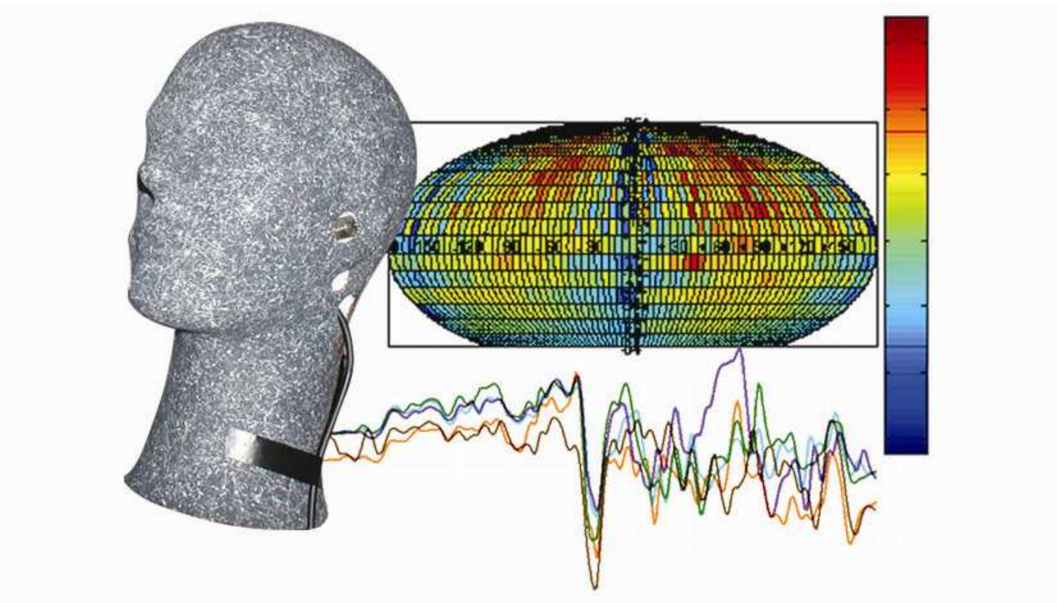
CHALMERS



Simplified Binaural Measurement Systems for Interior Noise Evaluation of Truck Vehicle Compartments

Master's Thesis in the International Master's programme in Sound and Vibration

MADELENE PERSSON AND PETER TORSTENSSON

Department of Civil and Environmental Engineering
Division of Applied Acoustics
Chalmers Room Acoustics Group - MSA
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2007

Master's Thesis 2007:32

MASTER'S THESIS 2007:32

Simplified Binaural Measurement Systems for Interior Noise Evaluation of Truck Vehicle Compartments

Master's Thesis in the International Master's programme in Sound and Vibration

MADELENE PERSSON AND PETER TORSTENSSON

Department of Civil and Environmental Engineering
Division of Applied Acoustics
Chalmers Room Acoustics Group - MSA
CHALMERS UNIVERSITY OF TECHNOLOGY

Göteborg, Sweden 2007

Simplified Binaural Measurement Systems for Interior Noise Evaluation of Truck Vehicle Compartments
Master's Thesis in the International Master's programme in Sound and Vibration
MADELENE PERSSON AND PETER TORSTENSSON

Master's Thesis 2007:32
Department of Civil and Environmental Engineering
Division of Applied Acoustics
Chalmers Room Acoustics Group - MSA
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone: + 46 (0)31-772 1000

Cover;
Plot of interaural level difference calculated from head related transfer function measurements,

Chalmers Reproservice
Gothenburg, Sweden 2007

Simplified Binaural Measurement Systems for Interior Noise Evaluation of Truck Vehicle Compartments
Master's Thesis in the International Master's programme in Sound and Vibration
MADELENE PERSSON AND PETER TORSTENSSON
Department of Civil and Environmental Engineering
Division of Applied Acoustics
Chalmers Room Acoustics Group - MSA
Chalmers University of Technology

ABSTRACT

This Master Thesis was carried out in cooperation with Volvo 3P to design a simple system capable to record, playback and measure levels of noise in truck compartments. The desire was to achieve good enough quality of the output data to replace the currently used binaural recording system. It was requested that the recording system should be able to measure correct absolute levels, be portable as well as provide a binaural recording fidelity comparable to that of the equipment used at the company today. The binaural recording fidelity of five different binaural-similar recording systems has been studied, with commercially produced artificial head and headset as reference. For level measurements, free hanging microphones were used as reference.

The binaural recording quality was investigated using spectral analysis, measurement of head related transfer functions, and listening tests. From the head related transfer function we obtained knowledge of the directional properties of the recording systems, while the listening tests provided us with the perceived differences of the recordings. To judge the binaural quality of the recordings in their intended acoustical environment, the final evaluation was performed with in-cab noise measurements. The two systems composed of a spherical smooth head, "Sphere", and the display head, "Friggo" performed well in the listening test regarding reality of in-cab noise recordings. These two heads were manufactured at the Volvo Trucks.

Because of its good binaural recording quality compared to the artificial head and reliable level measurements, Friggo is recommended to be used for recordings of in-cab noise. It performed well in all listening tests. Moreover, Friggo measured in comparison to the artificial head, in-cab noise levels more similar to those of measurements done with free hanging microphones, in comparison to the other systems. Thus, further investigation is recommended to obtain loudness levels in playback corresponding to the real situation.

Contents

ABSTRACT I

CONTENTS 1

PREFACE 3

1 INTRODUCTION 4

1.1 Goal 4

2 BACKGROUND 5

2.1 Human localization of sound 5
2.1.1 Interaural level difference, ILD 5
2.1.2 Interaural time difference, ITD 6
2.1.3 Confusion in localization 8
2.1.4 Binaural loudness 8
2.1.5 Binaural masking level difference 9
2.1.6 Cocktail party effect 10

2.2 Head related transfer functions 10

2.3 Stereo reproduction and recording systems 11
2.3.1 Recording systems 11
2.3.2 Reproduction 11

2.4 Sound fields 12
2.4.1 Free field 12
2.4.2 Diffuse field 12
2.4.3 The sound field inside a truck compartment 13
2.4.4 Interior noise 14

3 MEASUREMENT SYSTEMS 16

3.1 ORTF-configuration 16

3.2 Optimum stereo signal - Jecklin Disc 17

3.3 Angular spherical patent – Wedge 18

3.4 Spherical shape with built-in microphones – Sphere 19

3.5 Head with built-in microphones – Friggo 19

3.6 Artificial head – Manikin 20

3.7 Binaural headset – Headset 21

3.8 Microphones 22

4 PHYSICAL MEASUREMENTS 23

4.1 Reference source 23
4.1.1 Results from measurements with reference source 24

4.2 Measurement of HRTFs 27
4.2.1 Measurement results 29

4.3 Measurements of in-cab noise 35
4.3.1 Comparisons to free hanging microphones 35
4.3.2 Spectral analysis 38

5 PERCEPTUAL SENSATION OF THE MEASUREMENT SYSTEMS 43

5.1 Localization 43
5.1.1 Results of the Localization test 45

5.2 Source separation ability 48
5.2.1 Results of the Source separation ability test 49

5.3 Cocktail party effect 51
5.3.1 Results from the perception of numbers 52
5.3.2 Results from the intelligibility of numbers 54
5.3.3 Results, disregarding deviant answers 56
5.3.4 Spectral analysis of the sounds in the numbers test 60

5.4 Recorded truck sound 65
5.4.1 Results from in-cab test 66

6 DISCUSSION 67

7 CONCLUSIONS 70

REFERENCES 72

APPENDIX A – MEASUREMENT SYSTEMS SPECIFICATIONS 74

APPENDIX B - ABSORPTION MATERIAL 75

APPENDIX C - CALIBRATION 76

APPENDIX D - ADDITIONAL MEASUREMENT EQUIPMENT: 77

APPENDIX E – REFERENCE FAN 78

APPENDIX F – RESULTS FROM HRTF MEASUREMENTS 79

APPENDIX G – LISTENING TESTS 85

APPENDIX H – BARK SCALE 89

APPENDIX I – STATISTICAL ANALYSIS 90

Preface

This project has been carried out from September 2006 to February 2007 in cooperation with Volvo 3P and the Chalmers Room Acoustic Group at Applied Acoustics (CRAG-MSA), Chalmers University of Technology. The main part of the thesis has been done at the Noise and Vibration Laboratory at Volvo Trucks, Gothenburg.

Throughout the master thesis we received an immense support from our supervisors MSc Christina Keulemans, Volvo 3P and Tech Lic Anders Genell, CRAG who always could find time for our thoughts and questions. We also like to give special thanks Eugen Joss and Dr. Kaj Bodlund for their commitment and help with proofreading the report.

Special regards to the staff both at Volvo 3P and Chalmers that did their best of answering our questions and help us with small bits and pieces along the way.

Finally we also would like to thank Dr. Mendel Kleiner, examiner at Applied Acoustics, for the guidance and critical review of the thesis work.

Göteborg March 2007

Madelene Persson and Peter Torstensson

1 Introduction

Recording equipment rendering good binaural fidelity is needed to enable an evaluation of a recorded sound field. Sound waves travelling inside an enclosure are reflected and shattered. Hence the resulting reverberant sound field will be a superposition of waves travelling in different directions. All sounds together create the acoustical environment and contribute to the listener's perception of the sound.

Measurements of sound pressure levels are preferably performed using a single omnidirectional microphone. This diminishes the influence from the recording equipment on the recorded sound field. Measurements of absolute levels with a two-channel binaural recording system are associated with a number of difficulties. One obvious difficulty is to find a strategy for combining the two channels, and establishing a correct resulting level comparable to that obtained with single microphone measurement. Also, the presence of the recording equipment in the recorded sound field will influence on the sound level measurements.

This Master Thesis was carried out in cooperation with Volvo 3P in order to help designing a simple system capable to record, playback and measure levels of noise in truck compartments. The desire was to achieve good enough quality of the output data to replace the currently used binaural recording system (HMS.II.1). It was requested that the recording system should be able to measure correct absolute levels, be portable as well as provide a binaural recording fidelity comparable to the equipment used at the company today.

1.1 Goal

The goal of this thesis work was to investigate the possibility of developing a microphone system having a binaural recording quality comparable to that obtained using the recording equipment available at Volvo today. Therefore an investigation regarding if simpler binaural-similar measurement systems, including stereo techniques, can be used instead of the binaural recording system is of interest. The recording system should also be able to provide correct absolute levels in accordance with the present equipment. Moreover, as the recording system is intended to be used for on-road measurements at the passenger side in the truck compartments, it is desired that the recording system is easily portable.

2 Background

2.1 Human localization of sound

Sound localization is the process by which humans perceive the apparent spatial position of an acoustical source. We are able to differentiate between sound sources that are in front and behind, behind and below, to the right and to the left. Human hearing is provided with the ability to evaluate a variety of cues which enables this perception in ways yet not entirely understood. The auditory system processes and correlates the sounds coming to both ears continuously. The concept of the process, from the moment when the sound arrives at the ears to the perceived sensation is called binaural hearing.

Listening to a binaural recording will ideally give the same auditory experience as if the listener experienced the situation in real life. In a binaural recording it is possible to capture the acoustical room environment and the different source locations in the recorded sound field. This is not possible when listening to monaural recordings done with a single microphone.

There are many acoustical cues and properties which enable humans to localize sound. Some of them have been selected to be more thoroughly presented in this section of the report.

2.1.1 Interaural level difference, ILD

Using both ears when listening to a sound will give the ability to acquire a sensation of from which direction the sound arrives. One important property of hearing is to sense the level difference between the ears. This is known as the Interaural level difference (ILD) and is a complex quantity depending on head and ear shadowing effects, frequency and the angle of incidence. Figure 1 shows a principal sketch of the ILD concept.

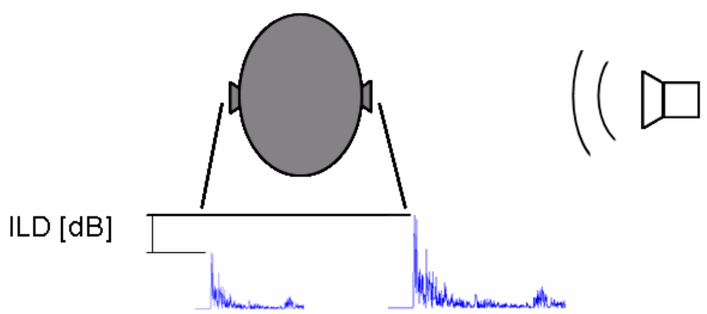


Figure 1: Interaural level difference is the difference in level of a sound at both ears.

For sound sources at relatively long distance from the head (> 1 m), the path length difference from the source to the ears is small compared to the distance from the source to the listener's head. The difference in path length between the two ears is negligible, and the ILD is caused by the diffraction and reflection of the head. This is often referred to as acoustical shadowing, which reduces the sound pressure level at the far ear. When the source is within 1m of the head, however, the path length difference is no longer negligible, and the ILD is also created by the additional attenuation incurred along the path to the ear farther from the source. At frequencies with a wavelength longer than the shortest distance between the ears the ILD will be small. For higher frequencies the scattering due to the head will increase significantly and the ILD will increase as a consequence.

2.1.2 Interaural time difference, ITD

The difference in time of arrival of sound at the two ears from a sound source displaced from the median plane informs the auditory system about the position of the source. This time delay is referred to as interaural time difference, ITD, and constitutes a very important binaural cue for localization ability. The function of ITD is presented in Figure 2.

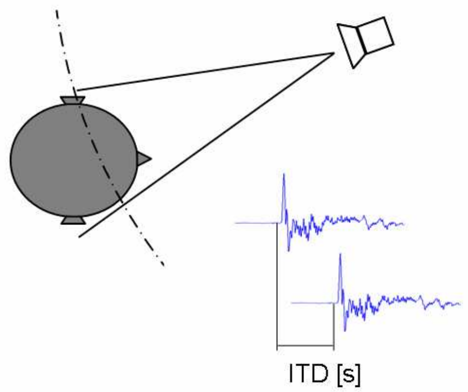


Figure 2: The interaural time difference describes the difference in time of arrival at both ears.

The ITD is exploited by the auditory system in two different ways depending on the frequency of the sound. In the frequency range up to approximately 800 Hz the phase delay between the right and left signal is the dominating attribute whereas for frequencies higher than approximately 1.5 kHz, the group delay between the left and right signal becomes of greater importance. At frequencies above 800 Hz the ability to use the phase delay in localization decreases and it is not used at frequencies higher than approximately 1.5 kHz. In the frequency range 0.8-1.5 kHz both phase- and group delay are used as cues for localization. Based on the interaural transfer function, ITF, defined as the ration between the right and left HRTFs (explained later in this section) the ITD based on the phase and group delay can be computed using the following expressions:

$$ITD = \frac{\arg(ITF)}{\omega} \qquad f <\sim 1.5\,kHz \tag{1}$$

$$ITD = \frac{d(\arg(ITF))}{d\omega} \qquad f >\sim 800\,Hz \tag{2}$$

As mentioned before ITD calculated with phase delay, (eq. 1) is only relevant at low frequencies and ambiguity occurs at higher frequencies, where the phase delay is greater than the half-period T/2 of a tone.

For higher frequencies where the ITD becomes ambiguous, ITD can be estimated with the group delay (eq. 2). The ITD is evaluated by the time shift between envelopes. ITD based on the group delay has been proven in psychoacoustical experiments to be an important cue for localization [16] f > 800 Hz.

The ITD has been studied by evaluating the perceived lateral displacement from the median plane as a function of ITD. The result for impulse content (noise, speech) was found to be linear until approximately ITD = 630 μs [16]. In extension this implies that for frequencies with a half-period greater than 630 μs will attain a full lateral displacement. Another reason for the upper frequency limit of lateral displacement of tones is due to refractory period: the rest period for which the neurons can not send signals. This period is 1-2 ms, which indicates a lateral displacement induced by phase delay not noticeable at 1.5-1.6 kHz [16].

Humans can detect a change in the ITD as small as 10 μs from an original ITD of 400 μs [25]. The variation of ITD caused by a human head is in the range of the phase delays that are perceptible. Experiments have shown that prominent changes of about 80 μs (over a frequency range of 500 Hz to 1 kHz) in ITD can be found for a single azimuth between 30° and 60° of the arriving sound. In the frequency range up to approximately 1 kHz it has been shown in previous studies that ITD dominates the cues of localization. In this frequency range good resemblance of ITD estimated using spherical heads and real or artificial heads has been observed [26], [27].

To produce a just-noticeable difference of ITD for frontal incident the relative time delay between the ears has to be about 50 μs at frequencies below 1.5 kHz [2], which corresponds to an angle of 5°. It varies though a lot from subject to subject; research has measured values of ITD between 30 and up to 200 μs for just-noticeable interaural delay.

2.1.3 Confusion in localization

ITD and ILD are powerful cues for the localization of a source, but they have important limitations. For a source moving in the median plane (the perpendicular bisector of a line drawn trough both ears), the binaural differences ILD and ITD are zero. This can cause front/back confusion which arises as the listener cannot differentiate between sources in the front and in the back. The cone of confusion also describe a situation where the sound arriving from different angles still give rise to equal ILD and ITD. As an example all sounds arriving from the angles along the circumference of the sphere cone in Figure 3 will have the same relation of ITD and ILD at both ears.

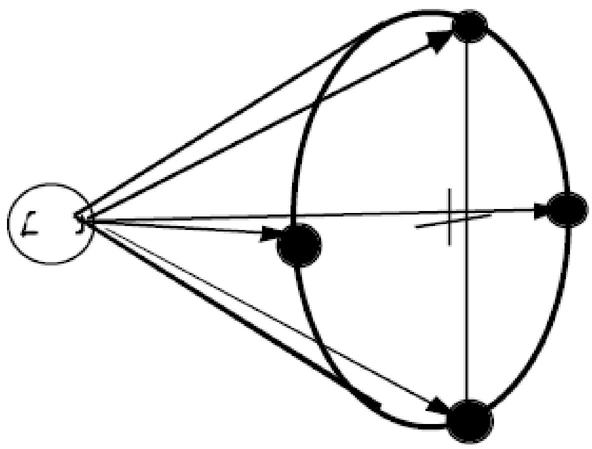


Figure 3: Sound arriving from all positions along the circumference of the cone will have equal ILD and ITD if the head is seen as spherical.

Evidently, there are other properties of the human hearing which enable us to localize a sound source. These additional properties might be related to how the individual listener's outer ears, head, shoulder and upper torso affect the sound [6]. The body of the listener leads to scattering of the sound, an acoustical filtering of the signal appears at the left and right ear.

2.1.4 Binaural loudness

Binaural loudness is a property of hearing which describes the perceived loudness when listening with both ears. You might have personal experience of the difference in perceived loudness between listening to one and two earphones. Changing from one earphone to two earphones give an increased perceived loudness level. The increased level depends on the present sensation level (SL) and loudness. Binaural loudness is not linear dependent of either frequency or sound pressure level. Switching from monaural to binaural listening of a soft sound (20dB SL) give rise to an increase of the perceived loudness with a factor 2. This corresponds to a level increase in (monaural) sound pressure level of 8dB. At higher sensation levels 80dB (SL), the increase from monaural to binaural loudness is only a factor of about 1.4 (SPL increase of 6dB) [2].

2.1.5 Binaural masking level difference

Binaural Masking Level Difference (BMLD) can be described as an improvement of signal detection in noise when either the phase or level differences of the signal at the two ears are not the same as for the masking signal, [2]. BMLD can be understood as a reduction of the masking effect if the signal and masker originate from different locations in space.

BMLD can arrive from various conditions and are investigated in many ways. Usually one studies a masker with same phase at both ears to which a signal is added at both ears but with a phase shift of 180°. Opposite conditions is also used; the signal has constant phase at both ears and the masker has a phase shift. Generally the effect of BMLD is higher in the low frequency range [3]. Figure 4 shows one example of BMLD where the comparable monaural and binaural thresholds [2] are investigated.

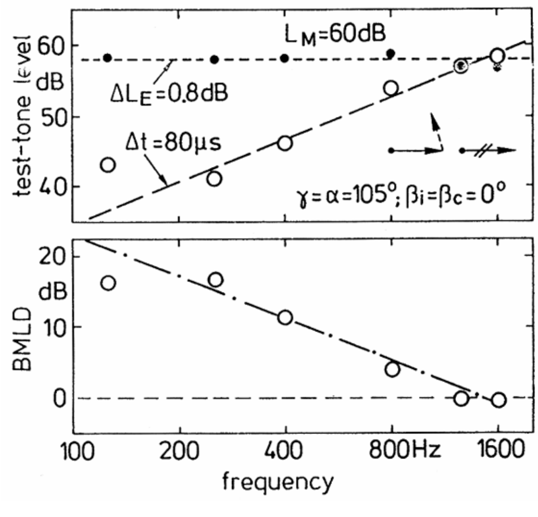


Figure 4: The upper panel shows the median test tone thresholds as a function of the test tone frequency. The masking level is 60 dB SPL. The dots show the monaural results while the open circles are the binaural result. The dotted line shows the expected values according to just-noticeable difference in level. The lower panel shows the corresponding BMLD (open circles), as a function of test frequency and predicted values (dash-dotted line), [2].

Using a sinusoidal tone as masker and the detectable tone phase shifted 105° the threshold has been found both in the monaural case (masker + tone in one ear, level 60dB) and binaural case (masker at one ear and masker + tone in the other ear both at level 60dB). There is a noticeable level difference between monaural and binaural threshold. For binaural hearing at lower frequencies below 400 Hz, the threshold is 15dB lower than in the monaural case while no improvement is found above 1 kHz in comparison to the monaural threshold. The phase shift of 105^{o} is comparable to a time delay of 80 µs and the dashed line shows the expected level.

2.1.6 Cocktail party effect

The cocktail party effect refers to the ability to focus one's listening attention on a single talker among a cacophony of conversations and background noise. It is taken for granted and is used on a daily basis by most people. The cocktail party effect is a consequence of the binaural hearing and is mainly derived from that sound origin from different locations will not be as effectively masked as sound coming from the same location.

The cocktail party effect has been researched since the mid 1960s at MIT. Their studies have for example concerned intelligibility of certain messages read out simultaneous with other disturbing messages. The results showed that generally the context is understood and only short phrases of 2-3 words are incorrect [19].

2.2 Head related transfer functions

The shape and size of the head, torso, and in particular of the pinnae have a great influence on the perception of sound. The head related transfer function (HRTF) somehow gives an analytical measure of those influences. It accounts for diffraction around the head, reflections from the shoulders and most significant, reflections inside the pinnae. As the HRTF are measured for many angles of sound incidents they can be used to study the direction-dependent perception of sound of recording system or individual. The anatomical features vary greatly across individuals and hence the HRTF is highly personalized. To achieve the best possible binaural playback one should therefore measure a personal HRTF to filter the recorded sound with. If one then uses an appropriate playback configuration, the experience at playback is as close as possible to the experience during recording. There are of course still limitations if a high number of sources are present with sound arriving from several directions in the recorded environment. Pre-measured HRTF can also be used for binaural sound synthesis, where anechoically recorded sound is filtered with transfer functions corresponding to the desired spatial position of the source [22].

2.3 Stereo reproduction and recording systems

2.3.1 Recording systems

There are a number of basic systems for stereo recordings using different arrangement and directivity characteristics of microphones. Basically all systems rely on the use of various directivity pattern microphones mounted in an angle from each other or spaced symmetrical from a center line. Sound sources on the center line will produce equal signals at both microphones which produces a phantom image in the middle of a stereo playback. Sound sources off center will produce a higher-level signal from the microphone with direction towards the source which will result in an off center phantom image [5].

In one study two main categories of stereo recording systems were investigated [2]; near- coincident pair and baffled-omni pair. The near-coincident pair method uses two directional microphones angled apart with the diaphragms some decimetres apart horizontally. The spacing gives the stereo spread and adds a sense of ambient warmth or air to the recording. The greater the angle or spacing between the microphones, the greater the stereo spread [5]. The directional pattern of the microphones produces level differences between the channels and the spacing produce time differences. A recording system in this category is the ORTF (Office de Radiodiffusion Television Francaise) configuration, which consists of two angled microphones with cardioid pattern (further explained in section 3.1), and is among the recording systems studied here. The Baffled-Omni Pair uses two omnidirectional microphones, usually spaced at a distance equal to the distance between the ears, and also separated by a hard or padded baffle. The spacing between the microphones creates time differences (corresponding to ITD) and the baffle creates level differences (corresponding to ILD), most prominent for high frequencies. Among the recording systems concerned in this report, the artificial head as well as the recording systems using microphones mounted on baffles belongs to the Baffled-Omni Pair category.

2.3.2 Reproduction

It is a basic purpose of stereophonic microphones to pick up the directional information of the sound scene. This is in order to achieve a corresponding directional distribution and dimension of the sound sources during reproduction [4]. During playback of a stereo recording the listener perceives virtual sound sources contained in the sound. The image location ideally corresponds to the sound source location during the recording session [5].

Two methods commonly used to reproduce binaural recordings are playback using loudspeakers with cross-talk cancellation and direct headphone playback. Cross-talk cancellation consists of two loudspeakers which together form the binaural sound field by ILD, ITD and include also interference such as cancellation. It only reproduces the recording at a correctly specific point and has high demands to be adjusted properly. The cross-talk cancellation has a clear advantage to generate frontally located auditory images in comparison to headphones. A refined version of cross-talk cancellation is known as stereo dipole cross-talk cancellation [28]. Playback with stereo dipole cross-talk cancellation in a loudspeaker setup has the advantage of allowing small head movements in order to localize sounds. In stereo dipole cross-talk cancellation the loudspeaker are placed at an angle of 10° which is larger than the angle generally used in cross-talk cancellation setups. The reason to use playback through headphones is naturally its simplicity. The fact that it is portable makes it very suitable to use during measurements and listening tests.

2.4 Sound fields

Sound waves emitted by a sound source inside a room, enclosed space, or outdoors, are reflected, diffracted, scattered, and might be partly absorbed by obstacles as they travel in space. The sound field then refers to the acoustical environment which is created from all this contributions. Different sound fields will be perceived differently by the listener. The perception of sound depends on the spectral, temporal and spatial properties of the sound field.

2.4.1 Free field

The free field condition is easiest described as an open space sound field. The sound travels without any interruption in space and is therefore not affected or influenced of the surrounding. The anechoic chamber is a room used to simulate free field condition. It is designed to absorb most of the power in sound waves hitting the boundaries of the room. Measurements made in an anechoic environment present the sound pressure level of a sound source with little interference of the environment.

2.4.2 Diffuse field

A perfectly diffuse sound field has a uniformly distributed energy density at any point in an enclosure independent of the direction. Another definition of a diffuse sound field can be formulated that each point on the boundary of the enclosure will be hit by the same amount of energy per unit area. In reality such a field does not exist but it will normally be simulated in reverberant rooms. According to ISO standard (EN ISO 354:2003), it is recommended to have a volume of at least 200 m^3, but not larger than 500 m^3 (then air absorption start to influence) to enable acceptable measuring conditions. The sound field is then seen to be diffuse at positions more than 1 m from the walls, floor, and ceiling, at frequencies above the shröder frequency, f_s, of the room.

$$f_s = 2000\sqrt{\frac{T}{V}} \quad (3),$$

V = volume of the room (m^3)
T = reverberation time (s)

2.4.3 The sound field inside a truck compartment

In a vehicle or a room the sound field is neither a free- nor a diffuse field. The sound field in an enclosure is composed by modes which depend on the size of the enclosure and the speed of sound.

