Improving Two–Way Loudspeaker Directivity

Olli Kantamaa

School of Electrical Engineering

Thesis submitted for examination for the degree of Master of Science in Technology.

Iisalmi 24.2.2020

Supervisor

Prof. Vesa Välimäki

Advisor

DSc Aki Mäkivirta

Aalto University
School of Electrical Engineering

Aalto University
School of Electrical Engineering

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi
Abstract of the master's thesis

Author Olli Kantamaa

Title Improving Two–Way Loudspeaker Directivity

Degree programme Master's Programme in Computer, Communication and Information Sciences

Major Acoustic and Audio Technology **Code of major** ELEC3030

Supervisor Prof. Vesa Välimäki

Advisor DSc Aki Mäkivirta

Date 24.2.2020 **Number of pages** 73+6 **Language** English

Abstract

Compact conventional two–way loudspeakers radiate omnidirectionally typically at frequencies below 500 Hz and the acoustic directivity of increases with frequency. Such loudspeakers couple with room acoustics efficiently at low frequencies, evoking room mode resonances and early reflections, and this tends to decrease the perceived sound quality in room.

In the present work, functional model of a compact a two–way loudspeaker with increased directivity was implemented. Increasing the low–middle frequency directivity also reduced the variation in directivity across audible frequencies. Because the directivity in conventional loudspeaker designs depends primarily on the physical dimensions and shapes, compact two–way loudspeakers naturally have low directivity at low frequencies and particularly compact loudspeaker designs cannot generate significant directivity at low frequencies. In this Thesis directivity at low frequencies was increased actively, using a secondary transducer and signal processing. The secondary transducer cancelled sound behind the loudspeaker creating directivity in the frequency range 80–600 Hz. Anechoic measurements, listening room measurements and listening tests studied the performance of the loudspeakers with typical directivity and actively increased low frequency directivity.

Objective room measurements demonstrated reduction in the decay energy and the decay started at a lower level, causing less psychoacoustic masking. The subjective test data was analysed using the analysis of variance, principal component analysis and factor analysis. The statistical analysis demonstrated that the increase of the loudspeaker directivity in the low–middle frequencies can improve the subjective sound quality significantly. Increase of the low–middle frequency directivity improved clarity, reduced sound colouration, improved virtual sound image definition and transient reproduction compared to the conventional loudspeaker directivity.

Keywords Audio Systems , Psychoacoustics , Low–Frequency Localization , Reflections

Aalto University
School of Electrical Engineering

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi
Diplomityön tiivistelmä

Tekijä Olli Kantamaa

Työn nimi Kaksitiekaiuttimen suuntaavuuden parantaminen

Koulutusohjelma Tieto-, tietoliikenne- ja informaatiotekniikan maisteriohjelma

Pääaine Akustiikka ja Ääniteknologia **Pääaineen koodi** ELEC3030

Työn valvoja Prof. Vesa Välimäki

Työn ohjaaja TkT Aki Mäkivirta

Päivämäärä 24.2.2020 **Sivumäärä** 73+6 **Kieli** Englanti

Tiivistelmä

Tavanomainen kooltaan pieni kaksitiekaiutin on yleensä ympärisäteilevä alle 500 Hz taajuuksilla ja kaiuttimen suuntaavuus lisääntyy taajuuden kasvaessa. Matalilla taajuuksilla ympärisäteilevä kaiutin kytkeytyy voimakkaasti huoneen akustiikkaan, jolloin se herättää huoneresonansseja ja tuottaa aikaisia heijastuksia. Yleensä vahva huoneen akustiikkaan kytkeytyminen huonontaa havaittua äänenlaatua.

Tässä työssä suunniteltiin ja toteutettiin kootaan pieni matalilla taajuuksilla suuntaava kaksitiekaiutin. Suuntaavuuden lisääminen matalilla taajuuksilla vähentää samalla suuntaavuuden vaihtelua kuulotaajuusalueella. Tavanomaisen kaksitiekaiuttimen suuntaavuus määräytyy pääasiallisesti kaiuttimen mittojen ja muotojen perusteella, eikä kooltaan pieni kaiutin pysty tuottamaan suuntaavuutta matalilla taajuuksilla. Tässä työssä suuntaavuus saatiin aikaan aktiivisesti käyttäen digitaalista äänenkäsittelyä ja toista kaiutinelementtiä, joiden avulla suuntaavuutta lisättiin 80–600 Hz taajuuksilla vaimentamalla kaiuttimen takasäteilyä. Äänenlaatua verrattiin tavanomaiseen kaksitiekaiuttimeen tekemällä vapaakenttämittauksia, huonemittauksia ja kuuntelukokeita.

Huonemittauksissa nähtiin huoneen jälkisoinnin vähenevän. Jälkisointi myös käynnistyi alemmalta tasolta ja näin jälkisoinnin tuottama psykoakustinen peittoilmiö väheni. Kuuntelukoedataa tutkittiin varianssi–, faktori– ja pääkomponenttianalyysillä. Tilastolliset analyysit kertoivat suuntaavuuden lisäämisen matalilla tajuuksilla parantavan koettua äänenlaatua merkittävästi. Suurempi suntaavuus lisäsi äänen selkeyttä, selkeytti äänikuvan rakennetta ja tarkkuutta, vähensi äänen värittymistä ja paransi transienttitoistoa tavanomaiseen kaksitiekaiuttimeen verrattuna.

Avainsanat Audiojärjestelmä, Psykoakustiikka Matalien Taajuuksin lokalisaatio, Heijastukset

Preface

I bought my first proper loudspeakers when I was 15 years old. Most teenagers prefer to have extreme amount of bass, I bought a pair of compact Genelecs. They produced impressive stereo image on my desktop. Buying the compact studio monitoring loudspeakers was the first step into music production hobby. I had played the piano and the saxophone mere 12 years and audio engineered tens of events or recordings when I visited career advisor in high school. Our high school study advisor pushed me towards acoustics and signal processing studies in Aalto University since I was both musically and mathematically talented. I was sold. Over 9 years after the one single meeting with career advisor I have arrived at the end of the journey, beginning of the next one.

First of all, I want to thank Genelec Oy for providing me a chance to work on my Thesis. I want to thank DSc Aki Mäkivirta for all the guidance and conversations we had during the Thesis process. My gratitude for guidance during this Thesis and courses in Aalto University goes to Prof. Vesa Välimäki. I want to thank my colleagues 2 × Juha, Jussi, Ilkka and Giles for all the brainstorming, humour and buzzing in our office cubicle. My gratitude goes also to my family who has supported me all the way from elementary school, without the values I received from them I would not be where I am at the moment. I want to thank my girlfriend Ninni for enduring me during the final sprint of my studies and the long Master's Thesis project. Special thanks goes to the people from PELMU, Teekkarispeksi, Lång Johnsons, Fii Fighters, Pöh(i)nä and many others who offered me well aimed distractions from studies even after moving from Otaniemi.

Iisalmi, 21.2.2020

Olli Kantamaa

Contents

Symbols

A_i	m^2	absorptive surface area
c	m/s	speed of sound in air
C		constant
e		euler's number
d	ms	delay
D	m	distance
f_n	Hz	eigenfrequency
f	Hz	frequency
F_t	Hz	transition frequency
H		transfer function
$H_m(z)$		modal resonance transfer function
I	W/m^2	sound intensity
k		wavenumber
l	m	length
p		p–value
p_s	Pa	sound pressure
P	W	acoustics power
Q		directivity factor
R_ϕ		radiation pattern
T_m	s	average reverberation time
T_{60}	s	room reverberation time
V	m^3	room volume
V_0	m^3	reference volume
α		absorption coefficient
λ	m	wavelength
τ	s	inter–aural time difference

Acronyms

ANOVA analysis of variance

AS auditory scene

ASW apparent source width

BEM boundary element method

dB decibel

DI directivity index

DRR direct-to-reverberant energy ratio

DUT device under test

EQ equalization

FDTD finite-difference time-domain

FEM finite element method

HF high–frequency

HRTF head related transfer function

IACC inter–aural cross correlation

ILD inter–aural level difference

IPD inter–aural phase difference

IR impulse response

ITD inter–aural time difference

ITU International Telecommunication Union

LEV listener envelopment

LF low–frequency

LMF low–middle frequency

MF middle–frequency

PA Public Address

PCA Principal component analysis

Q–value quality value

SI spatial impression

SNR signal–to–noise ratio

SPL sound pressure level

1 Introduction

Most loudspeaker configurations are stereophonic. Each recorded and reproduced sound has auditory attributes such as location, tone, loudness and the feeling of space. Loudspeaker configuration, listening room and the recording itself contribute to the perceived sound image and envelopment. Human auditory system has limitations but it can be surprisingly accurate. Properties of the listening room define the path from the vibrating loudspeaker cone to the ear drum of the listener. Radiated sound waves reflect on the room surfaces, sound wave transforms depending on the material property and size of the surface. Standing waves form between parallel surfaces in–room, frequencies of the standing waves are determined by the room dimensions. Radiating sound source contributes to the auditory attributes. A flat and smooth on–axis frequency response is the single most significant factor in creating high–quality sound reproduction. Conventional two–way loudspeaker has coloured off–axis response. Loudspeaker radiation pattern determines loudspeaker directivity.

Directivity of loudspeakers has been a subject of research for almost 70 years (Olson 1957). Many schools of directivity exist among the researchers, motivation for the research has varied from diminishing room effect and improving stereo imaging (Salmi & Weckström 1982, Salmensaari et al. 1993, Lipshitz & Vanderkooy 1985, Flindell et al. 1991, Ferekidis & Kempe 2005, Kamaris & Mourjopoulos 2018, Kates 2002). Optimal directivity increases the listening sweet spot size (Rodenas & Aarts 2001, Boucher et al. 2006, Hill 2012, Davis 1987). Perception differences on the vastly changing directivity pattern are small, (Evans et al. 2009). Small off–axis magnitude response fluctuations decrease sound quality (Olive 2004*a*,*b*). Less directivity improves the subjective sound quality by widening the sound image (Olive et al. (1995)). Literature includes opposite views on the desired directivity.

Low frequency directivity has been considered subsidiary to the auditory attributes due the natural radiation pattern of the recorded instruments. Directional hearing is not very sensitive in low frequencies (Blauert 1983, Griesinger 1998, 1999, Borenius 1985). Increasing low–frequency (LF) directivity has not been considered necessary. Low frequencies are very important to sound quality perception, low frequencies are localizable in certain conditions (Subkey et al. 2005, Kelloniemi et al. 2005, Griesinger 2018). Studios and sound control rooms use compact two–way loudspeakers for sound monitoring, compact loudspeakers are naturally omnidirectional in low and low–middle frequencies. Low frequency directivity control is substantially more difficult than high or middle frequency directivity control due the long wavelengths. The directivity control in long wavelengths needs either large mechanic structures or multiple driving units. Utilizing switchable directivity with compact 2–way loudspeakers, directivity performance is compared between conventionally radiating loudspeaker and low–middle frequency (LMF)–directive loudspeaker with measurements and listening tests.

This Thesis studies the effect of directivity on the sound quality perception. Sound quality improvements are pursued by increasing the directivity in the low–middle frequency band. In–room measurements and listening tests evaluate the effects of the increased directivity. Statistical analysis is conducted to confirm the significance of sound quality perceptions.

Section 2 describes the human hearing system. Section 3 discusses the effect of room for the sound quality perception. Section 4 covers the audibility of directivity and the hypothesis on the perceived sound quality when LMF–directivity is increased. Section 5 lays an overview specifically to the directivity measures of a loudspeaker and discusses the currently available loudspeaker directivity and covers the methods used to create directivity and discusses the design philosophy of designing directivity. Section 6 covers the properties of the prototype designed in this Thesis. Free–field and in–situ measurements are conducted to define the performance of auditioned loudspeaker and reduced room–coupling of more directive sound source. Results of the listening tests and their statistical analysis are presented in Section 7. Section 8 discusses the results and their meaning. Section 9 concludes the Thesis.

2 Auditory Event

When sound enters the human auditory system, it goes into ear canal, through eardrum and ossicles into cochlea where physical wave is coded to neural signals. Neural signal proceeds to brain via auditory nerve. Brain decodes the neural signal into an auditory event. Auditory event includes direction, loudness, frequency and time information. (Blauert 1983, Pulkki & Karjalainen 2015)

Non–damaged human hearing frequency range extends from 20 Hz to 20 kHz. The human ear is more sensitive to some frequencies than others, equal loudness curves visualise the frequency and sound pressure dependency on hearing threshold. Human auditory system has limited spectral resolution. Two equally loud narrow band signals are perceived as one or two auditory event depending their temporal and spectral difference. (Blauert 1983, Zwicker & Fastl 1999, Pulkki & Karjalainen 2015)

2.1 Terminology

Subjective attributes for describing the perceived sound field spatial information have been used to describe concert hall perception (Beranek & Martin 1996). Perception in small room adopts terminology concerning concert hall acoustics (Griesinger 1997).

Spaciousness is a relation between actual sound of sound source and auditory event width in concert acoustics (Norcross et al. 2003). "The perception of being surrounded by a large and enveloping space" (Toole 2017). "Spreading of the auditory events, in particular the apparent enlarged extensions of the auditory image compared to that of the visual image" (Lehnert 1993). "Start to think in terms of 'preference', 'spaciousness', 'low interaural cross-correlation (IACC)', and 'lateral reflections' as positively correlated with each other" (Toole 2017). Spaciousness and envelopment are used to describe low frequency sound field in a concert hall, even though envelopment is usually considered as the sensation of being surrounded by music. Envelopment is included in the good recording (Griesinger 2018). Spaciousness and envelopment are considered simultaneously as properties of programme, room, loudspeaker and human hearing system. In this Thesis, spaciousness is used to describe the authenticity of recorded space reproduced in small control room with loudspeakers.

Spatial impression (SI) describes the sound–field perceived inside room (Griesinger 1997, Lehnert 1993). "(SI is) The concept of type and size of an actual or simulated space to a listener arrives when he/she is exposed to an appropriate sound field" (Lehnert 1993). SI increases when level of reflections increases (Toole 2017). SI increased when more directive loudspeakers were used and reflection magnitude was decreased, Zacharov 1997. SI has many meanings and motivation, SI is a measure of sound stage credibility and authenticity (Zacharov 1997). Space size and image width increase SI in Toole (2017). SI is adopted to describe sensation of being in a virtual acoustic space, SI is used widely to quantify sound field created by immersive loudspeaker set–up or headphone spatialization. SI consists of apparent source width (ASW) and listener envelopment (LEV) (Pulkki & Karjalainen 2015, Toole 2017).

ASW has been referred as the sound image width (Toole 2017). Term ASW originates from concert hall acoustics where single instrument sounded a lot wider than point–like source due strong early reflections (Beranek & Martin 1996). ASW is "the width of a sound image fused temporally and spatially with a direct sound's image" (Morimoto 1997). Early reflection are attenuated in sound control rooms and ASW does not appear (Walker 1994).

LEV is "Envelopment – the perception of being surrounded by a beautiful acoustic space – is one of the joys of concert halls and great recordings" (Griesinger 2018). LEV "is the directional distribution of the auditory object associated with reverberant sound. LEV is judged to be high when the reverberant sound is perceived to arrive from all directions" (Beranek & Martin 1996). LEV is described "the degree of the fullness of sound images around the listener, excluding a sound image composing ASW" (Morimoto 1997).

The Figure 1 presents the head–related coordinate system used in this Thesis. Localization is the relation of direction between sound and perceived auditory event. (Blauert 1983, Lehnert 1993, Griesinger 1997. Localization includes information about direction, distance and spatial width of the event.

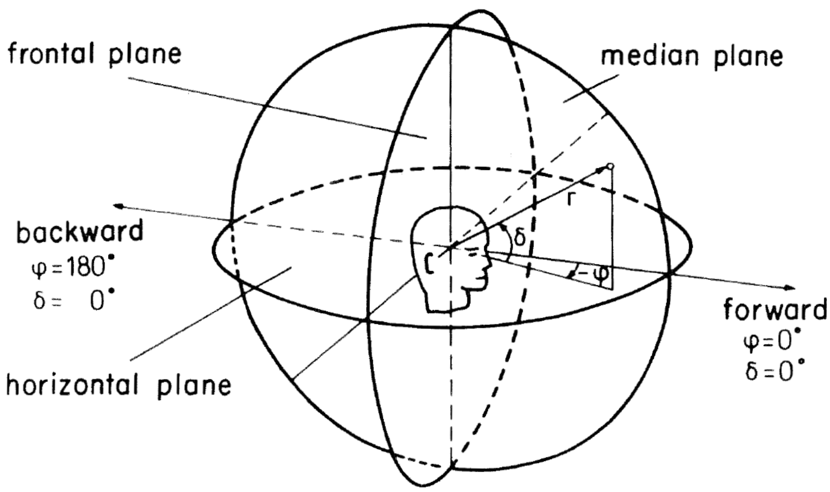

Figure 1: Head–related coordinate–system (Blauert 1983).

Sound image width and localization are used to describe the directional attributes. Timbral attributes are sound balance and sound colouration. Sound balance describes the balance across audible frequency band. Bass–heavy magnitude response is perceived as warm sound and vice versa. Sound colouration is used when describing tone change in one frequency band. Auditory coordinate system is described with lateral and sagittal directions. Lateral directions corresponds to left–right direction and sagittal corresponds to front–back direction. Azimuth is the angle in horizontal plane and elevation is the median plane angle. (Blauert 1983)

Frequency ranges are described with low, middle or high frequency. LF band corresponds to frequencies below 100 Hz, middle–frequency (MF) corresponds to frequencies between

100 Hz and 1 kHz. Frequency is considered high above 1 kHz. Low−middle frequency (LMF) range is 80 − 600 Hz in this Thesis.

