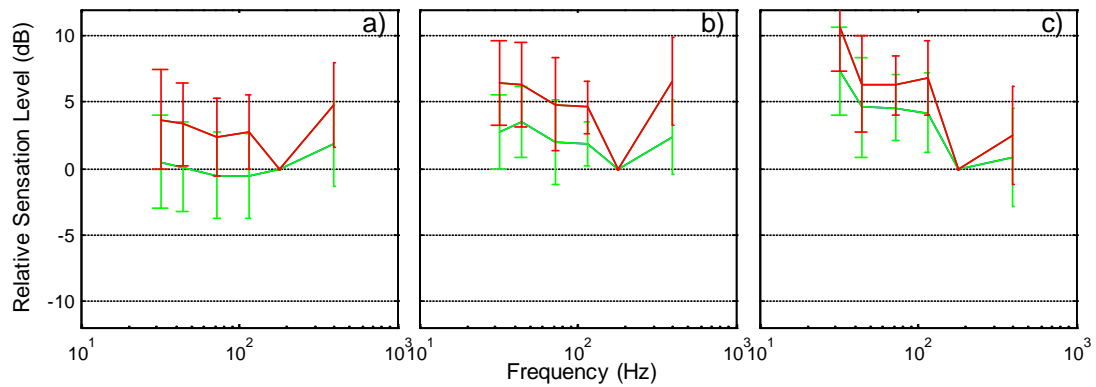


Figure 17.2 Indoor scenario: Relative sensation level for equal annoyance in the outdoor scenarios with various masking noises at tone prominence levels of 5 dB (green) and 10 dB (red). a)  $L_{Aeq,WT}$  39 dB(A), b)  $L_{Aeq,WT}$  44 dB(A), c)  $L_{Aeq,WT}$  49 dB(A), Error bars show 95% confidence intervals.

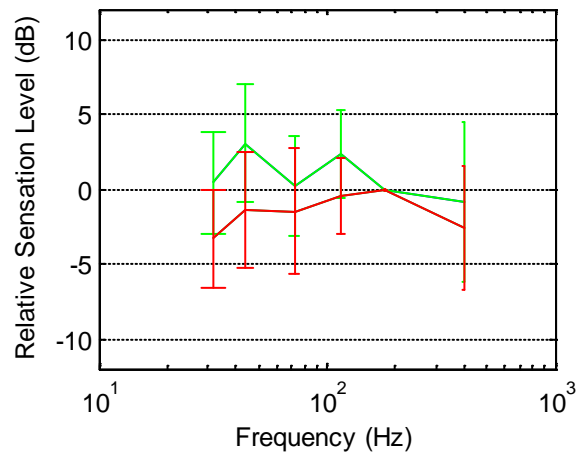


Figure 17.3 Relative sensation level for equal annoyance without masking noise at tone prominence levels of 5 dB (green) and 10 dB (red). Error bars show 95% confidence intervals.

Source	df	F	Sig.
GN	1	.118	.736
L <sub>AeqWT</sub>	2	.594	.557
<b>L<sub>R</sub></b>	<b>1</b>	<b>52.934</b>	<b>.000</b>
Freq	2.59	1.535	.221

Table 17.2 ANOVA<sub>GN</sub> Tests of Within-Subjects Effects (significant effects are emboldened).

Source	df	F	Sig.
Scen	1	.519	.48
L <sub>AeqWT</sub>	2	2.764	.076
<b>L<sub>R</sub></b>	<b>1</b>	<b>46.671</b>	<b>.000</b>
Freq	4	2.18	0.79

Table 17.3 ANOVA<sub>Scen</sub> Tests of Within-Subjects Effects (significant effects are emboldened).

# 18Appendix IX: Statistical evaluation of masking comparison

## 18.1 Difference in tone SPL

Two ANOVAs were conducted on the changes to 180 Hz pure tone level to achieve equal annoyance to the outdoor reference sounds with 3 factors. Table 18.1 details the factors, levels and DVs for each ANOVA. Both ANOVAs showed the effect of  $L_R$  to be highly significant, (see details below).