If a truck compartment is approximated by a rectangular space, the mode frequencies can be calculated as [20]:

$$f_{a,b,c} = \frac{c}{2\pi}\sqrt{\left(\frac{a\pi}{x}\right)^2 + \left(\frac{b\pi}{y}\right)^2 + \left(\frac{c\pi}{z}\right)^2} \quad (4),$$

where c is the speed of sound in air (ms^{-1})
a, b, c = integer mode number 1, 2,
x, y, z = dimensions of space (m)

Calculation with width x = 2.27, length y = 2 and height z = 1.52 m give the first modes at approximately 75, 85, 112, 113, 135 and 140 Hz.

These numbers coincide well with those obtained in simulations of truck compartments. A study has been made of the modes in the truck. The strongest modes in the empty truck compartment are close to 80 Hz when excited at the drivers and passenger heads position respectively. Two modes are dominant; one extending over the width (maximum at doors and minimum node in between the seats) and one stretching diagonally from the roof at the back wall down to the front (maximum at the roof and the front floor). In the study simulations of a several truck compartments all resulted in with the previously mentioned dominating 80 Hz mode. Higher up in frequency there appeared larger shifts of the modes for the different compartments

The simulations were done assuming empty compartments, therefore additional obstacles will of course help to achieve higher diffuseness of the compartment. According to Head Acoustics, the sound field in a passenger car can be described with a sound field having propertied in between free and diffuse field. The sound field is then called ID by Head Acoustics and stands for independent of direction [23].

2.4.4 Interior noise

Noise in a vehicle is a combination of:

- engine noise
- road noise
- intake noise
- exhaust noise
- aerodynamic noise
- noise from components and ancillaries
- brake noise
- squeaks and rattles

Noise can either be air borne or structure born [21]. As an example, airborne sound can be reduced by sound insulation, while the structure borne sound can be controlled by resilient mounts and radiation limiting.

To avoid high levels of noise in the truck compartment coming from the engine it is important to limit the acoustical coupling between the engine and the modes of the compartment. Generally in Volvo trucks the engine is a six cylinder engine which is fired in pairs of two, i.e. three firings each rotation. The firing of the cylinders, an engine speed of 1000 rpm or 1700 rpm would generate strong radiation of sound at a frequency of 50 Hz and 85 Hz respectively (corresponding to the 3rd engine order). The harmonics of those frequencies that coincide with the modes of the compartment will therefore result in a standing wave pattern and consequently high noise levels. The figure below shows results from an in-cab noise level measurement (interior noise measurement in the truck compartment) at the passenger side. One can clearly see the resonance at 50 Hz and its harmonic at 100 Hz.

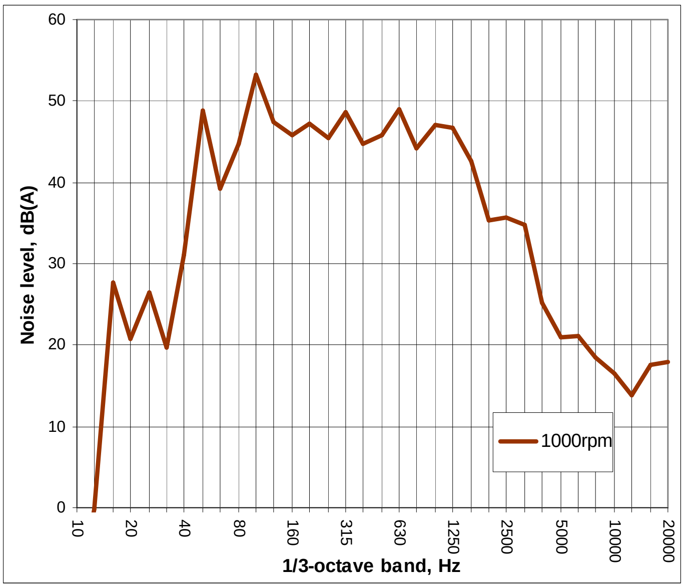


Figure 5: Noise level measurement in truck compartment for the rotational speed of 1000 rpm.

3 Measurement systems

Five binaural-similar recording systems were designed with a complexity ranging from a simple stereo-configuration to a manikin head more similar to that used in conventional binaural recording. In addition, commercial equipment such as an artificial head and binaural headset was part of the study too.

Simple stereo recording systems are frequently used in concert auditoria, concert halls etc, in order to record concerts. The requirement on the equipment in such measurements is more of an emotional matter than a correct reproduction. In our case the realism of the playback is of high importance along with measuring correct sound pressure levels.

The measurement systems are presented in the next sections and additional description of the measurement systems and microphones can be found in Appendix A.

3.1 ORTF-configuration

The ORTF configuration (Office de Radiodiffusion Television Francaise) represents a standard method for recording stereo signals. Two microphones, with cardioid directivity, are used spaced 165 mm and diverging in an angle of 112.5^{o}. This system will enable reproduction of level and phase difference between right and left channel with minimum phase cancellation [9, 10]. Since the cardioid directivity rejects sounds from certain directions, less of the ambient room characteristics are recorded in comparison to a binaural recording [10]. The ORTF configuration is often considered to be the closest simulation of how human ears perceive sound if one only uses two microphones [7]. Therefore it was chosen to be one of our measurement systems. Figure 6 shows a picture of the ORTF configuration used in the study.

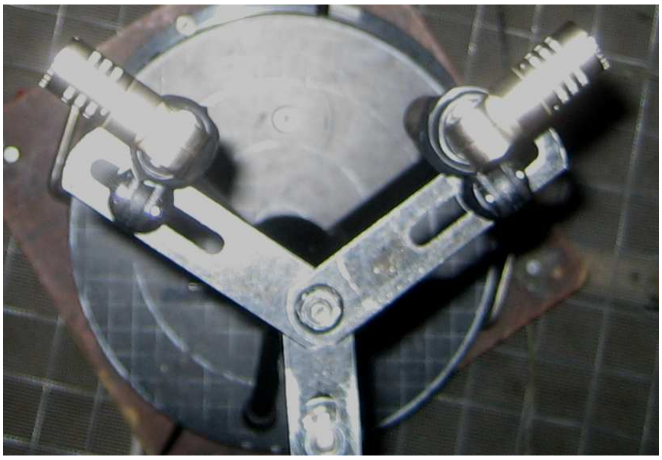


Figure 6: ORTF-configuration, seen from above.

3.2 Optimum stereo signal - Jecklin Disc

The Jecklin disc was developed by Jürg Jeckling, Switzerland [11] in order to both record and reproduce a relative satisfying stereo sound with a simple setup. The Jecklin disc, see Figure 7 has an increased stereophonic separation in comparison to ORTF. It is made of 10 mm plywood and has a diameter of 300 mm. A foam sound absorber of 30 mm thickness was added on both sides (Material data in Appendix C). The distance between the omnidirectional microphones is 170 mm. The microphones are symmetrically placed on both sides of the disc. Windshields were used.

Figure 7: Jecklin disc, front view.

The disc, inserted between the microphones, ensures a difference in frequency response at the two microphones due to the angle of incidence. A side effect is the angle dependent disturbance due to the edge of the disc.

3.3 Angular spherical patent – Wedge

In order to achieve a closer correspondence to the human head than the Jecklin disc, a simple three dimensional shape with microphones placed parallel with the shape was investigated. The idea of the method was taken from the patent "Apparatus for picking up sound waves" [12]. In the patent it was described: "It is a principal object of the present invention to provide an apparatus which allows a pick-up and recording of sound waves which is as true to nature as possible." It can be described as two circular discs combined by a cylindrical shell to a wedge shaped arrangement.

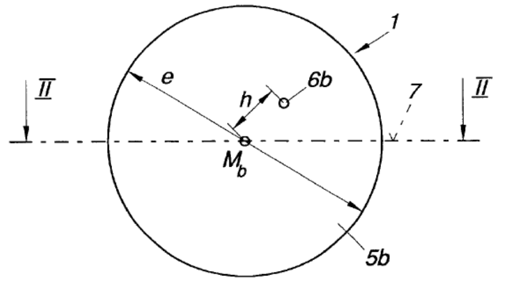


Figure 8: View form the side: M_b is the center of the disc, the distance h = 52 mm from M_b to the microphone position. 2a+2b = 210 mm and the nose width 60 mm [12]. Both sides were

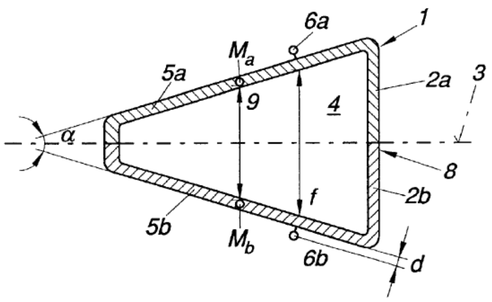


Figure 9: View from above: 6a and 6b are positions of the microphones [12].

The dimensions of the shape can be seen in Figure 8 Figure 9 with the angle α = 17°. To avoid the lateral reflections and interference in the interesting frequency range the microphones are placed at a distance of 5 mm from the shape a foam sound absorption was also attached on both sides of the shape; see Figure 10 (Material data in Appendix C).

Figure 10: Wedge, front view.

3.4 Spherical shape with built-in microphones – Sphere

This measurement system uses a fishing float, with built-in microphones, see Figure 11. The Sphere is symmetric and measures the sound pressure level where the outer ear normally is placed, using omnidirectional microphones. The material of the float is a ~20 mm thick plastic shell which was sawed into two halves. In order to avoid resonances inside the Sphere it was filled with foam absorption material and the upper part was fixed with thick tape. There already exists a commercial product known as the "Sphere". The commercial "Sphere" microphone was developed to achieve convincing results for loudspeaker playback in the same way as artificial heads are used for playback in headphones [13].

Figure 11: Sphere, front view.

3.5 Head with built-in microphones – Friggo

A further refinement in accordance to the human head is to use a display head made by expanded polystyrene, see Figure 12. Two microphones are placed at the positions of the right and left ear by drilling holes from the opposite side of the head.

Figure 12: Friggo, side view.

3.6 Artificial head – Manikin

Head Acoustics artificial head (referred to as Manikin in the report) is the current measurement system used by Volvo. It is placed at the passenger position for measurements of interior noise in truck compartments. Therefore comparisons of Manikin and designed measurements device were of interest. The Manikin used was HMS II.1, see Figure 13 which is a quite old version and consequently differences in quality to newer versions can occur. In order to reproduce the signal in headphones it has to be equalized for the aural cavity and ear canal for which the Manikin has an analogue filter. The equalization used for playback of measurements in vehicle compartments is a filter which is called ID (independent of direction), in between free field and diffuse field. Head acoustics recommend using the ID filter when measuring in vehicles, because the filter is “designed” according to the sound field in a car. However, as this work concerns truck compartments, one must consider possible influences on the results from the differences in the sound field between a car and a truck. When measuring with the Manikin two filters are used; first the built-in (analogue) filter and then the digital filter added in the software. This second filter is used to achieve better resemblance with single microphone recordings.

Figure 13: Manikin, front view.

3.7 Binaural headset – Headset

The Headset equipment from Head Acoustics, (BMH III.3) when worn correctly can replace a Manikin for binaural recording, in situations where the Manikin cannot be used, for example when measuring at the driver's position in a truck.

Figure 14: Headset placed on Manikin.

One difficulty when measuring with a headset is that the results will differ for different persons wearing the headset. As always, when doing measurements the positioning of the microphones is very important. This is even more important when using a headset, since the measurement will also incorporate influence from the torso [14]. All measurements in this report were done with the Headset placed on the Manikin to ensure comparability, see Figure 14.

The Headset, with additional earplugs, can record the sound pressure level with characteristics corresponding to the ID equalization. The ID filter was created using results obtained from many measurements and will, according to Head Acoustics, give results comparable to those from measurements made with Head's artificial head. Still, the influence on binaural recording due to the person's anatomic properties i.e. head and ear shape should not be underestimated, when evaluating and comparing results.

3.8 Microphones

The noise levels were also measured with single microphones. Such microphones can be used as a complement to binaural recording systems in order to measure the absolute noise levels. In this study, a setup called "free hanging microphones" has been used, as a complement to the other measurement systems presented earlier, in order to compare the absolute noise levels measured. This setup is used within Volvo to do comparable noise level measurements. The "Free hanging microphones" consists of two single omnidirectional ¼" microphones, with windshields, placed at 0.1 m distance from the ears of the Manikin, see Figure 15.

Figure 15: Free hanging microphones with Manikin.

The omnidirectional microphones used in these measurements were G.R.A.S. prepolarized condenser microphone type 40AE. Both the microphones in the headset and the omnidirectional microphones had preamplifiers that operate using a constant current power supply.

Microphones with a cardioid pattern where used in the measurements for the ORTF configuration. Further information regarding the calibration can be found in Appendix B.

4 Physical measurements

The performance of the different recording systems was evaluated using both physical- and perceptual measurements. The perceptual characteristics were studied in a number of listening tests. The purpose, design and results from these are found in next section of the report. Any recording system must be able to measure correct absolute levels. Partly for this reason, and also to analytically support conclusions drawn from the listening tests, physical measurements were made as well. Spectral- and limited directional properties in the various recording systems were investigated by measurements in an anechoic chamber. The directional properties of the recording systems were investigated in detail by measuring both conventional and head related transfer functions as appropriate. It can be expected that the level measurement results will vary with the properties of the measured sound field. Therefore it was also important to do measurements inside truck compartments.

4.1 Reference source

Measurements with a reference fan were performed to get a first impression of the measurement system's directional properties as well as the possible differences in measuring absolute noise levels. The setup can be seen in Figure 16. From the reference measurement we could also verify the symmetries of the measurement systems, ensuring that both channels measured equal levels during a 360˚ sweep. The results were compared to those from measurements using free hanging microphones.

Measurements of 360˚ sweep were done with a reference fan placed on the net floor at 3 m distance from the measurement system, for further specifications see Appendix E.

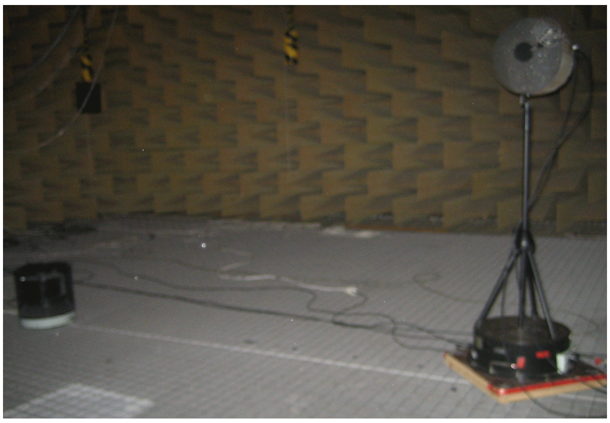


Figure 16: Measurements in anechoic chamber with reference fan, Wedge.

4.1.1 Results from measurements with reference source

The signals were analysed spectrally in comparison with free hanging microphones. The measurement results are plotted in smoothed 12th octave bands. The peak at 50 Hz is due to electrical noise. Both Manikin and Headset signals were filtered with recommended filter from Head Acoustics. In Figure 17 and Figure 18 the results from the measurements with all measurement systems for frontal incidence are presented.

Frontal incidence

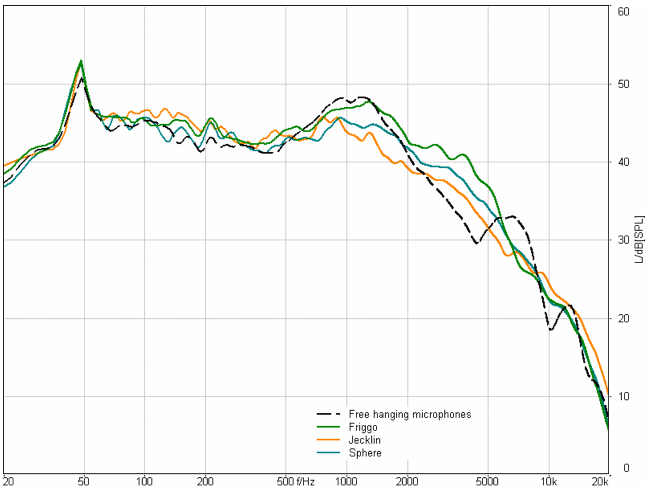


Figure 17: Frontal incidence with reference source: fan, right channel.

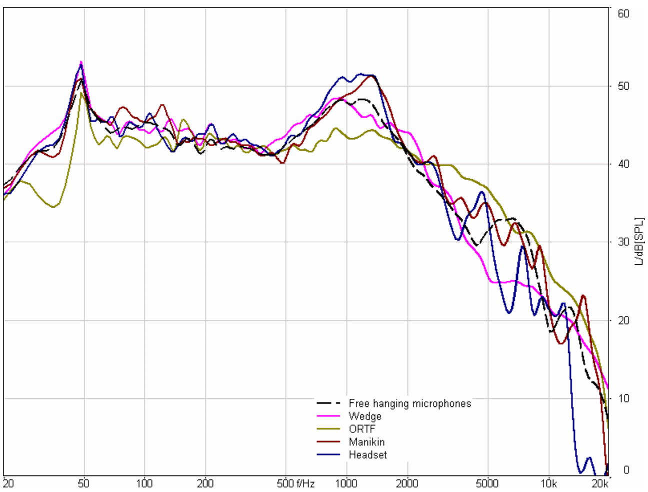


Figure 18: Frontal incidence with reference source: fan, right channel.

Above Figure 17 and Figure 18 shows results obtained by the different measurement systems in comparison to the results using free hanging microphones (black dashed curve). The Manikin, Sphere, Wedge and Jecklin system all show quite similar result to those using the free hanging microphones, except for the Wedge's lack of the peak at 6 kHz. Note the resonances above 2.5 kHz for Headset. The peaks and dips, which are due to the construction of the Headset design, are even more prominent before the filtering with recommended filter.

The figures below show the difference between the free hanging microphones and the measurement systems. Average of two channels with the reference fan at frontal incidence is shown.

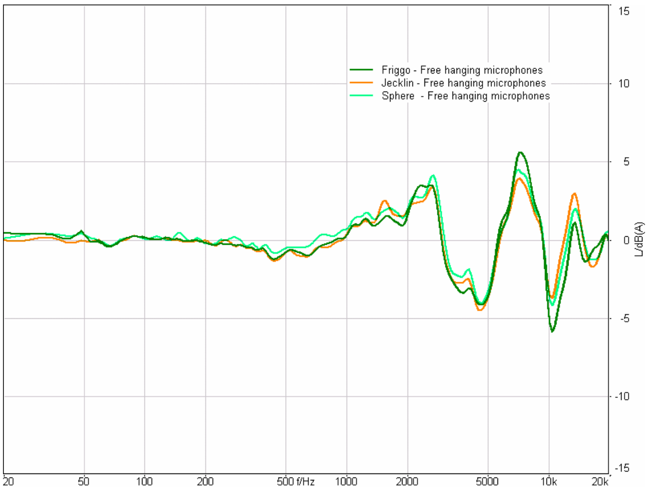


Figure 19: Comparisons between Friggo, Jecklin and Sphere vs. free hanging microphones.

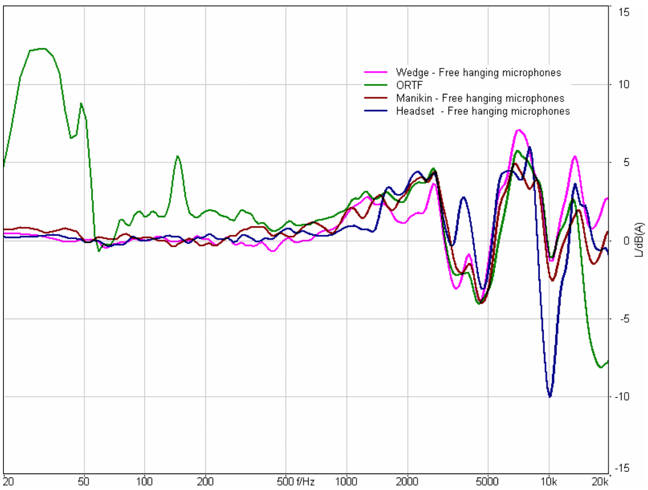


Figure 20: Comparison between Wedge, ORTF, Manikin and Headset.

Figure 19 and Figure 20 show the differences between Wedge, ORTF, Manikin and Headset versus free hanging microphones. The low frequency noise in Figure 19 for the ORTF is a result of problems with the measurement setup. Generally all systems obtain similar spectra in comparison to those obtained using free hanging microphones in the lower frequency range, up to 500 Hz. At high frequencies, between 3-5 kHz, a similar behaviour can be seen for the measurement systems but with individual differences. The Manikin, Jecklin and Sphere have the best similarity to the free hanging microphones.

Incident angle of 180°

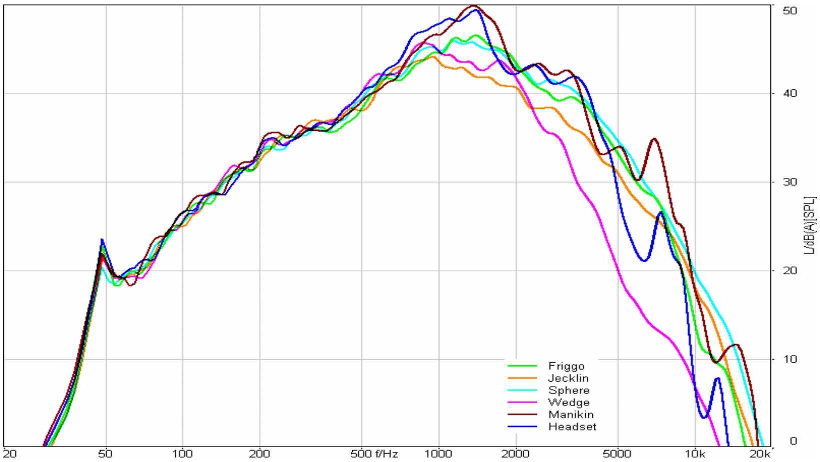


Figure 21: Frequency dependency with 180° incidence, right ear.

The levels shown in Figure 21 are average over a 10 s interval from the sweep of 360°. The approximate angle of incidence is 180° ± 5°. Above 700 Hz the measurement systems start to separate more clearly. Because of the shape of Wedge, with the widest part at the back, its measured sound level (A-weighted) decreases faster above 2 kHz than the other measurement systems, which are closer to the Manikin and Headset. All systems measure lower average SPL than Manikin and Headset.

The conclusion of the measurements with the reference fan is that Sphere and Jecklin are the two measurement systems that have closest resemblance to the free hanging microphones. But the differences in-between the measurement systems are small compared to free hanging microphones.

In addition, the average sound pressure levels from the 360° sweep with the turntable, were analysed. Most measurement systems measured equal levels for the both channels. The 0.1 dB difference for Friggo can probably be explained by two noise peaks that can be seen in the measurement. Additional results from the measurements with the reference fan can be found in Appendix E.

Table 1: Average SPL over 360º rotation in the anechoic chamber.

[dB]	Headset	Manikin	Friggo	Sphere	Wedge	Jecklin	ORTF
Left channel	75.3	75.8	74.2	73.3	74.0	71.8	76.3
Right channel	75.3	75.6	74.3	73.3	74.0	71.8	74.3

4.2 Measurement of HRTFs

Because of the expected different directivity properties of the recording systems, head related transfer functions (HRTF) was measured. The intention was to compare the directivity properties of the recording systems and possibly explain tendencies seen in listening test results. We could also investigate the symmetry of the two channels.

Measurements of HRTF were performed in a setup that consisted of an array of 32 loudspeakers, placed in a sector of a sphere, as shown in Figure 22. The radius of the sphere sector of loudspeakers was 1.25 m which reasonably corresponds to far-field measurements. The rig enabled a resolution of about 8° in the horizontal plane and 10° in the vertical plane. Measurements were done for both channels each 18˚ rotation [33].

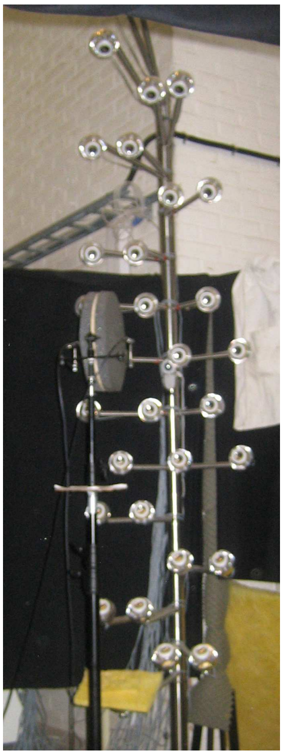


Figure 22: HRTF measurement rig.

HRTF measurements were performed for all the different recording systems. The HRTF for a particular sound source location was created by taking the ratio of the measured complex sound pressure with the recording systems to the same measurement using a single omnidirectional microphone. The results are presented in terms of ILD and ITD for the two channels of the recording systems.

The binaural cues, ILD and ITD, were computed from amplitude and phase of the interaural transfer function, ITF. The ITF was computed as:

$$ITF = \frac{HRTF_{ch1}}{HRTF_{ch2}} \tag{5}$$

ILD was calculated from the following expression:

$$ILD = 10\log\left(|ITF|^2\right) \tag{6}$$

The ITD was computed with the phase delay (see eq. 1) in the frequency range below 800 Hz and the group delay (see eq. 2) for higher frequencies. Special care was taken to "unwrap" the phase, i.e. to avoid errors caused by the abrupt phase changes going from $-\pi$ to π.

Both quantities were analysed by averaging into critical bands and presented in mollweide projection plots, which are map projections. Due to limitations in the measurements we did not present results below 1.6 kHz. The coordinate system in the figures has its origin in the cross section of the horizontal, frontal and median plane.

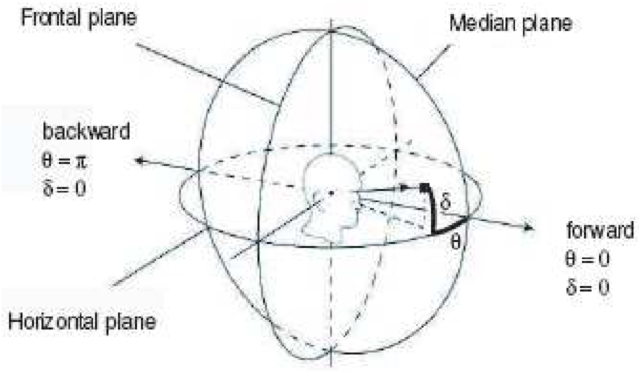


Figure 23: The coordinate system used in the presentation of the HRTF.

The azimuth angle θ is the angle between a vector to the sound source and the median plane, it varies between –180° and 180°. An azimuth of 90° corresponds to the left side of the subject while –90° corresponds to the right side. The elevation angle δ varies from –64° (below the horizontal plane) to 86° (above the horizontal plane).

To reduce the influence of noise in the measurements, all measured impulse responses were windowed using a Hanning window function. Moreover, possible influence from the measurement rig, such as non-flat frequency response of the loudspeaker, was reduced by normalising the measurements obtained using the different recording systems to those of measurements with a single microphone, in accordance with the definition of HRTF.