Human physiology offers the binaural cues to help brains translate auditory event interpretation.

2.2 Human Auditory System

Transition path between sound and auditory event in human auditory system causes the signal magnitude and phase properties to change. Relation

$$H_{\mathrm{HRTF}}(f) = H_{\mathrm{ec}}(f)/H_{\mathrm{ff}}(f) \tag{1}$$

is called head related transfer function (HRTF), where free−standing transfer function is $H_{\mathrm{ff}}(f)$ and in−ear transfer function is $H_{\mathrm{ec}}(f)$. HRTF magnitude and phase varies d in function of sound wave arrival direction and frequency. (Pulkki & Karjalainen 2015).

HRTFs has effect in frequencies, where the structures in the outer ear or the torso have dimensions similar to the wavelength and and do effect magnitude and phase of the incident sound wave by reflections, scattering, shadowing and diffraction. HRTF describe colouration to the sound depending on the direction of arrival. Brain automatically interprets the colouration as the direction of arrival. Fluctuation in HRTFs occurs with both azimuth and elevation angles. HRTFs contribute to the directional hearing in both median and frontal plane. Localization in median plane has other binaural cues since the ears are located in sides of head. (Pulkki & Karjalainen 2015) HRTF represents the whole performance of the human auditory system. Diffraction, shadowing and reflecting defined the inter−aural level difference (ILD) and inter−aural time difference (ITD) are combined in the HRTF. The directional hearing is characterized with HRTF (Zacharov et al. 2000).

ILD is the sound pressure level difference between ears, measured at the ear drum. ILD is caused by shadowing, reflections and diffraction of the head between ears, ILD is effective in short wavelengths. Short wavelengths scatter and reflect from the head and outer ear structures (Pulkki & Karjalainen 2015). ILD contribution to directional hearing increases with frequency (Feddersen et al. 1957). Figure 2 presents the perceived sound direction in function of ILD.

ITD is the difference in the time of arrival of sound at the left and right ear drums. ITD depends on the distance of the ears and angle of arrival:

$$\tau = \frac{D}{2\mathrm{c}}(\phi + \sin\ \phi) \tag{2}$$

where D is distance between ears, c is the speed of sound and ϕ is the arriving direction. Arrival time difference causes frequency depended phase shift between the signals called inter−aural phase difference (IPD). Phase and time differences cause sound colouration

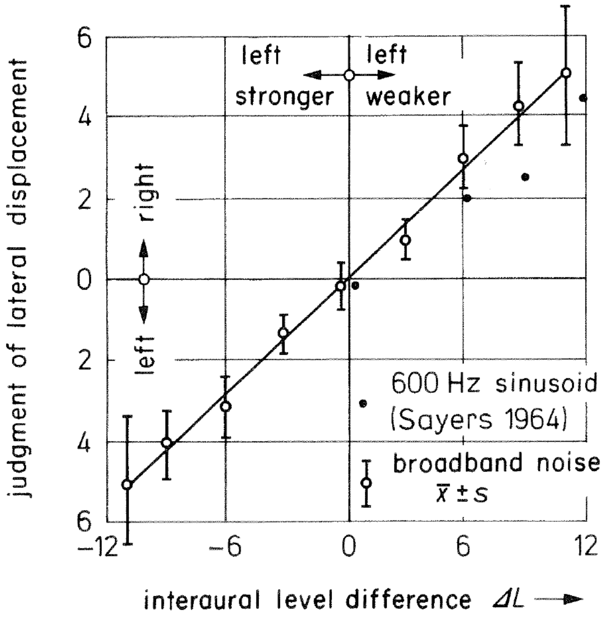

Figure 2: Perceived sound direction in function of ILD (Blauert 1983).

due cancellation and comb filtering. IPD can be perceived as lateral displacement (Pulkki & Karjalainen 2015, Blauert 1983). ITD varies from 0 μs to approximately 600-700 μs depending on the direction of arrival of sound (Blauert 1983). Figure 3 presents the perceived sound direction in function of ITD.

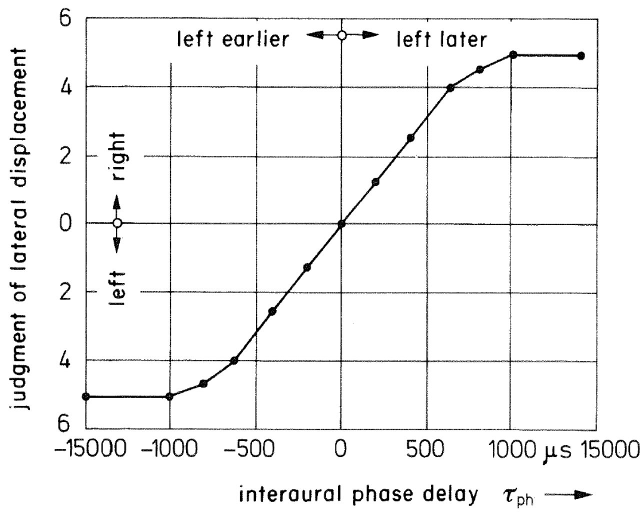

Figure 3: Perceived sound direction in function of ITD (Blauert 1983).

ILD is the main cue for directional hearing in high frequency (Pulkki & Karjalainen 2015). ITD is the directional cue in the low frequencies, the direction perception has been thought to be weaker in low frequencies (Borenius 1985).

Inter–aural cross correlation (IACC) describes the similarity of the sound arriving at the ear drums. IACC is the difference between sagittal and lateral reflections. IACC varies between 0 to 1, where 0 responds to uncorrelated signals and 1 responds to fully correlated signals. IACC can be measured artificial head. Pulkki & Karjalainen 2015

2.3 Virtual Sound Image

Stereo loudspeaker set–up forms virtual sound image of directly between left and right loudspeakers. Precedence effect, Haas effect or law of the first wavefront enables the perception phantom source and virtual sound image. (Haas 1972) Stereophonic listening set–up is capable of fading out the actual sound sources and create uniform sound image or auditory scene (AS) (Linkwitz 2009, Clark 2010). Virtual sound image includes the reproduced sound image width and localization of the auditory objects. "Speakers Disappear, Local Acoustics not heard, Images Lateral Localization, Images Depth Localization, Ambience Non–Localized and Freedom of Movement" describe AS (Clark 2010). Virtual sound image perception allows human to sense wide sound stage rather than sound coming from individual sound sources (Linkwitz 2007).

Directional hearing is the main requirement for human being able to form virtual sound image. Stereo imaging and localization in–room is complex and multidimensional problem since sound wave propagating from sound source to ears has infinite amount of transfer paths. Single reflection causes sound location displacement, sound colouration and secondary auditory event perception. (Haas 1972, Barron 1971, Pulkki & Karjalainen 2015, Blauert 1983, Olive & Toole 1989)

Directional hearing and localization are more accurate on the middle and high frequencies (Blauert 1983). ITD and phase coherence enable directional hearing below 500 Hz (Griesinger 1997). Low frequencies are localizable above 120 Hz (Miller III 2005). LF localization transforms into envelopment below 120 Hz (Griesinger 1998, 1999). Subwoofer is localizable with crossover frequency higher than 120 Hz (Kelloniemi et al. 2005). Ability to localize LF depends on the listening conditions. Low frequencies are localizable below 60 Hz in free–field conditions (Subkey et al. 2005). Low frequencies are localizable in acoustically treated listening rooms (Benjamin 2006, Wang et al. 2016). Spatial information on low frequency band contributes to the envelopment, localization and sound source broadening (Griesinger 1997, 1999, 2012, 2018).

Large sound image size is connected with low IACC (Toole 2017, Subkey et al. 2005, Perrott & Buell 1982). Low IACC is caused by strong lateral reflections (Toole 2017). IACC increases with the presence of sagittal reflections (Ando 1977). Spatial perception models predict perceived sound stage and spatial sound reproduction quality (Theiss & Hawksford 1998, Kamaris & Mourjopoulos 2018).

2.4 Sound Colour Perception

Sound colour describes the timbre of an auditory event. Flat and smooth in–room or in–situ magnitude response is the most important factor for sound quality rating (Schulein 1975). Smooth free–field on–axis response leads to smooth in–situ magnitude response. (Olive 2004*a*,*b*)

Predicting perceived timbre of system were implemented including effect of HRTFs using ear–microphones (Staffeldt 1984). Spatially averaged model predicts the perceived magnitude response of loudspeaker systems (Devantier 2002). Free–standing microphone measurements without HRTFs predict predicting the perceived timbre (Lipshitz & Vanderkooy 1985). Free–field measurements do not include directional information. Two different kinds of sound colour –related attributes can be described, tonal balance and sound colouration. Tonal balance is the balance across the audible frequency range. Signal including more LF and LMF than high–frequency (HF) content translates to warm tone. Sound colouration differentiates sounds with same level, duration and pitch different to each others (Pulkki & Karjalainen 2015).

Secondary sound source with variable delay and level simulate early reflections in the laboratory experiments. Delaying secondary sound source modifies perceived timbre, one subject perceives only one auditory event. Reflections modify the perceived sound structure and single auditory event is perceived despite multiple sound sources (Haas 1972).

2.5 Masking

Spectral and temporal masking modify the perceived sound colour and clarity (Zwicker & Fastl 1999). Masking effect occurs when multiple sound sources produce sound waves closely in the same frequency band, simultaneously or slightly offset in time domain. Masking can be divided to two classes, temporal and spectral. Masking changes the sound structure of auditory event on each hearing band. (Pulkki & Karjalainen 2015)

Spectral masking occurs when auditory event is present. Continuous auditory event masks spectrally close signals. Spectral masker width dependents on the frequency content and level. Louder pure sine tone creates wider masker than spectrally equal attenuated signal (Pulkki & Karjalainen 2015). Wideband signal creates wide masker attribute. Even pure sine tones create masker that considerably wider in frequency than the tone itself, harmonics of fundamental tones create very wideband masker (Zwicker & Fastl 1999).

Temporal masking refers to masking occurring in time–domain. Temporal masking occurs after finished sound. Temporal masking hides sounds that occur closely to the masker in time and frequency. Temporal mask is effective long period after perceived masker with large dynamic range. Increased masker duration and level increase the effective masker duration (Pulkki & Karjalainen 2015)

2.6 Clarity

Clarity is defined as quantity how well a sound from single instrument can be distinguished from the rest of musical performance when there is multiple sound sources present (Beranek 2012). Room acoustics definition for clarity in is the magnitude difference between early and late arriving acoustic energy (Pulkki & Karjalainen 2015). Clarity is separability of auditory events in this Thesis. Clarity is decreased with masking and strong early reflections (Hermes 2019, Imamura et al. 2014, Jensen & Welti 2003).

3 How Room Modifies Perception

Room Room influence on the perceived sound stage is substantial. Room influence on perceived sound can be 80 % of the total performance available in sound reproduction system (Salmi & Weckström 1982). This section reviews room effect mechanisms and concludes the effect on localization, sound image structure and sound colour perception.

Radiated sound reflects from the walls in room. Reflections can be perceived as de–localization, sound colour change or echo (Toole (2017), Haas (1972)). Parallel surfaces in–room form standing waves or room modes. Sound source inside room excites the room modes. Small rooms have pronounced room effects, the modal density is large in small room (Celestinos & Nielsen 2008). Sound has infinite amount of transfer paths from sound source to the listener (Backman 1995).

The in–room situation with infinite amount of reflections and directions is very complicated. Despite reflective energy existing in–room, stereophonic or multichannel loudspeaker set–up is able to create virtual sound image reproducing the original recorded sensation of sound in space. (Zacharov et al. 2000)

Loudspeaker location in room dominates the loudspeaker fidelity rating (Bech 1994). Wakened room mode frequency and magnitude depends on the location inside room (Groh 1974). Optimum placement of a loudspeaker can improve the flatness of magnitude response significantly (Ballagh 1983). Stereo image formation can be improved significantly with proper in–room placement (Linkwitz 2009). Finite-difference time-domain (FDTD) simulations model standing wave behaviour in–room (Celestinos & Nielsen 2008). Perception of sound quality depends significantly on the sound source location in–room (Bech 1994, 1995, Salmi & Weckström 1982, Olive et al. 1995).

3.1 Diffuse Field

Sound field where sound waves travel all directions simultaneously is a diffuse sound field (Beranek 2012). Diffuse sound field is a property of large concert hall (Toole (2017)). Modal sound field consists of sound waves that are too long for the room measures, both modal and diffuse sound fields exist in small rooms (Schultz 1963). In–room magnitude responses depends on reverberation time under diffuse sound field field conditions (Kuttruff & Thiele 1954).

Crossover or transition frequency between diffuse and modal sound fields is

$$F_{\mathrm{S}} = 2000 \times \sqrt{\frac{RT_{60}}{V}} \tag{3}$$

in large room or concert hall where RT_{60} is reverberation time and V is room volume in cubic meters. Translation frequency exists in small room, Equation 3 do not represent the transition frequency accurately in small room, transition frequency is typically higher than calculated transition frequency values in small room (Toole 2017).

Below the transition frequency is modal region. Room properties dominate the in–room magnitude response below transition frequency (Toole 1999). Room parameters and excitation define the in–room magnitude response in modal sound field (Toole (2017)). Figure 5 presents the in–situ magnitude response of one loudspeaker measured in 3 different rooms. LMF range magnitude responses fluctuate depending on the room. MF and HF magnitude responses remain similar between rooms despite the changed room parameters.

3.2 Reflections

Reflection is a sound wave which has been emitted from sound source and encountered one or more solid surfaces. Reflection on any surface follows the law of reflection which states that the incident angle and reflection angle are equal. Sound wave can encounter infinite amount of boundaries on the way to listening position. Each reflection absorbs energy from the travelling sound wave. (Olson 1957) Sound control rooms have attenuated early reflections (EBU 1998).

Absorption is a material property, every time sound wave encounters boundary, some energy arbsorbs and rest of the energy reflects (Sabine 1927). Absorption coefficient

$$\alpha = \frac{I_\mathrm{i} - I_\mathrm{r}}{I_\mathrm{i}} \tag{4}$$

describes frequency dependent material property where I_i is the intensity of incident sound wave and I_r is the intensity of reflected sound wave (Fahy & Salmon 1990). Absorption coefficient α depends on frequency, reflections have different spectral content than the direct sound (Olive & Toole 1989). Low frequencies need significantly thicker materials to attenuate. Low order reflections have greater magnitude than high order reflections (Hongisto 2015)

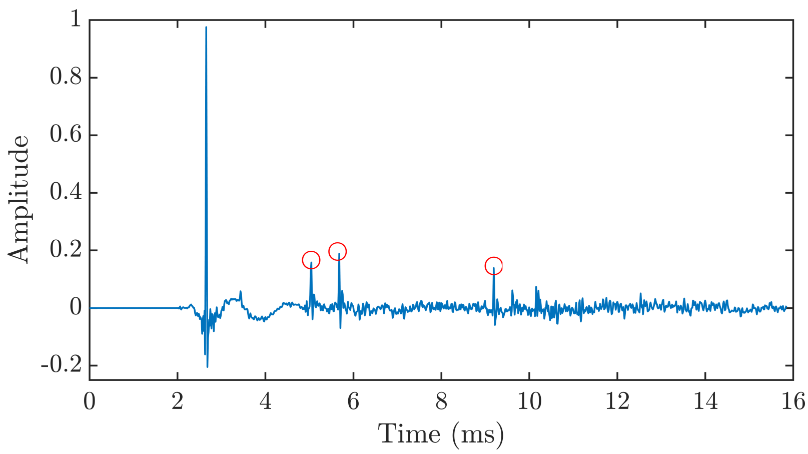

Figure 4: Room impulse response (IR), red circles indicate early reflections.

Early reflections are visible in room IR after the main impulse (see Figure 4). Arrival time between 15 –80 ms is used as the limit for early and late reflections in the literature (Jensen & Welti 2003, Haas 1972). Reflection intensity decreases as a function of time due the air and reflection attenuation. Reverberation is the sum of late reflections.

Early reflections in room modify sound colouration and localization. Sound image structure was modified by modifying time delay and level of the simulated reflection. (Haas 1972) Long time delay with equal sound level sound sources was needed to form two auditory events, shorter delays yielded changes in sound image structure (Haas 1972). The magnitude, delay and spectrum of single lateral reflection has been examined with simple experiment in free-field, where reflection has been simulated with secondary sound source (Olive & Toole 1989, Toole 2017).

Comb filtering changes timbre and sound structure in experiment (Haas 1972). The acoustical sum of direct sound wave and reflected wave does not sum, waves do not sum. Frequency dependent cancellation occurs at single frequency and comb filtering corrupts the magnitude response (Bücklein (1981)). Dips in magnitude response are less audible than peaks (Barron (1971), Case (2001)).

Early reflections cause localization shift when auditing pink noise and speech signals, perceived localization change can be stronger than sound colour modification (Olive & Toole 1989). Tests were conducted in both anechoic chamber and lively listening rooms. Image shift threshold was the same in all rooms with time delay from 0 to 25 ms, test conducted in anechoic chamber yielded significantly lower image shift thresholds with 25 ms or longer delays. (Olive & Toole 1989) Widened sound image is the auditory attribute for significant amount of early reflections (Bradley & Soulodre 1995). Early reflections in the lateral direction widen the perceived stereo image (Toole 2017). Sagittal reflections can disrupt the stereo image integrity (Ando 1977). Early lateral reflections below 700 Hz affect greatly on the stereo image perception (Angus 2013). Subjects can adapt to large amount of lateral reflections (King et al. 2011).