Label	IV Factors			DV
ANOVA <sub>GN</sub>	GN	$L_{AeqWT}$ dB(A)	$L_R$ dB	180 Hz pure tone relative to reference tone level (dB)
Levels	No GN With GN	39 44 49	5 10	
ANOVA <sub>Scen</sub>	Scenario	$L_{AeqWT}$	$L_R$ dB	
Levels	Outdoors Indoors	39 44 49	5 10	180 Hz pure tone relative to reference tone level (dB)

Table 18.1 Table of the details of the 2 ANOVAs on masking comparisons data.

Another effect was produced by  $L_{AeqWT}$ . Figure 7.6 suggests that the 180 Hz tone needs to be at higher levels to be equally annoying when the masking included lower  $L_{AeqWT}$  values. Again, this was significant for both the ANOVAs. Paired comparisons showed that for both ANOVAs this effect was only significant ( $p < 0.05$ ) between 39 and 49 dB (A).

Source	df	F	Sig.
GN	1	.762	.394
<b>LWT</b>	<b>2</b>	<b>4.560</b>	<b>.017</b>
<b>LR</b>	<b>1</b>	<b>29.390</b>	<b>.000</b>

Table 18.1 ANOVA<sub>GN</sub> Tests of Within-Subjects Effects (significant effects are emboldened).

Source	df	F	Sig.
<b>Scenario</b>	<b>1</b>	<b>7.958</b>	<b>.011</b>
<b>LWT</b>	<b>2</b>	<b>5.029</b>	<b>.012</b>
<b>LR</b>	<b>1</b>	<b>35.802</b>	<b>.000</b>

Table 18.2 ANOVA<sub>Scen</sub> Tests of Within-Subjects Effects (significant effects are emboldened).

(I) $L_{AeqWT}$	(J) $L_{AeqWT}$	Mean Difference (I-J)	Std. Error	Sig.
39 dB(A)	44 dB(A)	1.740	.687	.061
<b>39 dB(A)</b>	<b>49 dB(A)</b>	<b>2.692</b>	<b>1.003</b>	<b>.044</b>
44 dB(A)	49 dB(A)	.952	.986	1.000

Table 18.3 ANOVA<sub>GN</sub> Pairwise comparison of  $L_{AeqWT}$  levels (significant effects are emboldened).

(I) $L_{AeqWT}$	(J) $L_{AeqWT}$	Mean Difference (I-J)	Std. Error	Sig.
39 dB(A)	44 dB(A)	.596	.540	.851
<b>39 dB(A)</b>	<b>49 dB(A)</b>	<b>2.115</b>	<b>.713</b>	<b>.024</b>
44 dB(A)	49 dB(A)	1.519	.787	.206

Table 18.4 ANOVA<sub>Scen</sub> Pairwise comparison of  $L_{AeqWT}$  levels (significant effects are emboldened).

## 18.2 Difference in $L_{Aeq}$

Two ANOVAs were conducted on the changes to 180 Hz pure A-weighted level compared to the A-weighted level of the outdoor reference sounds with the same factors and DVs used in Table 18.1 Both ANOVAs showed the effect of  $L_r$  to be significant, (see details below). In comparing the indoors and outdoors scenarios all three factors show a significant effect.

Source	df	F	Sig.
GN	1	1.252	.277
LWT	2	2.001	.149
<b>LR</b>	<b>1</b>	<b>5.662</b>	<b>.028</b>

Table 18.5 ANOVA<sub>GN</sub> Tests of Within-Subjects Effects (significant effects are emboldened).

Source	df	F	Sig.
<b>Scen</b>	<b>1</b>	<b>12.899</b>	<b>.002</b>
<b>LWT</b>	<b>2</b>	<b>3.202</b>	<b>.052</b>
<b>LR</b>	<b>1</b>	<b>5.662</b>	<b>.002</b>

Table 18.6 ANOVA<sub>Scen</sub> Tests of Within-Subjects Effects (significant effects are emboldened).

## 19 Appendix X: Statistical evaluation Part B results

Two three-way GLM repeated measures ANOVAs were performed similar to Part A, the details of which are given in Table 19.1.

Label	IV Factors			DV
ANOVA <sub>gn</sub>	GN	WT	$L_{AeqWT}$ , dB (A)	Level, dB(A) that TN had to be for EA
Levels	No GN With GN	Large Small	39 44 49	
ANOVA <sub>scen</sub>	Scenario	WT	$L_{AeqWT}$ , dB (A)	
Levels	Outdoors Indoors	Large Small	39 44 49	Level, dB(A) that TN had to be for EA

Table 19.1 Details of the 2 ANOVAs on TN changes for EA to recorded WTN.