4.2.1 Measurement results

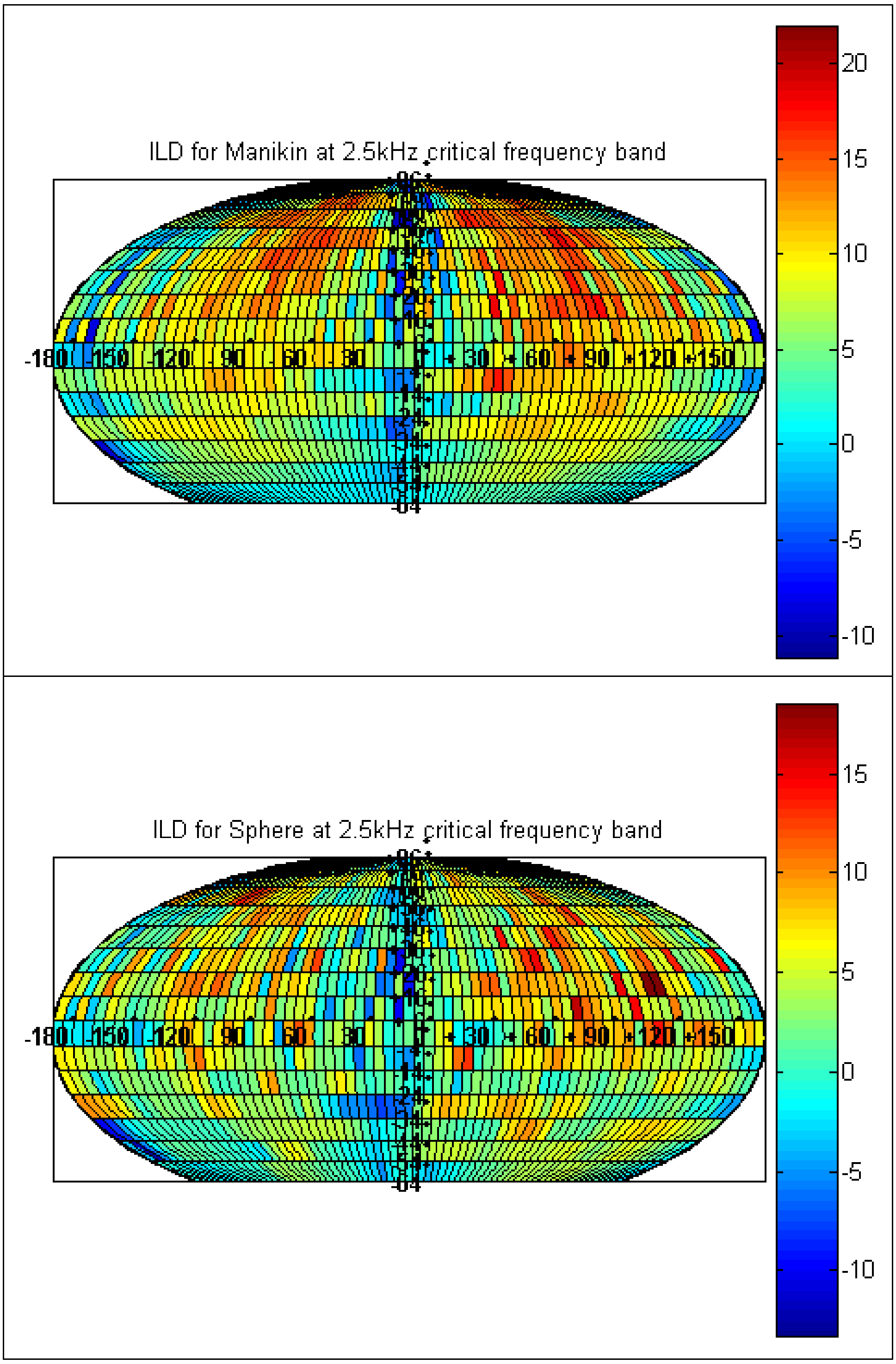
The Manikin is the only recording system that incorporates the influence on the localization due to the torso. Reflections from the torso can appear as interference between the direct and the reflected sound at the pinnae. These effects are likely to appear in the frequency range below 3 kHz [30]. Influence from the pinnae itself can be expected at frequencies above approximately 5 kHz. The only recording systems that, to any extend, resembles the pinnae are the Manikin and Friggo, possible this can be seen as diverging results compared to the other recording systems at frequencies above 5 kHz [30].

As for the Sphere, one can not presume it to be a good approximation of the human characteristics over the whole audible frequency range. At frequencies where the wavelength is comparatively short in comparison to the structure, such as the torso, head, and the pinnae, the correspondence to human hearing will be reduced [29]. In earlier studies a Sphere was found to be a reasonable approximation to a real head up to a frequency of around 2.5 kHz [29] in terms of frequency response characteristics. In terms of localization cues, ITD dominate ILD and spectral cues at low frequencies below approximately 1.5 kHz. Higher in frequency, ILD dominates the localization cues. Good correspondence to measurements with artificial heads was obtained for Sphere up to 2.5 kHz. Higher up in frequency, an influence on the localization from the pinnae can be expected and the ILD produced for sphere results in a large cone of confusion.

Results from the HRTF measurements were decided to be presented for the recording systems Manikin, Friggo and Sphere. The scales in the figures have individual ranges for each recording system. All results from the HRTF measurement show a specific zigzag pattern. This pattern has been found to correspond to the arrangement of loudspeakers at different elevations in the measurement rig (see Figure 22). The measurements for Friggo can be seen to have a shift due to incorrect adjustment along the median plane.

Interaural level difference

According to earlier studies, presented above, reasonably good resemblance could be expected between the three recording systems for frequencies below ~4 kHz (where the pinnae start to have a significant influence). In Figure 24 the ILD is presented at the 2.5 kHz critical frequency band.


ILD for Manikin at 2.5kHz critical frequency band
-180 -150 -120 -90 -60 -30 0 +30 +60 +90 +120 +150
20
15
10
5
0
-5
-10
ILD for Sphere at 2.5kHz critical frequency band
-180 -150 -120 -90 -60 -30 0 +30 +60 +90 +120 +150
15
10
5
0
-5
-10

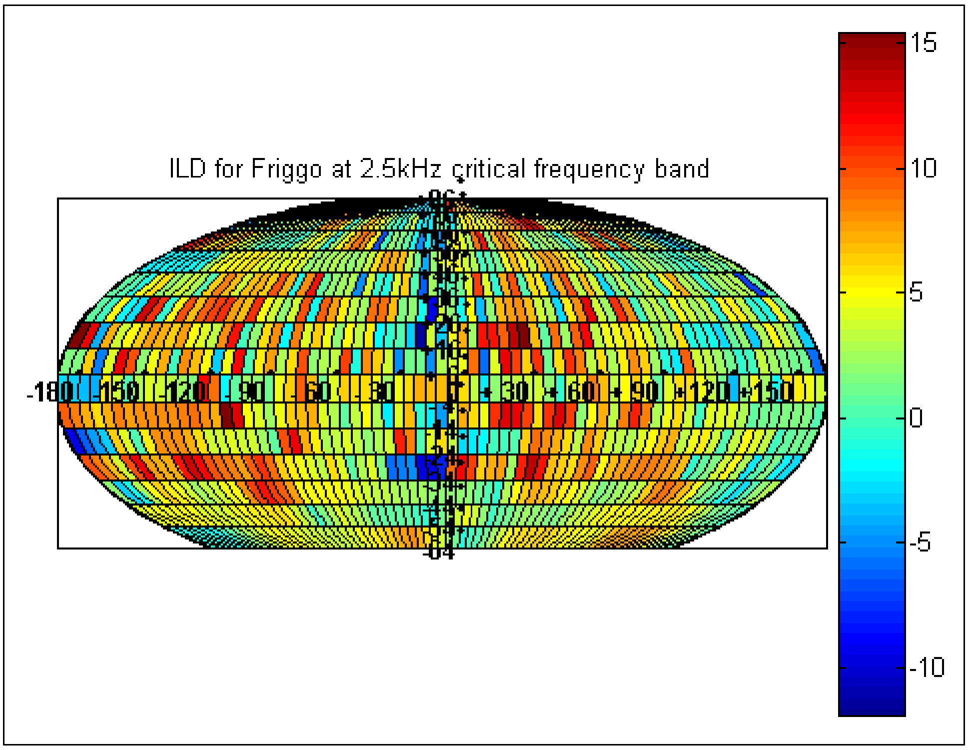


Figure 24: ILD [dB] for Manikin, Friggo and Sphere at the critical frequency band with center frequency of 2.5 kHz.

The ILD has a strong dependency of frequency and angle of incidence as seen in Figure 24 and the figures in Appendix F. Compared to measurements of ILD found in literature, our measurement fall into the same range [31]. It can be noticed that the range of ILD corresponds better for Sphere and Friggo, while the values for Manikin are higher. The Sphere and Friggo reach approximately 15dB, while for Manikin values reach 20dB. The ILD for Manikin is not symmetrical over elevation angles. The explanation for the increased ILD above the median plane may be due to torso reflections, which should be most prominent for positive elevations. The torso might also lower the levels of ILD for negative elevations because of scattering and interference between the direct sound and reflected sound from the torso.

In Figure 24 the ILD for Manikin appear to be more directional, i.e. more concentrated to the angle of incidence close to the ears. The same appearance can be seen up to a frequency range of 13.5 kHz (See Appendix F). Since the frequency range 2.5 kHz in Figure 24 is below where the pinnae starts to have an influence, the explanation is rather connected to the shape of the head and torso.

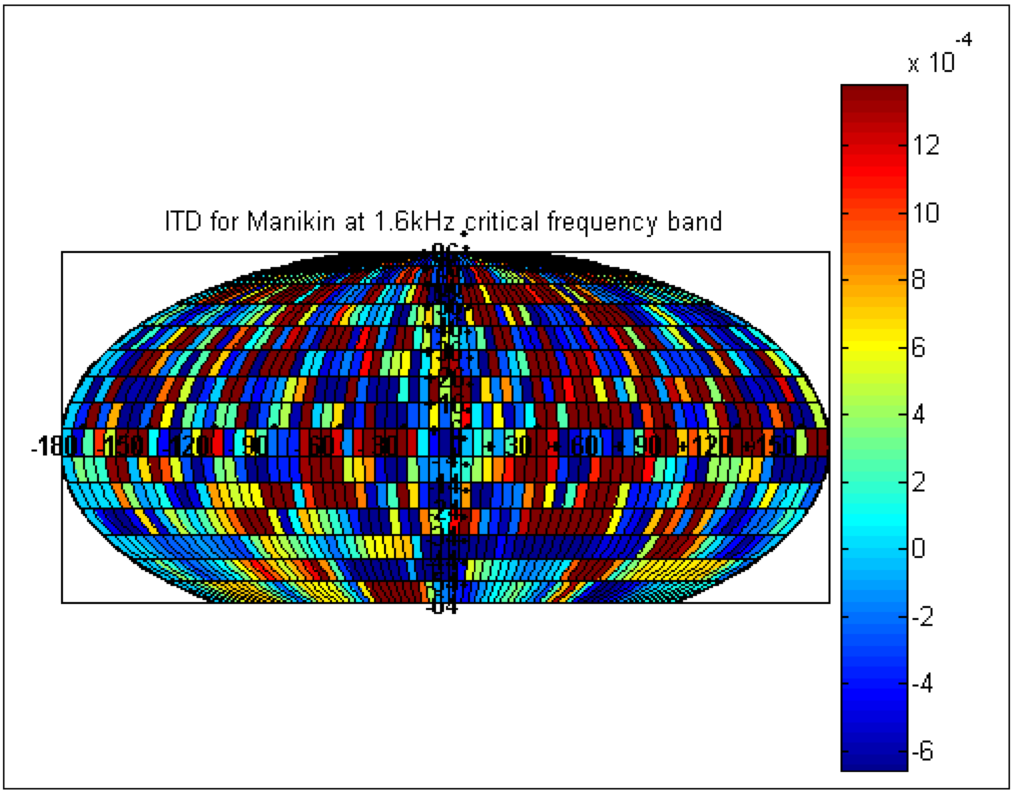
In the frequency range below 1.6 kHz, the principal diameter of the head shape of the measurement systems are smaller than the wavelength, which should indicate a relatively good correspondence between them. However, the same relations are shown as for 2.5 kHz (see Appendix F), the presumed similarities could not be found. Thus in this frequency range the ITD is dominating and ILD is less important as a binaural cue.

The ILD measured for Friggo (Figure 24) can be seen to be shifted in the direction of negative azimuth as mentioned earlier. Moreover, it can be noticed that Friggo measures higher ILD levels for angle of incidence with negative azimuth compared to the Manikin and Sphere.

Negative values of ILD can be seen in the figures of ILD. This is counter-intuitive, as it implies a higher pressure at the far side of the head than on the side closer to the sound source. Negative ILD can be measured on humans, and are often explained by pinnae notches. Thus, as the pinnae is quite simply modelled in all recording systems concerned in this study, including the Manikin, the main reason probably is lack of precision in the measurement setup.

Interaural time difference

In the frequency range below 1.6 kHz the ITD is the most important binaural cue for localization. Therefore, in Figure 25 ITD is presented for the 1.6 kHz critical frequency band.



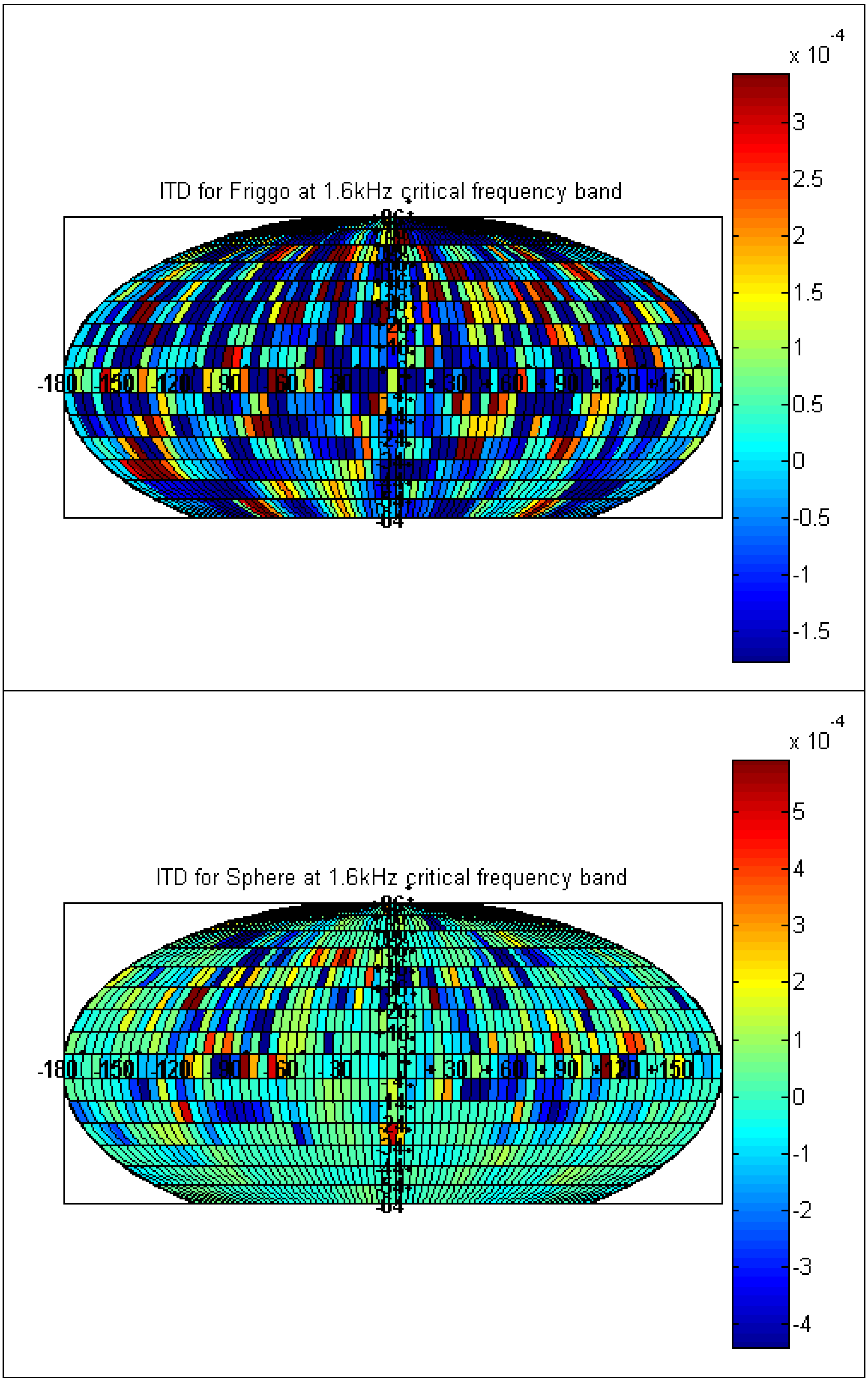


Figure 25: ITD [s] for Manikin, Friggo and Sphere at the critical frequency band with center frequency of 1.6 kHz.

The range of ITD used in Figure 25 is determined for each recording system individually over the entire measured frequency range. Based on that, it can be noticed that the range of ITD is larger for Manikin (up to 1.2 ms). An ITD of 1.2 ms corresponds to a sound path length of approximately 30 cm, which seems as rather long in comparison to the circumference of Manikin. These unrealistic results may have been caused by frequency dependent phase shifts introduced by the measurement system. Moreover, negative values of ITD can be seen in the plots. This does not agree with reality as it implies that the sound wave reaches the ear at greater distance from the source before it reaches the closer ear. Most likely this also is due to imprecision in the measurement of phase delay. Compared to studies of ITD in literature, one can expect ITD extending up to approximately 0.8 ms. This corresponds rather well to the results of ITD for Friggo and Sphere.

Judging from the appearance of the plots in Figure 25, Sphere shows ITD values remarkably levelled compared to Manikin and Friggo. This indicates that concerning ITD, the Sphere is less sensitive to changes in angle of incidence than the other two recording systems. The wavelength at 1.6 kHz correspond to the diameters of the head shapes of the recording systems. This clearly viewable characteristic of Sphere is rather surprising. Moreover, the ITD for Sphere in Figure 25, as well as in the entire measured frequency range (see Appendix F), shows ILD-values most focused to the angle of incidence close to the ears. Most likely this has to do with the smooth design of Sphere. Obstacles and cavities causes phase shifts dependent on the frequency and the angle of incidence, this behaviour appears in the figures of ITD for Friggo and Manikin.

The relation between the wavelength of sound and the size of the measured head shape seems to have a considerable influence on the ITD. In Figure 25, the ITD at 1.6 kHz for Manikin can be seen to be quite varying. However, already at the 2.5 kHz critical frequency band, the Manikin shows values of ITD considerable more focused to the areas close to the ears. It can be noticed, in correspondence with the figures of ILD, that the Manikin shows larger values of ITD for positive angles of elevation compared to negative (appear more clearly for frequencies higher in range, see Appendix F). This may be due to the relatively larger dimensions of the upper half of the Manikin.

The ITD results for Friggo deviate slightly over the measured frequency range compared to the other recording systems. Firstly it is not until at the 8.5 kHz critical frequency band where Friggo shows the levelled out ITD characteristics centered at the ears which is recognized at considerable lower frequencies for Manikin and Sphere. This can partly be explained by the size of Friggo which is smaller than the other recording systems. Thus, at frequencies above 3.4 kHz the wavelength approaches the size of the radius of Friggo and accordingly above this frequency the size should not serve as an explanation. Another deviation regarding Friggo is that negative values of ITD can be seen at 8.5 kHz critical frequency band for angles of incidence close to the ears, i.e. for angle of incidence where large ITD could be presumed. One explanation could be that sound waves are able to excite and travel within the polystyrene plastic and because of the high speed of sound in polystyrene, cause small values of ITD. This in combination with the possible influence from the measurement system discussed earlier could serve as an explanation. However, what causes this effect to be most prominent at specifically the 8.5 kHz critical frequency band and for certain angel of incidence is not explained from this. The explanation for this could rather be due to interference of sound waves at the ear on the far side.

Graphs of ILD and ITD can be found in Appendix F for a frequency range reaching from the 1.6 – 13.5 kHz critical frequency band. For frequencies above 4 kHz, where the cavities and obstacles of Manikin and Friggo would influence on the characteristics, some observations can be made. The range in which the ILD varies has a better correspondence between the different recording systems when the frequency is increased. It can be noticed from studying the figures that the maximum ILD is found in the 8.5 kHz critical band for all three recording systems. One could expect the ILD to increase with frequency for the entire measured frequency range and an obvious explanation for this is hard to find. However, the explanation may be found in the properties of measurement rig. The measured ITD for Friggo and Sphere increases for increasing frequencies. This agrees with what could be expected and implies that the phase shift caused by the recording systems increases with frequency and the travel distance (number of wave lengths).

4.3 Measurements of in-cab noise

Measurements were performed at Hällered proving ground to decide how well the measurement systems finally could be used for the purpose of recording in-cab noise at the passenger seat.

The measurement were done according to GDI 964-24 but included only the following driving cases:

- Full load acceleration
- Constant speed 85 km/h
- Low idle speed, without compressor working
- Low idle speed, with compressor working

The measurements were done during a day with stable weather conditions and medium wind. The temperature was 4°C.

4.3.1 Comparisons to free hanging microphones

The results from the In-cab noise measurements are presented in $L(A)_{eq}$ according to Volvo standard (method I). The $L(A)_{eq}$ is a time averaged A-weighted SPL. Narrow band spectra are studied to detect differences in the characteristics of the measurement systems.

Table 2 and Table 3 show the difference in averaged dBA level difference between the measurements performed with the measurement systems and those with free hanging microphones for all concerned driving cases (measurement system – free hanging microphones). The noise levels presented are energy averages of the two channels.

Table 2: Differences in overall levels between the measurement systems and the free hanging microphones. Recordings done at the passenger side.

	Full load acceleration 990-1620 rpm (level difference in dBA)	**Constant speed 85km/h (level difference in dBA)**
Friggo – Free mic.	-0.2	-0.3
Jecklin – Free mic.	0.6	1.1
Sphere – Free mic.	-0.6	-0.5
Wedge – Free mic.	0	0.4
Manikin – Free mic.	1.4	1.4

Table 3: Differences in overall levels between the measurement systems and the free hanging microphones. Recordings done at the passenger side.

	Low Idle 600 rpm without compressor working (level difference in dBA)	**Low Idle 600 rpm with compressor working (level difference in dBA)**
Friggo – Free mic.	-0.1	-0.2
Jecklin – Free mic.	0.8	0.8
Sphere – Free mic.	-0.6	-0.6
Wedge – Free mic.	0.6	0.5
Manikin – Free mic.	0.9	0.3

As can be seen in Table 2 and Table 3 the overall minimum differences between free hanging microphones and the measurement systems are obtained for Friggo (maximum of -0.3 dBA). Furthermore it can be noticed that Jecklin and the Manikin measure higher levels compared to free hanging microphones, while Friggo and Sphere measure lower levels. It can be seen that the maximum differences is obtained for Manikin (1.4 dBA).

Additional, a study of the repeatability of the measurements done with Headset was made. Table 4 and Table 5 show the average and maximum differences between the measurements performed with the Headset and those with free hanging microphones. The values in the tables are calculated from series of five recordings.

Table 4: Average level differences between Headset and free hanging microphones. Calculations based on five recordings done at the driver side.

	Full load acceleration 960-1620 rpm (level difference in dBA)	**Constant speed 85km/h (level difference in dBA)**
Average difference (Headset – Free mic.)	-1.2 ± 0.2	-0.6 ± 0.1
Max difference (Headset – Free mic.)	-1.4	-0.8

Table 5: Average level differences between Headset and free hanging microphones. Calculations based on five recordings done at the driver side.

	Low idle without compressor working 600 rpm (level difference in dBA)	**Low idle with compressor working 600 rpm (level difference in dBA)**
Average difference (Headset – Free mic.)	0.2 ± 0.1	-0.2 ± 0.1
Max difference (Headset – Free mic.)	0.3	-0.3

A large difference between measurements performed with Headset and free hanging microphones for the driving case full load acceleration can be noticed in Table 4. The Headset measures levels lower than free hanging microphones in both the full load acceleration and constant speed driving cases.

Lower average- and maximum differences can be noticed in Table 5 for both driving cases compared with values in Table 4. It can be noticed that the average differences are positive for low idle without compressor working but negative for the case with compressor working. However for all driving cases, the variation in the measurements of ± 0.1-0.2 dBA is small and accordingly the repeatability when measuring with Headset is relatively good. Moreover, the variation in measured level might be caused by differences in the driving conditions rather than by the measurement equipment.

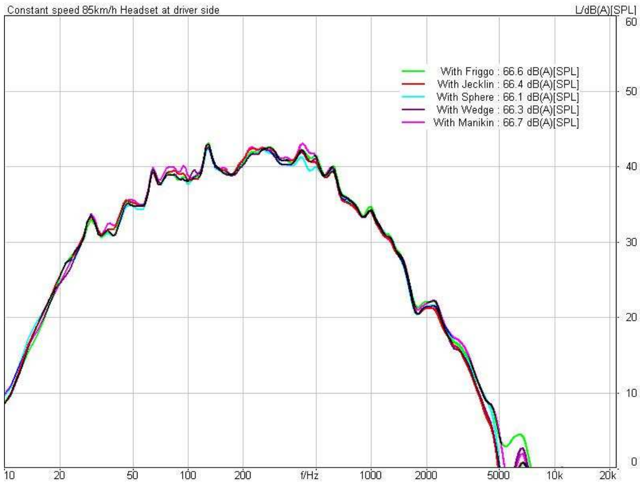
4.3.2 Spectral analysis

Constant speed 85 km/h:

In Figure 26 A-weighted spectra from the constant speed 85km/h driving case are presented for all measurement systems and free hanging microphones. Energy averages of the two channels are observed.

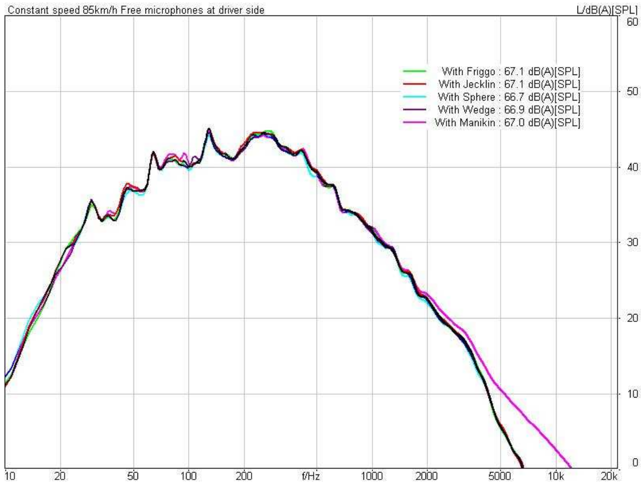
<table>
<tr><th>Binaural/ Similar-binaural recording System</th><th>Free hanging microphones</th></tr>
<tr><td>

Driver side – Headset



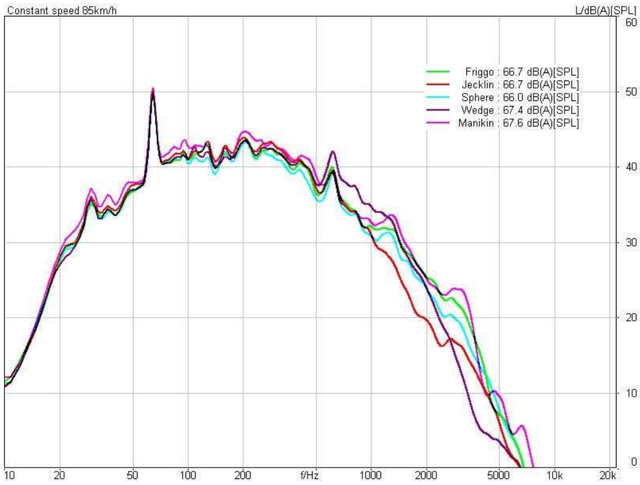
</td><td>

Driver side – Free hanging microphones



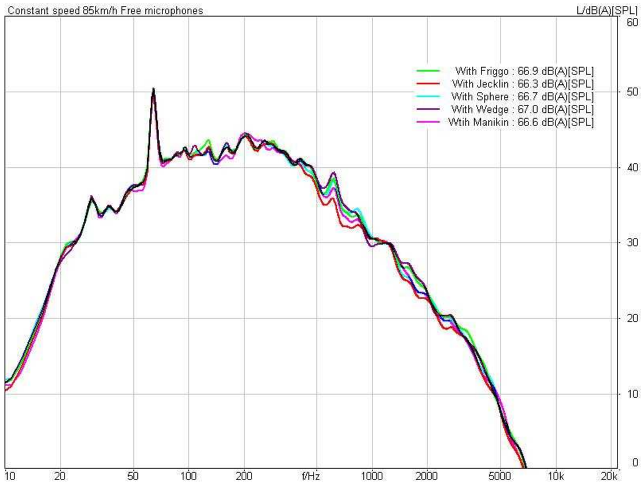
</td></tr>
<tr><td>

Passenger side – Measurement systems



</td><td>

Passenger side – Free hanging microphones



</td></tr>
</table>

Figure 26: A-weighted spectra from the constant speed 85km/h driving case. Average of the two channels is presented.