3.3 Reverberation

Reverberation is the sum of late reflections. Absorption coefficients of the in–room surfaces define reverberation timbre and time (Toole 2017). Reverberation time RT_{60} is an objective measure of reverberation duration. RT_{60} describes the amount of time for 60 decibel (dB) attenuation of a room impulse response.

Acoustic energy bouncing in room attenuates in each reflection, reflected frequencies depend on the surface absorption coefficients (Toole & Olive 1988). Reverberation time RT_{60} depends on the volume of the room and absorption with relation

$$RT_{60} = 0.161 \frac{V}{\sum \alpha_i A_i} \tag{5}$$

where α. A_i is the surface area with surface absorption of α_i (Sabine 1927). Sabine's formula is not accurate in small room but it represents the reverberation time being function of room parameters. (Beranek (2006)).

Average reverberation time (T_{m}) is defined:

$$T_{\mathrm{m}} = 0.25 \times (V/V_0)^{1/3} \tag{6}$$

where V_0 is reference room volume of 100 m^3. ITU (2015) standardized reverberation limits are relative to T_{m}. The required reverberation time is short and similar to domestic small room reverberation time (Pulkki & Karjalainen 2015).

Reverberation influences perceived auditory attributes such as timbre, loudness, source width (Kaplanis et al. 2014). Reverberant sound defines the masking pattern after the direct sound. Sound decay colours the auditory attribute Reverberation time and spectrum modify the in–room magnitude response above the transition frequency in conjunction with loudspeaker radiation pattern. (Kuttruff & Thiele 1954, Thiele & Kügler 1992)

3.4 Direct to Reverberant Sound Energy Ratio

Direct-to-reverberant energy ratio (DRR) is the sound energy ratio between of direct sound field and reverberant sound field. Increasing the DRR means more direct sound and less reverberant sound. DRR diminishes when listening distance increases or sound source directivity decreases producing more sound energy to reflect. DRR Room effects are stronger with longer listening distances since the relation between the direct sound and reflected sound decreases. Amount of reflections will decrease and DRR will increase when the directivity of sound source increases. (Laitinen et al. 2015)

Increasing DRR improves directional hearing, increasing reverberation time and decreasing DRR affects the accuracy of directional hearing. Increasing DRR produces more accurate localization. Auditioning only the reverberation do not completely impair listeners capability of detecting direction. Increasing the reverberation time decreases the accuracy of directional hearing. (Gómez et al. 2011)

3.5 Room Modes

Loudspeaker couples to the listening room in the whole frequency bandwidth produced by the sound source. Reflections with suitable wavelengths form forced vibrations depending on the room measures (Hill 2012). Small rooms used for sound reproduction and critical listening include high modal density, plethora of room modes have significant effect on the in–room magnitude response and masking (Fazenda 2004).

Eigenfrequencies have been quantified with equation

$$f_{n_{\mathrm{x}}n_{\mathrm{y}}n_{\mathrm{z}}} = \frac{\mathrm{c}}{2}\sqrt{\frac{n_{\mathrm{x}}}{\mathrm{l_x}} + \frac{n_{\mathrm{y}}}{\mathrm{l_y}} + \frac{n_{\mathrm{z}}}{\mathrm{l_z}}} \tag{7}$$

where c is speed of sound in air, $\mathrm{l_x}$, $\mathrm{l_x}$ and $\mathrm{l_x}$ are room measures, integers

$$n_{\mathrm{x}}, n_{\mathrm{y}}, n_{\mathrm{z}} \geq 0 \in \mathbb{Z} \tag{8}$$

(Toole 2017). Equation 7 suggests that infinite amount of room modes exist in any room. Modal structure in–room is complicated, simulations and approximations predict in–room modal behaviour (Hill 2012).

Room modes have properties in time–domain. Envelopment and magnitude depend on the sound source excitation and room parameters. Single room modes can be described as transfer function. (Mourjopoulos & Paraskevas 1991) Transfer function

$$H_{\mathrm{m}}(z) = \frac{1}{(1 - re^{j\theta}z^{-1})(1 - re^{-j\theta}z^{-1})} \tag{9}$$

presents single modal resonance. Transfer function has a pair of poles in z–domain with radius r and pole angle θ. Reducing the radius r decreases quality value (Q–value) and shortens the decay time (Karjalainen et al. 2003).

Passive absorption or resonators control room modes (Antsalo et al. 2004, Noy et al. 2003). Absorption needed for long wavelength low frequencies is achieved with large damping material thickness and density (Kashani & Wischmeyer 2004). Small rooms used for critical listening have high mode density (Fazenda 2004). Mode density is not directly linked to better sound quality (Fazenda & Wankling 2008). Modifying loudspeaker radiation pattern changes the induced in–room magnitude response (Salmi & Weckström (1982), Ferekidis & Kempe (2005)). Using one or more cardioid subwoofers improve the LF magnitude response (see Figure 6). Space is limited in a small room, large amount of absorbing material or large volume resonator bass trap are not feasible solutions.

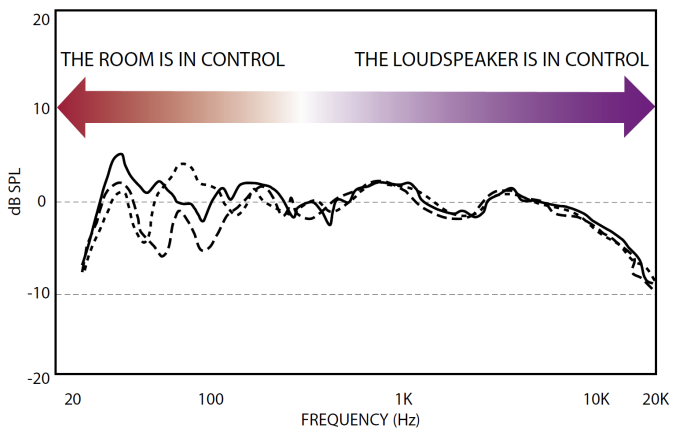

Figure 5: A loudspeaker measured in 3 different rooms (Toole 1999).

Active room mode control is implemented with room equalization (EQ) (Schulein 1975, Genereux 1990). Magnitude EQ attenuates the room mode steady state and decay linearly. Modal EQ controls the gain and decay of the room mode (Mäkivirta et al. 2003). Loudspeaker do not control room decay, it is a property of the room (Beranek & Martin 1996). Acoustic energy exciting the room mode depends on the radiation pattern of the loudspeaker placed inside room (Ferekidis & Kempe 1996).

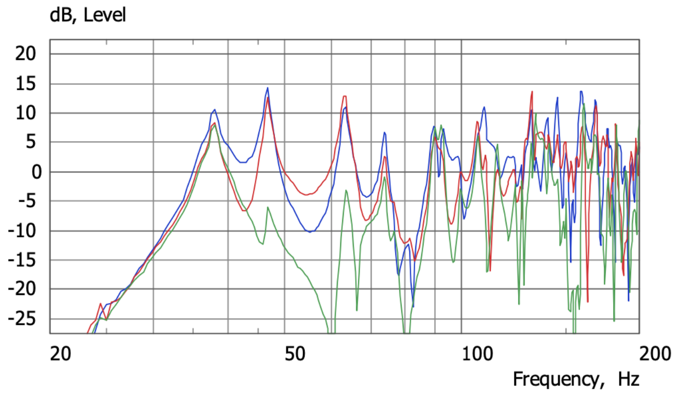

Figure 6: 3 loudspeakers measured in one room, cardioid, 4 x cardioid, omnidirectional (Ferekidis & Kempe 2005).

Single room modes are audible below transition frequency (Toole 2017). Loudspeaker position defines the excited room modes and perceived sound colour (Bech (1994)). Room modes are audible when constructive resonance or peak occurs, destructive resonance or dip are less audible than constructive resonance (Bücklein 1981.) Large Q–value increases audibility of room resonance (Fazenda 2004). Room mode decay creates masker which affects the perceived clarity and sound colouration. Small differences in room energy attenuation are audible due masker magnitude and spectrum change (Antsalo et al. 2004).

4 Audibility of Directivity

Area coverage of sound do not define desired directivity, perceived sound quality defines the desirable directivity (Queen 1979). Smoothly behaving off–axis responses are significant factors for the loudspeaker preference ranking (Devantier 2002, Olive 2004*a*,*b*, Gentner et al. 2007). Modifying directivity changes loudspeaker–room coupling and sound quality perception (Olive & Martens 2007).

4.1 Audible Effects of Loudspeaker Directivity

4.1.1 Changes in Early Reflection Character

The precedence effect enables single auditory event perception despite strong early reflections and reverberation (Norcross et al. 2003). Strong early reflections must be avoided in control rooms (EBU 1998). Early reflections cause linear distortion and significant image shifts (Walker 1994). Decreased lateral reflections improves localization below 700 Hz (Angus 2013). Early reflections should have similar spectral content to the direct radiated sound to be advantageous for the sound quality (Olive et al. 2013). The room absorption affect magnitude and sound colour of the arriving reflections (see Section 3.2). Loudspeaker radiation pattern and surface absorption define the spectral content and magnitude of arriving reflection (Walker 1994). Strong early reflections cause masking and reduced clarity (Jensen & Welti 2003). Subjects ability to perform critical listening tasks is impaired with strong lateral reflections (King et al. 2011). Directivity effects virtual sound source location and image integrity, timbre, sweet spot, localization accuracy and general perception of sound quality, small directivity radiates more sound energy to the room. More radiated sound energy increases early reflection and rooms mode magnitudes (Evans et al. (2009)).

Increasing directivity improves localization, spectral balance and transient response (Kantor & de Koster 1985). Magnitude response and sound colouration from strong early reflections improves with more directivity in LF range by radiating less energy to the room (Salmensaari et al. 1993). Unidirectional loudspeakers induce less sagittal reflections and yield better listener envelopment (Gudvangen 2014). Increased directivity improves the directional hearing accuracy (Gómez et al. 2011). Higher directivity with more precedent sound leads to more authentic sound reproduction (Zacharov 1997). Small directivity causes strong early reflections that de–localizes sound (Hartman 1983, Toole 1986*b*). Dipole radiation pattern diminishes magnitude response irregularities at 500 – 2000 Hz (Kates 2002). Directive loudspeaker increases DRR (see Section 3.4), sound production engineers prefer larger DRR (Griesinger 2009).

4.1.2 Changes in Decay Spectrum

Room modes are the single most significant factor affecting LF and LMF sound quality (see Section 3.5). Reverberation time and timbre have significant influence on the perceived sound quality (see Section 3.3). Sound decay describes energy after direct Temporal masking caused by room modes and reverberation has effect on the perceived clarity and sound colour (see Section 2.5). Equalization controls static sound wave magnitude, equalization do not control the sound decay (Antsalo et al. 2004, Toole 2015).

Resistance enclosure creates LF directivity (Ferekidis & Kempe 2004). Loudspeaker couples to the room modes differently depending on the radiation pattern, more directive LF radiator improves magnitude response and reduces decaying acoustic energy significantly (Ferekidis & Kempe 2004, 2005). Low–frequency cardioid sound source with significantly increased LF directivity reduces magnitude response fluctuation caused by rear radiation and is less dependent on the room position in room mode excitation point of view (Ferekidis & Kempe 2004). Dipole radiation pattern yields improved magnitude by controlled room mode excitation (Ferekidis & Kempe 1996, Kates 2002).

Room resonances excited by dipole sound source are significantly less audible up to 200 Hz (Linkwitz 1998). Subjectively more accurate reproduction was achieved with dipole LF radiation pattern at 30–200 Hz (Linkwitz 2003). Loudspeaker set–up with LF directivity capability can reproduce the recorded LF envelopment (Griesinger 2018). More directive loudspeakers excite less room and and receive higher preference ratings when used in multichannel system Zacharov (1998). Simulated wideband cardioid loudspeaker yields less fluctuation in room magnitude response (Backman 2003).

4.2 Hypothesis to be Tested

The working hypothesis in this Thesis is very simple: Increasing loudspeaker LMF range directivity improves sound quality. More uniform directivity radiates less energy in room with less coloured off–axis magnitude response which attenuate the the early reflection level and room mode resonances. This Thesis concentrates on increasing the directivity in LMF range and monitoring the perceived sound quality. Attenuated room mode magnitudes and decreased decay yield less masking and improved room response, sound colouration decreases and clarity increases. Less reflections will create narrower and more focused stereo image with less localization errors. Essential questions are the magnitude of sound quality improvement and magnitude of the subjective perception of sound quality improvement.

5 Loudspeaker Directivity

Most currently available monitoring loudspeakers are so called conventional loudspeakers with the moving coil transducers. In this context conventional loudspeaker stands for two– or three–way loudspeaker where a single driver is responsible for single frequency band and the drivers are pointed towards the listening position. Conventional two–way loudspeaker has one woofer responsible for the low and middle frequency bands, one transducer reproduces the high frequencies.

All sound emitting devices have non–uniform 3–dimensional sound power radiation pattern. Figure 7 presents the polar coordinate system used in this Thesis. Azimuth or horizontal angle ψ and elevation or vertical angle θ describe directions from the loudspeaker point of view.

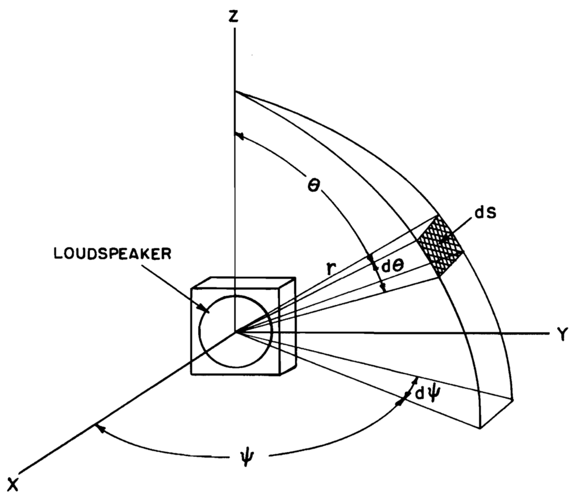

Figure 7: Coordinate system of loudspeaker (Olson 1957).

The of non–uniform radiation pattern of sound source varies reverberant and reflected sound spectrum, in–room listening situation is very complex with multiple mechanism of sound transfer and perception (see Sections 2 and 3). This chapter focuses on describing loudspeaker directivity with measurements.

5.1 Measures of Directivity

Directivity of a loudspeaker is measured under free–field conditions in the anechoic chamber without any room effects affecting the measurements. Presenting accurate 3–D directivity presentation is difficult, simplified measures have been developed to describe directivity adequately.

5.1.1 Dispersion

Dispersion is a direction where the magnitude has attenuated by 6 dB. Dispersion can be plotted as function of frequency but it is often announced as one number. Dispersion describes the beamwidth of Public Address (PA) system. Dispersion defines the areas covered with sound. (d & b Audiotechnik 2015)

5.1.2 Power Response

Omnidirectional sound source emits total acoustic power

$$P_{\mathrm{AN}} = \frac{4\pi r^2 p_s^2}{\rho c} \tag{10}$$

where p_s is sound pressure in $\mathrm{Pa}/10$ and r is distance in centimeters (Olson 1957). Sound power emitted by directional sound source is

$$P_{\mathrm{AD}} = \frac{r^2}{\rho c} \int_0^{2\pi} \int_0^{\pi} p^2(\theta, \psi, r) \mathrm{sin}\theta \, d\theta \, d\psi \tag{11}$$

where θ and and ψ are the polar coordinates (Olson 1957). θ responds to the vertical direction and ψ responds to the horizontal direction (see Figure 7).

Power response do not present the relation between on–axis magnitude response and off–axis radiated energy. Directivity index (DI) represents the relation between power response and on–axis magnitude response. Inverted power response equals directivity index for a radiator with flat on–axis magnitude response.

5.1.3 Directivity Index

Point source radiation and directional power response define directivity factor Q (Olson 1957, Beranek 2012).

$$Q = \frac{P_{\mathrm{AN}}}{P_{\mathrm{AD}}} \tag{12}$$

Q is relation between point source power response (see Equation 10) and directive source power response (see Equation 11). DI in dB is calculated from Q using equation

$$DI = \log_{10} Q \tag{13}$$

DI is defined as a sound power relation between on–axis magnitude response and the power response (Toole 2017). Omnidirectional source has equal P_{AN} and P_{AD} and Q = 1, which yields DI of 0. P_{AD} decreases and Q increases for directional sources. DI describes the overall sound power radiating from sound source in function of frequency. (Olson 1957)

5.1.4 Polar Plot

Polar plot presents radiation pattern in polar coordinates as function of θ or ψ. Polar plot have been used describing constant directivity horns (Klippel 1995, Sinclair 1978). Single polar plots describes point frequency directivity (Vanderkooy 2006, Di Cola et al. 2009, Günel et al. 2007, Gómez et al. 2011, Panzer & Ponteggia 2011). Single polar plot does not present full picture of loudspeaker directivity. Off–axis response have wide dynamic range across directions and frequency. Polar plot requires coarse magnitude scale when all data is presented, coarse magnitude scale hides details in directivity.

5.1.5 Off–Axis Responses

Manufacturers can present the off–axis responses in the loudspeaker data sheet (Genelec Oy 2014). Off–axis present magnitude responses corresponding to the angles 0°, 15°, 30°, 45° and 60° or their relation to on–axis magnitude response. Measurements are conducted in free–field conditions separately for horizontal and vertical angles ψ and θ. Frequency smoothing is used to improve readability of the off–axis response figures. Figure 8 presents an example of off–axis response plot. Off–axis responses give approximation of the off–axis colouration of the examined loudspeaker. Presented off–axis responses do not present the rear radiation spectrum (see Figure 8).