The marginal estimated means suggest that with increasing  $L_{AeqWT}$ , TN level has to increase to be equally annoying. This effect of  $L_{AeqWT}$  can be seen in Figures 7.1 and 7.2 and was found to be significant in both ANOVAs. Pairwise comparisons of the estimated means in both ANOVAs between the three levels of  $L_{AeqWT}$  found that the difference was only significant between 44 and 49 dB(A) to a significance of  $p < 0.001$  in ANOVA<sub>scen</sub>, but significant between each level in ANOVA<sub>gn</sub>.

It can be seen in the marginal estimated means that, for equal wind turbine levels, the large wind turbine usually had lower equal annoyance levels ( $M = 47.6$ ) than the small wind turbine ( $M = 49.1$ ) both indoors and outdoors when no garden noise was present. This was a significant main effect. However, this was only true without garden noise. Garden noise had a significant effect of raising the traffic noise level needed for equal annoyance. The mean was 48.23 dB(A) without garden noise and 50.44 dB(A) with garden noise.

Source	df	F	Sig.
<b>GN</b>	<b>1</b>	<b>6.287</b>	<b>.022</b>
WT	1	3.016	.100
<b><math>L_{AeqWT}</math></b>	<b>2</b>	<b>35.305</b>	<b>.000</b>

Table 19.3 ANOVA<sub>gn</sub> Tests of Within-Subjects Effects (significant effects are emboldened).

(I) $L_{AeqWT}$	(J) $L_{AeqWT}$	Mean Difference (I-J)	Std. Error	Sig.
<b>39 dB(A)</b>	<b>44 dB(A)</b>	<b>-2.422</b>	<b>.595</b>	<b>.002</b>
<b>39 dB(A)</b>	<b>49 dB(A)</b>	<b>-5.727</b>	<b>.842</b>	<b>.000</b>
<b>44 dB(A)</b>	<b>49 dB(A)</b>	<b>-3.305</b>	<b>.584</b>	<b>.000</b>

Table 19.2 ANOVA<sub>gn</sub> Pairwise comparison of  $L_{AeqWT}$  levels (significant effects are emboldened).

Source	df	F	Sig.
Scen	1	.408	.531
<b>WT</b>	<b>1</b>	<b>17.760</b>	<b>.001</b>
<b><math>L_{AeqWT}</math></b>	<b>1.631</b>	<b>18.902</b>	<b>.000</b>

Table 19.4 ANOVA<sub>scen</sub> Tests of Within-Subjects Effects (significant effects are emboldened).

(I) $L_{AeqWT}$	(J) $L_{AeqWT}$	Mean Difference (I-J)	Std. Error	Sig.
39 dB(A)	44 dB(A)	-2.224	1.350	.350
<b>39 dB(A)</b>	<b>49 dB(A)</b>	<b>-6.551</b>	<b>.861</b>	<b>.000</b>
<b>44 dB(A)</b>	<b>49 dB(A)</b>	<b>-4.327</b>	<b>.980</b>	<b>.001</b>

Table 19.5 ANOVA<sub>scen</sub> Pairwise comparison of  $L_{AeqWT}$  levels (significant effects are emboldened).



## Appendix XI: Traffic recordings

Traffic recordings were carried out to the southwest side of the motorway M6, about 12 km west of Manchester airport midway between junctions 19 and 20a<sup>††</sup>.

The surrounding area is predominantly farmland, with the field itself being a corn field owned by the farmer.

The only other major source of noise evident was the occasional plane pass – though these were reasonably infrequent – and have been noted down during measurements. Other sources include the occasional passing vehicle over the Cann Lane Bridge to the east, though this is essentially a very small back country road with minimal traffic.

Weather conditions were overcast and calm with occasional drizzle.

Measurements were carried out at approximately 100m, 75m and finally 50m from the edge of the motorway and are shown below, determined later on by a number of points of reference noted within the field.



Figure 20.1 Map of measurement side including distances from motorway.