Figure 26 shows that Jecklin's response decreases more quickly above 1 kHz than the other measurement systems. Wedge is higher in level between 0.8-1.2

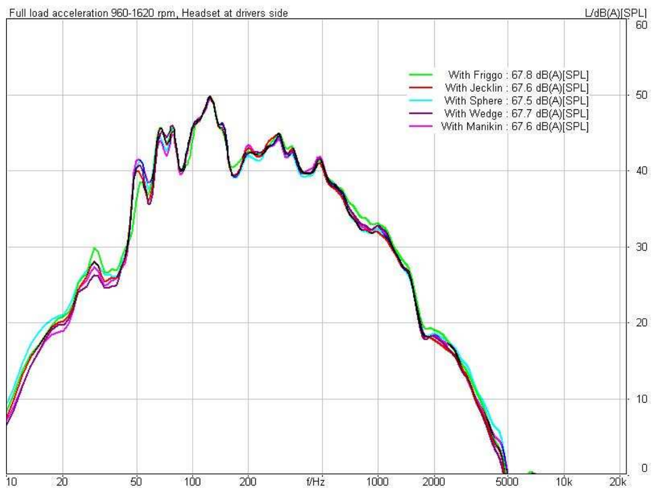
kHz and decreases prominent above about 3 kHz in comparison to the other systems. Furthermore, one can see how well the shape of both Sphere and Friggo curves coincide with Manikin, especially the peaks at approximately 1.2 kHz and 3 kHz. Moreover, shifting the curves of Sphere and Friggo a few decibels up would generate a very good overall correspondence with Manikin both in shape and level.

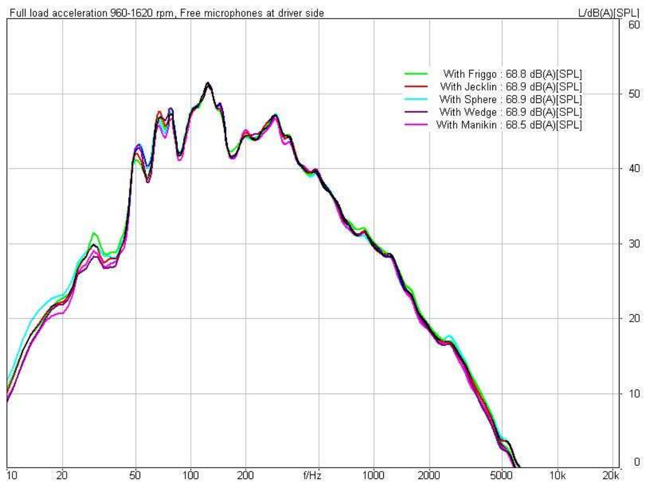
In accordance with what could be concluded in the comparisons of levels in Table 4 and Table 5, the variations of the measurements with Headset are small.

Full load acceleration:

A-weighted spectra from the full load acceleration are presented for all measurement systems and free hanging microphones in Figure 27. The respective left and right channels of the measurement systems and the free hanging microphones are averaged together.

Binaural/ similar-binaural recording System	**Free hanging microphones**
Driver side - Headset	***Driver side – Free hanging microphones***





Passenger side – Measurement systems | **Passenger side – Free hanging microphones**

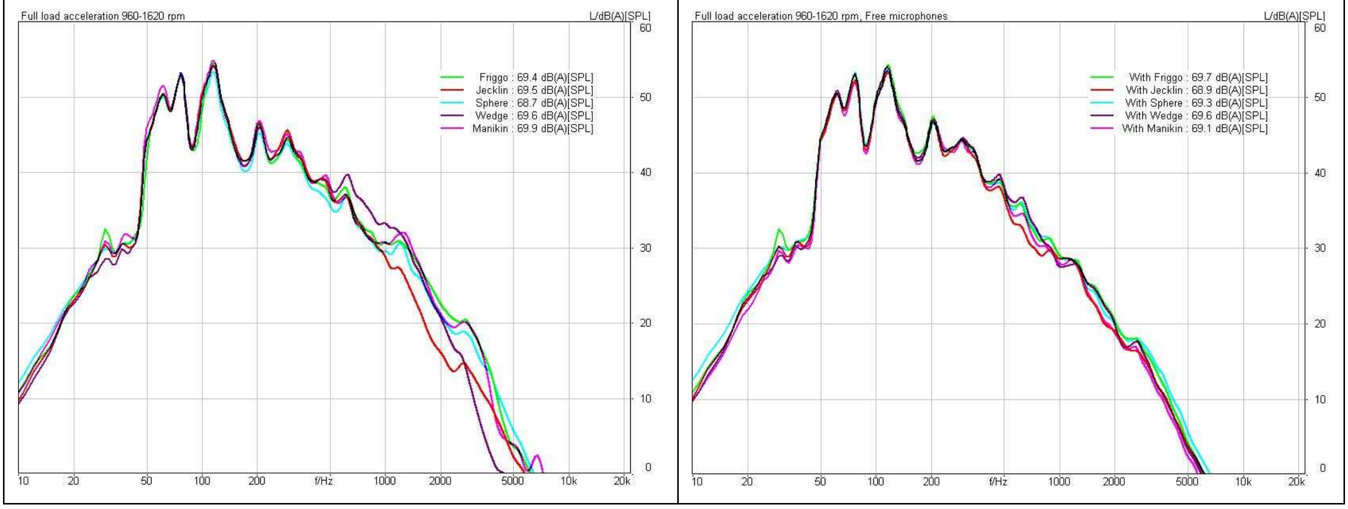


Figure 27: A-weighted spectra from the full load acceleration. Average of two channels is presented.

As shown in Figure 27, the spectra from the full load acceleration case have a slightly different shape compared to the constant speed case. However, the mutual relations of the measurement systems within the test group are similar.

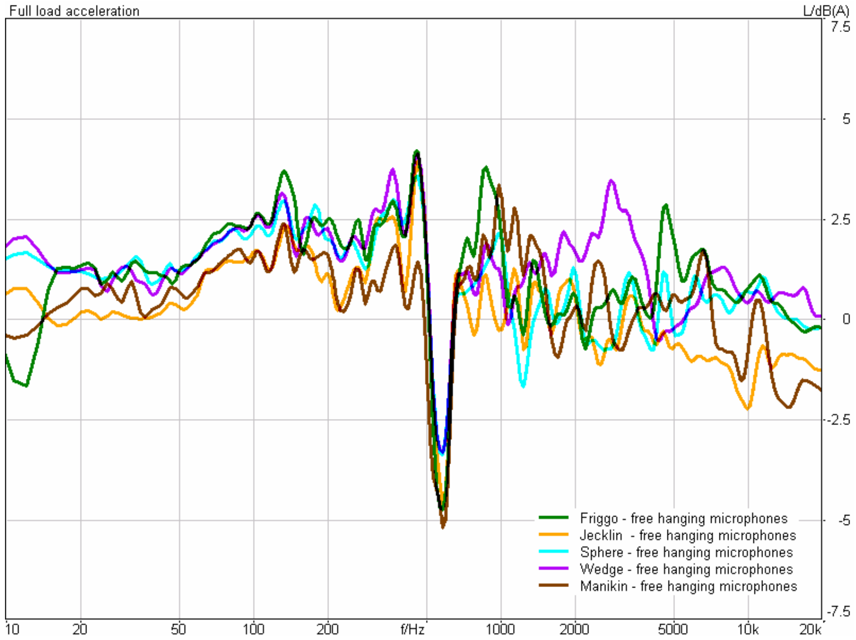
Studying the frequency spectra of the measurements at constant speed 85 km/h and full load acceleration one finds that both Friggo and Sphere have similar response as Manikin while Wedge and Jecklin show lower levels in the high frequency range.

The difference in measured level between the measurement systems and the free hanging microphones was calculated. In Figure 28 this difference is presented for the driving cases: full load acceleration and constant speed 85km/h. The reason for plotting this is to estimate a filter for processing the data measured with the measurement systems to better correspond to that from measurements with free hanging microphones.

Difference in measured level between measurement system and free hanging microphone, average of both channels

($L_{eq,\ measurement\ system}$ - $L_{eq,\ free\ hanging\ microphones}$)

Full load acceleration



Constant speed 85 km/h

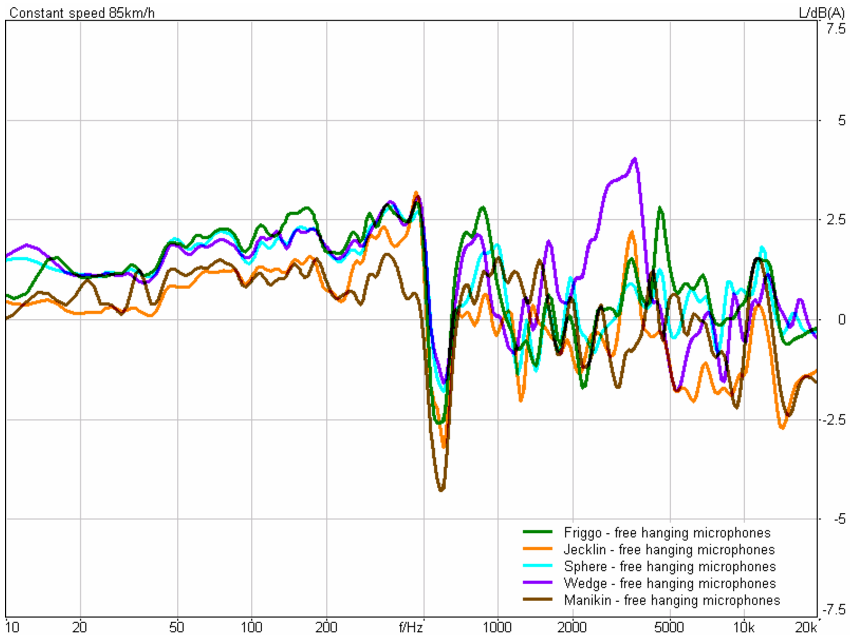


Figure 28: Difference in measured L_{eq} between the measurement systems and free hanging microphones for full load acceleration and constant speed 85km/h. Average of two channels is presented.

A very similar shape of the driving cases for each measurement system showed in Figure 28 is desired. This would mean that a filtering of the recordings done with the measurement systems could give reasonable results agreement with measurements using free hanging microphones for all driving cases. Moreover, if agreements are found for the different driving cases this also should indicate similarities in the measured sound fields. However, Figure 28 shows quite clear differences in shapes both for the two driving cases and for the measurement systems.

In Figure 29 the difference in measured level between Friggo and the free hanging microphones are presented for all driving cases.

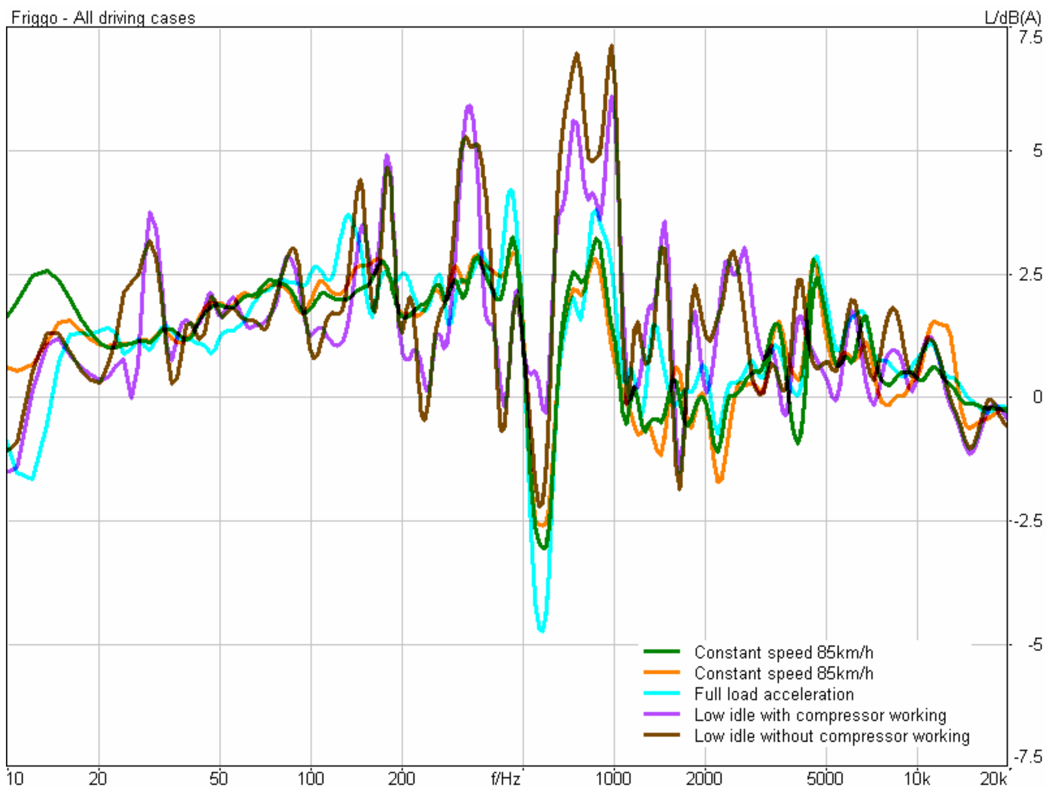


Figure 29: The level differences between Friggo – Free hanging microphones, all driving cases.

As seen in Figure 29 the differences are approximately within ±2.5 dB for both the full load acceleration and the constant speed driving cases. This range is somewhat larger for the two low idle driving cases, mostly because of the pronounced peaks that they have in common. These peaks occur for example at 30, 90 and 180 Hz, where the measurement systems measured higher levels than free hanging microphones. The peaks in the low frequency range are most likely due to the mode shapes of the compartment interior. With a constant source such as that at low idle the measurements, the microphones are more sensitive to positioning than in a more fluctuating sound field. Still, Friggo measured overall levels in good correspondence to free hanging microphones for all driving cases, see Table 3.

5 Perceptual sensation of the measurement systems

The recording systems used in this study were evaluated in several listening tests. In the first test, the binaural quality of the recordings was investigated by examining how well the test participants could make use of the cocktail party effect. Questions regarding the ability of sound source separation in a recorded sound field consisting of several sources were also asked. In the second test the binaural quality of the recordings were investigated using a localization test. Only the directional location in the horizontal plane was examined.

Throughout the test integer response scales without any mean number was used. The reason why the mean value was not written in the answering scales, was to avoid over representation of answers in that region.

The test participants performed the test one by one. The listening test was designed in ArtemiS analysis software and was played back from a laptop through headphones, type Sennheiser HD 600. In the listening test regarding in-cab noise the Sound Quality Simulator at Volvo was used for reproduction. The instructions and questions used in the test are found in Appendix G.

The statistics used in the analysis of the listening tests can be found in Appendix I.

5.1 Localization

To investigate the psychoacoustical spatial characteristics of the measurement systems, a localization test was used. The purpose of this test was only to examine the directional localization; accordingly the distance to the sound source was constant, and only the horizontal plane was considered. Recordings were done in an anechoic environment placing the measurement system on a rotating table. A recording of a male counting “one two” was used as basic stimuli and the angles of incidence used are presented in Figure 30.

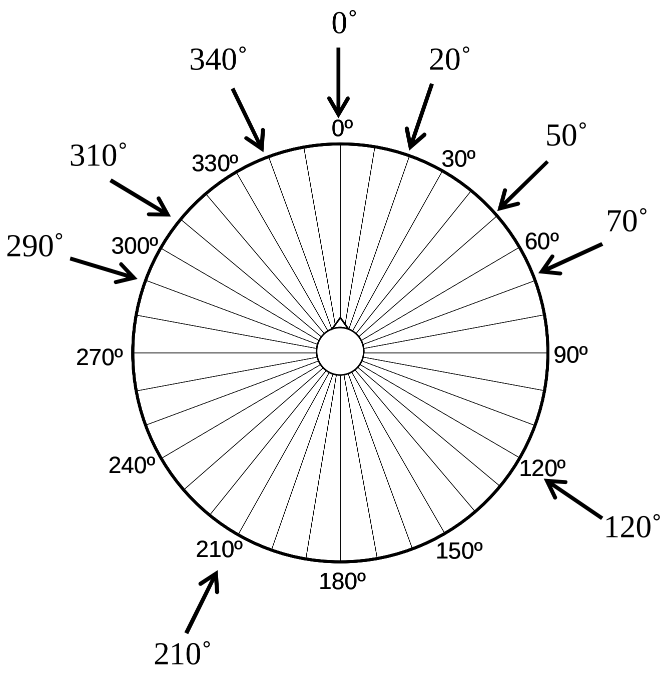


Figure 30: Angles of incidence chosen for the localization listening test.

As can be seen in Figure 30 it was decided to investigate frontal incident more thoroughly than sounds arriving from the back. The reason for this was mainly to reduce the number of samples in the test. Still, in the chosen design it should be possible to estimate the front-back confusion that may arise in the recordings. The angles for frontal incident were chosen closely positioned to be able to evaluate the localization performance of the measurement systems for small angle differences. To further decrease the number of samples in the listening test, only four measurement systems were chosen for the localization test; Friggo, Sphere, Wedge and Manikin. To possibly improve the ability to decide if the sound is coming from the front or back uncorrelated A-weighted white noise was added to the recordings. The playback order was randomly chosen for all systems and angels. The order was not changed among the test participants. All 13 participants listened to all measurement systems and angles. At the end of the listening test some additional questions were asked regarding the recordings at 0˚, frontal incident. The questions concerned the experienced source distance and if the sound was experienced as coming from inside or outside the head. The sound pressure levels in playback were calibrated manually by tuning the levels of the recordings done at frontal incidence to be perceived as equally loud.

The stimulus was reproducing with a loudspeaker at 5.30 m distance from the measurement systems. The loudspeaker and the measurement system were at the same lateral level.

Figure 31: Measurement setup from the localization measurement in the anechoic chamber. Stimuli were reproduced from the loudspeaker outlined in the figure.

5.1.1 Results of the Localization test

Human hearing is able to detect deviations of about 2° in the horizontal plane from a forward direction with sinusoidal signals. For sound sources facing one of the ears the localization ability is reduced and deviations of about 10° are detectable [2]. In the following analysis the localization error, which is calculated as the difference between the actual direction and the answered direction was regarded as a random variable. The distribution was assumed to be a normal distribution. Unbiased estimators for sample mean $\bar{x}$, and sample variance s^2, were calculated and sample standard deviation, s, of the localization error were calculated as the square root of the estimated variance [16]. A 95% confidence interval centered at the estimated global mean was introduced having a width equal to 3.92 times the standard deviation.

In Figure 32 all reported answers from the listening test are presented together with actual sound directions for the four studied measurement systems.

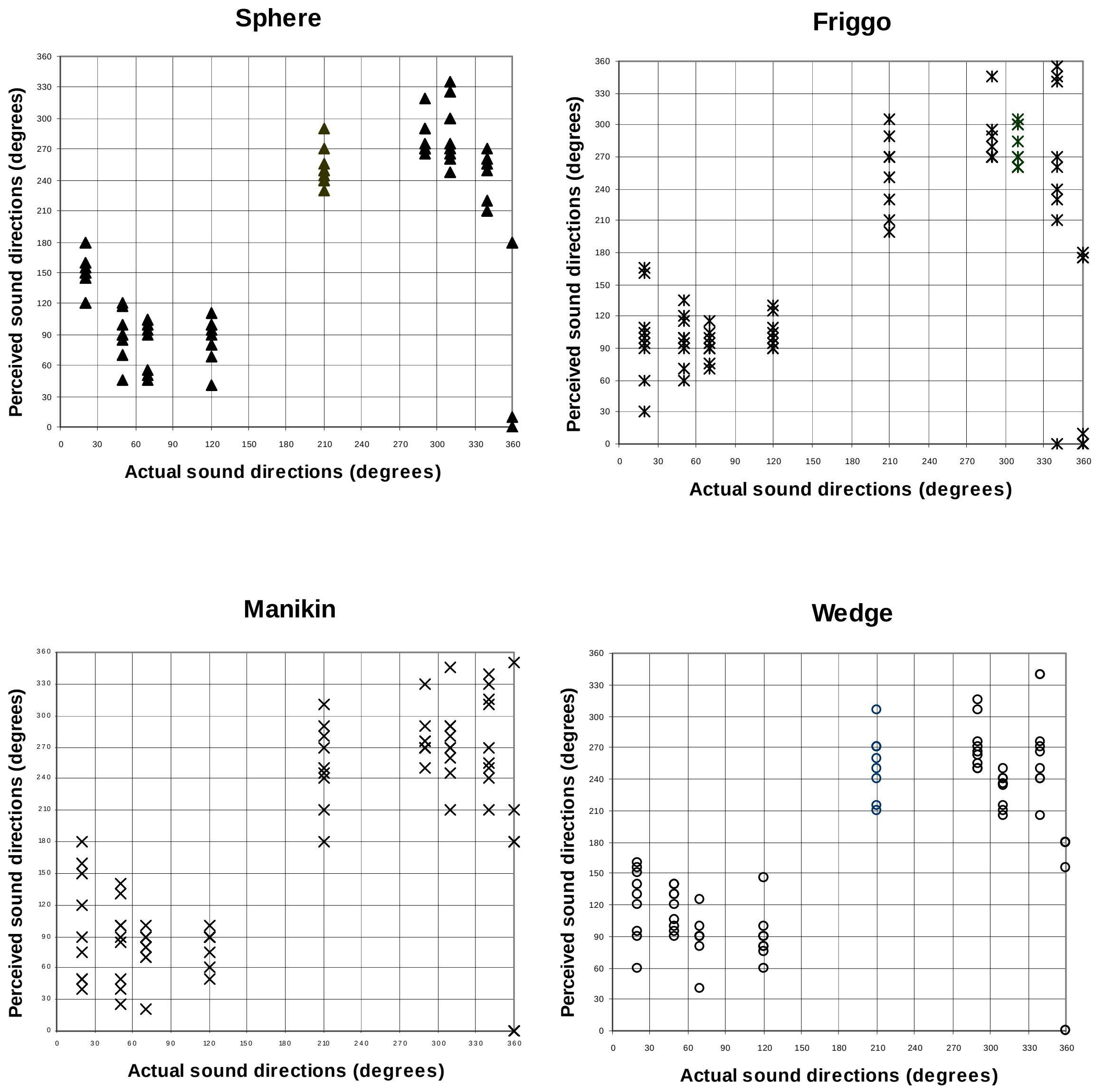


Figure 32: Judged directions in the horizontal plane for the four measurement systems under study: reported sound directions vs. actual sound stimuli directions.

Figure 32 shows that the front-back ambiguity is a considerable problem in the test. A special analysis to uncover this problem has been performed, where the localization error was calculated to be the smallest difference between the perceived direction and the stimulus direction. Table 10 gives the results from this analysis, where front/back localization error is a percentage measure of how many times the mirrored position was used compared to the total number of positions (also called front-back confusion). The 95% confidence intervals centered at the respective measurement system's mean value were calculated for the case of front-back confusion mirrored and not mirrored (normal).

Table 6: : 95% confidence intervals centered at the measurement system's mean value for normal and front/back compensated localization errors. Front/back reversals are mirrored back to positions closer to stimuli directions.

	Normal	Mirrored	Front/back localization error
Wedge	[-184.2˚; 144.5˚]	[-53.1˚; 44.3˚]	68%
Sphere	[-168.3˚; 133.3˚]	[-49.7˚; 50.4˚]	54%
Friggo	[-139.3˚; 108.2˚]	[-55.8˚; 45.3˚]	47%
Manikin	[-130.7˚; 123.7˚]	[-45.7˚; 47.6˚]	44%

A highly pronounced bias in the front-back confusion is clearly visible in Table 6. The least percentage of mirrored answers was obtained for the Manikin, the smallest 95% confidence interval for mirrored answers were also obtained using the Manikin. However, the results regarding the front-back confusion are so poor and closely positioned that it suggests that the listeners in the test reported the front-back perception by mere chance. Moreover no conclusions of which measurement system that performed best, with respect to front-back confusion, can be drawn.

Figure 34 shows the mean values of the front/back compensated localization errors and associated standard deviations for each measurement system. Actual stimuli directions are plotted on the x-axis. In Table 7 statistical characteristics used in the Figure 33 can be found. Positive localization error equals an underestimated angle and consequently a negative localization error equals an overestimated angle of perception.

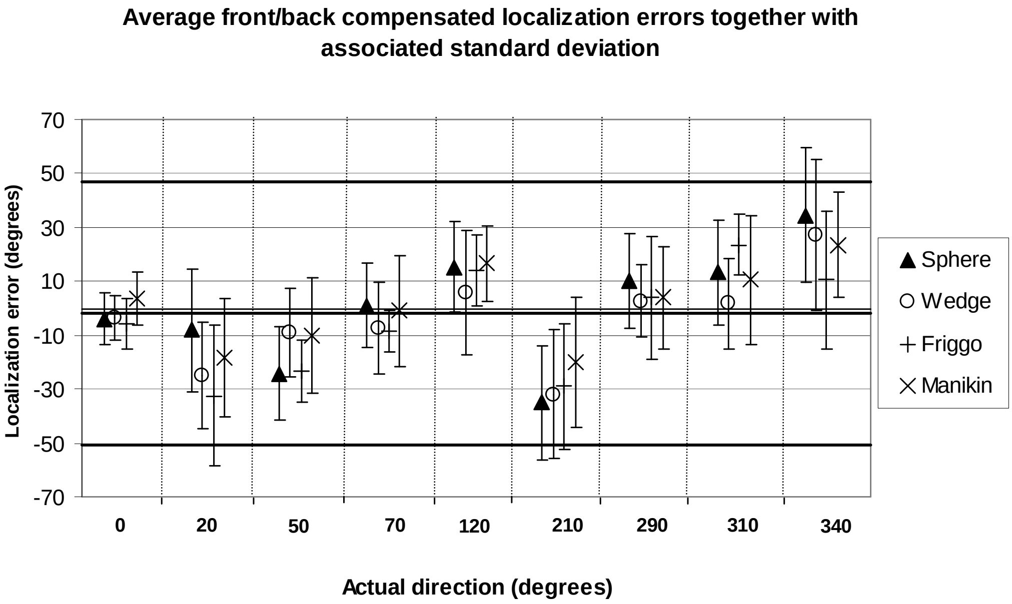


Figure 33: Mean values of the localization errors with standard deviations for the four concerned measurement systems plotted over actual directions. The black horizontal line indicates the global average localization error and the red horizontal lines indicate its associated 95% confidence interval. Values are compensated for front/back localization errors.

Table 7: Statistical characteristics for the localization error obtained for all answers in the localization listening test.

average error ($\bar{x}_{global}$)	-2.1°
Standard deviation (s)	24.9°
95% confidence interval	[-51.0° ; 46.8°]

To conclude if any of the measurement systems could be considered as performing better than the others an ANOVA was conducted. However, the ANOVA showed that no measurement system could, at any relevant confidence level, perform better than the others in the localization test.

5.2 Source separation ability

Recordings were performed in the lecture room of Applied Acoustics at Chalmers to uncover the source separation ability. This was done in a comparison test in which the Manikin was used as reference. The test participants were asked to judge if their ability to separate sound sources, i.e. to perceive the different locations of sound sources in a recorded sound field, was better or worse for sound samples recorded with the measurement systems compared to a reference. An interactive listening test design was used which allowed the test participants to listen to the samples an arbitrary number of times. The advantage of this system is that it reduces the influence of the playback order of the measurement systems. All measurement systems except the free hanging microphones were considered in the test.