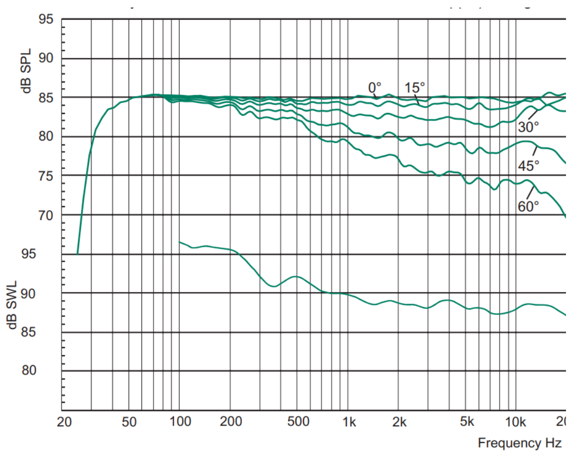

Figure 8: Off axis response of professional studio monitoring loudspeaker (Genelec Oy 2014).

5.1.6 Directivity Contour Plot

Directivity contour plot presents directivity accurately (Di Cola et al. 2009, Mäkivirta et al. 2017, Sridhar et al. 2016, Tylka et al. 2015). Directivity contour plot describes sound pressure level (SPL) zone in function of off–axis angle and frequency. Directivity contour plot presents off–axis colouration in reference to on–axis responses. Directivity plot combines the angle resolution of polar plot and frequency resolution from off–axis magnitude response plot. Directivity plot is drawn normalized to the on–axis magnitude response. Contour lines mark attenuated zones with 3 dB resolution, colour fills aid the visual interpretation of the contour plot (see Figure 9).

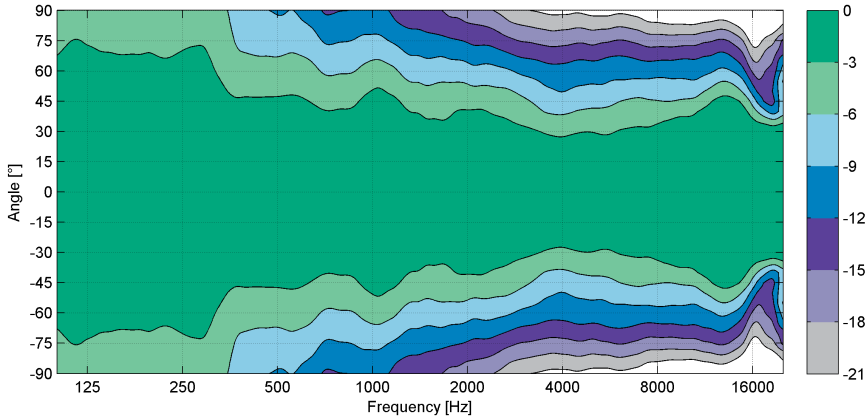

Figure 9: Example of Directivity plot with 3 dB resolution and 1/3rd octave smoothing (Genelec Oy 2014).

Angle resolution, frequency smoothing and magnitude zone size define directivity contour plot readability and accuracy. Small frequency smoothing produces blurry images. Large magnitude and angle resolutions hide details. This Thesis uses 5° angle resolution, 3 dB magnitude zone resolution and 1/3–octave smoothing.

5.2 ITU Recommendation

International Telecommunication Union (ITU) has defined the desired radiation pattern using DI as the directivity measure (see Section 5.1.3). Required DI in (ITU 2015) is

$$6\ \mathrm{dB} \leq DI \leq 12\ \mathrm{dB} \tag{14}$$

in the frequency band of

$$500\ \mathrm{Hz} \leq f \leq 10\ \mathrm{kHz} \tag{15}$$

Off–axis magnitude response at ψ = 10° and θ = 10° should not differ more than 3 dB from on–axis magnitude response. Off–axis response at ψ = 30° and θ =30° should not differ more than 4 dB from the on–axis magnitude response in the full produced frequency band. ITU standard requires the directivity to increase with frequency (ITU 2015). Conventional two–way loudspeaker directivity increases with frequency (see Section 5.3).

5.3 Typical Loudspeaker Directivity

Seven two–way loudspeakers A–G and one large three–way loudspeaker H were measured in anechoic chamber. DI for each loudspeaker was calculated (see Section 5.1.3). All measured loudspeakers fulfilled the ITU (2015) requirement of smooth off–axis behaviour and increasing directivity with frequency. Most of the loudspeakers did not have enough directivity in the low–limit frequency for International Telecommunication Union (ITU) compliance. Table 1 presents discrete frequency DIs with ITU (2015) compliance. Table 2 presents transducer configurations of the measured loudspeakers. Directivity of the conventional loudspeaker depended significantly from the physical measures of the driver (see Section 5.6.1). Larger driver corresponded to more LMF directivity.

Baffle and driver size control MF and LF directivity (see Tables 1 and 2). Waveguide controls high frequency directivity (see Section 5.6.2).

	Directivity Index				
	125 Hz	**500 Hz**	**1 kHz**	**10 kHz**	**ITU-R BS.1116-3**
A	0.6	4.6	7	8.4	Not compliant
B	1	3.8	6	8.7	Not compliant
C	1.9	4.4	6.2	10.6	Not compliant
D	1.7	4	6.5	8.1	Not compliant
E	1.9	5.3	8	9.7	Not Compliant
F	1.8	5.5	7.1	8.7	Not Compliant
G[1]	3.9	5	7	8.4	Not Compliant
H	3.0	6.3	10	11	Compliant

Table 1: DI values for herd of commercially available professional monitoring loudspeakers. [1] Loudspeaker G is DUT B constructed for this Thesis (see Section 6.1.)

	Woofer Size(s)	**Baffle Width**	**Crossover Frequency F_{co}**
A	5 inch	185 mm	1.7 kHz
B	5 inch	189 mm	2.7 kHz
C	5 inch	190 mm	3 kHz
D	6.5 inch	237 mm	3 kHz
E	8 inch	286 mm	1.7 kHz
F	10 inch	320 mm	1.8 kHz
G	2 × 5 inch	189 mm	2.7 kHz
H	15 and 5 inch	480 mm	420 Hz & 3 kHz

Table 2: Table of loudspeaker A–H driver configurations and baffle sizes.

None of the currently available two–way loudspeaker exploit possibilities of creating directivity actively. Loudspeaker G is the device under test (DUT) B constructed for this Thesis and created directivity actively (see Section 6.1). Despite the size difference across the observed loudspeaker, all but the large three–way loudspeaker fail the ITU requirement. None of the two–way loudspeakers have DI of 6 dB at 500 Hz. Large three-way monitoring loudspeaker H with large than 15 inch driver reaches the ITU requirement with natural directivity. Compact loudspeaker with large natural directivity is not possible (see Section 5.6), active methods must be used to increase directivity.

5.4 Design Philosophies: Omnidirectional to Highly Directive

5.4.1 On–Axis Response

On–axis free–field magnitude response is the most important factor for producing flat in–situ magnitude response, flat in–situ magnitude response corresponds to good perceived sound quality (Toole 1986*a*,*b*, Olive 2004*a*, Bridges 1980, Devantier 2002). Highest ranked loudspeaker had flat free–field response and smooth off–axis magnitude responses in Olive (2004*a*,*b*). Typical listening rooms have low reverberation time (see Section 3.3). DRR in typical listening room is low and mostly direct sound is perceived.

5.4.2 Power Response Flatness

Loudspeaker directivity pattern has sometimes been optimized by using musical instrument directivity pattern as target (Warusfel et al. 1997, Lipshitz & Vanderkooy 1985). Musical instruments are omnidirectional in LMF band. Loudspeakers radiation pattern mimics the musical instrument loudspeaker pattern. Omnidirectional LMF radiation in loudspeaker is commonly accepted (Warusfel et al. 1997). Most of the recorded instruments are close–miked with single microphone and natural radiation pattern information is lost. Recording engineer will place the instruments into the auditory scene and fix the localization errors occurred during the recording in the recording studio. (Benade 1985)

Power response predicted in–situ spectral balance in listening position (Chapelle 1973). Almost ideal monopole has flat power–response and flat on–axis magnitude response (Lipshitz & Vanderkooy 1985).Flat power response with flat on–axis yields unpleasant listening experience, which is accentuated in small room (Lipshitz & Vanderkooy 1985). Measured in–situ magnitude responses are more similar to free–field on–axis response than power response (Goldberg et al. 2006).

5.5 Off–Axis Colourations

Off–axis colouration and smoothness are important factors for perceived sound quality (Toole 1986*a*,*b*, Olive 2004*a*,*b*). Off–axis magnitude response defines the sound colour for reflecting sound wave. High Q–value fluctuations in off–axis magnitude responses are audible (Toole 1986*a*). Conventional loudspeaker has coloured off–axis response due the natural LMF radiation pattern (see Section 5.6). Enclosure geometry affect the off–axis colourations (see Sections 5.6 and 5.5.1). Diffraction and lobing are effects that occur even in the most sophisticated enclosure solutions (Backman 1995, van der Werff 2001).

5.5.1 Diffraction

Diffraction induces linear and non–linear distortion to the loudspeaker sound output (Backman 1989). Discontinuities in the surface shape causes diffraction (Holm & Mäkivirta 2013). The diffraction point acts as sound source which arrives later to the listener creating cancellation (see Figure 10). On–axis response can be always equalized to have flat frequency response despite the diffraction. Compensating on–axis fluctuations caused by diffraction creates ripple and colouration to the off–axis response since the EQ affects loudspeaker power response in all radiated directions.

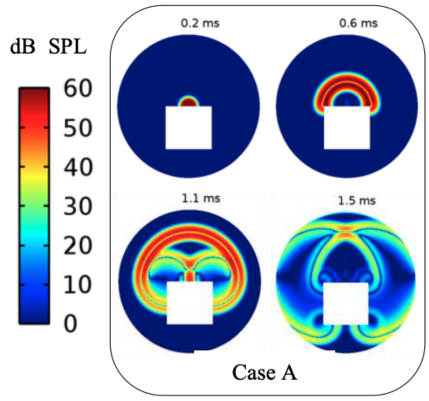

Figure 10: Rectangular corners induce diffraction. Diffracted sound wave departs later compared to direct sound wave (Holm & Mäkivirta 2013)

Geometrical measures of the loudspeaker enclosure define affected frequency and magnitude band of diffraction(Rasmussen & Rasmussen 1994). Diffraction from smooth and constantly changing shape needs very fine geometrical resolution, complex geometries are heavy to compute (Backman 1995). finite element method (FEM) and boundary element method (BEM) quantize diffraction from complex shapes (Holm (2010), Holm & Mäkivirta (2013)).

5.5.2 Lobing

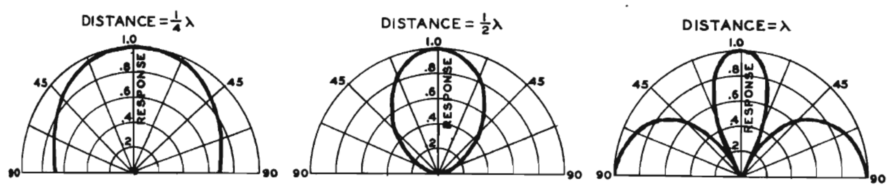

Figure 11: Lobing of small equal sources with distance of $\frac{\lambda}{4}$, $\frac{\lambda}{2}$ and λ (Olson 1957).

Otherwise smooth radiation pattern has a null which appears as lobe in polar plot. Lobing occurs when the radiating units are more than half wavelength apart. A radiator can behave as multiple sound source in very high frequencies. Directivity can be created with multiple radiators producing the same frequencies. Lobing occurs in two–way design if the crossover frequency wavelength is more than twice the distance of the radiators (see Figure 11). Lobing occurs in the cross–over frequency of multi–way loudspeakers. (D'Appolito 1983, van der Werff 2001)

5.6 Passive Methods to Create Directivity

5.6.1 Natural Radiation Pattern

Single direct radiator is omnidirectional in the low frequency and the directivity increases with frequency (Olson 1957). Attaching any driver to baffle increases directivity, the baffle sizes defines the effective frequency range (Allison 1974, Salmensaari et al. 1993). Woofer geometry controls LMF radiation pattern in conventional loudspeaker (Olson 1957).

Increasing diameter of the cone in single–cone loudspeaker increases directivity towards higher frequencies (Beranek 2012). Increasing baffle size increases the directivity (Salmensaari et al. 1993). Analytical solution for the ideal piston source radiation pattern in infinite baffle is

$$R_{\psi} = \frac{2 \times J_1(ka\,\sin\psi)}{ka\,\sin\psi} \tag{16}$$

where J_1 is a first order Bessel function and ka is defined as

$$ka = \frac{2\pi r}{\lambda} \tag{17}$$

where r is the radius of the radiator and λ is the wavelength (Olson 1957). Conventional piston–like moving coil transducer radiation pattern depends on ka. Radius of the radiating piston defines the directivity of the conventional two–way loudspeaker. Direct radiators found in real–life do not follow the Equation 16 since they radiate full space instead of half space.

Section 5.3 addresses the directivity of conventional two–way loudspeakers.

5.6.2 Waveguides

The enclosure mechanics can be designed to have directivity control features (Mäkivirta et al. 2017). High frequency directivity has been created with horn in the past but the current monitoring loudspeakers have been equipped with waveguide (Howze & Henrickson 1983, Geddes 2006, Mäkivirta et al. 2017).

Waveguide affects to the baffle gain and radiation pattern of the driver (Geddes 1986). Waveguide effects frequencies where geometric feature sizes are comparable to radiated wavelengths (Holm 2010). Affected frequency range is calculated in Geddes (1986). Waveguide is very feasible directivity control solution in the medium and high frequencies (Mäkivirta et al. 2017). Controllable frequency range is wider with lower cross–over frequency. Wider frequency band increases complexity of waveguide design (Holm 2010). Waveguides design process utilizes FEM or BEM for waveguide radiation pattern optimization (Dodd & Oclee-Brown 2007, Holm 2010, Hayashi et al. 2012).

5.6.3 Dipole Pattern Source

Dipole sound source radiates equally to the front and rear of the loudspeaker, rear radiation is 180° out of phase compared to the front radiation. Dipole loudspeaker has radiator nulls on the side of the loudspeaker (Linkwitz 1992). Dipole loudspeaker has a wideband uniform directivity with significantly higher directivity than omnidirectional sound source (Linkwitz 2007). Dipole creates strong rear radiation, which creates strong sagittal reflections. Dipole loudspeaker has no side radiation which leads to diminished lateral reflections (Kates 2002). Conventional dipole does not have closing enclosure, dipole loudspeakers have poor LF sensitivity since enclosure gain is not utilized.

5.6.4 Acoustical Resistance Enclosure

Acoustical resistance can be used in a loudspeaker enclosure to control directivity (Holmes 1986, Iding 1971). Frequency dependent resistance controls the radiation pattern. Leaking negative signal attenuates the rear radiation. Material with frequency dependent resistance covers leaking structure in the enclosure. Enclosure acts as sealed enclosure for higher frequencies where the acoustic resistance is high. Resistance enclosure produces cardioid radiation pattern in LF range (Gunness 2018, Balogh 1977).

Frequency dependent acoustic resistance defines the leaked frequency band and the directivity pattern. Material used to resist the airflow defines the acoustic resistance (Iding 1971). The acoustic resistance enclosure causes the system cut–off frequency to rise. Acoustic resistance enclosure has poor LF sensitivity compared to sealed of reflex enclosures (Iding 1971). Controlling the directional pattern is difficult due the non–linear acoustic resistance found in typically used materials (Backman 1999).

5.7 Active Methods to Create Directivity

High and middle frequency wavelengths are short and passive methods to create directivity offer easy and feasible methods to shape the radiation patterns (see Section 5.6). Single direct radiator radiates omni–directionally in LF range. Low frequency directivity control is considerably more difficult since LF wavelengths are longer and need physically large transducers and enclosures to control them with passive solutions. Active signal processing can be used to control LMF directivity.

5.7.1 Array of Drivers

Array of radiators or drivers are used for beamforming to create directivity. The target have been to create omnidirectional, constant or variable directivity loudspeaker (Møller et al. 2010, Taylor & Keele 2017). High frequency drivers do not displace air enough to disturb each other so large arrays can be mounted to single enclosure. Adjusting phase modes creates directivity in the MF and HF bands (Møller et al. 2010). Circular array create different radiation patterns, radiation patterns are switchable with active signal processing (Poletti & Betlehem 2013, Møller et al. 2010, Sato & Haneda 2017).

Side by side placed radiators simulate a line source when the transducer distance is small compared to the radiated wavelength in line arrays (Fazi et al. 2015, Sato & Haneda 2017, Taylor et al. 2017, Taylor & Keele 2017). Lobing occurs when the distance between driver is a lot larger than the reproduced frequency wavelength (Fazi et al. 2015). Combining dipole and omnidirectional radiation creates the controllable cardioid polar pattern (Boone & Ouweltjes 1997).

5.7.2 Gradient System

When the wavelength is long in the low frequencies, the directivity can be created with multiple driver units positioned in–line each other. Distance between drivers defines the frequencies where the directivity can be created. Most common application on the multiple sources directivity in the low–frequency band is directive subwoofer consisting of two or more discrete subwoofer. Directivity is created by shifting the phase of secondary monopole unit by 180° and delaying the unit with constant time delay delay to create cancellation of the rear radiation. (Backman 2003, Ferekidis & Kempe 2004, Olivier 2010, Olson 1973) Figure 12. presents the topology of first order unidirectional gradient.