<sup>††</sup>

[http://maps.google.co.uk/maps?f=d&source=s\\_d&saddr=53.326421,-2.464542&daddr=&geocode=&hl=en&mra=mi&mrsp=0&sz=15&sll=53.325875,-2.462525&spn=0.012355,0.033174&ie=UTF8&ll=53.325875,-2.462525&spn=0.012355,0.033174&t=h&z=15](http://maps.google.co.uk/maps?f=d&source=s_d&saddr=53.326421,-2.464542&daddr=&geocode=&hl=en&mra=mi&mrsp=0&sz=15&sll=53.325875,-2.462525&spn=0.012355,0.033174&ie=UTF8&ll=53.325875,-2.462525&spn=0.012355,0.033174&t=h&z=15)

All measurements were 11 minutes long, with the intention that the first and last 30 seconds would be deleted due to leaving and approaching the microphone. They were carried out using 01dB Symphonie field measurement kit No. 4, and recorded at 51200Hz sampling frequency and 16 bit.

Measurement No.	Approx Start Time	Approx Distance	Notes
1	12:08	100m	Occasional spots of rain – light.
2	12:53	100m	-
3	13:07	100m	Plane pass in approx first 2 mins. Second plane pass at approx 6-7 mins though quiet. Tonal clang – lorry going over nearby bridge at approx 8-9 mins.
4	13:26	75m	Plane pass at approx 3 minutes. Small spots of rain towards end.
5	13:37	75m	Drizzle. Plane pass right at end (maybe missed off recording).
6	13:49	75m	Plane pass at end.
7	14:02	75m	-
8	14:25	50m	Plane pass in approx first 2 mins. Second plane pass at approx 5 mins though very quiet.
9	14:42	50m	Plane pass at approx 8-9 mins.
10	14:53	50m	-

Table 20.1 Overview of traffic noise recordings.

Note: Traffic became noticeably more 'continuous' from around measurement 6. Earlier traffic was slightly more sporadic. Due to this being a daytime measurement, a large proportion of the traffic consisted of lorries and other heavy goods vehicles.



Figure 20.2 a) Traffic recording position 75m away from the motorway. b) Dense motorway traffic consisting of a combination of cars, vans and lorries.

## 20 Appendix XII: Results from first test series: April 2010

Listening tests were conducted in April 2010 with 20 participants. Technical problems occurred such that the sub-woofer levels were played at about 20 dB below the mid and high frequency levels. The cross-over region of the two loudspeaker system affected the tones at 80 Hz and 115 Hz. The A-weighted spectra levels of stimuli in masking noise were too high. Some tonal artefacts were also detected in the tonal analysis of quiet stimuli.

For comparison the results of these tests are shown below and it is discussed how they relate to the results shown in Chapters 6 and 7.

### 20.1 Part A

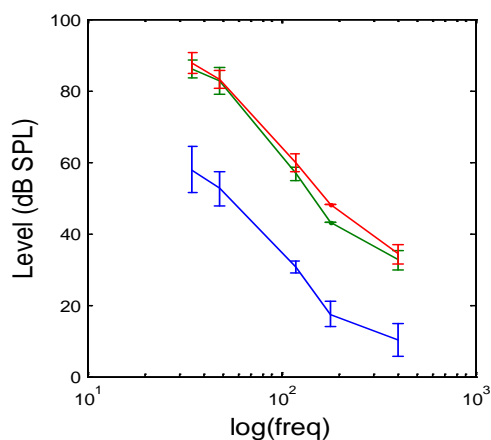


Figure 21.1 Room background without masking noise  
- Audibility threshold (blue), equal annoyance to 180 Hz tone at 5 dB audibility (green) and 10 dB audibility (red) of low frequency tones. Error bars denote standard errors.

Figure 21.1 shows the measured tonal audibility threshold in quiet in analogy to Figure 6.3. The audibility threshold follows the same trend as the final results with values ranging between just below 60 dB to about 12 dB. These values agree well with the final results however the level at 115 Hz is about 5 dB higher than expected. The 400 Hz audibility is slightly higher than expected from literature because the same loudspeaker system was used in both tests with the consequences discussed in Chapter 6. Equal annoyance contours follow the same general trends as seen in Figure 6.3. The levels are generally higher compared to the final results due to the high sound level of the stimuli.

Figure 21.2 shows equivalent results for the outdoor scenario with and without masking noise. The corresponding indoor results are shown in Figure 21.3. Trends are in general similar to the ones in the final tests. Audibility thresholds are less linear with frequency due to sound reproduction artefacts. However the indoor trends look at lot more linear than in the final test because the stimuli were played at unrealistically high tone levels.