Five loudspeakers were used to reproduce speech, placed at various distances in the room (9 x 6 m) and at various directions relative from to the measuring point. To create a murmuring acoustical environment the stimuli chosen were Danish, American English and British English voices both male and female talking about various subjects. The speech played back in the test where uncorrelated and adjustments in levels where done subjectively in order to have a separation ability in the measurement. The same sound was recorded for all measurement systems.

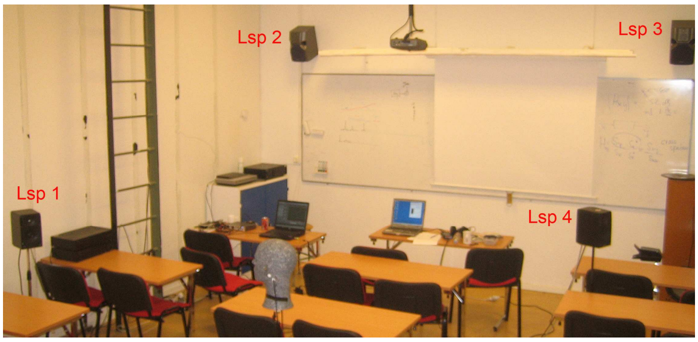


Figure 34: Measurement setup, seen from the rear of the room. Friggo is placed in the measuring location.

Figure 35: Measurement setup, seen from the front of the room. Friggo is placed in the measuring location.

5.2.1 Results of the Source separation ability test

Figure 34 shows the results regarding the source separation of the measurement systems in comparison with the Manikin. The scale ranged from 1 (much lower source separation ability) to 6 (much higher source separation ability) and a rate of 3.5 implied equal source separation ability. Only one recording position was used in the listening test.

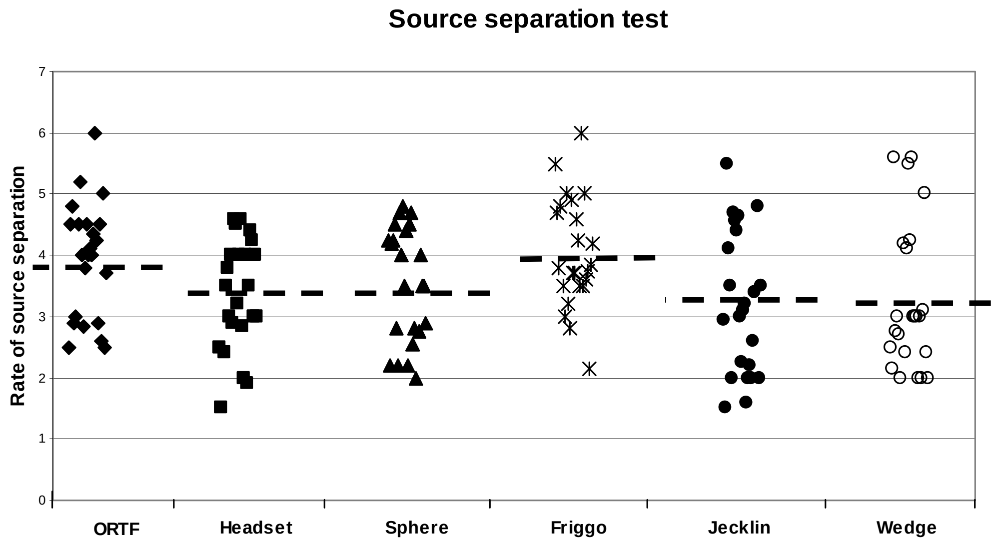


Figure 36: Ratings of the source separation ability in comparison to the Manikin presented together with mean values (dashed-lines).

In Figure 36 the highest mean value of the source separation ability can be seen for Friggo. Friggo and ORTF were the only measurement systems that were the test persons reported as providing a better source separation ability than the Manikin, however this is not statistically ensured. The lowest mean value was received for Wedge. Great variances can be noticed in the answers.

ANOVA showed significant difference in source separation ability $(F = 2.76, p < 0.05)$ for the measurement systems in comparison to Manikin.

From Tukey's multiple comparison test (see Table 10) with range statistic, q (6.132) = 4.1 and *Tukey's limit*: 0.9 no significant differences was found.

No statistically significant differences could be found between the recording systems in the listening test. However, Friggo received a mean rate that differed from Jecklin with the same amount as the Tukey's limit.

Table 8: Differences between mean values of the answers in the listening test regarding the source separation ability. Significant differences are enframed. μ_{tot} is the average response and (μ_i – μ_j) the difference between the column μ_{tot}- row μ_{tot}

	ORTF	Sphere	Jecklin	Headset	Friggo	Wedge	
μ_{tot}	3.9	3.5	3.2	3.4	4.0	3.3	
(μ_i – μ_j)	*	0.4	0.7	0.6	-0.1	0.7	ORTF
		*	0.3	0.2	-0.5	0.3	Sphere
			*	-0.2	-0.9	-0.1	Jecklin
				*	-0.7	0.1	Headset
					*	0.8	Friggo

5.3 Cocktail party effect

The cocktail party effect was investigated because of that the ability to use it in playback, to great extend, relies on the binaural recording qualities of the recording systems. The cocktail party effect was examined by studying the ability to perceive a certain combination of numbers in the background noise of human speech. The reason for using human speech is to achieve a well recognized acoustical environment.

The recordings used for examining the cocktail party effect were done in the lunch room at Volvo. In total six loudspeakers at various distances, heights and directions from the recording systems were used. Five loudspeakers reproduced speech, while the sixth reproduced a set of numbers. The same loudspeaker was used in all recordings to reproduce the combination of numbers. The locations and levels of all loudspeakers were held constant for all recording systems. The recordings were done having the recording system in the same position and directed equally in all recordings. The test participants were asked to write down the numbers that were spoken from a male voice. The combination of numbers was repeated once for every system. Different combinations of the numbers (1-9) were used for the different recording systems. The order of which the different recording systems were played back in the test was randomly chosen and arranged into six test groups. To further reduce the influence of the playback order of the recording systems, all the test participants were able to familiarize themselves with the sound by listening to a recording done with free hanging microphones which were not included in the test. A total of 23 subjects participated in the test, about four persons in each test group.

Besides writing down the interpreted number combination, the test participant was also asked to mark the experienced intelligibility of the male voice on a scale ranging from one to six.

Figure 37: Measurement setup of the cocktail party effect recording, two loudspeakers is outlined in the figure.

5.3.1 Results from the perception of numbers

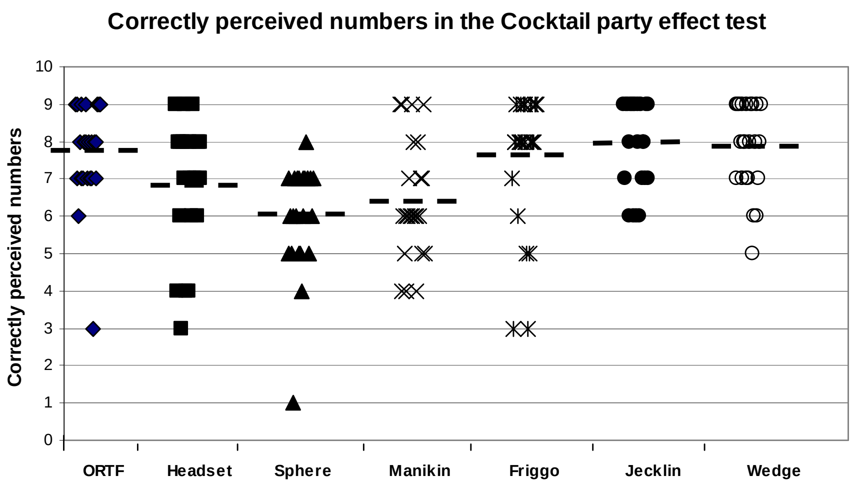


Figure 38: All listening test answers from the perception of numbers in the cocktail party effect test presented together with the mean value. The maximum numbers able to perceive was nine.

Jecklin received the highest mean value of correctly perceived numbers. The resulting mean values for Wedge, Friggo and Wedge were all very close in level just below Jecklin. Sphere got the lowest mean value in the test. Large spreads can be noticed for all recording systems, particularly for Headset, Manikin and Friggo.

To be able to discover if the playback order had an influence, we studied the answers divided into test groups. The result can be found in Figure 39.

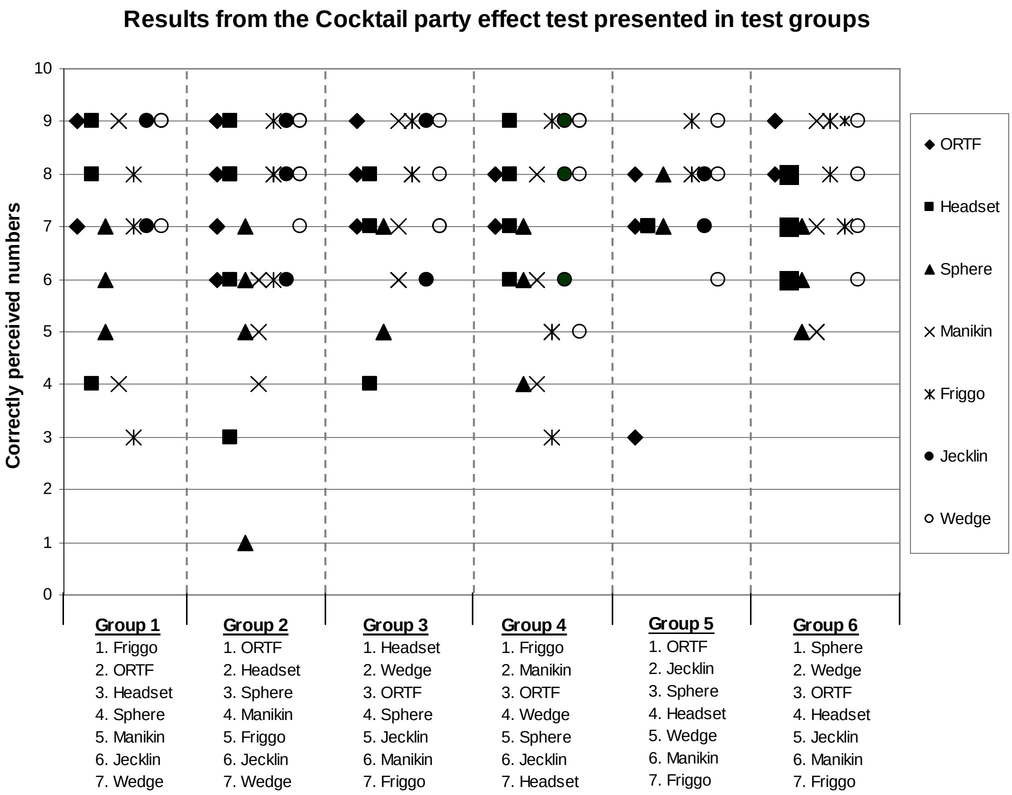


Figure 39: All answers from the perception of numbers in the cocktail party effect test presented in test groups.

Figure 39 suggests that it can be suspected that some answers are affected negatively by early play order positions. For example unexpected answers can be seen for Friggo in test groups 1 and 4. In test group 5 one outlier for ORTF can be noticed. These outliers for ORTF and Friggo can presumably be explained by them being the first measurement system in the playback. Further investigation neglecting the outliers can be found in a later section.

From ANOVA, significant difference was found for the reported numbers $(F = 5.94, p < 0.05)$ for the measurement systems.

Further, a test according to Tukey's procedure was done. The differences exceeding the Tukey's limit are marked in red. To calculate the Tukey's limit the range statistic $q_{k,\,v}$ for $k = 7$ mean values and $v = 154$ degrees of freedom at a significance level of 0.05 were used. The range statistic can be found in tables, *q (7.154) = 4.2.*

Tukey's limit: *1.3*

From Tukey's multiple comparison (Table 9) it was found with 95% confidence that:

- ORTF, Friggo, Jecklin and Wedge perform better than Sphere.

- Jecklin and Wedge perform better than Manikin

Table 9: Differences between the mean values of the measurement systems for listening test regarding the perception of numbers. Significant differences are enframed. μ_{tot} is the average response and (μ_i – μ_j) the difference between the column μ_{tot}- row μ_{tot}

	ORTF	Headset	Sphere	Manikin	Friggo	Jecklin	Wedge	
μ_{tot}	7.7	6.9	6.0	6.4	7.5	8.0	7.9	
(μ_i – μ_j)	*	0.8	**1.7**	1.3	0.2	-0.3	-0.2	ORTF
		*	0.9	0.4	-0.7	-1.1	-1.0	Headset
			*	-0.4	**-1.5**	**-2.0**	**-1.9**	Sphere
				*	-1.1	**-1.5**	**-1.4**	Manikin
					*	-0.4	-0.4	Friggo
						*	0.1	Jecklin

5.3.2 Results from the intelligibility of numbers

Ratings of the intelligibility in the cocktail party effect test are presented in Figure 40 and Figure 41. In Figure 41 the results are divided into test groups. The scale used ranged from 1 to 6, a rating of 1 meant that the numbers were hardly intelligible, 6 meant that the numbers were clearly intelligible.

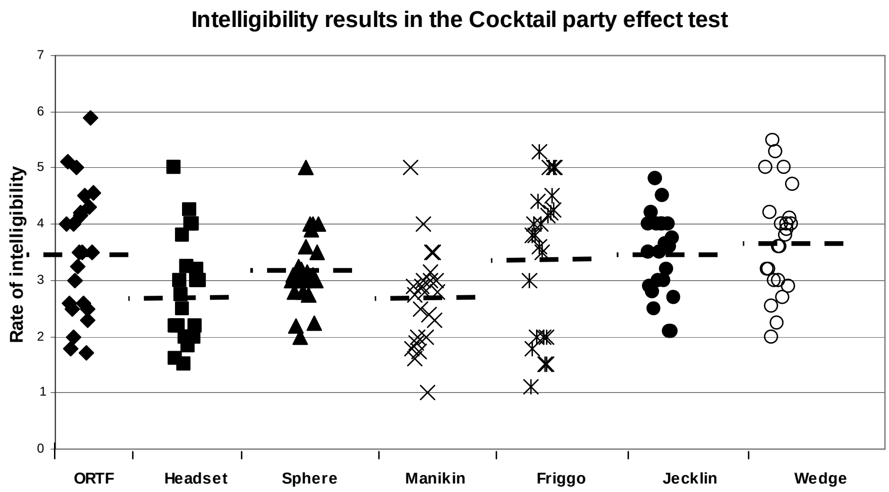


Figure 40: Ratings of intelligibility in the cocktail party effect test presented together with mean values.

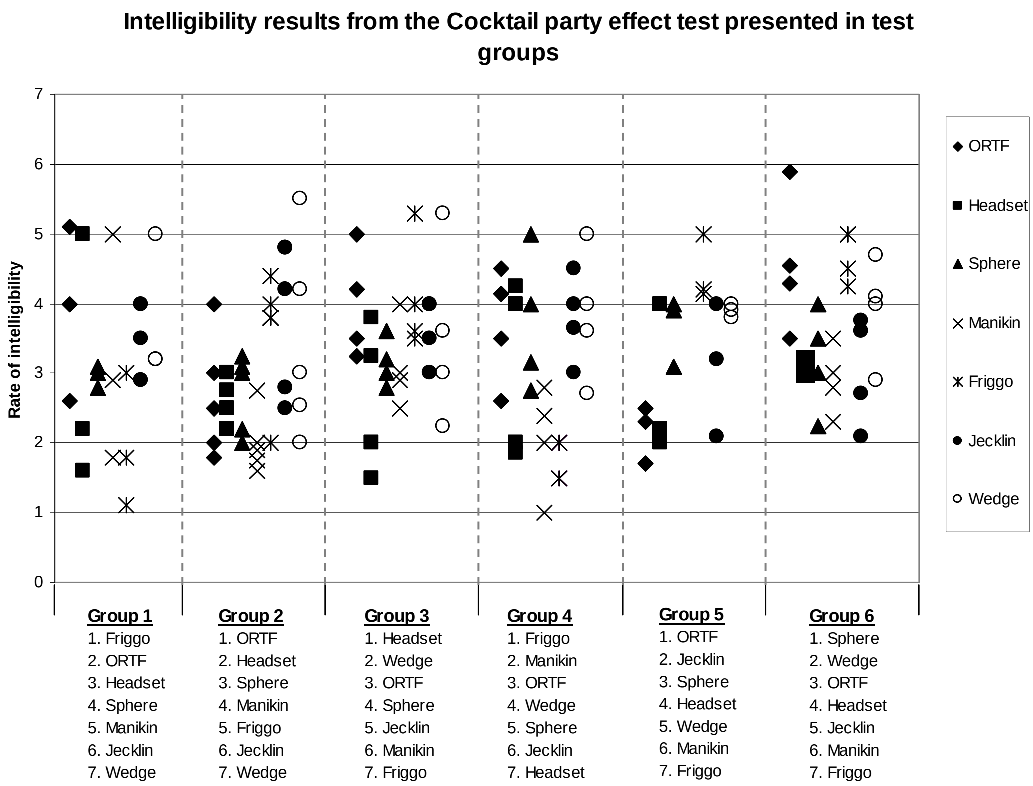


Figure 41: Ratings of intelligibility in the cocktail party effect test presented in test groups.

When studying the results in Figure 40 and Figure 41, one can note a considerable variation in the answers. It could be presumed that a good result from the perception of numbers test would imply a good performance of the recording system in the intelligibility test as well. However, the results from the two tests were found not to be entirely correlated. In agreement with the perception of numbers, the mean values for ORTF, Friggo, Jecklin and Wedge are closely levelled. Thus, in contrary to the test of perception of numbers, where Jecklin got the highest mean value, Wedge received the highest ratings of intelligibility. The lowest mean value in this test was received for Manikin and Headset got a mean value just slightly higher. It can be noticed that Sphere performed better in this test than in the test of perception of numbers. Apparently, in spite of the poor result in previous test, the intelligibility with Sphere was experienced as better relative to the other recording systems.

From ANOVA, significant difference was found for intelligibility of number combinations $(F = 2.158, p < 0.05)$ for the measurement systems.

From Tukey's multiple comparison test (see Table 10) with range statistic: q (7.154) = 4.2 and *Tukey's limit*: 0.9, it was found with 95% confidence that:

- Wedge performs better than Manikin.

Table 10: Differences between the mean values of the measurement systems for the listening test regarding the intelligibility of numbers. Significant differences are enframed. μ_{tot} = average response and $(\mu_i - \mu_j)$ = the difference between the column μ_{tot}- row μ_{tot}

	ORTF	Headset	Sphere	Manikin	Friggo	Jecklin	Wedge	
μ_{tot}	3.5	2.8	3.2	2.7	3.5	3.5	3.7	
$(\mu_i - \mu_j)$	*	0.7	0.3	0.8	0.1	-0.0	-0.2	ORTF
		*	-0.4	0.1	-0.7	-0.7	-0.9	Headset
			*	0.5	-0.3	-0.3	-0.5	Sphere
				*	-0.8	-0.8	$\boxed{\mathbf{-1.0}}$	Manikin
					*	-0.1	-0.3	Friggo
						*	-0.2	Jecklin

5.3.3 Results, disregarding deviant answers

In Figure 42 the answers from the perception of numbers is shown again. Some encircled deviant answers can be found and in many cases these deviants can be concluded to be connected with an early position in the playback order of the samples. The playback order position for the outlying answers is outlined in the figure. To estimate the influence of outliers these were removed from the analysis and a new ANOVA was done.

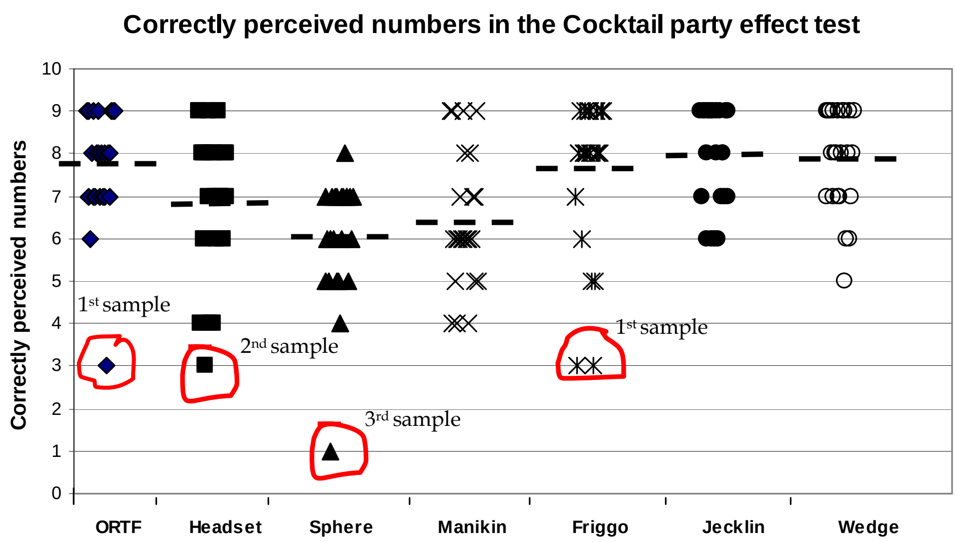


Figure 42: Answers from perception of numbers, cocktail party effect listening test. Deviant answers are circled with red and the playback order position is outlined.

Mean values and standard deviations of the answers in the part of the listening test regarding the perception of numbers are presented in Figure 43,.

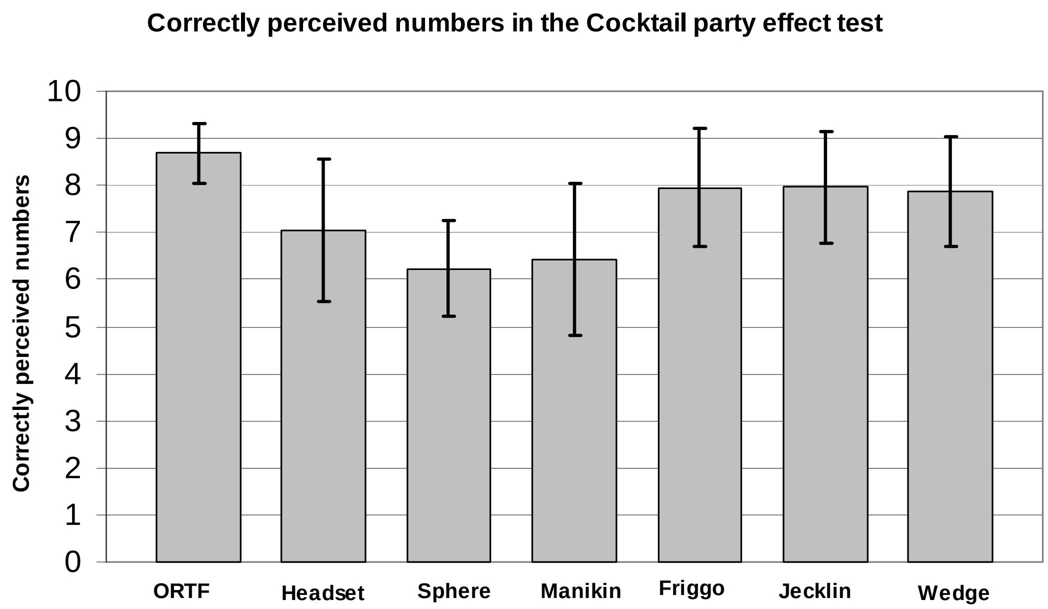


Figure 43: Number of perceived numbers in the cocktail party effect test, with 9 as maximum of correct answers. Deviating answers disregarded.

Outliers have been removed from the analysis shown in Figure 43 and some changes can be noticed compared to earlier results. The best perception of numbers is received for ORTF instead of Jecklin that got the best result in earlier analysis. The standard variation associated with the ORTF is small in comparison to the other recording systems.

The ANOVA shows, as suspected, significant differences for perception of numbers for the recording systems even when the deviant answers were disregarded.

From Tukey's multiple comparison test (see Table 11) with range statistic from table, q (*7.149*) = 4.2 and *Tukey's limit*: 1.2 it was found with 95% confidence that for perception of numbers:

- ORTF performs better than Headset, Sphere and Manikin.

- Friggo, Jecklin and Wedge perform better than Sphere and Manikin.

Table 11: Differences between mean values, disregarding deviants, in the listening test regarding the perception of numbers. Significant differences are enframed. μ_{tot} = average response and $(\mu_i - \mu_j)$ = the difference between the column μ_{tot}- row μ_{tot}

	ORTF	Headset	Sphere	Manikin	Friggo	Jecklin	Wedge	
μ_{tot}	8.7	7.0	6.2	6.4	8.0	8.0	7.9	
$(\mu_i - \mu_j)$	*	**1.6**	**2.5**	**2.3**	0.7	0.7	0.8	ORTF
		*	0.8	0.6	-0.9	-0.9	-0.8	Headset
			*	-0.2	**-1.7**	**-1.7**	**-1.7**	Sphere
				*	**-1.5**	**-1.5**	**-1.4**	Manikin
					*	0.0	0.1	Friggo
						*	0.1	Jecklin

In Figure 44 the mean value of the intelligibility ratings and standard deviations eliminating outliers for the cocktail party effect test are presented.

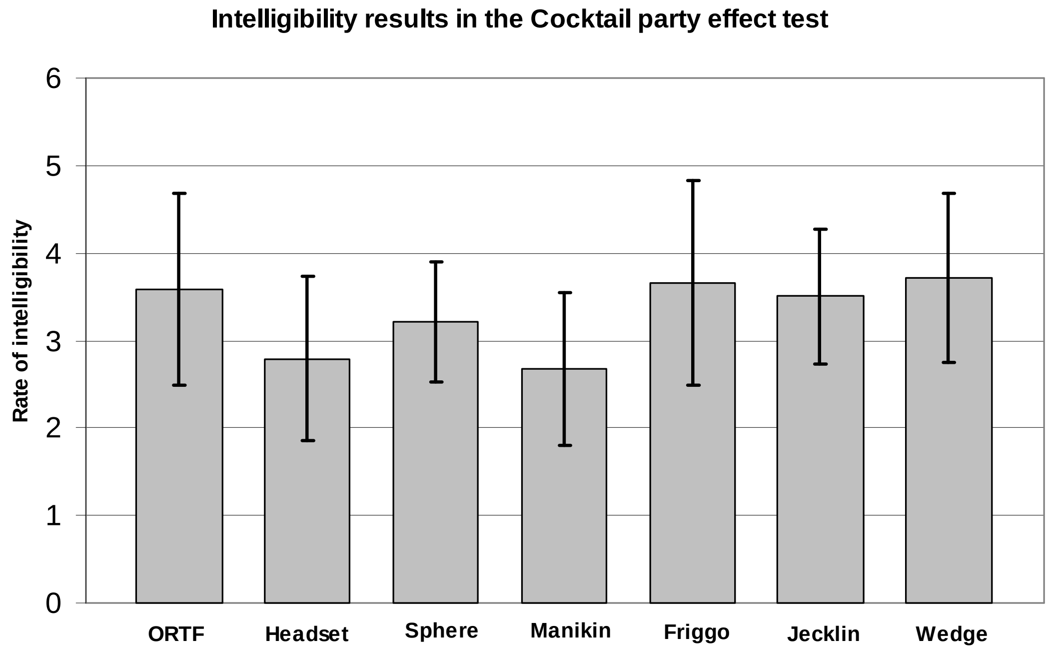


Figure 44: Answers from the ratings of intelligibility in the cocktail party effect test (scale 1-6), eliminating outlying answers.