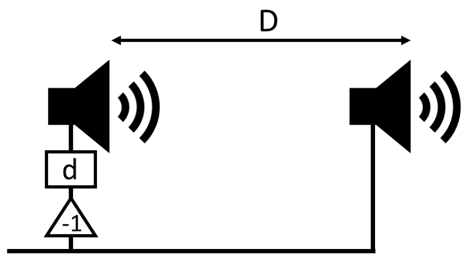

Figure 12: Typical single order unidirectional gradient loudspeaker arrangement.

Same principle applies to the higher frequencies since polar pattern R_θ is defined by equation

$$R_\theta = \sin(k\frac{d}{4} + k\frac{D}{4}\cos\theta) \tag{18}$$

for two omnidirectional point sources. d is the delay in meters and D is the distance in meters. If frequency dependent delay is designed, bandwidth of the directivity control in gradient system can be widened. Utilizing frequency dependent magnitude and delay filters, wider directive frequency band can be achieved while retaining the original sensitivity. (Olson 1973, Olivier 2010)

6 Improving Low Frequency Directivity

Current two–way loudspeaker can be improved by increasing the LMF directivity (see Section 4). Improved two–way monitoring loudspeaker should couple less to the listening room and produce more accurate reproduction of the programme. Current compact loudspeakers radiate almost omnidirectionally in the low–middle frequencies (see Section 5.3). Gradient method with frequency dependent delay was implemented to create directivity in the LMF band. (see Section 5.7). Tested devices were measured under free–field conditions to define their acoustic performance. In–situ measurements quantify the interaction between loudspeakers and room.

6.1 Devices Under Test

Loudspeaker B from Section 5.3 represented the compact conventional 2–way loudspeaker. Conventional loudspeaker corresponds to the DUT A. DUT or loudspeaker A was a commercial high quality active monitoring loudspeaker with controlled directivity in the HF range, full band frequency response and low distortion. Audiovisual studios and productions use DUT A as a monitoring loudspeaker.

Two DUT A Loudspeakers formed DUT B, loudspeaker units were placed in line, they had identical acoustic axis. DUT A was the primary radiator of DUT B. DUT B created LMF directivity in the frequency band of 80 – 600 Hz. Equation 18 presents the analytical solution for simple gradient loudspeaker. The analytical solution presumes the sound sources to be point sources. Non–linear least squares optimization method optimized the frequency–dependent delay and magnitude parameters from measured IRs. Target for optimization was the radiation pattern of DUT A at 850 Hz. Difference between target and actual radiation patterns was minimized. The digital frequency–dependent delay and magnitude filters were implemented in real–time. Passively controlled HF directivity remained the same between the loudspeakers. Figure 13 and Appendix C visualize the arrangement of the loudspeakers.

Created directivity is symmetrical in both vertical and horizontal planes since the primary and secondary sound sources shared the same acoustical axle. Only DUT A (and primary radiator in DUT B) produced sound above 600 Hz and both DUTs had the same radiation pattern above 600 Hz. Created directivity frequency band limitation was due the physical size of original DUT A, hence constant directivity was not achieved.

Increased woofer surface area and amplifier power did not increase the perceived headroom in DUT B. Lack of increased headroom was due using the secondary loudspeaker unit for cancelling the rear radiation rather than amplifying. DUT B had slightly more pre–ring due the digital processing introduced in the signal chain. Pre–ringing was not audible to the author in extensive listening conducted along with the listening test design process Pre–ringing was not audible while listening the DUT B in the anechoic chamber or in the listening room.

Directivity performances of the used loudspeakers were measured under anechoic conditions.

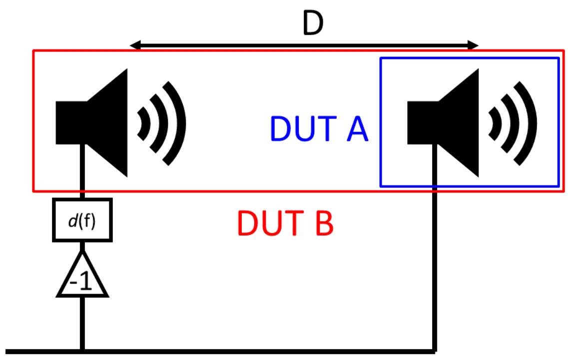

Figure 13: Arrangement of Loudspeakers used to create switchable directivity configuration. Acoustic axes of both loudspeakers point to listening position.

6.2 Free–Field Performance of Devices Under Test

Performances of the loudspeaker A and B were determined with a series of measurements described in Section 5.1. Dispersion and polar plot do not give comprehensive understanding of the directivity performance and are not presented. DI presents inverted power response when the on–axis response is flat. In this scenario, power response and DI represent the same data, only DI is presented.

6.2.1 On–Axis Frequency Response

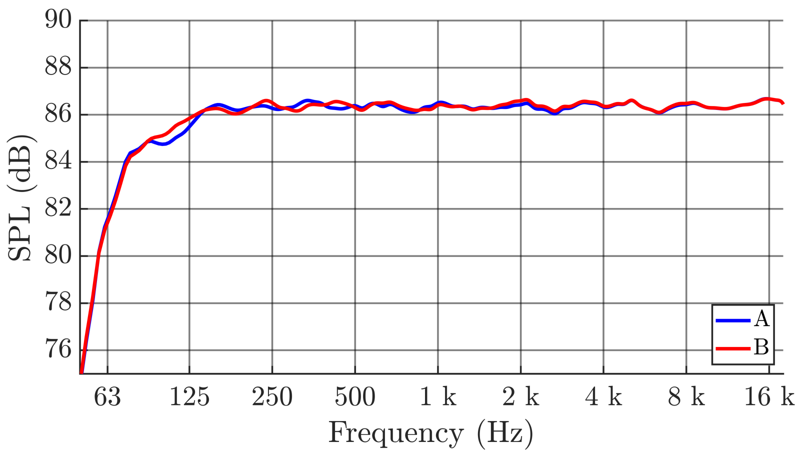

Figure 14: On–axis responses of DUTs. Blue line presents DUT A and Red line presents DUT B

On–axis frequency response is a significant factor for the fidelity and sound colour perception (see Section 5.4). Free–field frequency responses of the DUTs were closely matched in anechoic conditions. 60 Hz A 2nd order high–pass filter was applied. Figure 14 presents the implemented frequency responses. Figure 15 presents the free–field on–axis magnitude response difference between DUTs.The difference between the DUTs is under 1 dB. Both DUTs fulfil the free–field magnitude response flatness $\pm$ 1 dB requirement (EBU 1998, ITU 2015).

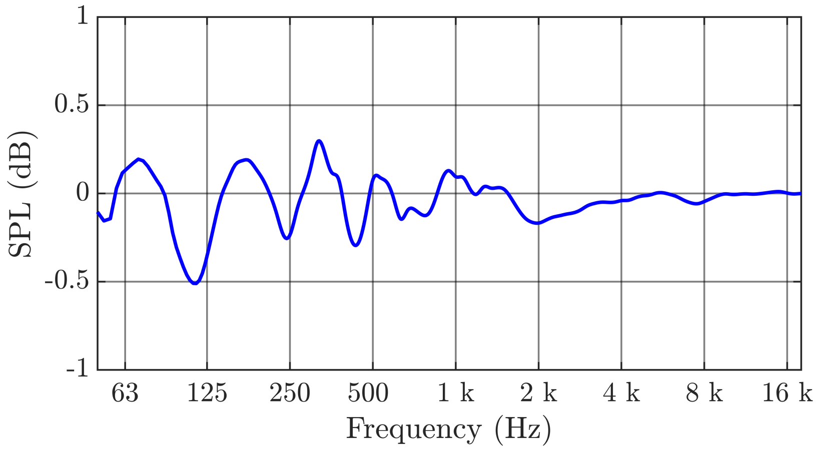

Figure 15: Frequency response difference DUT A-B

6.2.2 Off–Axis Frequency Responses

The 1/3–octave smoothed horizontal off–axis magnitude responses of DUTs were plotted from ψ = 0° to ψ = 120° while the vertical angle θ remained zero. Directivity between 80 and 600 Hz is symmetrical about acoustical axis on both DUTs and measured vertical and horizontal off–axis responses would be identical below 600 Hz.

Figure 16 presents horizontal off–axis response for the DUT A. The off–axis curves behave smoothly and no sudden fluctuation is noted in the set of measurements. Directivity increases significantly with frequency. Maximum attenuation in the 120° measurement is 5 dB below 600 Hz.

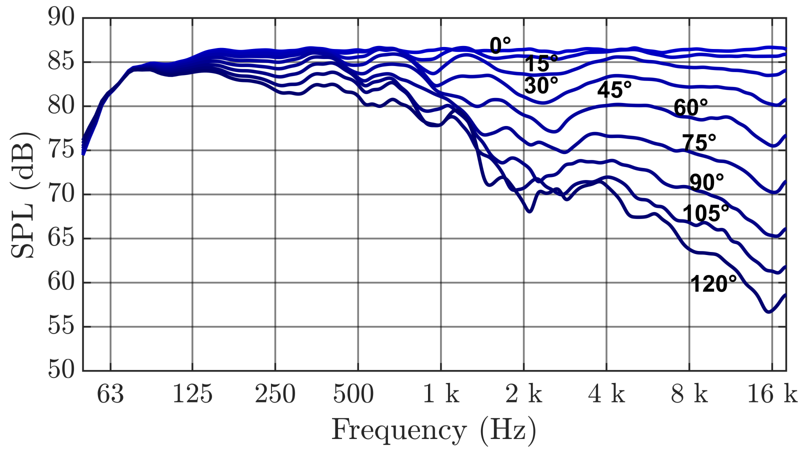

Figure 16: DUT A Frequency responses from ψ ψ =120° in 15° increments.

Figure 17 presents the DUT B horizontal off–axis response. The off–axis curves behave smoothly and no sudden fluctuation is noted in the set of measurements. The increased directivity is visible with off–axis angles between 30° and 120° at 80 – 600 Hz. Directivity is increased slightly in the frequency range of 80–600 Hz. Off–axis magnitude responses are identical between the DUTs above 600 Hz.

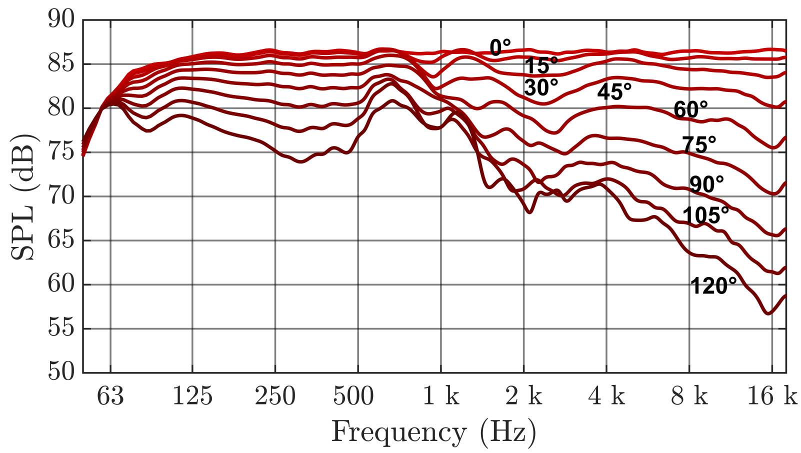

Figure 17: DUT B Frequency responses from ψ = 0° to ψ =120° in 15° increments.

6.2.3 Directivity Index

The directivity index plots were calculated from the measurement data for both DUTs according to the section 5.1.3.

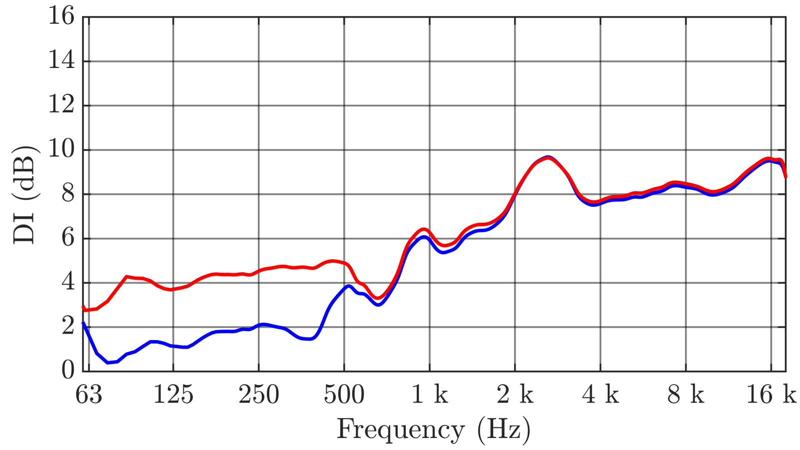

Figure 18: Directivity indexes of DUTs. Blue line presents DUT A and Red line presents DUT B

DUT A behaved as a typical 2–way loudspeaker, DI increases with frequency. DI of DUT B was increased by over 2 dB compared to the DUT A (see Figure 18). Despite the increased DI, the ITU (2015) requirement of DI = 6 dB at 500 Hz was not achieved. DI did not increase smoothly with frequency as requested in the ITU (2015). 1.5 dB dip in DI was present at 600–800 Hz. The relation between averaged high and low frequency DI was reduced from 8 dB to 5 dB. The relative DI increase is high. Total energy radiated into the listening room is reduced significantly in the LMF band.

6.2.4 Directivity Contour Plots

The horizontal directivity plots were drawn for both DUTs. 1/3–octave smoothing, 3 dB magnitude resolution and 5 degree angle resolution were used as the measurement parameters. The vertical and horizontal directivity plots for both DUTs behave symmetrically below 1 kHz due the radiator arrangement in the DUT B (see Section 6.1). New information is not presented in the vertical directivity contour plot.

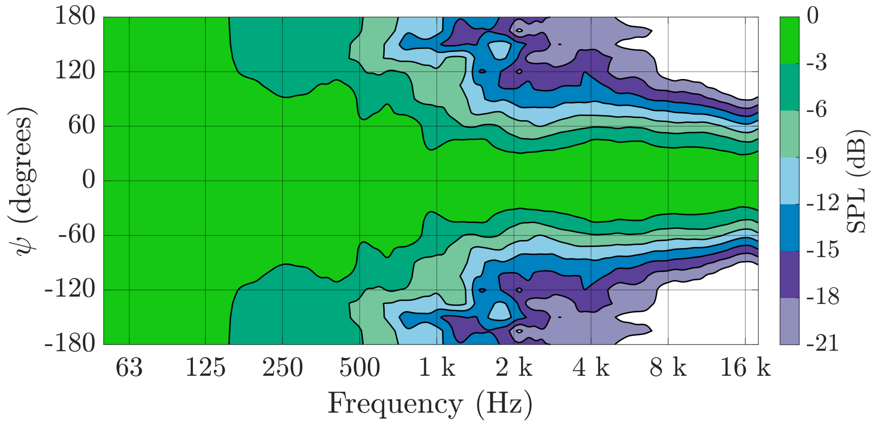

Figure 19: DUT A directivity contour plot, large off–axis colouration.

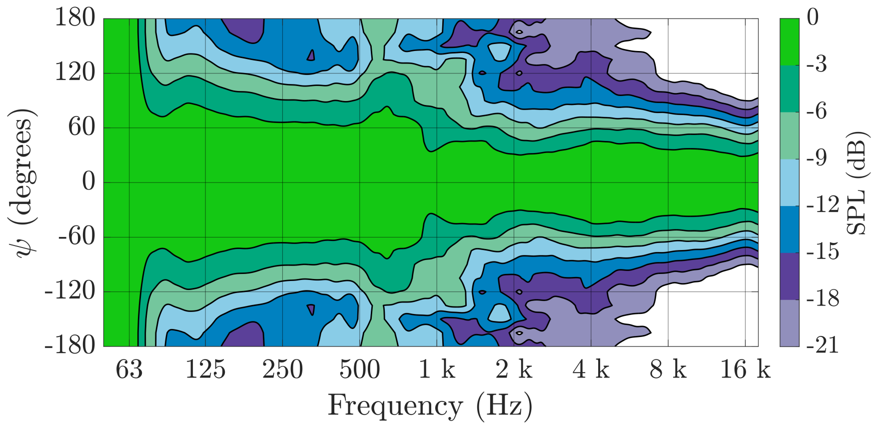

Figure 20: DUT B directivity contour plot, reduced off–axis colouration.

Figure 19 presents the horizontal directivity plot for DUT A. Directivity in the low frequencies is low, wide dispersion is observed below 500 Hz. The directivity increases with frequency. Off–axis lobing or beaming is not observed. The directivity is increasing smoothly with frequency. Off–axis magnitude colouration is significant in the conventional loudspeaker.

Figure 20 presents the horizontal directivity plot for DUT B. Transition from the directivity controlled frequency band to the naturally radiating frequency band results in a narrow discontinuity at 650 Hz where the directivity decreases momentarily. Off–axis colouration is significantly reduced compared to the DUT A. Increased directivity at the frequency band 80 – 600 Hz is clearly visible when comparing Figures 19 and 20. Front–to–back level difference for the DUT B is 15 – 18 dB compared to 3 to 6 dB front–to–back level difference of the DUT A, 12 dB improvement is achieved in the front–to–back radiation.

6.3 In-Room Performance of Devices Under Test

In-situ measurements were performed to quantify the room coupling difference between the DUTs.

6.3.1 Room Magnitude Responses

Room magnitude responses were measured to adjust the sound balance between HF and LF. The sound balance was adjusted flatter with equalization (EQ), EQ was applied for both loudspeakers to preserve the relation between loudspeakers. Shelving was utilized to flatten out the high frequencies above 1 kHz and intensive 70 Hz resonance was tamed with a single notch filter. 70 Hz notch filter gain was adjusted between listening distances while Q–value and filter centre frequency remained the same. Both DUTs coupled with the 70 Hz room mode equivalently and DUT specific magnitude equalization was not applied since the principle was to observe increased directivity effect on room coupling and magnitude response.