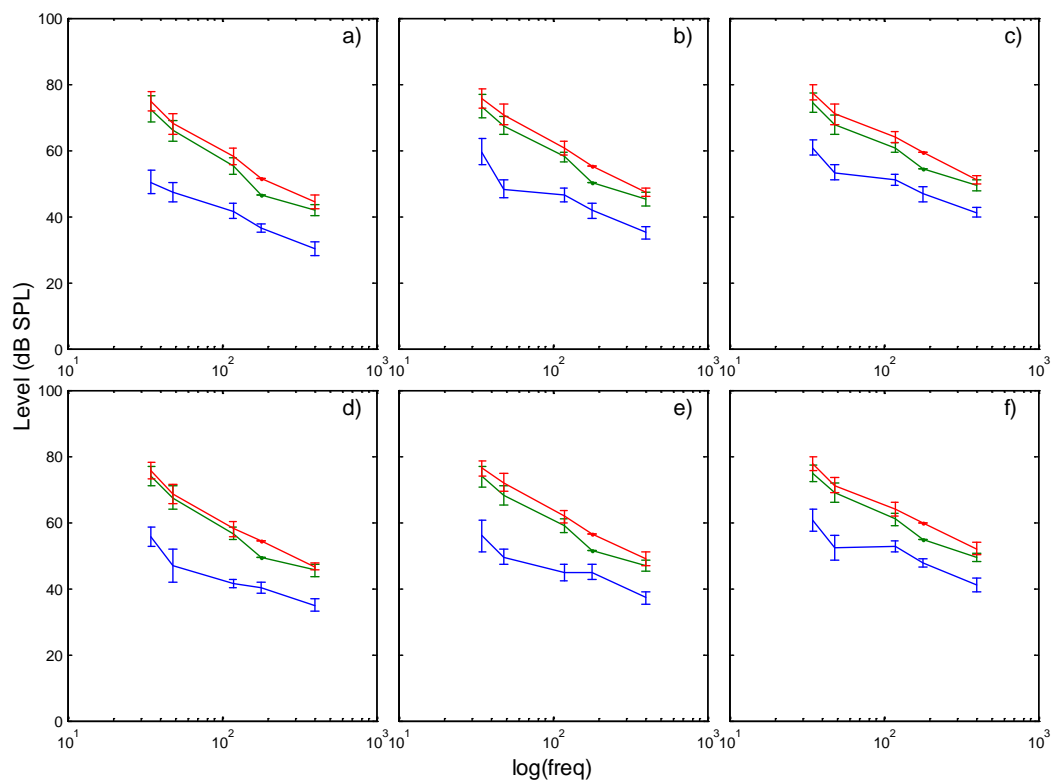


Figure 21.2 Outdoor scenario – Masking thresholds (blue), equal annoyance to 180 Hz tone at 5 dB audibility (green) and 10 dB audibility (red) of low frequency tones within masking noise a) 39 dB(A), b) 44 dB(A), c) 49 dB(A), d) 39 dB(A) and garden noise, e) 44 dB(A) and garden noise and f) 49 dB(A) and garden noise. Additional green and red dots at 180 Hz denote the results from the masking comparisons discussed in Section 6.4 (for 5 dB and 10 dB above  $L_{pn}$  respectively). Error bars denote 95% standard error.