No particularly changes can be noticed in Figure 44 compared to the case including all answers in Figure 40.

The ANOVA shows, as suspected, significant difference between the recording systems concerning the intelligibility of numbers also when deviant answers are disregarded.

From Tukey's multiple comparison test (see Table 12) with range statistic from table, q (*7.149*) = 4.2 and *Tukey's limit*: 0.9 it was found with 95% confidence that:

- Friggo and Wedge perform better than Manikin.

Table 12: Differences between mean values, eliminating outliers in the listening test regarding the intelligibility of numbers. Significant differences are marked in read. μ_{tot} = average response and ($\mu_i - \mu_j$) = the difference between the column μ_{tot}- row μ_{tot}

	ORTF	Headset	Sphere	Manikin	Friggo	Jecklin	Wedge	
μ_{tot}	3.6	2.8	3.2	2.7	3.7	3.5	3.7	
($\mu_i - \mu_j$)	*	0.8	0.4	0.9	-0.1	0.1	-0.1	ORTF
		*	-0.4	0.1	-0.9	-0.7	-0.9	Headset
			*	0.5	-0.4	-0.3	-0.5	Sphere
				*	**-1.0**	-0.8	**-1.0**	Manikin
					*	0.2	-0.1	Friggo
						*	-0.2	Jecklin

5.3.4 Spectral analysis of the sounds in the numbers test

Nearly all information of speech is contained in the frequency range of 200 Hz - 6000 Hz i.e. 2-13 Bark. The frequency dependency of the hearing is dependent on the position along the basilar membrane where a wave oscillation is excited by an incoming sound wave. Observations have shown that this frequency selectivity of the basilar membrane can be assumed to follow the critical band scale (Bark scale). The Bark scale provides for example information of the masking conditions of a recorded sound and it has a closer connection to the human perception of sound than other frequency-band analysis methods. A table of the Bark scale can be found in Appendix H.

In order to understand the results from the listening test, a study of the spectra of the numbers on a critical band scale was done. The analysis concerned only the sound files used for playback in the listening test, thus no influence from loudspeakers or the measurement systems were included.

In Figure 45 and Figure 46 we see the spectrum for each number in comparison to the average background “murmur”. It is important that the background “murmur” is time-averaged over a longer time than the length of the spoken number. Generally it can be seen that the numbers have a high spectral content in the low frequency range.

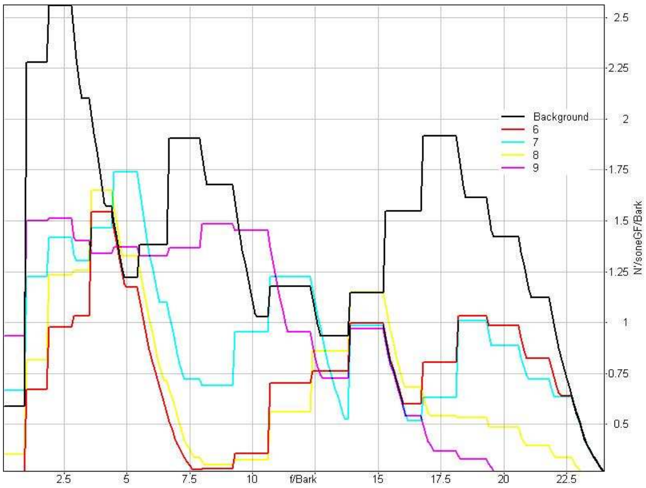


Figure 45: The number spectra in comparison to the background level.

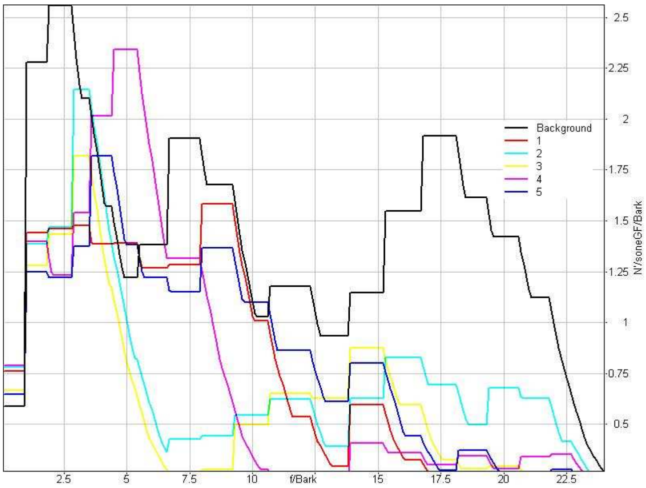


Figure 46: The number spectra in comparison to the background level.

A comparison of the perception of numbers is shown below. It is used as base for further investigation of masking phenomena.

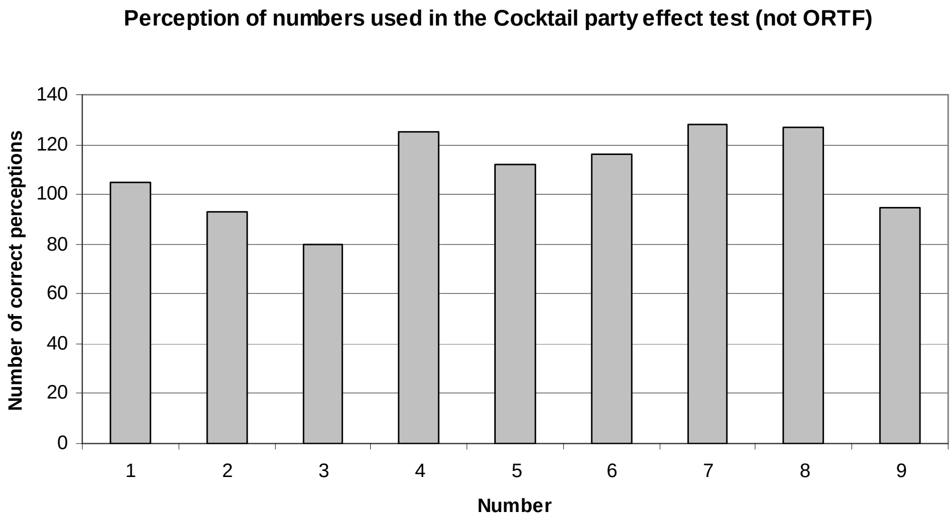


Figure 47: Number of correctly perceived numbers, 6 measurement systems (not including ORTF and Free hanging microphones).

Figure 47 show that number 7 is generally easy to hear as are 8 and 4, while it is more difficult to hear number 3. Figure 48 shows the results for each measurement system separately.

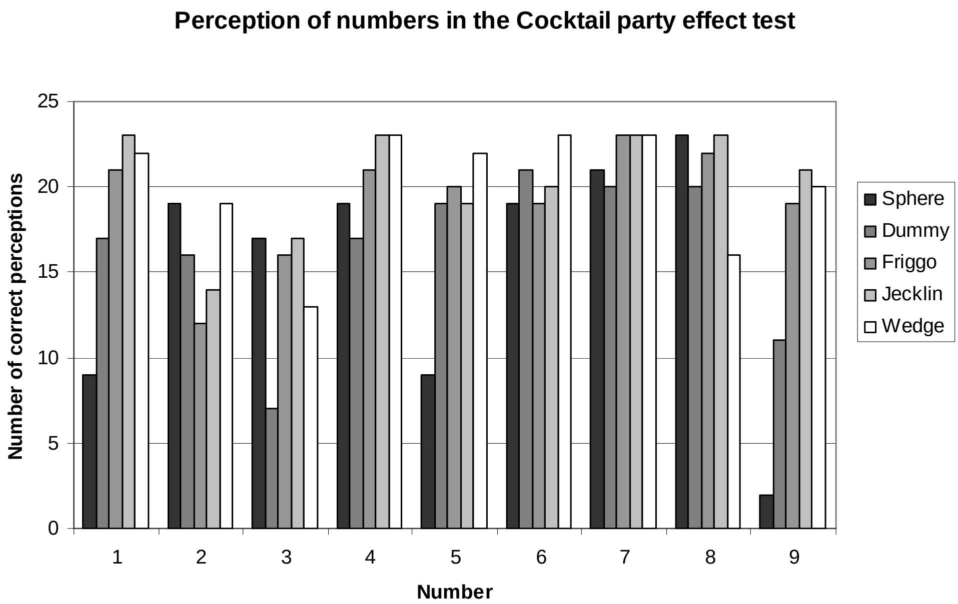


Figure 48: Perception of different numbers presented for each measurement system separately.

Figure 48 shows a high number of correct answers for both Jecklin and Friggo. Manikin has exceptionally low number of correct answers for number 3 which might be explained studying the spectra further.

Spectral investigation was performed for number 3 and 7. The measurement systems Jecklin, Manikin, Friggo and Wedge were analysed. In the following figures loudness levels divided into bark bands for the numbers and the simultaneous background noise are shown.

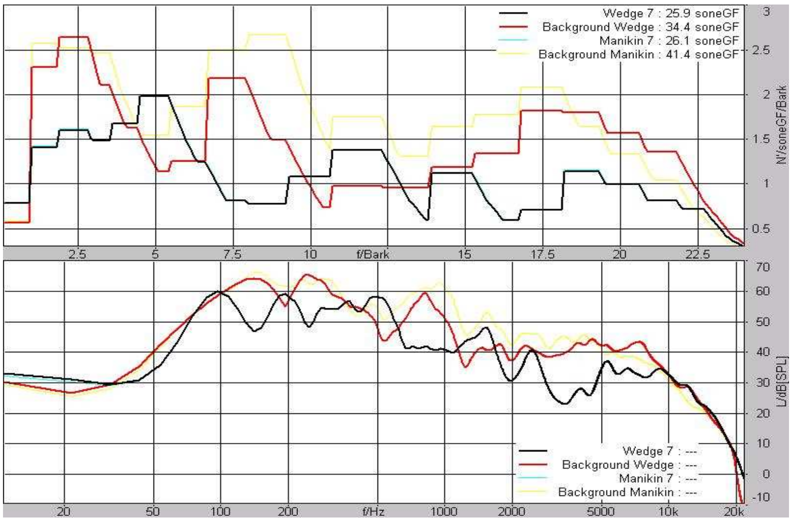


Figure 49: Number 7 in listening test for Wedge and Manikin together with background noise.

Since the analysis only concerns the original sound files there will be no difference between the frequency content of the same numbers, therefore the turquoise curve is hidden by the black curve. In Figure 49 the level of background noise is 6 sone higher for Manikin than Wedge, this can explain the better result of Wedge seen previously in the listening test. The 7 has slightly higher levels for Manikin than Wedge, 0.2 sone, but it can probably be neglected.

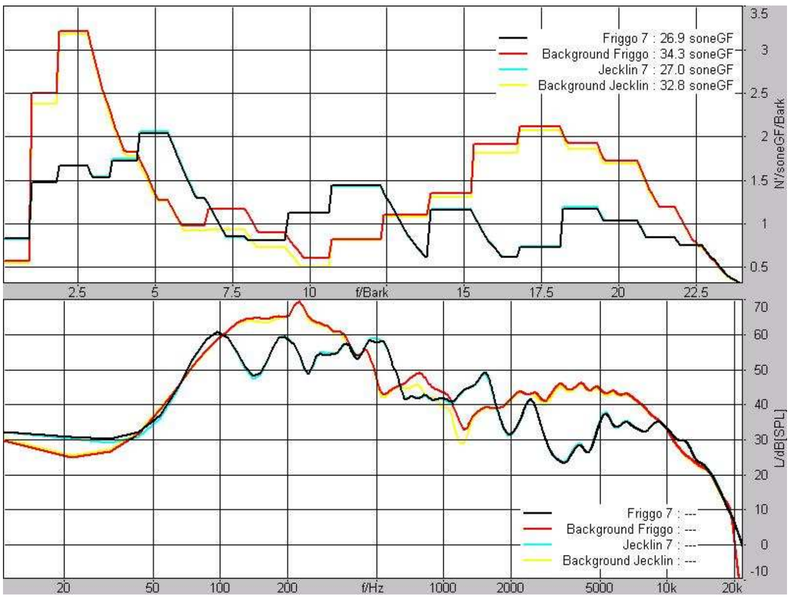


Figure 50: Number 7 for Friggo and Jecklin together with background noise.

The background noise level for Friggo and Jecklin coincide well for number 7. This can be connected to previously good results of number of correct answers in Figure 48.

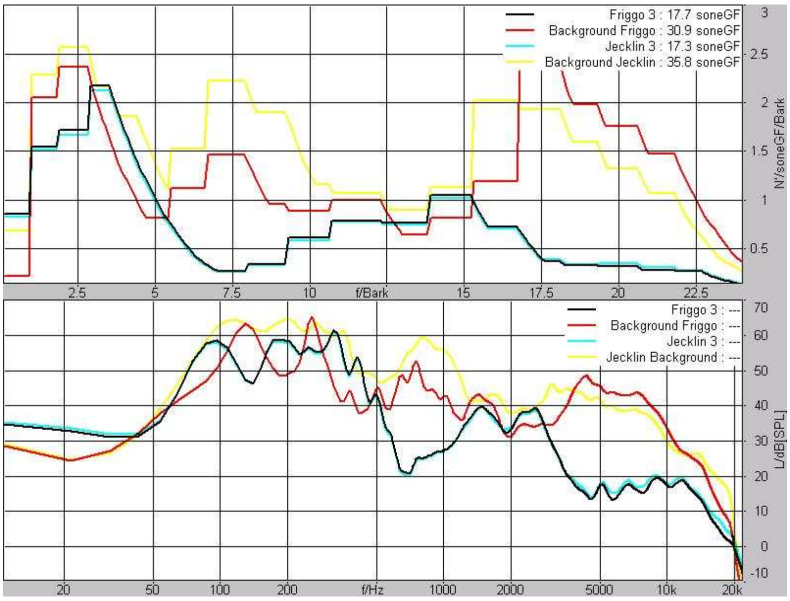


Figure 51: Number 3 for Friggo and Jecklin together with background noise.

When recording number 3 the background noise deviated more. Friggo had higher levels above 17th bark than Jecklin but lower levels between 7-10 barks, see Figure 51. Still their result in listening test did not deviate too much, Jecklin had slightly better results.

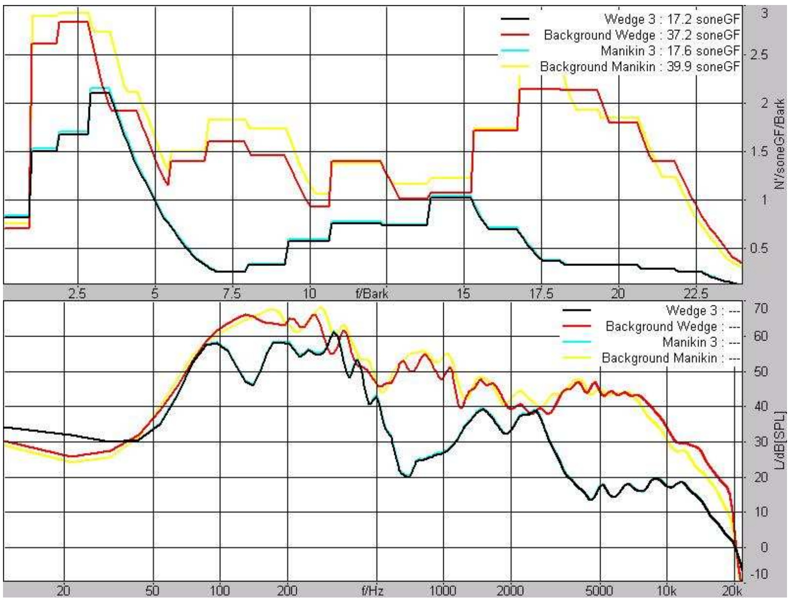


Figure 52: Number 3 for Wedge and Manikin together with background noise.

The background noise for Manikin is at some Bark bands higher than for Wedge, see Figure 52. The results from the listening test show that Wedge is the measurement system achieving most correct answers and Manikin is the measurement system with less number of correct answers regarding number 3. This can not be explained by differences in the background noise.

5.4 Recorded truck sound

To estimate the performance of the measurement system in their intended acoustical environment, recordings of in-cab noise at the passenger seat was done at the Hällered proving ground.

The measurement were done according to Volvo standards (GDI 964-24) but only included the driving cases:

- **Full load acceleration**
- **Constant speed** 85 km/h
- **Low idle speed**, without compressor working
- **Low idle speed**, with compressor working

The measurements were done during one day with acceptable weather conditions and low wind. The temperature was 4^{o} C.

Based on the results from the previous listening tests, the four best systems, Wedge, Jecklin, Sphere and Friggo were selected for measurement of in-cab noise. Simultaneous measurements were done with the Headset at driver position and free hanging microphones at both driver and passenger side.

The constant speed 85 km/h and full load acceleration driving cases were chosen for the listening test. In the test the participants were asked to rate the binaural quality of the recordings performed with the recording systems in comparison with recordings done with Manikin. The test participants were also asked to order the recordings done with the recording systems according to increasing binaural recording quality. The participants were free to listen to the sound samples as many times as they wished. The sound samples were played back with Head Acoustics playback system. In the playback, the recordings done with the developed recording systems were only compensated for the spectral influence of the headphones. The Artificial head and Headset recordings were played back using the Head Acoustic ID filter. The listening test was carried out at Volvo and nine persons among the sound engineering staff participated.

5.4.1 Results from in-cab test

Figure 53 shows the results for the in-cab noise listening test. The positioning of the measurements systems in relation to each other is linked to the number of times the recording systems were graded to a certain mutual position.

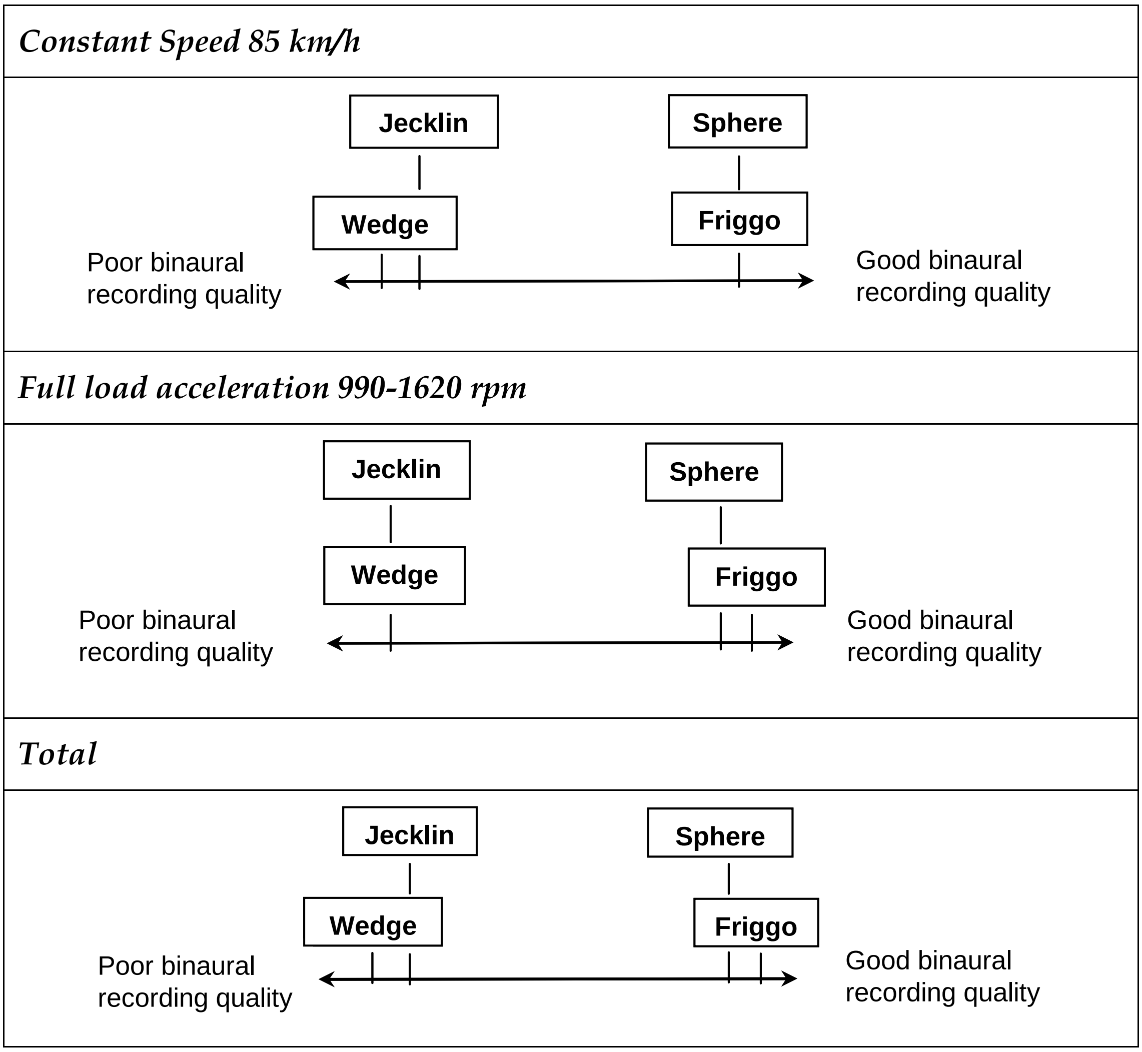


Figure 53: Results from the in-cab noise listening test. The uppermost and middle figures present the result from the constant speed andthe full load acceleration driving cases, respectively. The result for both driving cases weighted equally together are presented in the bottom figure. The positioning of the measurement systems is based on the mutual positioning in the tests.

The results of the listening test shows a clear difference of perceived binaural recording quality between the pairs Jecklin, Wedge and Sphere, Friggo. In total Friggo achieved the best results. However, Sphere and Friggo were very closely positioned and for the constant speed driving case they even received equal results. Jecklin and Wedge can be noticed to have received the same ratings for the full load acceleration driving case.

6 Discussion

The measurement with reference fan gave a first impression of directivity and symmetry of the measured level at both channels for each recording system. The frequency spectra for the different systems coincide well below 500 Hz. At higher frequencies the recording systems shows more diverging spectral behaviour. At frequencies where the wavelength is longer than the acoustic distance between the microphones positions, influence is seen for: Jecklin, Friggo ~ 30 cm => f ~ 1130 Hz, Wedge ~ 38 cm => f ~ 902 Hz. Consequently the shape of Wedge will influence the recordings less than the other recording systems. In the measurement from behind, the influence of the shape of the recording system was particularly seen for Wedge, for which the sound pressure level decreased clearly in higher frequencies. This is probably due to a shadowing effect caused by the wider back of Wedge.

If we compare our designed measurement systems with Manikin and Headset there are more resonances in the high frequency range for the commercial equipment. This can be explained by influence from the more detailed shape of the Manikin head and the existence of torso and pinnae which is not used in the other recording systems.

HRTF measurements were performed to discover possible similarities and differences between the binaural-similar recording systems and Manikin. Rather big differences could be observed regarding the directional properties of the recording systems. Generally, over the entire measured frequency range, the ILD for Manikin achieved its maximum values for angles of incident close to the ears. This is in agreement with what should be expected. However, this tendency also could be noticed for Manikin at low frequencies such as 1 kHz. Compared to Manikin, the ILD measured for Friggo and Sphere were considerably more scattered over angle of incidents, especially in the frequency range below 8 kHz. The ILD could be seen for all recording systems to reach its maximum at 8 kHz and then decrease for higher frequencies. This was not expected as ILD is supposed to increase for increasing frequency. The influence from the more detailed design of Manikin was clearly visible, as significantly larger changes in phase were observed for Manikin compared to Friggo and Sphere.

Designing the listening test regarding the cocktail party effect was difficult. It was particularly difficult was to adjust the levels of the loudspeaker in the setup to reach the limit of intelligibility. In the test it was considered as an advantage to use a masker which the test participants could relate to. Therefore the cocktail party effect was examined by studying the intelligibility of number combinations in a murmur of speech. Possible bias from that the numbers being perceived differently was avoided by using all numbers in all combinations. However, because of the non-static background noise numbers could possible be masked differently. From the investigation of the masking of numbers it seemed like it is not only the level of the background noise that has an influence on the perception. From that analysis it could be concluded that lower background noise does not necessarily give better result. This could indicate that there are other parameters than loudness that influence on the result, i.e. the binaural recording quality of the recording systems.

All results from the listening test regarding subjective ratings have to be analysed carefully, especially when the variances are as considerable as seen in our results. This was also noticed in the multiple comparison tests where we only could observe significant differences between a few measurement systems.

Generally throughout the listening test neither the Manikin nor Headset performed as well as one could expect, in comparison to the more simple recording systems. In many cases they even had worse results. From the statistical analysis of the perception of numbers in the cocktail party effect test one could say with 95% confidence that Jecklin, Friggo, ORTF and Wedge performed better than Sphere and that Jecklin and Wedge performed better than Manikin. Moreover the results regarding the intelligibility of numbers indicated a significantly better result for Wedge compared to Manikin. The good result in the cocktail party effect test by ORTF could partly be explained by the directionality of the microphones which was large in the direction of the incidence of the number combination.

A measurement system giving good source separation ability is desirable to enable perception of different sound sources in a sound field. Unfortunately, the results from the source separation test were not as clear as desired. The reason could be either that the questions asked were hard to understand or that there actually were only small differences among the recording systems concerning this acoustical property. Noticeable is that Jecklin and Wedge, which obtained god results in the listening test regarding perception of numbers, received comparatively low ratings in this test. However, this could not be supported by the ANOVA test.

As expected, front-back confusion caused a lot of errors in the localization listening test, which is a well documented problem in literature. Answers that were influenced by front-back confusion were seen for all recording systems, but least frequently occurring for the Manikin. According to the results, sounds with frontal and perpendicular angles of incident were less difficult to locate than intermediate angles. This might partly be because of that listeners tend to perceive sound as coming from angles close to 0^o and multiples of 90^o. In literature, frontal incidence is reported to have better accuracy than perpendicular incidence [20]. However, because of the large variances in our test nothing could be concluded regarding this.

From the in-cab listening test it was a clear difference between the pairs Jecklin, Wedge and Sphere, Friggo. Common for Jecklin and Wedge is that they include absorption material which might give difficulties in perceiving the binaural properties. When we listened to the recordings of Jecklin and Wedge we experienced them as more damped than the recordings with Sphere and Friggo. As in previous listening tests the binaural properties are difficult to examine, in this test we therefore decided to compare the binaural-similar recording systems and the Manikin. This simplified the test procedure in the way that instead of letting the test participants estimate the binaural quality of the recordings they could judge the similarity between the measurement systems. Also, in this way the test was independent on the participant's previous experience of trucks.

As it was suspected that the presence of absorption material on Wedge may have an influence on the perceived binaural recording quality a swift investigation was performed. Recordings of music in a truck compartment were made with Wedge having the absorption material added and taken away. When listening to the recordings an audible difference between the two cases could be noted, however it is hard to estimate the impact of this on the perceived binaural quality.

In Table 13 the result from the measurements that were experienced as most important, both instrumental and perceptual, for respective recording system can be found. The results in the table are based only on the average rate from the listening tests and hence not statistically proven. From the basis of this table it would be possible to designate the most appropriate recording system.

Table 13: The performance of the concerned recording systems in all decisive measurements performed in the thesis. The recording system which performed the best and second best respective worst and second worst are marked with ++ and +, respective – and -. The recording system that was found to perform best is circled in red.

Recording system	Source separation test	Cocktail party effect test	In-cab noise listening test	In-cab noise level measurements
ORTF	+	++	not concerned	not concerned
Sphere		--	+	-
Wedge	--	-	--	+
Friggo	++	+	++	++
Jecklin	-		-	--

In Table 13 it can be seen that Friggo achieves the best result including all performed measurements.

7 Conclusions

By performing recordings over a 360º sweep, all measurement systems could be concluded to measure symmetrically at both channels. In the reference measurements Wedge and Jecklin obtained the best spatial resemblance with recordings done with free hanging microphones.