Figures 21 and 23 present the in–situ left and right channel magnitude responses for 1 m listening distances. The corresponding magnitude differences are presented in Figures 22 and 24. Figures 25 and 27 present the in–situ left and right magnitude responses for 3 m listening distance. The corresponding magnitude differences are presented in Figures 26 and 28.

The measured magnitude responses behave similarly between DUTs, same features are distinguishable from the shapes of magnitude responses. Differences between DUTs do exist but the differences are small because of the in–room location do not change and on–axis responses are very similar. The overall shape of magnitude responses were noted to resemble magnitude responses gathered from 89 professional studios (Goldberg et al. 2006). 50 % percentile fluctuation in in–room magnitude is 5 dB for 89 professional audio reproduction installations (Goldberg et al. 2006). Magnitude response difference is very small compared to other studies where loudspeakers have been evaluated by listening tests (Staffeldt 1974, Salmi & Weckström 1982, Olive et al. 1995). Magnitude difference was greater for 3 m listening distance which is natural since the DRR is larger.

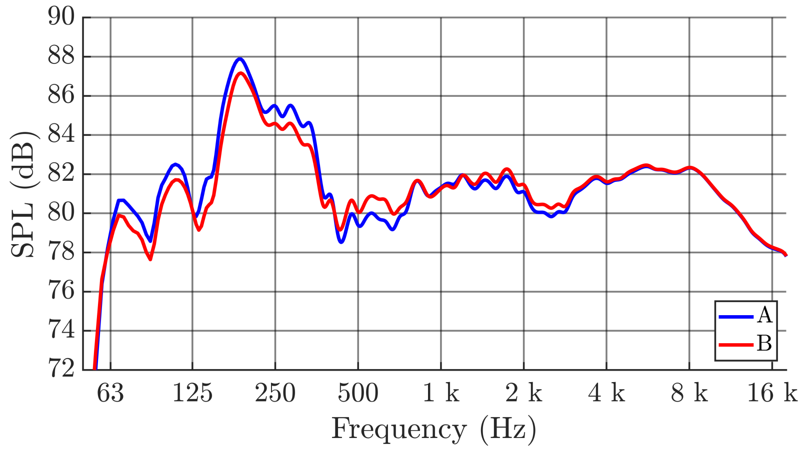

Figure 21: Left channel in–room magnitude responses of the DUTs, 1 m listening distance. Blue line presents the DUT A magnitude response and Red line presents the DUT B magnitude response

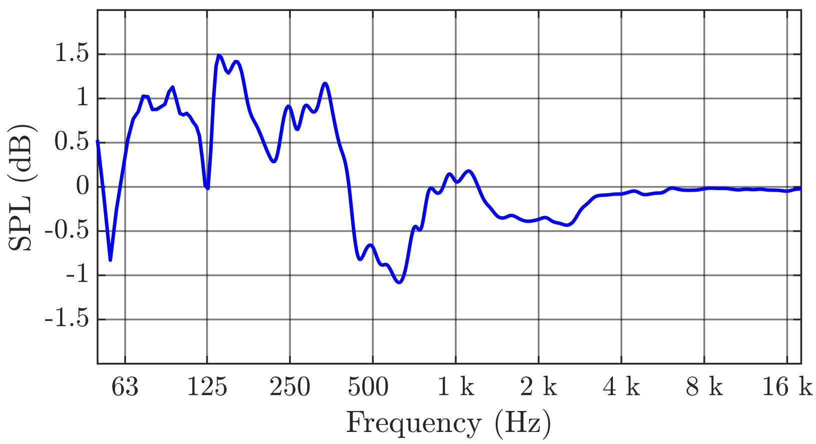

Figure 22: Left channel in–room magnitude response difference, 1 m listening distance.

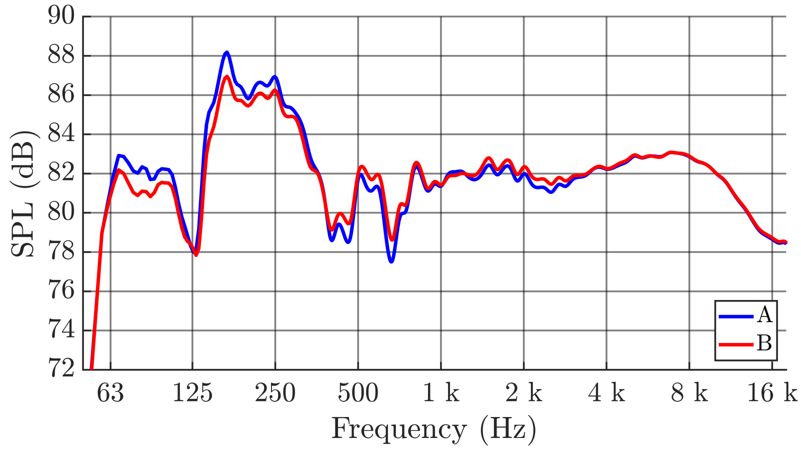

Figure 23: Right channel in–room magnitude responses of the DUTs, 1 m listening distance. Blue line presents the DUT A magnitude response and Red line presents the DUT B magnitude response

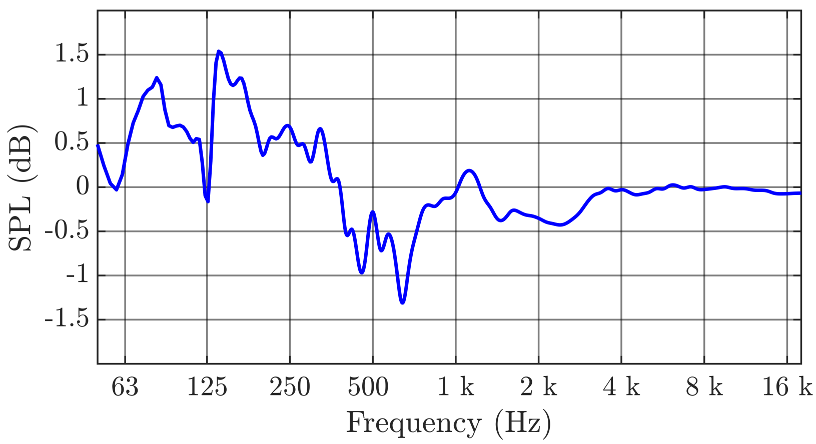

Figure 24: Right channel in–room magnitude response difference with 1m listening distance.

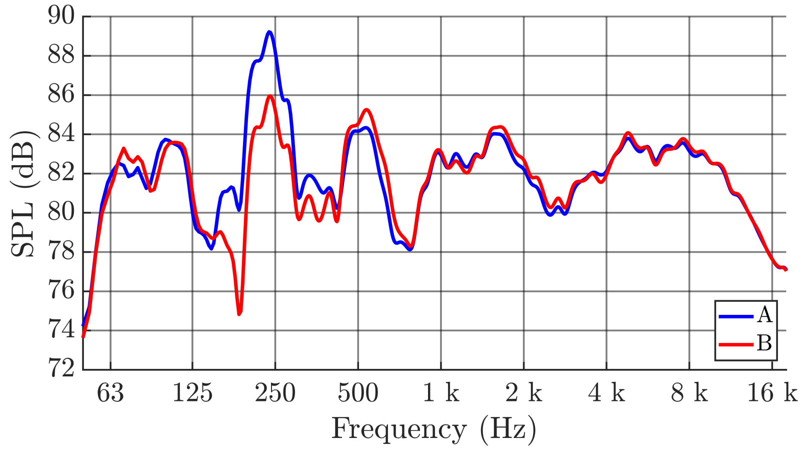

Figure 25: Left channel in–room magnitude responses of the DUTs, 3 m listening distance. Blue line presents the DUT A magnitude response and Red line presents the DUT B magnitude response

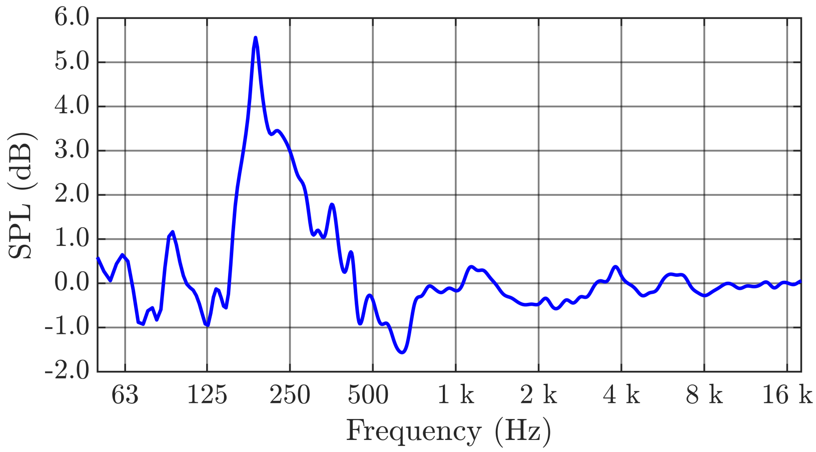

Figure 26: Left channel in–room magnitude response difference, 3 m listening distance.

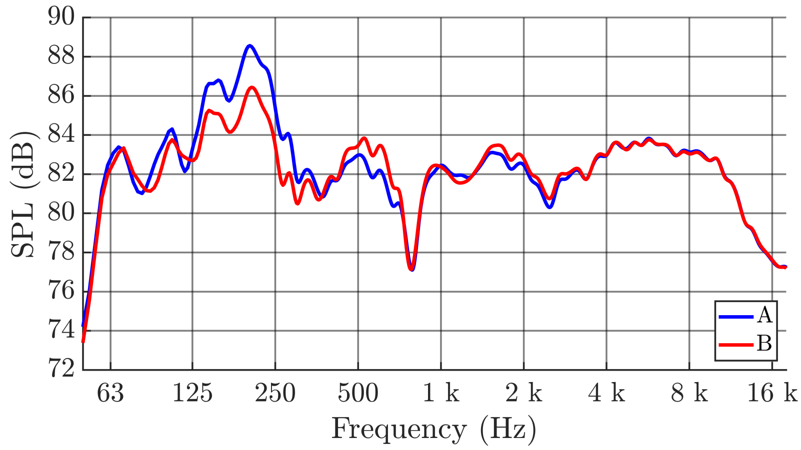

Figure 27: Right channel in–room magnitude responses of the DUTs, 3 m listening distance. Blue line presents the DUT A magnitude response and Red line presents the DUT B magnitude response

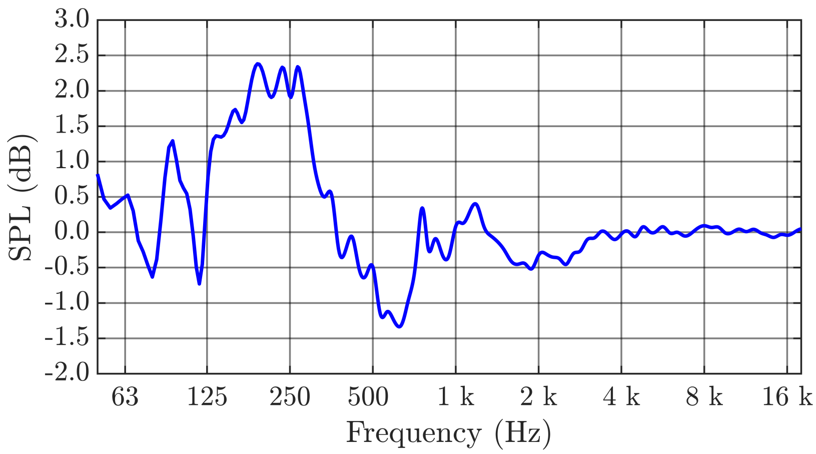

Figure 28: Right channel in–room magnitude response difference, 3 m listening distance.

6.3.2 Room Mode Decay

As discussed in Section3.5, the loudspeaker radiation pattern will modify the room decay envelopment. Room IRs were convolved with 4 second bursts of pure sine waves at various frequencies to simulate the steady state of room excitation. The resulted signal was measured and plotted. Measured resonances were not band passed, output of the room may include other frequencies than the pure sine wave fed into the system. Room IR signal–to–noise ratio (SNR) of 60 dB was reported for the measurements. Decay at 90 Hz for 3 m listening distance is presented in figure 31. 82 Hz and 200 Hz decays for 1 m listening distance are presented in Figures 29 and 30.

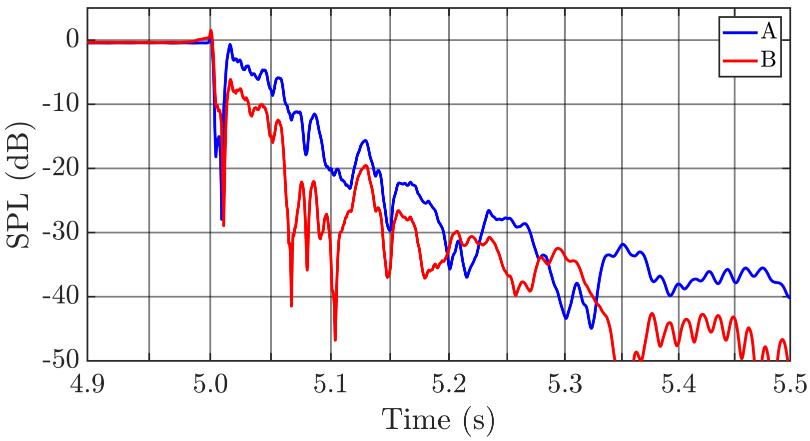

Figure 29: Room decay at 72 Hz with 1 m listening distance. Blue line presents DUT A and Red line presents DUT B

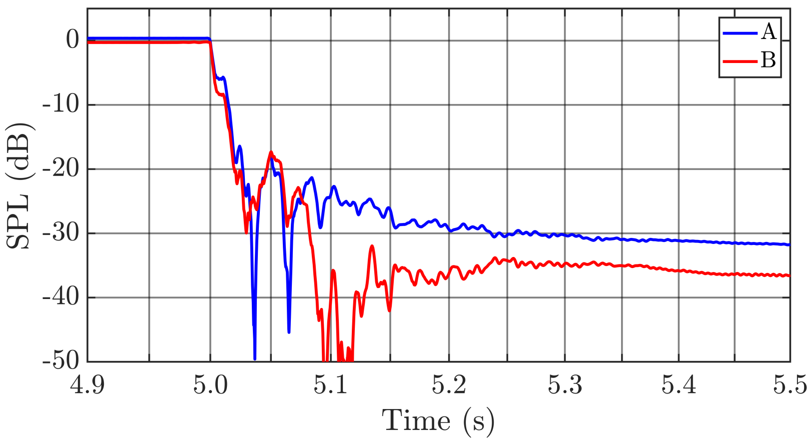

Figure 30: Room decay at 200 Hz with 1 m listening distance. Blue line presents DUT A and Red line presents DUT B

The more directive DUT B benefited in terms of mode decay attenuating. Despite equalised magnitude in static state of forced vibration, the sound energy decay was improved with directive loudspeaker. The sound decay envelopment varies across frequencies.

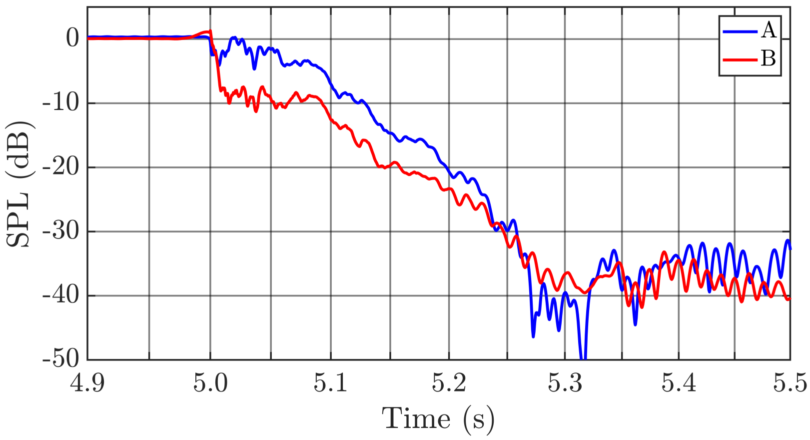

Figure 31: Room decay at 90 Hz with 3 m distance. Blue line presents DUT A and Red line presents DUT B

Constant differences in the steady state of room mode can be matched with the room EQ but the sound energy decay can not be compensated with the conventional magnitude equalization. The magnitude differences of over 10 dB over hundreds of milliseconds yields significantly less masking. Decay from the DUT B creates significantly less masking than from the DUT A.

6.3.3 Reverberation Energy

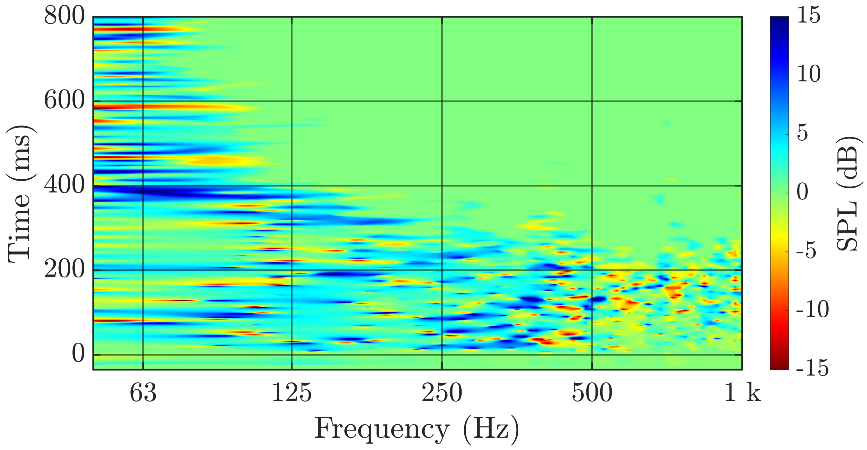

Figure 32: Difference wavelet representing reflection difference between DUTs. Blue colour indicates more reflected energy is induced by DUT A then DUT B. DUT B inducing more reflected energy than DUT A is indicated with red colour.