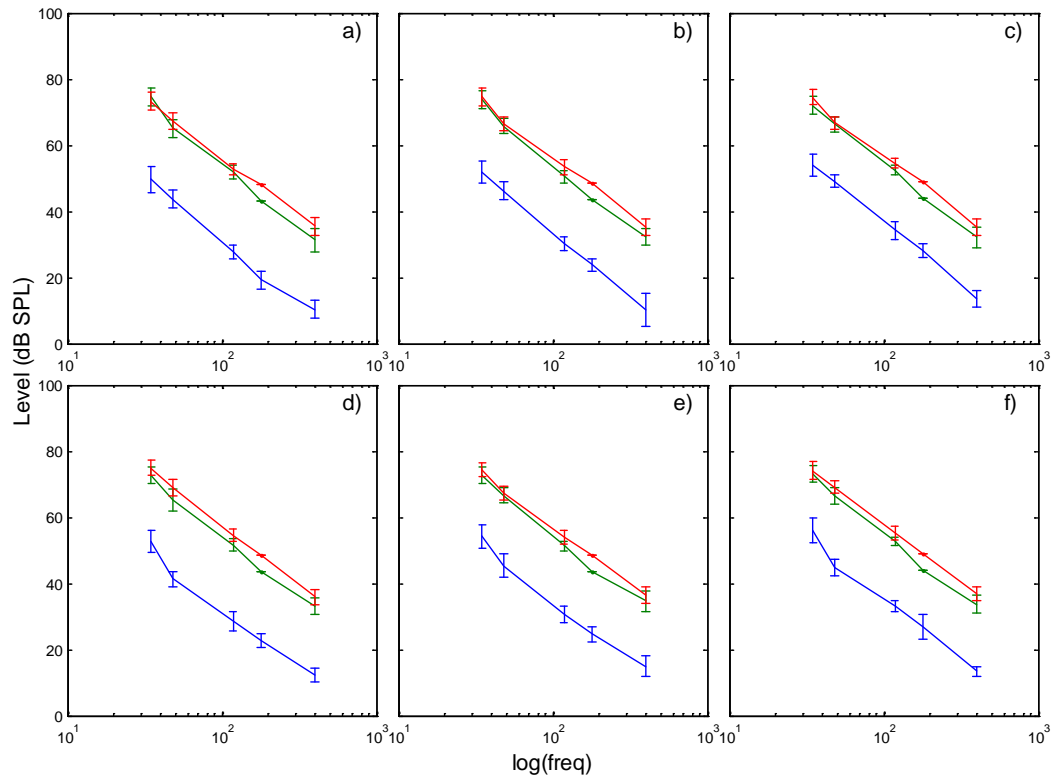


Figure 21.3 Indoor Scenario – Masking thresholds (blue), equal annoyance to 180 Hz tone at 5 dB audibility (green) and 10 dB audibility (red) of low frequency tones within façade attenuated masking noise comprising of  $L_{Aeq,WI}$  a) 39 dB(A), b) 44 dB(A), c) 49 dB(A), d) 39 dB(A) and garden noise, e) 44 dB(A) and garden noise and f) 49 dB(A).. Error bars denote standard error.

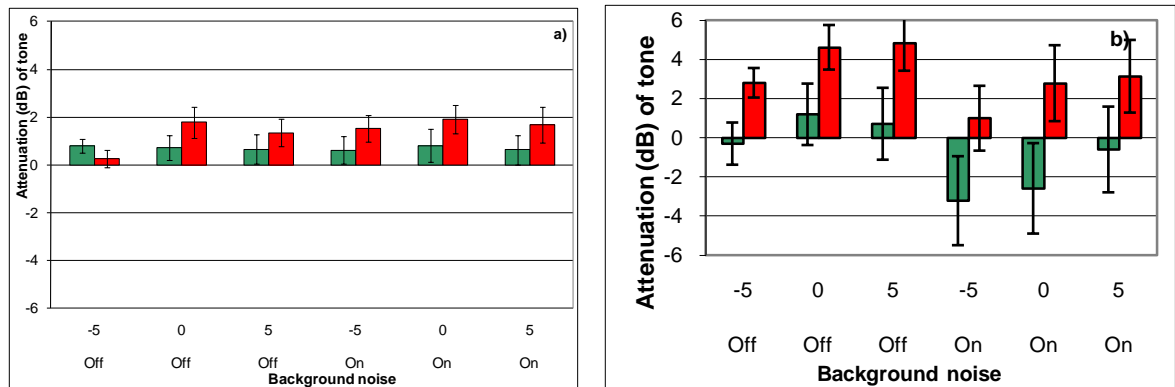


Figure 21.4 Equal annoyance level difference between a tone in quiet (room background noise) and a tone in masking noise. The tone frequency was 180 Hz and the levels within masking were played at 5 dB (green) and 10 dB (red) above  $L_{pn}$ . a) indoor, b) outdoor Error bars denote standard error.

Figure 21.4 shows the results for the masking comparisons. The results of the indoor scenario show clearly that the masking was not audible in that scenario. There is a consistent difference in rating for the two different reference levels  $L_R$ . The outdoor results are similar for the case without garden noise. The addition of garden noise led to completely different results than the ones seen in Figure 6.10.