In the HRTF measurements rather big differences could be observed regarding the directional properties between the binaural-similar recording systems and Manikin. A prominent feature of Manikin was that the largest values of ILD were achieved for angles of incidence close to the ears. This did not appear as clear for Sphere and Friggo.

In the listening test regarding the perception of numbers Wedge, Jecklin could be concluded with 95% statistically certainty to have performed better than the Manikin. In the intelligibility ratings the results were somewhat different, here only Friggo and Wedge were perceived as being better than Manikin at the same level of confidence. The best mean result in the listening test regarding the source separation ability was obtained for Friggo. There were no significant differences found from the source separation test.

The listening test regarding localization ability only showed the very pronounced source of error deriving from the front-back confusion. Moreover, the ANOVA results could not state any influence on the localization ability from the choice of recording system. The front-back confusion was least pronounced for Manikin, still 44% of the perceived locations were front-back confused.

The analysis of the recordings of in-cab noise showed that the Manikin measured a channel averaged L_{eq}-level which differed the most from measurements with free hanging microphones. The level measured with the Manikin was approximately one decibel higher than measurements with free microphones. The minimum difference was received for Friggo. The measurement stability of the Headset was also examined. An average deviation from the mean of ± 0.1 dB was concluded. Some conclusions could be drawn from the listening test regarding the binaural quality of in-cab noise recordings. Friggo showed the overall best ratings in the test, however very closely followed by Sphere. For recordings of in-cab truck noise, Friggo and Sphere were perceived as significantly better than Jecklin and Wedge.

Because of its good binaural recording quality compared to the Manikin and reliable level measurements, Friggo is recommended to be used for recordings of in-cab noise. It performed well in all listening tests. Moreover Friggo measured, in comparison to the artificial head, in-cab noise levels less diverging from measurements done with free hanging microphones in comparison to the other systems. Thus, further investigation is recommended to obtain loudness levels in playback corresponding to the real situation.

During the progress of the thesis work difficulties were experienced regarding the estimation of the binaural quality. The binaural quality is a complex acoustical property which especially has appeared during the performance of the listening tests. Another general conclusion is the pronounced influence from the properties of the sound field when measuring with binaural recording systems. Therefore, level measurements with binaural recording equipment can not be considered as equivalent to measurements performed with single microphone.

References

[1] Head Acoustics (2004): Data sheet, SQuadriga (Code 1369).

[2] Zwicker, E, Fastl, H (1999): Psycoacoustics – Facts and Models, Germany

[3] Ossietzky,C. Common principles of across - channel processing in binaural and monaural hearing, DTU, Oldenburg.

[4] Pulkki, V, (2002): Microphone Techniques and Directional Quality of Sound Reproduction Audio engineering society; 112th convention 2002 May 10-13, No. 5500.

[5] http://www.tape.com/cgi-bin/SoftCart.exe/Bartlett_Articles/stereo_microphone_techniques.html?E+cassette, (1999), Bartlett, B.

[6] Hartmann, W. (1999): How we localize sound, Physics Today.

[7] Reproduction of auditorium spatial impression with binaural and stereophonic sound systems, AES 118th Convention, Barcelona, Spain May 28-31 paper 6485

[8] Merimaa, J. (2006): Analysis, synthesis and perception of spatial sound. Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing report 77.

[9] Farina A, Tronchin, L. (2005): Measurements and reproduction of spatial sound characteristics of auditoia, Acoustic Science & Technology 26, 2

[10] http://www.wikirecording.org/ORTF, (2006)

[11] (2004), http://www.macmusic.org/articles/view.php/lang/en/id/71

[12] United States Patent, (2002): Patent No.: *US 6463158 B2*

[13] http://www.schoeps.de/E-2004/kfm6.html, Schoeps GmbH

[14] Comparability of Recordnings Made with Measurement Systems from HEAD Acoustics, HEAD's application notes.

[15] Beranek, L., (1996): Concert and opera halls: How they sound, J. Acoustical society of America, New York

[16] Blauert, J, (1995): Spatial Hearing The Psychophysics of Human Sound Localization, revised ed. The MIT Press, Cambridge.

[17] Gerzon, M., (1992):Signal processing for simulating realistic stereo images, 93rd AES convention, New York, 1-4 October, Preprint 3424,

[18] Kendall, G. S. (1995): The Decorrelation of Audio Signals and Its Impact on Spatial Imagery, Computer Music Journal, vol. 19(4), pp 71-87,

[19] Arons,(1992): A Review of the Cocktail Party Effect, http://citeseer.ist.psu.edu/cachedpage/10582/3

[20] Kleiner, M, (2005): Audio Technology & Acoustics, Department of Applied Acoustics CTH, Gothenburg

[21] Harrisson, M, (2004): Vehicle Refinement Controlling Noise and Vibration in Road Vehicles, Cranfiels University,UK,

[22] Väljamäe, A, (2004): Auditory Presence, Individualized Head-Related Transfer Functions, and Iiiusory Ego-Motion in Virtual Enviroments, Department of Signals and Systems, Chalmers University of Technology, Sweden

[23] Head Acoustics, (2006): Binaural_Measurement_Okt_06_e

[24] Head Acoustics, System Comparability

[25] Constan Z, Hartman W (2003): On the detection of dispersion in the head-related transfer function

[26] Khun, G.F. (1982): Towards a model for sound localization, in Localization of sound: Theory and applications, edited by R.W.Gatehouse (Amphora, Groton,CT)

[27] Brungart, D.S,, and Rabinowitz, W.M, (1999): Auditory localization of nerby sources - HRTFs, J.Acoust.Soc.Am.106 1465-1479.

[28] Cabera. D, (2006): Percieved room size and source distance in five simulated concert auditoria, ICSV12, 2005 Lisbon, School of Architecture, University of Sydney, Australia

[29] Kahana. Y, (1998): A multiple microphone recording technique for the generation of virtual acoustic images. Institute of Sound & Vibration Research, Southampton University, United Kingdom

[30] Raykar. V, (2004): Extracting frequencies of the pinnae spectral notches in measured head related impulse responses. Department of Computer Science, University of Maryland, USA

[31] Calamia. P.T, (1998): Three-dimensional localization of a close-range acoustical source using binaural cues. University of Texas at Austin, USA

Appendix A – Measurement systems specifications

MEASUREMENTS SYSTEMS, SPECIFICATIONS

- ORTF angle: 112°
  17cm distance between the microphone grills.
  Microphones AKG with cardioid directivity:
  Serial no. 15253 and Serial no. -

- Jecklin disc
  17cm distance between the microphone grills symmetrical divided with 10 mm plywood, diameter 300 mm.
  Foam absorber of 30 mm thickness.
  Microphones with omnidirectional directivity G.R.A.S. prepolarized:
  Microphone Type 40AE Serial no. 53557 Serial no. 53548

- Manikin HMS II.1
  44.1 kHz samplings frequency.
  Filter HP0, level settings 94dB
  Applied filter in ArtimiS; User1_EJ_N

- Headset BHM III on Manikin HMS II.1
  Filter ID, 94dB. Applied filter in ArtimiS: BMHIII.

- Wedge shaped dummy head
  20cm distance between microphone grills.
  Microphones with omnidirectional directivity G.R.A.S. prepolarized:
  Microphone Type 40AE Serial no. 53557 Serial no. 53548

- Simple Manikin, Friggo
  13cm distance between microphone diaphragms.
  Microphones with omnidirectional directivity G.R.A.S. prepolarized:
  Microphone Type 40AE Serial no. 53557 Serial no. 53548.

- Sphere
  23.5cm distance between microphone grills.
  Microphones with omnidirectional directivity G.R.A.S. prepolarized:
  Microphone Type 40AE Serial no. 53557 Serial no. 53548

- Single microphones, beside Manikin
  10cm distance from Manikin at ear height.
  Microphones with omnidirectional directivity G.R.A.S. prepolarized:
  Microphone Type 40AE Serial no. 53557 Serial no. 53548

Calibrator: B&K type 4228 serial no: 1756621, 1 kHz, Pistonphone type 4228 serial no. 1756621, 250 Hz

Appendix B - Absorption material

ABSORPTION MATERIAL

Table 14: Material parameters of absorption material

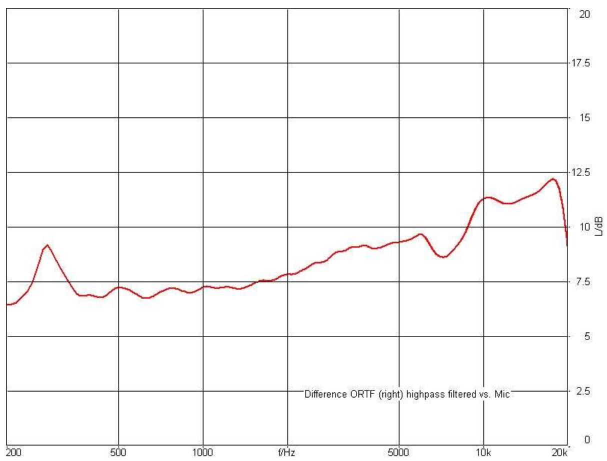
	Absorbent	**Type**	**Thickness**	
Wedge	Antiphon Skumabsorbent	LA 10 S	10mm	Material No. 2a
Jecklin	Antiphon Skum absorbent	LA 30 S	30mm	Material No 2c

Further information in “Documented materials of consumption Department 26747 folder” L47-A6.

Appendix C - Calibration

CALIBRATION

Calibration of the standard omnidirectional microphones was done with a 1 kHz calibrator, and a 250 Hz pistonphone was used for the Headset microphones. The ORTF configuration was calibrated trough simultaneous measurement with calibrated omnidirectional microphone. All microphones used in the measurement systems were calibrated individually before the performed measurements.



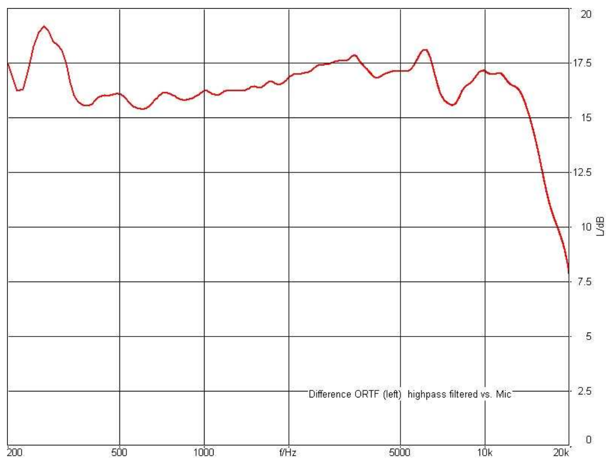


Figure 54: ORTF right vs. omnidirectional calibrated microphone, calculated with high pass filter 125Hz.

ORTF left vs. omnidirectional calibrated microphone, calculated with high pass filter 125Hz.

Appendix D - Additional measurement equipment:

SQUADRIGA

The SQuadriga is a mobile four-channel module used for all measurement equipment from Head Acoustics including a built in memory card. It allows up to 20 min recordings at the maximum sampling frequency of 48 kHz [1] without additional computer, which ensure reliable measurements in the audible range of 4 Hz - 20 kHz. The RPM or engine speed signals can be recorded by two additional electrical separated pulse channels which also can be used to trig start and stop of the recording.

REFERENCE SOURCE - FAN

The reference source was of B&K type 4204 serial no. 680816K.

Appendix E – Reference fan

Measurements with reference fan in the anechoic chamber over a 360°-sweep.

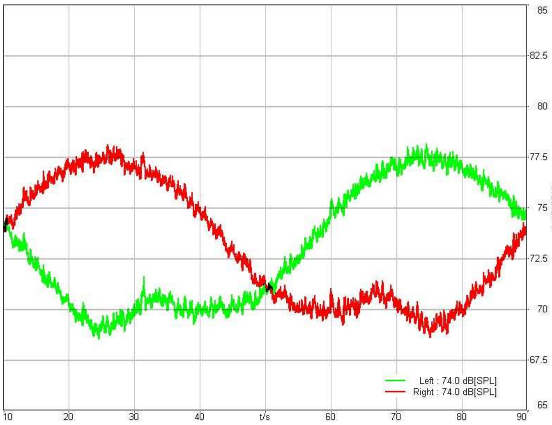


Figure 55: Wedge

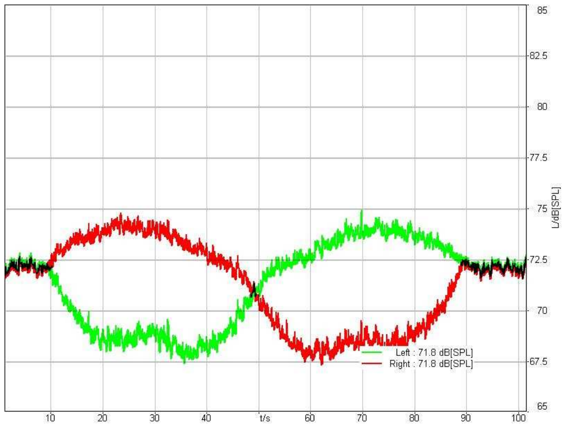


Figure 56: Jecklin

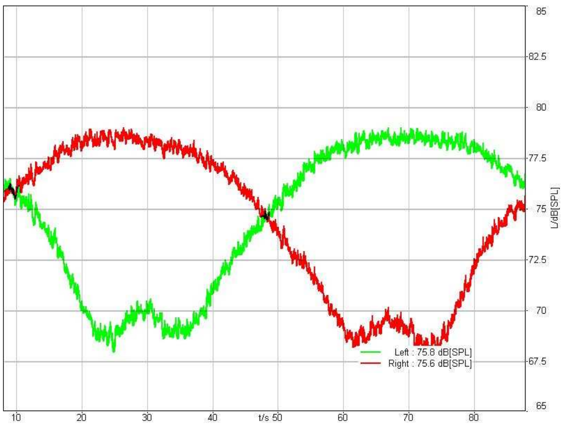


Figure 57: Manikin, filtered with recommended filter.

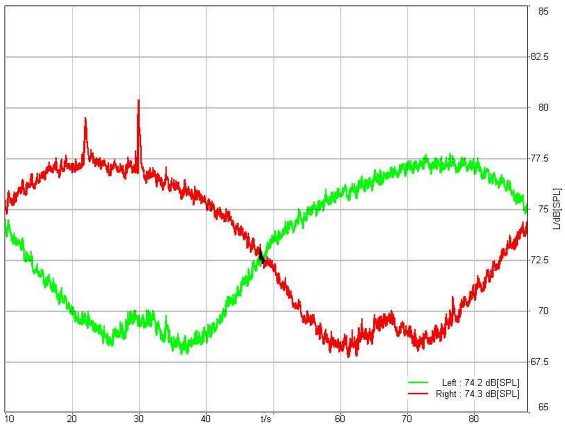


Figure 58: Friggo

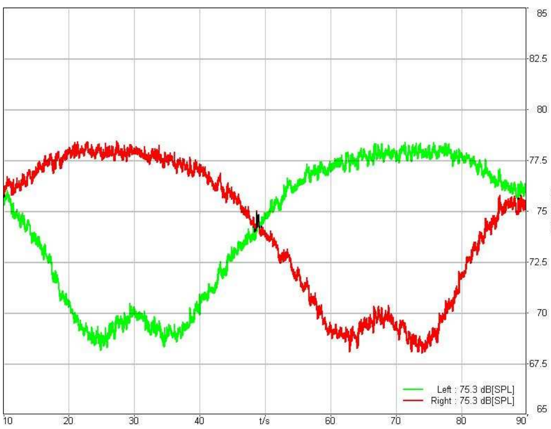


Figure 59: Manikin, filtered with recommended filter.

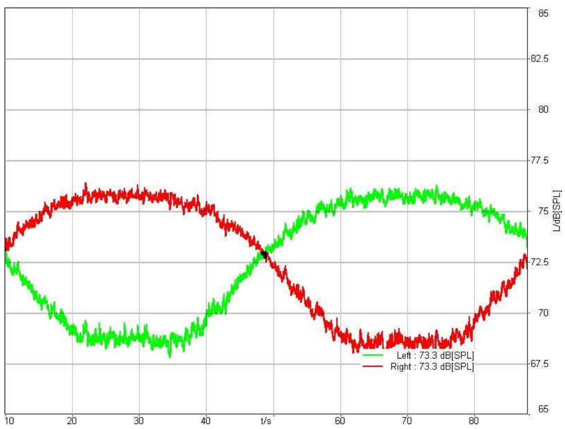
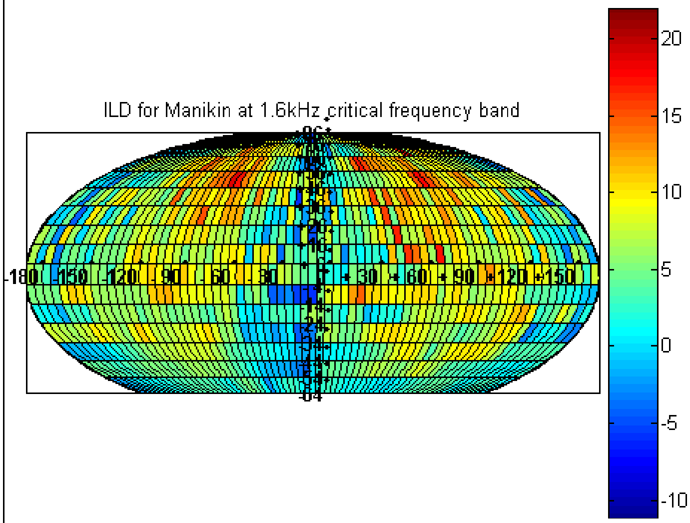
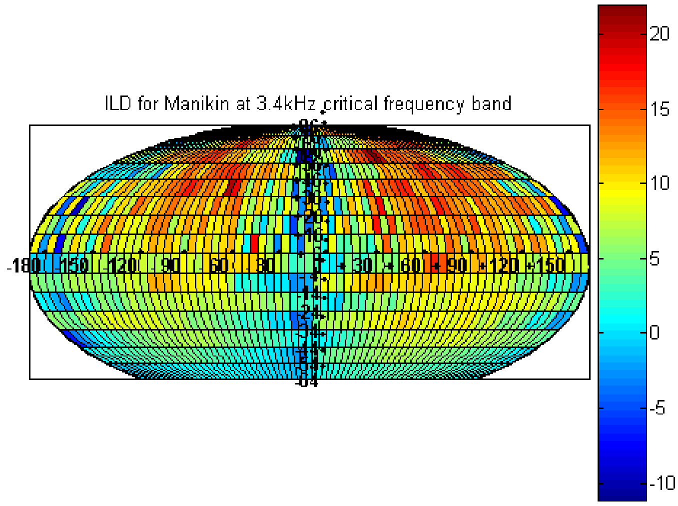


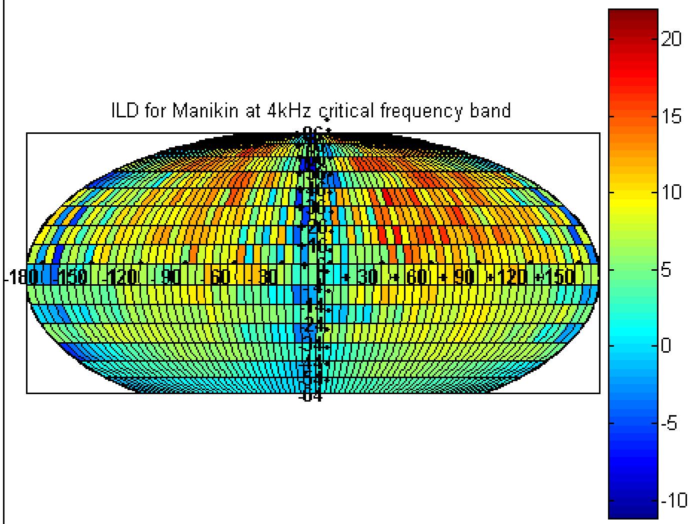
Figure 60: Sphere

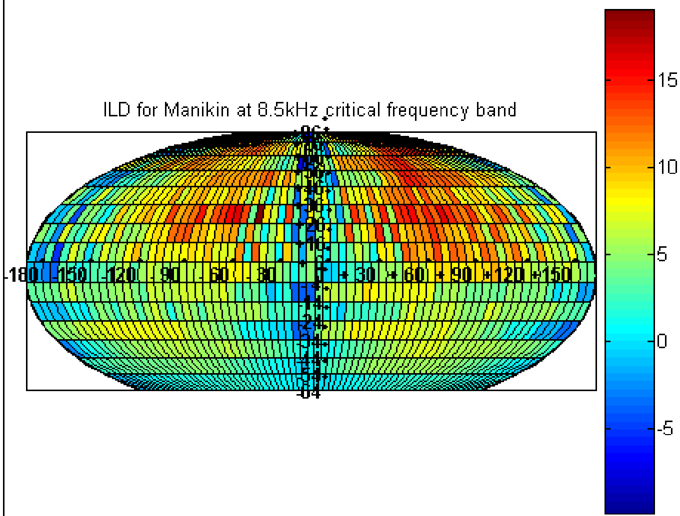
Appendix F – Results from HRTF measurements

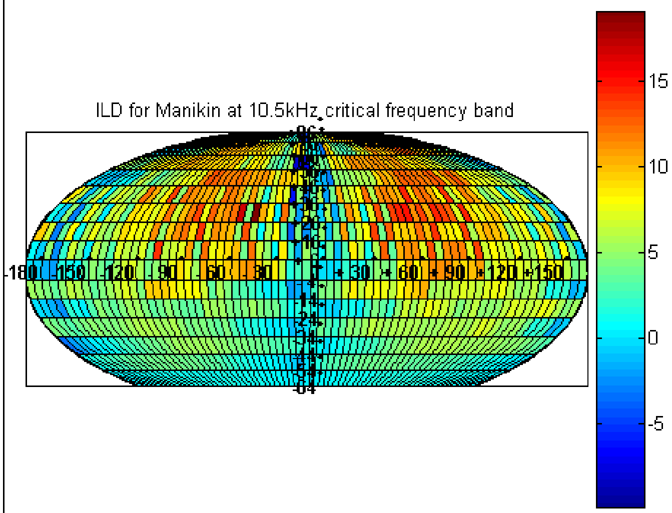
ILD [dB] for Manikin averaged into critical frequency bands (fc = 1.6 – 13.5 kHz)