Room reflections in the listening position can be observed with the wavelet transform. The wavelet transform visualizes the impulse response with a 3–dimensional time–frequency domain presentation. Room IRs were measured for both DUTs. The wavelet transform was performed for IR of the both DUTs with 1/3 octave smoothing between 50 Hz – 1 kHz. Both wavelets were normalized with the maximum intensity in time and frequency. The time–frequency normalized wavelets were deducted. Figure 32 presents the resulted difference plot indicating change of sound energy arriving to the listening position. The green colour indicates that acoustic energy in both impulses is exactly the same. The blue colour indicates that decayed acoustic energy induced from DUT B is weaker than from DUT A. The red colour indicates the opposite situation. Presumably the more directive source should always have less reflected energy in the listening room. However the envelope of the reflected sound energy varies and creates red spots around the blue zones.

The difference of reflected energy was integrated in respect to time at frequency band of 100 Hz – 600 Hz to visualize the total reverberant or decaying energy difference between the DUTs. 100 ms time periods was used as the integration limits. Decay energy differences are presented in Figures 33, 34 and 35. The DUT A impulse response decay being more energetic is denoted with blue bars. The DUT B impulse response decay being more energetic is denoted with red bars. The blue bars represent the improvement by more directive sound source. The red bars represent the regression by the more directive sound source.

Less coloured off–axis magnitude response in the DUT B reduces the in–room decay. DUT couples less to the room, the room decay is diminished and masking is decreased with the increased directivity.

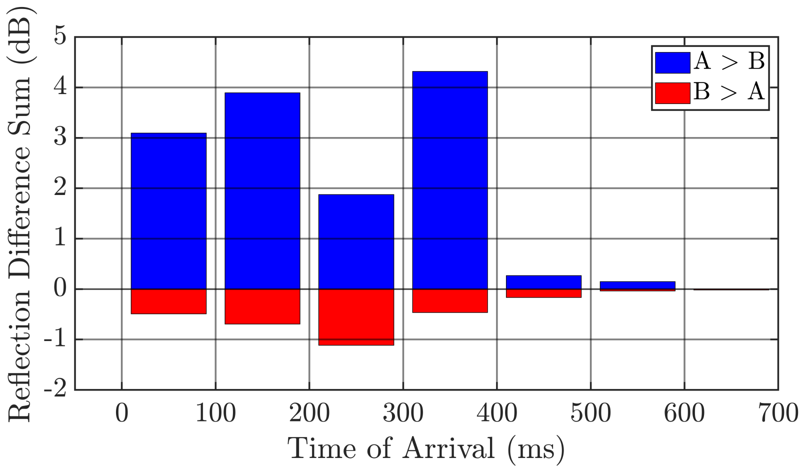

Figure 33: Integrated difference of reverberation at frequency band 100 – 150 Hz.

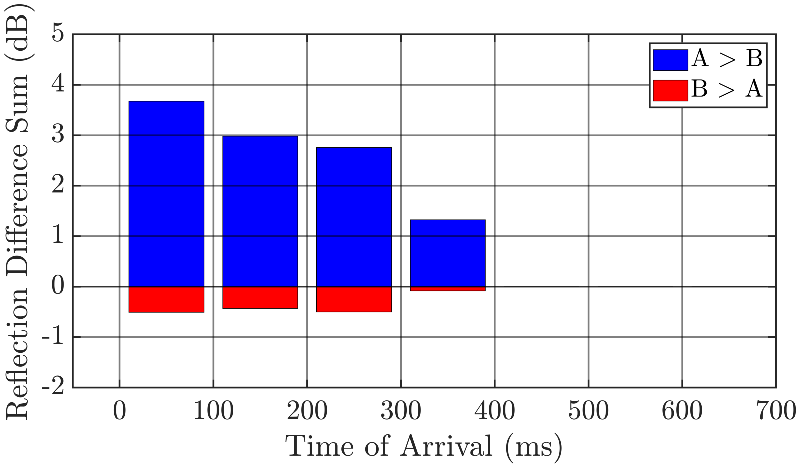

Figure 34: Integrated difference of reverberation at frequency band 150 − 300 Hz.

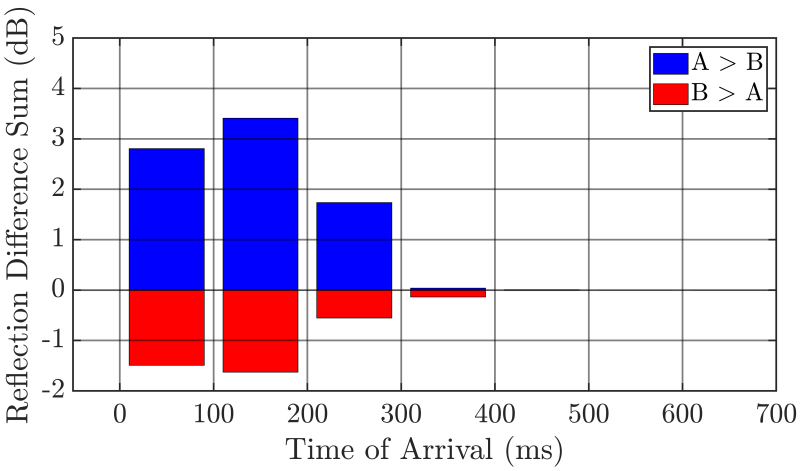

Figure 35: Integrated difference of reverberation at frequency band 300 − 600 Hz.

7 Subjective Testing

Listening tests measured the perceived sound quality caused by the changed in–room performance (see Section 6.3).

7.1 Listening Test Arrangements

7.1.1 Listening Room and Loudspeaker Positioning

Loudspeakers were placed into acoustically treated listening room. The listening room is 5 m wide, 7.7 m long and 2.5 m tall. The acoustical properties resemble acoustic properties of domestic room. Figure 36 presents the reverberation time of the used listening room.

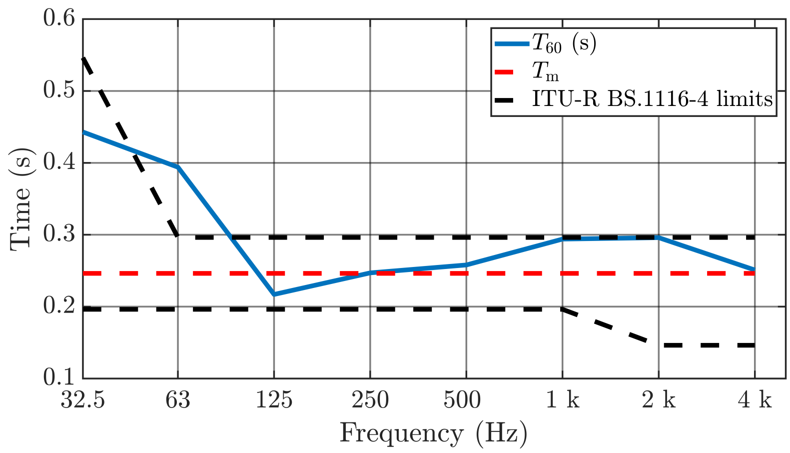

Figure 36: Listening Room T_{60} compared to T_{m} defined by ITU (2015)

Universally accepted equilateral listening triangle was formed for the 1 m listening distance, the loudspeakers were 2 m away from side walls and 1.2 meters from the front wall. The loudspeakers were positioned on the reverberant end of listening room. Floor, back and side walls were not heavily acoustically treated. No obstacles were between the subject and loudspeakers. The 3 m listening distance preserved the equilateral listening triangle, the loudspeakers were 1.5 m away from the side walls and 1.6 meters from the front wall. Appendix C presents an illustration of the loudspeaker arrangements.

The DUT A was used as primary unit for the DUT B (see Figure 13). When changing between the DUTs, signal processing changed accordingly, changing the auditioned loudspeaker pair took 150 ms. Acoustic axle was the same for both auditioned loudspeakers. The subject did not perceive the change of arrival direction when changing the auditioned DUT. Sound source location was identical for the both DUTs. The loudspeaker in–room location as variable was eliminated (see Section 3).

7.1.2 Listening Test Programme and Questionnaire

Author auditioned a broad selection of music and test signals. The selected programme is presented in Table 3. Sample 1 spectrum is concentrated on the frequency band where directivity was different between the DUTs. Sample 1 Samples 2 and 3 have very similar spectrum, sample 3 has a lot more dynamic variation and the envelope fluctuates more. Sample 4 is mostly consisted of electric bass melody bursts and a drum beat. Sample 4 spectrum is focused to low and middle frequencies, high frequency content consists of harmonics and transients. Samples 2 and 3 are repetitive compared to sample 4, where the melody and rhythm vary through the sample.

	Sample			
	#1	**#2**	**#3**	**#4**
Music Style	Noise pulse	Metal	Hard Rock	Fusion Jazz
Frequency Band	250 Hz Octave	Full Band	Full Band	Low and mid
Duration	1 s	31 s	56 s	1 min 32 s
Dynamic Variation	Pulse	Low	Average	High
Album Artist		Equilibrium	My Chemical Romance	Béla Fleck
Song Title		Unbesiegt	Famous Last Words	Bumbershoot
Sample Start Time	0:00	1:50	0:40	3:08

Table 3: Listening test programme descriptions.

Questions 1, 2 and 6 evaluated the stereo sound imaging and image width. Questions 4, 5, 8, 9, 10, 12, 13, 14 and 15 evaluated the perceived clarity in middle and low frequencies. Questions 3, 7 and 11 measured the timbre perception and sound colouration. Table 4 and Appendix A present the detailed questionnaire.

Question	Sample	Description
1	1	Sound image location
2	1	Sound image width
3	1	Middle frequency colouration
4	2	Low frequency Clarity
5	2	Middle frequency clarity
6	2	Sound image definition
7	2	Low–middle frequency colouration
8	3	Low frequency clarity
9	3	Middle frequency clarity
10	3	Vocal clarity
11	3	Low–middle frequency colouration
12	4	Low frequency clarity
13	4	Middle frequency clarity
14	4	Melody Clarity
15	4	Middle frequency transients

Table 4: Listening test questions

Continuous grading scale was used. Question–related descriptive adjectives were written in the both ends of the grading scale. Length of the grading line was 10 cm, corresponding to the scale of 0–10 in results (see Section 7.2).

7.1.3 User Interface

The user interface for the listening tests was simple: A wireless keyboard was used as selector between the DUTs, listener did not need to concentrate on operating any equipment and could focus on listening. The number keys "1" and "2" sound source, changing between the DUTs took 150 ms. The order of the sound sources was varied to avoid influencing subjects. The order of the DUTs was included as a variable in the statistical analysis (see Section 7.2.4).

7.1.4 Audience and Subjective Experience

The listening panel (N = 6 for 1 m and N = 8 for 3 m) possessed previous listening experience in critical listening, no reported hearing impairments existed among the subjects. The listening panel remained the same between listening test rounds of the two listening distance. Experienced audience for the listening test is important for obtaining significant and accurate results from the listening tests. Experience of listening panel is more important than number of the subjects (Olive 2003).

Perception is always subjective despite guiding questionnaire. Subject uses the answering scale subjectively. Subjects were encouraged to use most of the scale. Identical perception between DUTs was marked with same exact rating. Perceived sound quality of directivity was evaluated, DUT A is conventional two–way loudspeaker and DUT B has increased LMF directivity

7.2 Listening Test Results

Subjects filled the listening test questionnaire with pen and paper. The answering scale was 10 cm long, the data from questionnaire forms was gathered with ruler and transported into digital domain. After gathering the results, a diverse statistical analysis was conducted to ensure the received results are statistically valid, significant factors and correlation between variables were examined.

The gathered data was visualized with box plots. Result is significant when the boxes do not overlap, less than 25 % distribution of the results is overlapped. The bold question titles above the box plot indicate that the result is statistically significant (see Sections 7.2.1 and 7.2.2). Analysis of variance (ANOVA) determined the statistical significant (see Section 7.2.3). The dashed line in the figures represents the optimal answer for the question.

7.2.1 1 m Listening Distance

Figures 37, 38, 39 and 40 present the results for 1 m listening tests.

Figure 37 presents the results for the sample 1. Sound image was located in the centre between sound sources according to all subjects. One subject reported stereo image location to be blurry for DUT A, despite localization between the loudspeakers. The sound image width was significantly narrower for the DUT B than DUT A, the DUT B was perceived significantly less coloured compared to the DUT A in the middle frequencies.

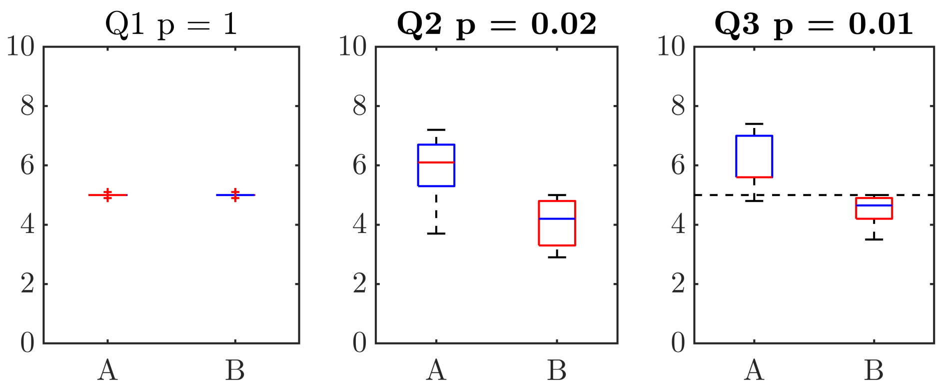

Figure 37: Sample 1 listening test results for 1 m listening distance. Statistically significant result subtitles are **bold**. Dashed line (- -) indicates bidirectional question midpoint.

Figure 38 presents the results for the sound sample 2. DUT B was significantly clearer in the LF and MF bands (questions 4 and 5). DUT B stereo imaging was significantly better to the DUT A by a significant difference in question 6. DUT B was significantly less coloured in question 7. One subject reported that the DUT B is lacking bass.

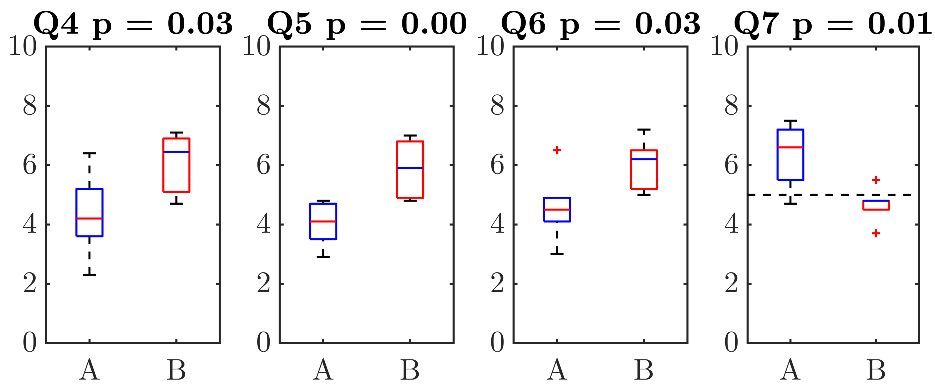

Figure 38: Sample 2 listening test results for 1 m listening distance. Statistically significant result subtitles are **bold**. Dashed line (- -) indicates bidirectional question midpoint.

Figure 39 presents results for the sound sample 3. LF clarity differences was not significant in the question 8. Statistical significance condition was not fulfilled for the questions 9 and 10. The DUT B was significantly less coloured than the DUT A in the question 11, DUT B deviation is closer to the optimal answer.

Figure 40 presents results for sample 4. Results for questions 12–14 were not statistically significant. DUT had significantly better transient reproduction than DUT A in question

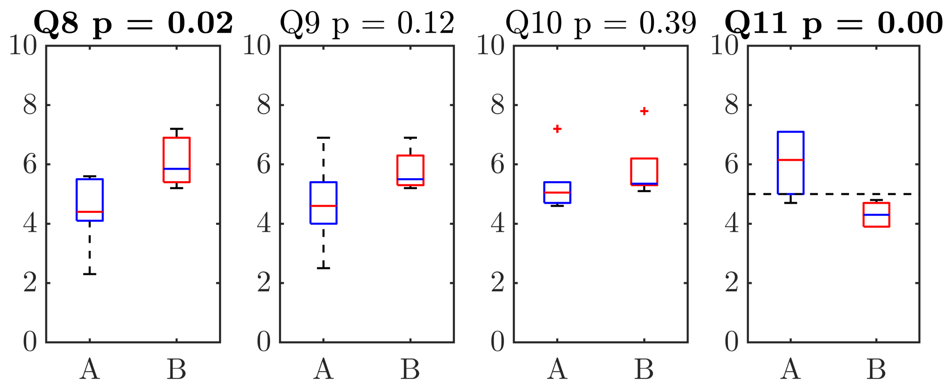

Figure 39: Sample 3 listening test results for 1 m listening distance. Statistically significant result subtitles are **bold**. Dashed line (- -) indicates bidirectional question midpoint.