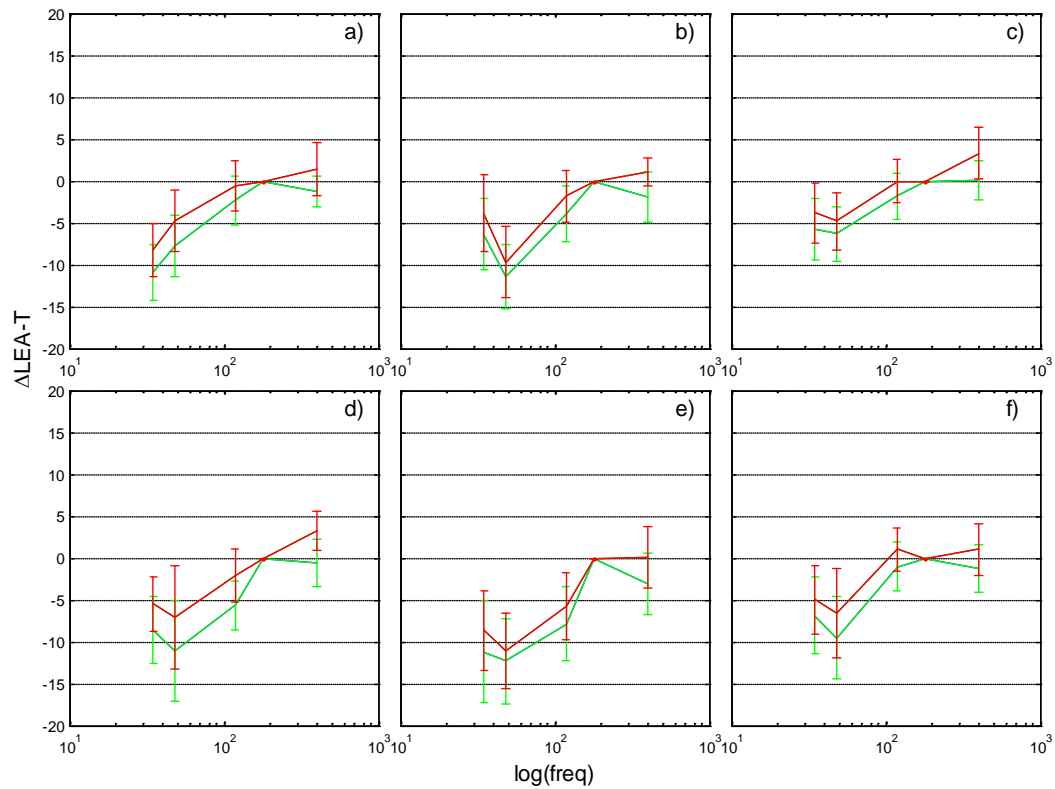


Figure 21.5 Outdoor scenario: Relative sensation level for equal annoyance with various masking noises at tone prominence levels of 5 dB (green) and 10 dB (red). a) LAeq,WT 39 dB(A), no garden noise, b) LAeq,WT 44 dB(A), no garden noise, c) LAeq,WT 49 dB(A), no garden noise, d) LAeq,WT 39 dB(A), with garden noise, e) LAeq,WT 44 dB(A), with garden noise and f) LAeq,WT 49 dB(A), with garden noise. Error bars show standard error.

Figure 21.5-7 show the results for the relative sensation levels that are equivalent to Figures 17.1-3. Results in quiet and in the indoor scenario show similar trends as the final tests with low frequency dependence. The standard errors are a lot higher than for the final tests. The results in the outdoor scenario are significantly different with a seemingly strong frequency dependence of the results and a trend that suggests that low frequencies are less annoying than higher frequencies. It is not clear what caused this pronounced difference in the results.