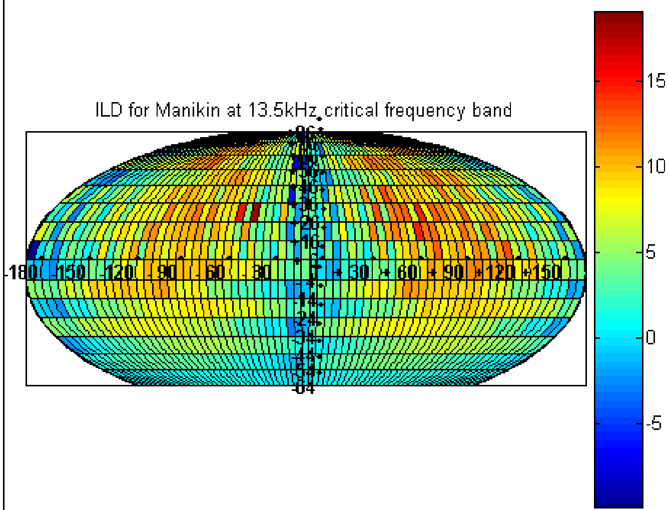




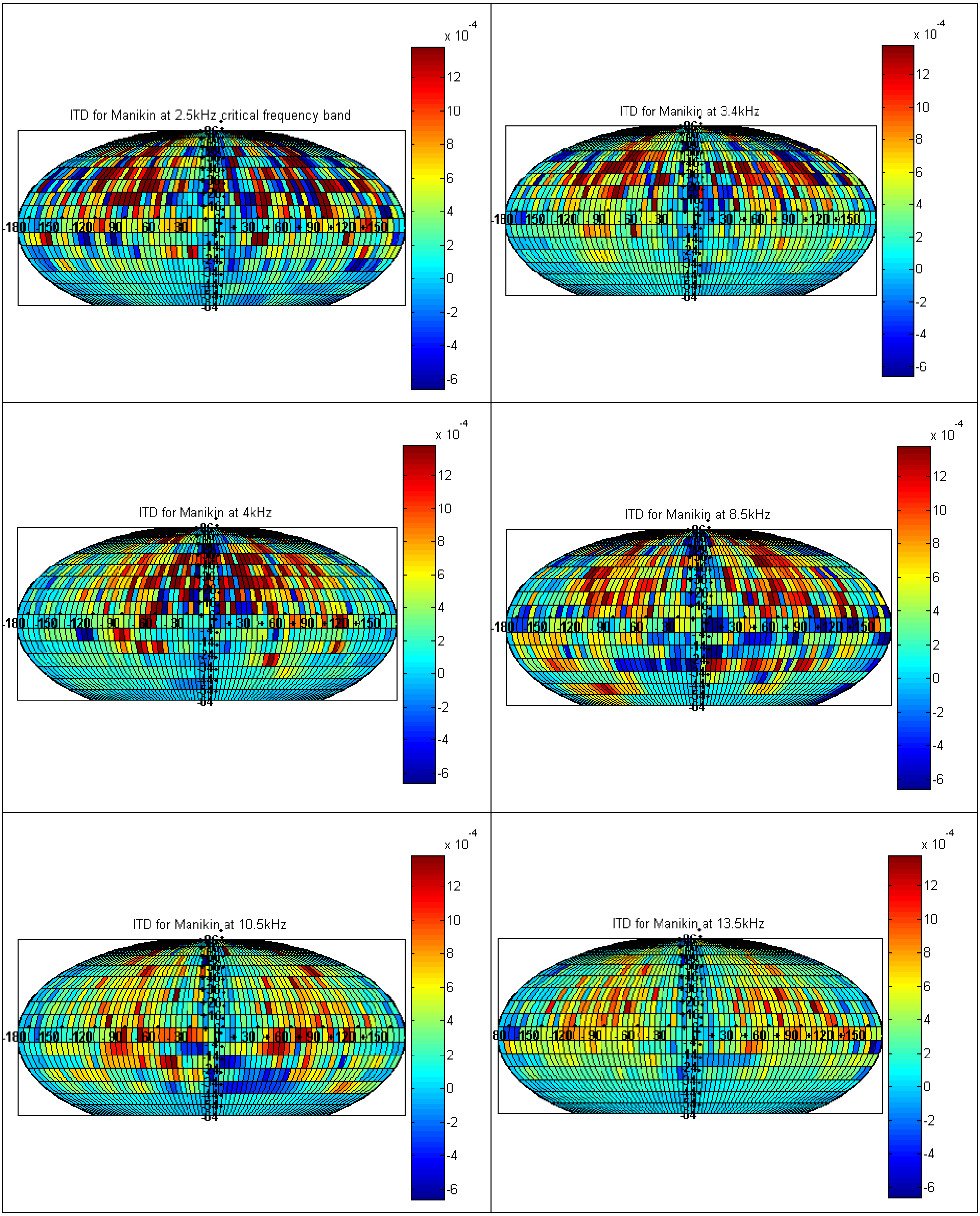




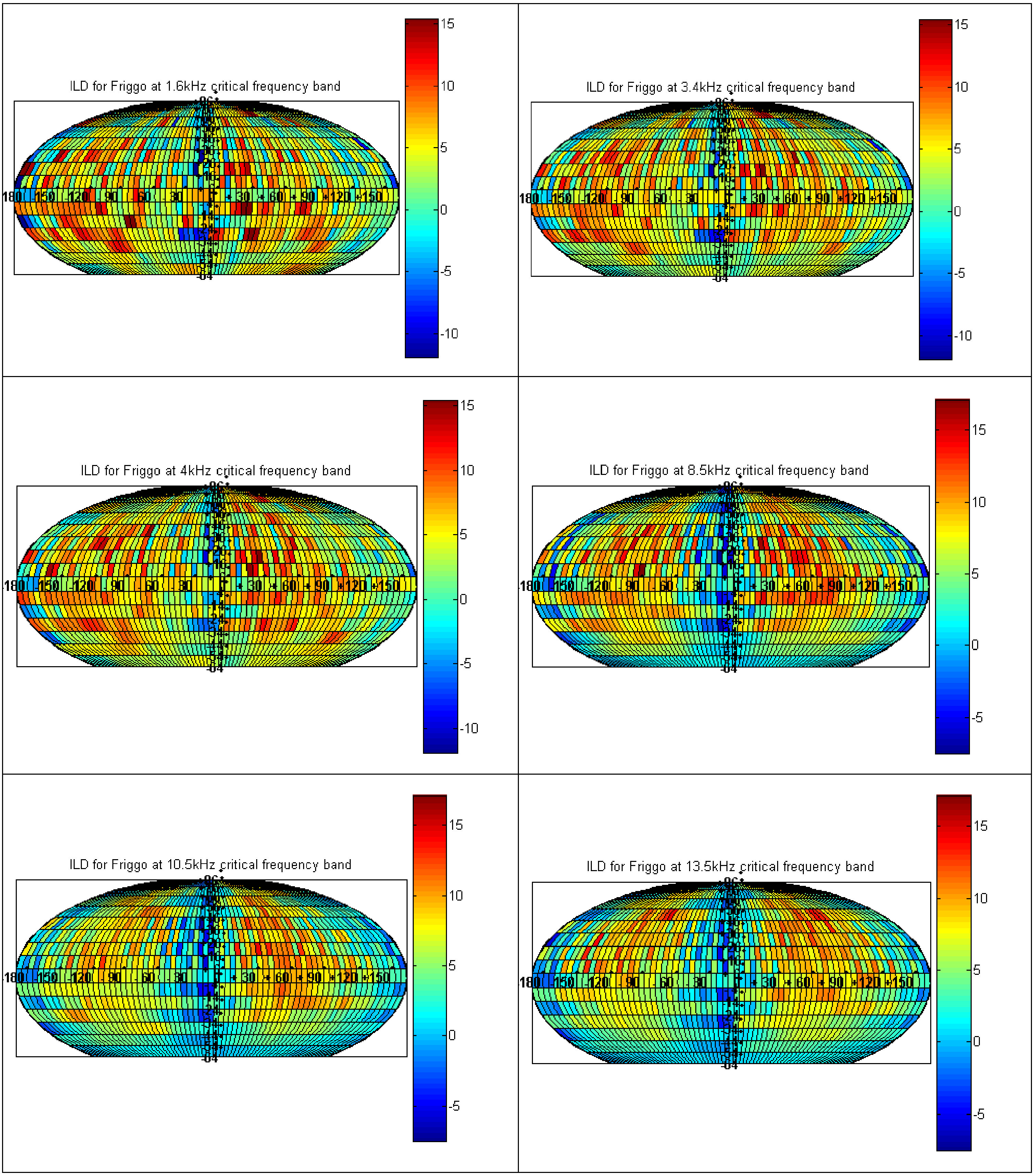




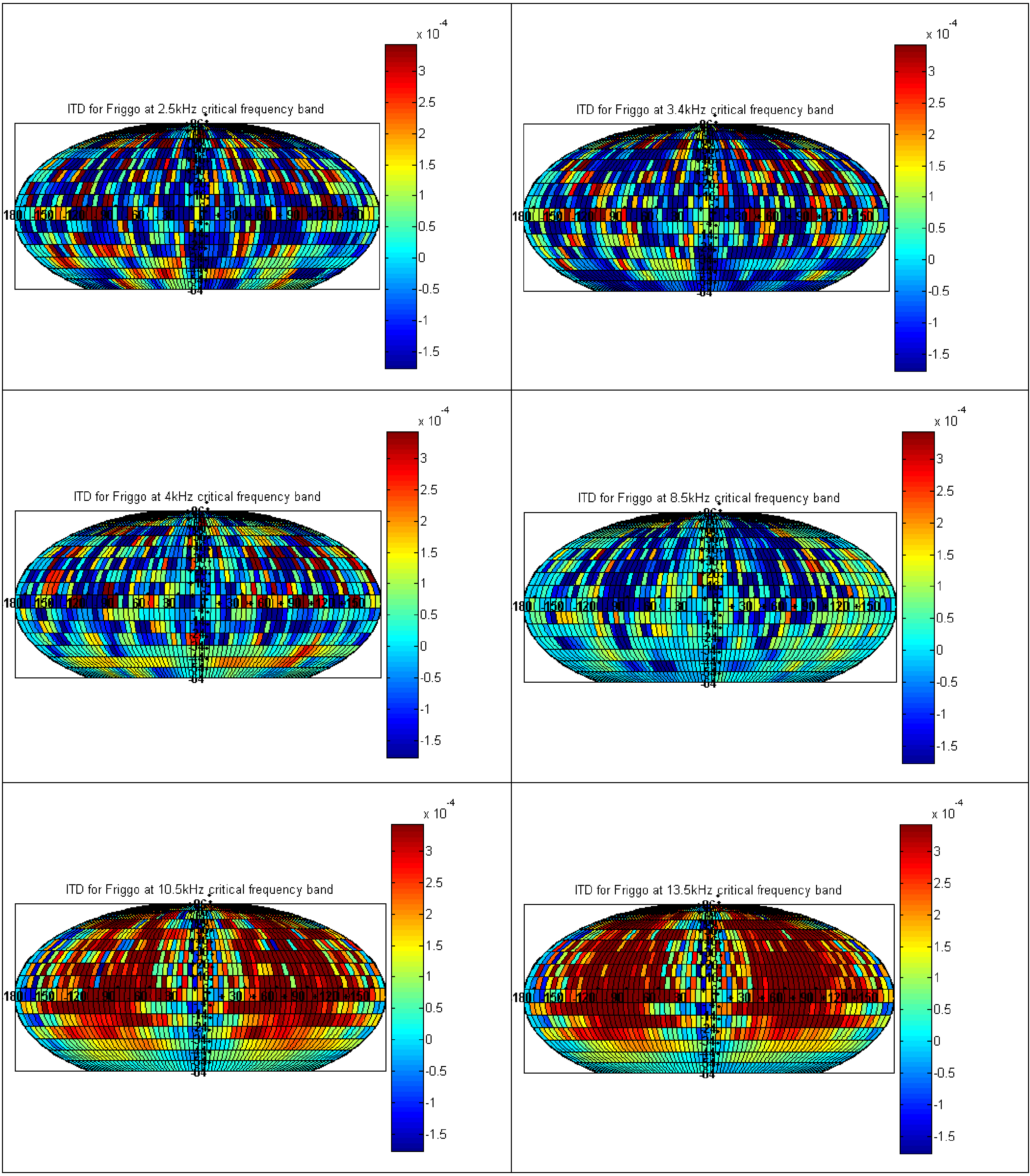
ITD [s] for Manikin averaged into critical frequency bands (fc=2.5–13.5 kHz)



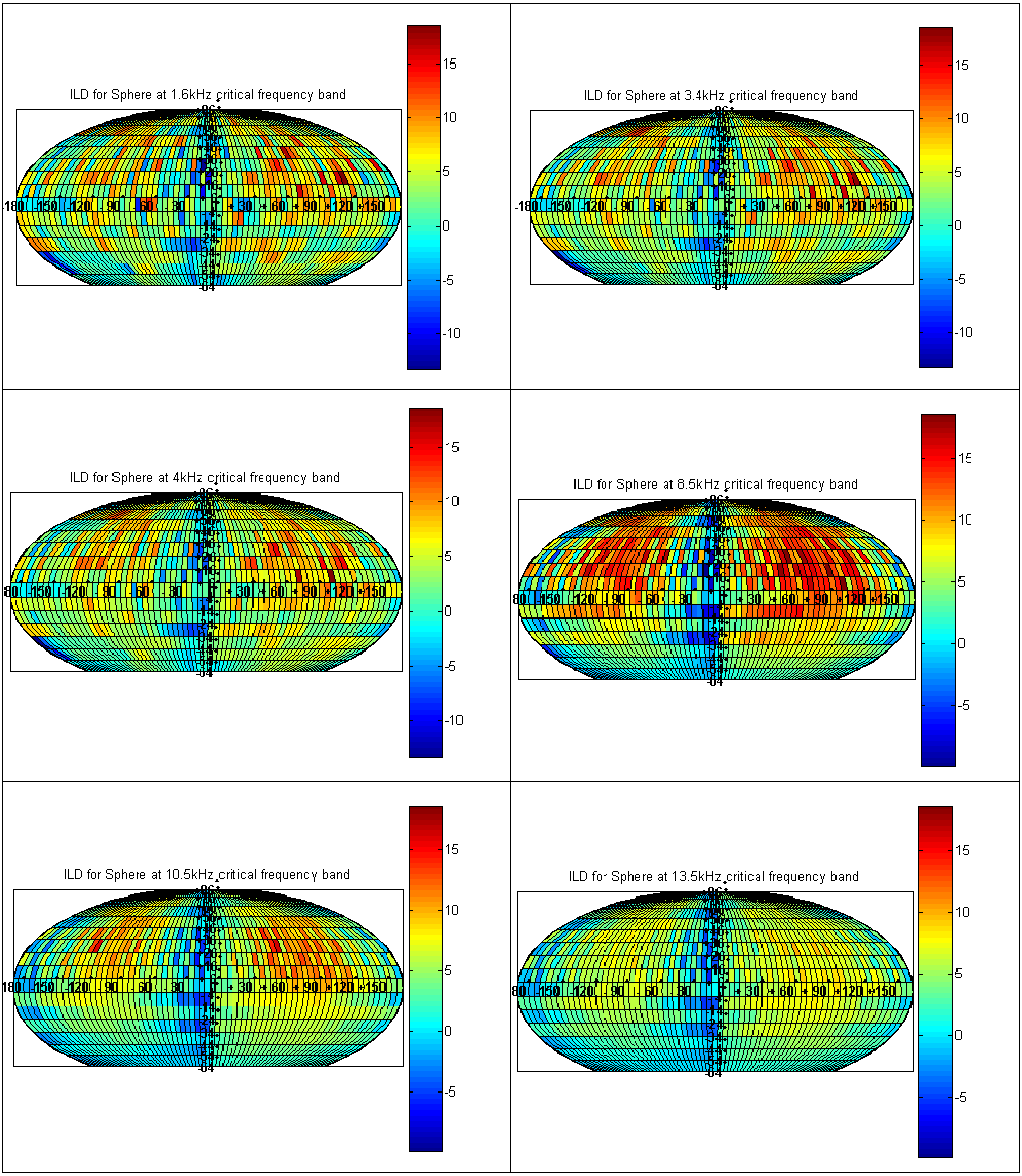
ILD [dB] for Friggo averaged into critical frequency bands (fc=1.6–13.5 kHz)



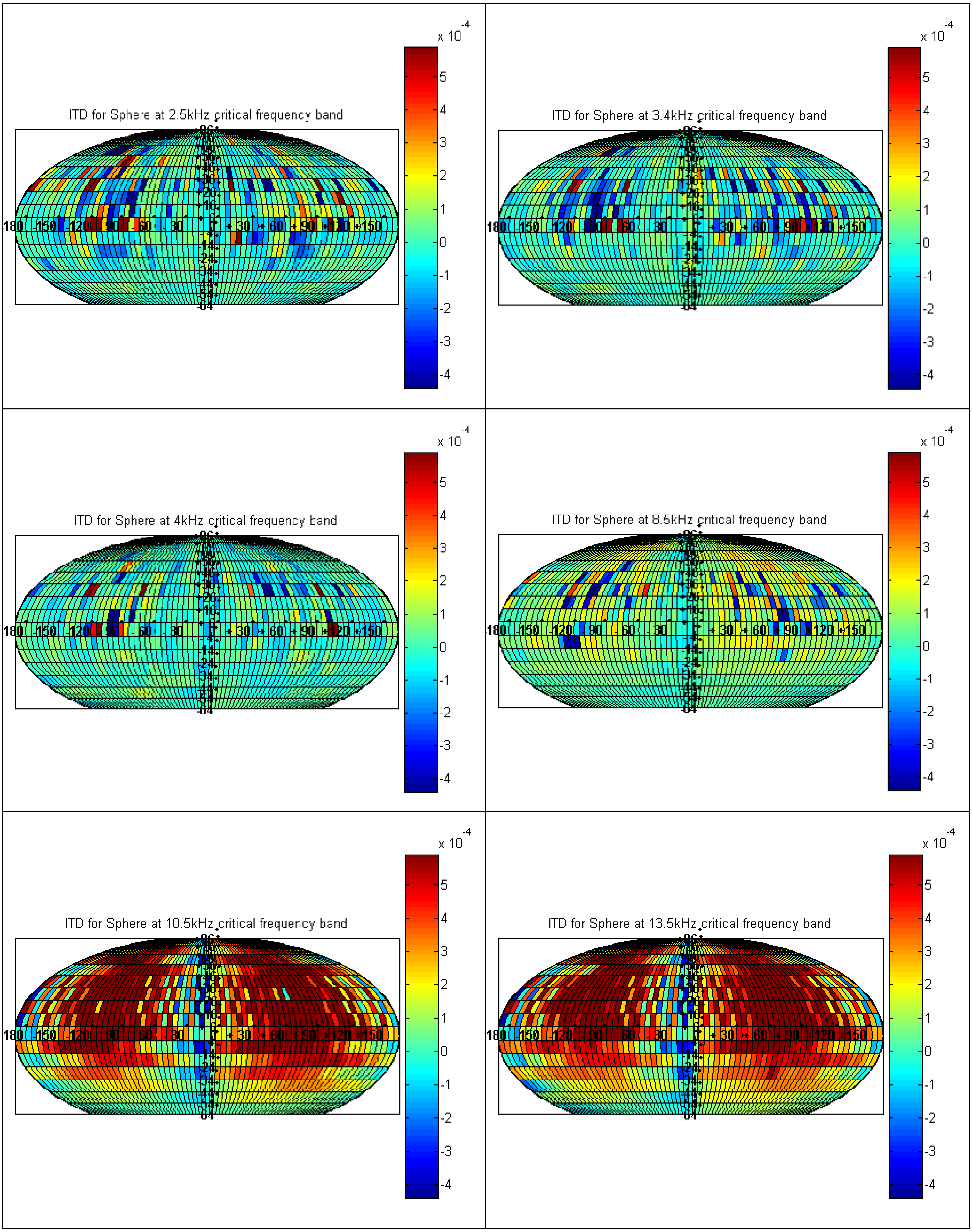
ITD [s] for Friggo averaged into critical frequency bands (fc = 2.5 – 13.5 kHz)



ILD [dB] for Sphere averaged into critical frequency bands (fc=1.6–13.5 kHz)



ITD [s] for Sphere averaged into critical frequency bands (fc = 2.5–13.5 kHz)



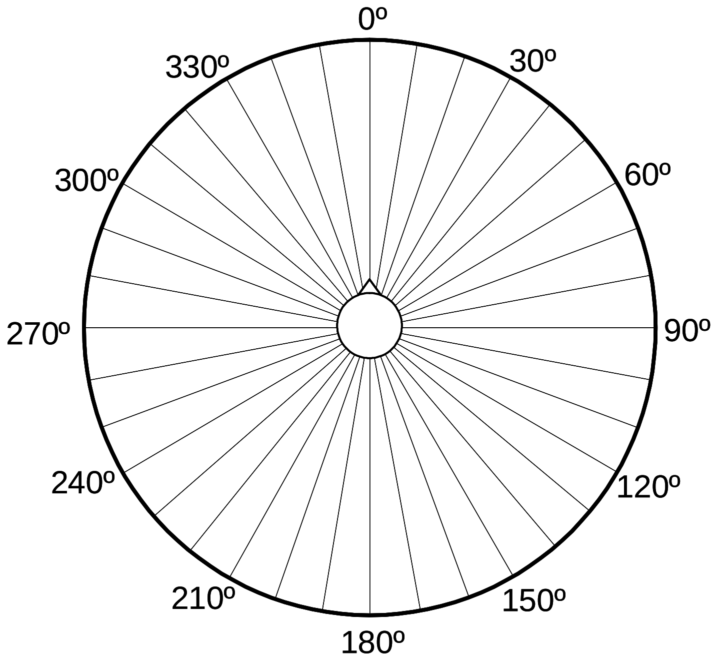
Appendix G – Listening tests

LOCALIZATION TEST

This localization test aims to investigate the spatial qualities of recordings done with binaural-similar recording systems. In the first part of the test, only the angular direction to a source is at question and hence the distance to the source is not concerned. In the second part, additional questions are written to investigate the perception of distance to a source.

The sound used is a male voice counting "one two". This is played back to you twice for every test sample direction. We want you to mark the experienced direction of the male voice in the angular charts below. There is one angular chart for every test sample and the listening test contains a total of 36 samples. An example can be seen below.

Example: The source is experienced as coming from the front, 35° to the right.



Part two:

Sound files: Test model 1 – Test model 4

Questions:

1. At what distance do you perceive the location of the source?

2. Do you experience the sound as being localized inside or outside your head?

Test sample	**Answers**
Test model 1	*Question 1*: …… *Question 2*: …………………………………..
Test model 2	*Question 1*: …… *Question 2*: …………………………………..
Test model 3	*Question 1*: …… *Question 2*: …………………………………..
Test model 4	*Question 1*: …… *Question 2*: …………………………………..

LISTENING TEST REGARDING THE COCKTAIL PARTY EFFECT AND SOURCE SEPARATION

You will hear eight different sound files, recorded with several people talking at the same time. Each sample will be played back to you two times.

One person will read out numbers 0-9 randomly.

Number combination:

Sample 1
First:………………………………   Second:……………………………

Please, try to **write down the number combination** on the lines to the left and mark the intelligibility with a cross on the scale to the right.

How easy was it to detect the numbers?

Sample 1					
Hardly intelligible		intelligible			clearly intelligible
1	**2**	**3**	**4**	**5**	**6**

Compare the sound samples 1 to 6 with the reference sample:

Has the sample higher or lower source separation ability than the reference?

Please mark with a cross on the axis below:

Sample 1:

Much lower Equal Much higher

LISTENING TEST REGARDING THE BINAURAL QUALITY OF RECORDINGS OF IN-CAB NOISE

Preface

We want you to:

- Rate the binaural quality of the recordings in comparison with the reference recording.

- Position the recording systems themselves in order corresponding to ascending binaural recording quality in comparison with reference.

The recording systems are referred to with numbers (1-4) and the scale used is reaching from "poor binaural recording quality compared to reference" to "good binaural recording quality compared to reference".

Position the samples in order corresponding to the perceived binaural quality in comparison with reference.

Constant speed

Poor binaural quality compared to reference	Good binaural quality compared to reference

|..|

Full load

Poor binaural quality compared to reference	Good binaural quality compared to reference

|..|

Appendix H – Bark scale

In the analysis of the listening test regarding the cocktail party effect the loudness are shown using the Bark scale. The band of the Bark scale is shown in Table 15.

Table 15: The Bark scale.

z Bark	f_l Hz	f_u Hz	Δf Hz	z Bark	f_l Hz	f_u Hz	Δf Hz
0	0	100	100	13	2000	2320	320
1	100	200	100	14	2320	2700	380
2	200	300	100	15	2700	3150	450
3	300	400	100	16	3150	3700	550
4	400	510	110	17	3700	4400	700
5	510	630	120	18	4400	5300	900
6	630	770	140	19	5300	6400	1100
7	770	920	150	20	6400	7700	1300
8	920	1080	160	21	7700	9500	1800
9	1080	1270	190	22	9500	12000	2500
10	1270	1480	210	23	12000	15500	3500
11	1480	1720	240	24	15500	22050	6550
12	1720	2000	280				

Appendix I – Statistical analysis

STATISTIC ANALYSIS: USED TO LISTENING TESTS

The analysis of variance (ANOVA) procedure was decided to be used in the analysis of the sample data from the listening tests. In ANOVA it is possible to combine sample data from several populations into a single test and it is capable of detecting when one or more of the population mean value differ from the rest. In our listening tests only one single factor (the choice of recording system) was investigated and therefore the results were analysed using single-factor (one-way) ANOVA. The one-way ANOVA is explained in next section.

If an ANOVA test is significant further investigations is needed before drawing conclusions. A significant ANOVA test only states that there is a significant difference between some means within the studied population means. Multiple comparison procedures are then used to determine which mean values that are different from each other. In this study the Tukey's paired comparison procedure was chosen to do this and it is explained later.

It was desired to estimate the inaccuracy associated with the results from the listening test. In statistics this is usually done by calculating confidence intervals indicating the range of deviation around a result. How to calculate confidence interval is also presented.

ANOVA

To determine whether population mean values differ, the ANOVA approach compares the variation between the sample means under study to the inherent variability within each sample. The more the sample means differ, the larger the between-samples variation will be. Figure 61 shows an imagined example of three populations with variances and mean values outlined. The test static that compares these two types of variation is the ratio of the between-samples variation to the within-samples variation.

$$Test\ statistic = \frac{Between - samples\ \mathrm{var}\,iation}{Within - samples\ \mathrm{var}\,iation}$$

The test statistic follows a continuous probability distribution called an *F* distribution. A null hypothesis assuming no difference between *k* population means $(H_0 : \mu_1 = \mu_2 = ... = \mu_k)$ and is true when the test statistic (*F* ratio) described previously follows the F-distribution. The variation measures used in an *F* ratio are based on certain sum of squares calculated from the sample data. The numerator and denominator sum of squares, each has an associated number of degrees of freedom. There is a different *F* distribution for every different combination of numerator (df_1) and denominator (df_2) degrees of freedom from the calculation of the *F* ratio.

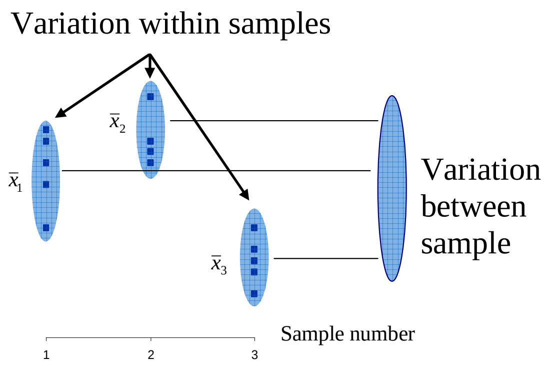


Figure 61: Example of between-samples and within-samples variation in ANOVA.

P-values are used in the hypothesis testing. A P-value associated with a calculated F ratio is the area under the F distribution to the right of the calculated F ratio. Figure 62 shows the P-value associated with a calculated F ratio of 4.53 based on $df_1 = 4$ and $df_2 = 6$.

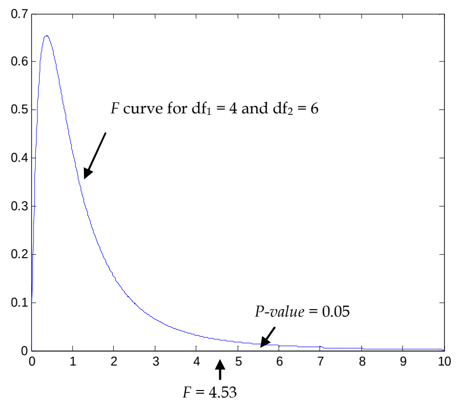


Figure 62: P-value for an upper-tail F test.

NOTATIONS AND FORMULAS

Sample sizes:	$n_1, n_2, ..., n_k$
Sample means:	$\bar{x}_1, \bar{x}_2, ..., \bar{x}_k$
Sample variances:	$s_1^2, s_2^2, ..., s_k^2$
Total sample size:	$n = n_1 + n_2 + ... + n_k$
Grand average:	$\bar{\bar{x}} = average\ of\ all\ n\ responses$

The treatment sum of squares (denoted SSTr) and error sum of squares (denoted SSE) are defined as

$$SSTr = n_1\left(\bar{x}_1 - \bar{\bar{x}}\right)^2 + n_2\left(\bar{x}_2 - \bar{\bar{x}}\right)^2 + \ldots + n_k\left(\bar{x}_k - \bar{\bar{x}}\right)^2$$
$$SSE = \left(n_1 - 1\right)s_1^2 + \left(n_2 - 1\right)s_2^2 + \ldots + \left(n_k - 1\right)s_k^2$$

SSTr and SSE form the basis of the between-samples variation and within-samples variation. Together, these two sources of variation comprise the total sum of squares (denoted SST)

$$SST = SSTr + SSE$$

HYPOTHESIS TEST

To be able to form a statistical procedure and to perform a hypothesis test a few assumptions about the population studied needs to be done. These are:

1. All k population variances are equal.
2. Each of the k population follows a normal distribution.

When sampling from normal populations, each sum of squares has its own unique number of degrees of freedom. For instance, in a one-way classification, the total degree of freedom (associated with SST) is n - 1, which equals the sum of k - 1 (associated with SSTr) and n – k (associated with SSE). This enables us to convert the sum of squares into mean squares by dividing each sum of squares by its associated df. To do this the following formulas were used:

$$MSTr = \frac{SSTr}{k - 1}$$
$$MSE = \frac{SSE}{n - k}$$

MSTr and MSE serve as measures of the between-samples and within-samples variation described previously.

To present all this information in this report the data has been organized into ANOVA tables. An example of an ANOVA table for one-way classification is shown below.

Table 16: ANOVA table for the one-way classification.

Source of Variation	**df**	**SS**	**MS**	**F**
Between samples	k - 1	SSTr	MSTr	MSTr/MSE
Within samples	n - k	SSE	MSE	
Total variation	n - 1	SST		

The hypothesis H_0 is rejected whenever the P-value of the test static F is less than or equal to the significance level α used for the hypothesis test.

MULTIPLE COMPARISON: TUKEY'S METHOD

Tukey's procedure allows us to conduct separate tests to decide whether $\mu_i = \mu_j$ for each pair of mean values in an ANOVA study of *k* population mean values. In correspondence with other multiple comparison procedures, Tukey's method is based on the selection of a significance level α. which says that it is at most a α % chance of obtaining a false positive among the entire set of pair wise tests. According to Tukey's procedure the distance between any two sample means $\left|\bar{x}_i - \bar{x}_j\right|$ can be compared to the threshold value *T*, *Tukey's limit* that depends on α as well as on the MSE from the ANOVA test. The formula for *T* is,

$T = q_\alpha \sqrt{\frac{MSE}{n_i}}$, where n_i is the size of the sample drawn from each population.

The value of q_α is found tables of statistics, *q*, and it follows the Studentized range distribution. The Studentized range distribution is a probability distribution that depends on two different degrees of freedom (*k, m*), where *k* is the number of population mean values to be compared and *m* is the error degrees of freedom in the ANOVA test. To determine whether two mean values μ_i and μ_j differ, we simply compare $\left|\bar{x}_i - \bar{x}_j\right|$ with *T*. If $\left|\bar{x}_i - \bar{x}_j\right| > T$, then it is concluded that $\mu_i \neq \mu_j$. Otherwise, it is concluded to be no significant differences between the two mean values.

CONFIDENCE INTERVALS

In analysis of experiments it is often of interest to calculate an interval within which the true value with some probability could be expected. Such an interval is called a confidence interval. The size of the confidence interval is dependent of the standard deviation (*s*) and the accuracy that is requested. In the analysis of the listening test it is of interest to now an interval in which a certain percentage of the population values are included. Assuming that a population is normal distributed, a single sample confidence interval centered at the sample mean value including 95% of the values in the population can be calculated with the following formula:

$$confidence\ interval\ limits = \bar{x} \pm 1.96 \cdot s$$