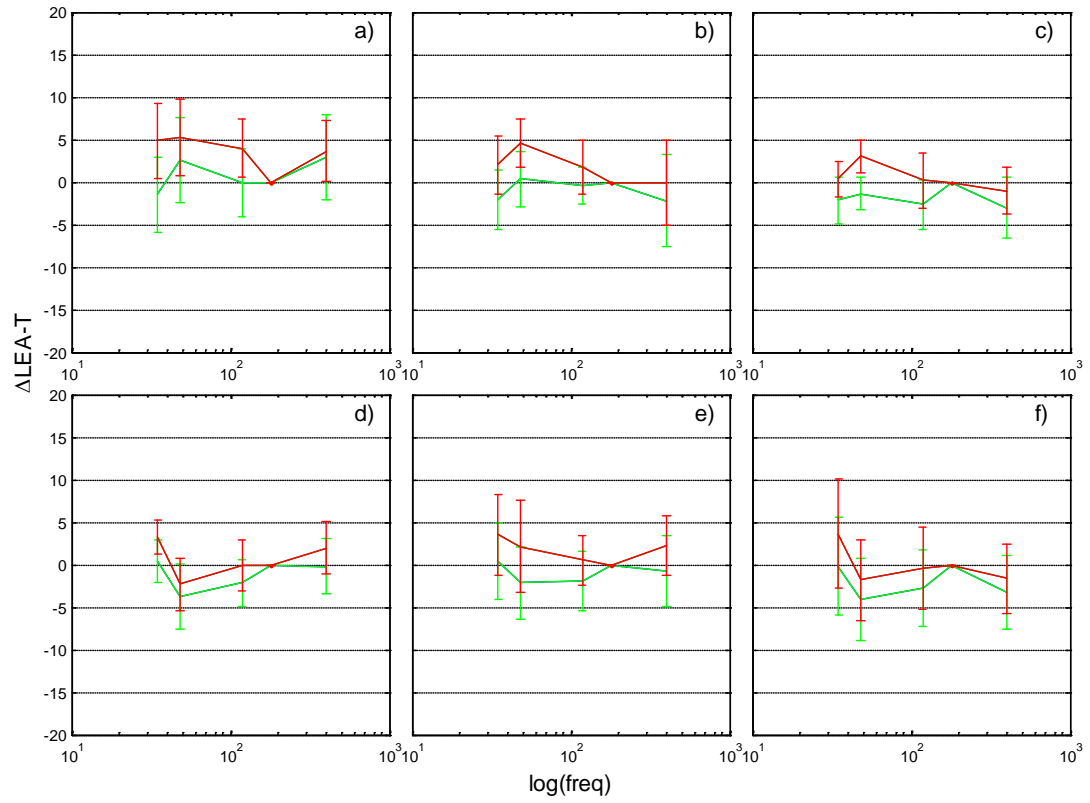


Figure 21.6 Indoor scenario: Relative sensation level for equal annoyance with various masking noises at tone prominence levels of 5 dB (green) and 10 dB (red). a) LAeq,WT 39 dB(A), no garden noise, b) LAeq,WT 44 dB(A), no garden noise, c) LAeq,WT 49 dB(A), no garden noise, d) LAeq,WT 39 dB(A), with garden noise, e) LAeq,WT 44 dB(A), with garden noise and f) LAeq,WT 49 dB(A), with garden noise. Error bars show standard error.

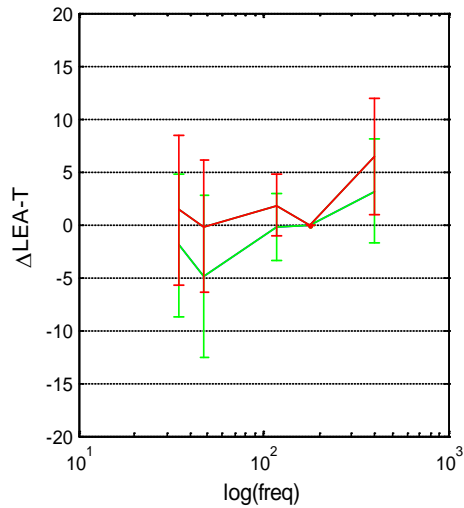


Figure 21.7 Relative sensation level for equal annoyance without masking noise at tone prominence levels of 5 dB (green) and 10 dB (red). Error bars show standard error.



## Bibliographic Data Sheet

---

Title and authors

Perception of Low Frequency Noise From Large Wind Turbines  
Jonathan Hargreaves, Andrew King, Andrew Moorhouse, Chris Plack, Sabine von Hünenbein

---

---

Pages	Tables	Illustrations	References
106	20	67	54

---

---

Descriptors INIS/EDB

LOW FREQUENCY NOISE, LISTENING TESTS, WIND TURBINE NOISE; WIND ENERGY, NOISE ANNOYANCE, NOISE PERCEPTION



**Acoustics Research Centre**  
The University of Salford,  
Salford, Greater Manchester  
M5 4WT. United Kingdom