15. Sound quality evaluation was difficult with sample 4. Sample 4 was perceived to lack repetition. Every time the subject changed sound source, different melody was reproduced. Overall sound timbre balance and loudness did not remain stable during the sample which made the task of evaluating loudspeaker performance even more difficult.

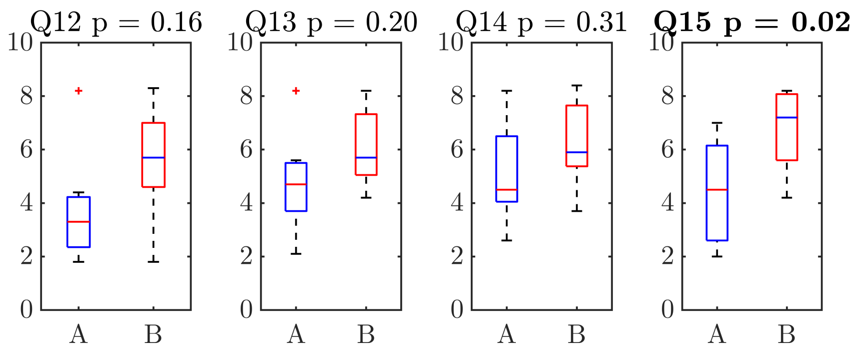

Figure 40: Sample 4 listening test results for 1 m listening distance. Statistically significant result subtitles are **bold**.

Author observed that once the difference was heard, identifying the more directive DUT B was easy. Subjective difference between DUT was significant. First impression with the 1 m listening distance was the timbre change between DUTs, the timbre difference was perceived larger than expected when compared to the in–situ magnitude responses (see Figures 21 and 23). Sample 2 offered the most distinguished difference between auditioned sound sources. Bass response was significantly "snappier" and clearer. Muddy sound was not perceived when listening away from listening position.

7.2.2 3 m Listening Distance

Listening test round 2 repeated the listening test with a longer listening distance of 3 m. Questionnaire, samples, listening panel, DUTs and listening test procedures remained the same between the listening test rounds, distance was varied and minor tweaks to the room equalization were made (see Section 7.1.1). Figures 41, 42, 43 and 44 the listening test results for 3 m listening distance.

Figure 41 presents results for the sound sample 1. Stereo image was located in the centre for both DUTs in he question 1. Sound image width was not significantly different. Sound colour for DUT B was significantly more neutral with the statistical significance condition fulfilled in the question 3.

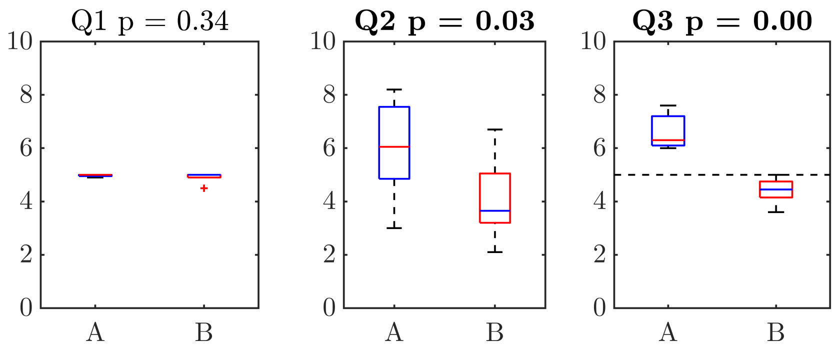

Figure 41: Sample 1 listening test results for 3 m listening distance. **Bold text** points out significant results and their p–values. Dashed line (- -) indicates bidirectional question midpoint.

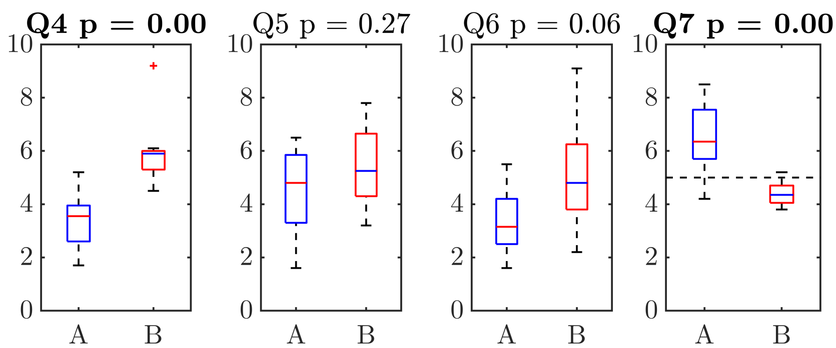

Figure 42: Sample 2 listening test results for 3 m listening distance. **Bold text** points out significant results and their p–values. Dashed line (- -) indicates bidirectional question midpoint.

Figure 42 presents results for the sample 2. DUT B had significantly improved low frequency clarity in the question 4. Results for middle frequency colouration and image definition were not statistically significant in the questions 5 and 6. Most subjects (7/8) still ranked the DUT B more clear and focused despite the large deviation in the results. DUT B was less coloured in the question 7.

Figure 43 presents results for the sample 3. DUT B had clearer bass reproduction in the question 8. Clarity perception is not reported to improve significantly with more directivity, the clarity related results are not statistically significant (see question 9 and 10 in Figure 43).

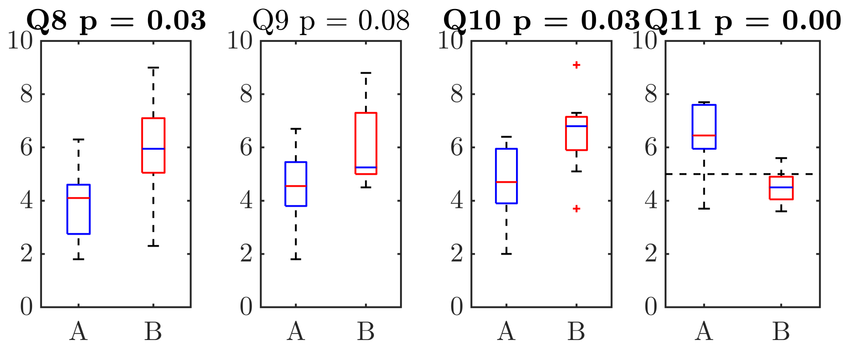

Figure 43: Sample 3 listening test results for 3 m listening distance. **Bold text** points out significant results and their p–values. Dashed line (- -) indicates bidirectional question midpoint.

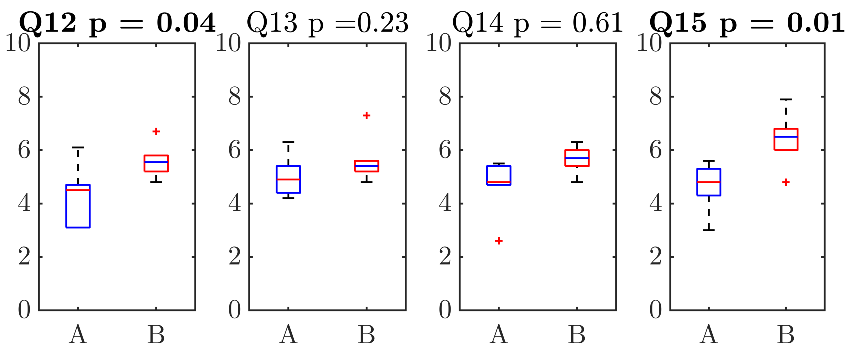

Figure 44: Sample 4 listening test results for 3 m listening distance. **Bold text** points out significant results and their p–values.

Figure 44 presents results for the sample 4. DUT B had improved low frequency clarity in the question 12. The DUT B had significantly better transient reproduction in the question 15.

The repeating observation between samples was the improved low frequency clarity and sound colour neutrality of the DUT B. The more directive loudspeaker improved transient reproduction significantly. Stereo imaging was similar between the DUTs. Less decaying energy yielded less masking (see Section 6.3), perceived sounds was less coloured, more accurate and less masked.

Some subjects noted the missing very low frequencies, correct sound balance between high, middle and low frequencies is a subjective matter (Olive et al. 2013). The subjects were noted on the high–pass filtering while briefing for the listening test. Author perceived larger difference between the auditioned loudspeaker with the 3 m listening distance compared to the 1m listening distance, though few subjects reported very small difference between the DUTs. Echoey sound was strongly present in the sample 2, guitars had tighter attack in the sample 3 and less modal ringing was perceived in the sample 4 when frequency masking was weak. Sound source was narrower in both lateral and sagittal directions when the sampled noise pulse was reproduced with the DUT B. Coloured echo was perceived as a separate auditory event along with the direct sound. The echo effect was significantly stronger with the DUT A. Overall critical listening experience with DUT B offered more details and nyances in the auditioned music.

7.2.3 Analysis Of Variance

The analysis of variance tells if the acquired data is valid for making conclusions. ANOVA compares the means and variances to define variables that are responsible for the reported results. Statistical significance is required for the results to be significant. Common consensus is that the difference of means between groups is statistically significant when

$$p < 0.05 \tag{19}$$

which stands for 95 % confidence interval for variable contributing to the gathered data. With p–value of 0.05, there is 5 % probability that observed variance could be explained by random fluctuation.

One–way ANOVA was performed separately for the each listening distance. Statistical significance was not found with any other variable than the auditioned loudspeakers for each listening distances in one–way ANOVA. Table 5 presents the p–values for 1 m and 3 m results in the one–way ANOVA

Question	p–Value 1 m	p–Value 3 m
1	1	0.340
2	**0.016**	**0.032**
3	**0.008**	**0.000**
4	**0.030**	**0.001**
5	**0.004**	0.272
6	**0.029**	0.059
7	**0.007**	**0.001**
8	**0.018**	**0.033**
9	0.120	0.077
10	0.392	**0.028**
11	**0.003**	**0.003**
12	**0.038**	0.163
13	0.232	0.199
14	0.061	0.309
15	**0.011**	**0.024**

Table 5: Table of p–values obtained from one–way ANOVA with 1 and 3 m listening distances. DUT was used as the grouping factor. Bold p–values indicate statistical significance for the acquired data.

Listener id and distance were added as variables by combining statistically significant results from the both 1 m and 3 m listening test results. The combined data is used for now on in the analysis.

The one–way ANOVAs were computed for listener identification, distance and DUT order. Listener identification was statistically significant for the questions 9, 10 and 14 (p–values 0.035, 0.018 and 0.025). Rest of the question had considerably big p–value. Listeners tend to use the scale similarly between the listener distances. Amount of correlation between the listening distances was therefore low. Other factors did not induce statistical significant to the results in one–way ANOVA

Multiple two–way ANOVAs were calculated to ensure the perceived differences in the listening tests are due to the increased directivity between the DUTs. Common effect of two separate variables was examined with a two–way ANOVA. The two–way ANOVA between was computed between the DUT and the listening distance. The more significant factor was the loudspeaker directivity. Distance was not significant factor for any of the questions. Combined factor of the distance and the DUT was not significant with 95 % confidence interval in any of the questions. The DUT and the listener identification were significant combined factors for the questions 5, 7 and 9 (p–values 0.043, 0.016 and 0.022).

7.2.4 Auditory Attributes

A factor analysis for explaining each variables contribution to the total variance was conducted. Statistically significant data from the one–way ANOVA was used in the factor analysis (see Section 7.2.3). The data was organized by the observed auditory attributes: Clarity, sound colouration, sound image width and MF transient reproduction. Contribution to the measured variance by each variable was analysed with the factor analysis, large variance stands for large influence on the perceived sound quality. Results for the factor analysis is presented in Table 6.

	Variance by Attributes			
Factor	**Clarity**	**Sound Colouration**	**Image Width**	**Transients**
Loudspeaker	**1.86**	**1.40**	**1.76**	**2.12**
Subject ID	0.40	0.16	-2.13[2]	0.01
Programme	0.21	-0.06[2]		
Listening Distance	0.15	-0.042[2]	-0.19[2]	0.09
DUT Order	0.16	-0.15[2]	-0.39[2]	0.12

Table 6: Factor analysis by attributes and variables. [2] Data includes outliers and therefore do not fit the constructed model, the true variance is likely to be 0.

The loudspeaker directivity caused the most variance in the listening test data. Contributed variances of the loudspeaker directivity: 1.40, 1.86, 1.76 and 2.1 are significant especially when compared to extracted variances by other variables. Variance of 2.0 is 25 % of the answering scale from 0 to 10. Other variables did not contribute on the perceived sound quality. Small contribution of the subject ID to the variance was unexpected since the auditioned loudspeakers, program or questionnaire did not change between the listening test rounds, each subject tends to use the answering scales similarly.

Outliers affect the image width results by reporting large negative variance. Scale usage between subjects was rough and therefore largely differentiating outliers can appear. Other negative values had small magnitudes. Negative variances propose that the data is not perfectly fit for the factor analysis model, negative values tra they can be though as zeros, not influencing the listening test results.

Principal component analysis (PCA) was conducted to find significant components and correlation between the variables. PCA is usually used to find the components that explain the data in smaller data size and reduce the raw data complexity. PCA in this Thesis is used to find correlations between the listening test variables. PCA do not point the significant component directly, it shows the correlations between variables. Rotation compensated pattern matrices with variable correlations are presented in the Tables 7, 8, 9 and 10. Two significant components were found for each attribute. The loudspeaker was noted with integers 1 and 2 corresponding to the DUT A and DUT B.

The Table 7 indicates the correlations between the listening test variables and clarity. The loudspeaker directivity correlated positively with clarity. The directive loudspeaker DUT B improved reproduced sound quality by increasing clarity. Other variables did not correlate with the clarity.

	Component	
Variable	**1**	**2**
Loudspeaker	**0.865**	0.011
Subject ID	0.04	0.894
Programme	-0.012	0.000
Listening Distance	-0.287	-0.004
DUT Order	-0.006	0.894
Clarity	**0.912**	-0.021

Table 7: Correlation between variables and LMF clarity. Correlation numbers extracted with PCA.

	Component	
Variable	**1**	**2**
Loudspeaker	**-0.907**	0.071
Subject ID	0.037	0.894
Programme	-0.032	0.003
Listening Distance	0.156	-0.001
DUT Order	-0.052	0.894
Sound Colouration	**0.913**	0.073

Table 8: PCA analysis for sound colour perception. Correlation between variables is presented.

The Table 8 presents the variable correlations between the listening test variables and sound colouration. The loudspeaker directivity correlated negatively with sound colouration, the conventional loudspeaker or DUT A was more coloured than the LMF directive loudspeaker. Other variables did not correlate significantly with the sound colouration.

The Table 9 presents transient reproduction correlation to other variables. Transient reproduction correlated with the auditioned loudspeaker, the LMF directive loudspeaker or DUT B had faster and clearer transient reproduction. Other correlations to the transient reproduction were not found.

	Component	
Variable	**1**	**2**
Loudspeaker	**0.898**	0.043
Subject ID	-0.073	0.894
Listening Distance	-0.138	-0.007
DUT Order	0.080	0.894
Transient Reproduction	**0.916**	-0.042

Table 9: PCA analysis for transient reproduction. Correlation between variables is presented.

Table 10 presents the image width correlation to other variables. Image width correlates negatively with the auditioned loudspeaker, the less directive DUT A has wider sound image. Significant correlations between other variables and the sound image width were not found.

	Component	
Variable	**1**	**2**
Loudspeaker	**-0.886**	0.026
Subject ID	0.115	0.894
Listening Distance	0.160	-0.005
DUT Order	-0.119	0.894
Sound Image Width	**0.927**	0.025

Table 10: PCA analysis for image width. Correlation between variables is presented.

The component 2 in all PCA Tables 7, 8, 9 and 10 suggested correlation to the order of auditioned loudspeakers and listener identification. Despite switching the order of DUTs randomly between subjects, many listeners auditioned loudspeaker with the same order in both listening test rounds. Relation between subject identification and the DUT order correlate accidentaly

8 Discussion

The target for a loudspeaker system is to reproduce the authentic sound stage with the correct localizations of signals. More early reflections may widen the stereo image and this can be interpreted as more pleasant (Klippel 1990*a*,*b*, Olive & Toole 1989). Increasing the loudspeaker LMF directivity improves the stereo imaging accuracy and the perceived quality of sound in a room (Linkwitz 1998, 2003, Ferekidis & Kempe 2004, 2005, Kang et al. 2008, Kantor & de Koster 1985, Weckström 1987, Salmi & Weckström 1982).

In the present work the active principle of creating directivity (Section 7), using multiple sound sources to steer acoustic radiation, was utilized to increase the LMF directivity in a compact form-factor loudspeaker. Objective measurements were conducted to quantify increase of the directivity, both in an anechoic chamber and in the listening room. The active method could increase directivity significantly. The directivity index DI in the LMF frequencies increased by over 2 dB compared to the natural directivity generated by a conventional loudspeaker design. No sound quality impairments were observed because of the active implementation. The physical size of the loudspeaker can remain relatively similar in size compared to a similar conventional loudspeaker design. The sensitivity of the active loudspeaker did not change significantly when the secondary transducer used for cancelling the rear radiation.

Listening tests (Section 7) show significant improvements in sound quality for the directive compact two–way loudspeaker. Clarity, lack of sound colouration and transient delivery improvements were due to the reduction of the spectral and temporal masking resulting from the higher directivity at LMF. The room sound decay time is similar to the masking threshold time (Section 2.5). Room measurements show reduction in the reflected energy level (Section 6.3). The room mode resonance decay time changes (Section6.3.2) show 10 dB improvements during the initial several hundred milliseconds into the sound decay, which significantly decreases the auditory spectral and temporal masking. These lower masking by the room reverberation. Improvements in sound colour were consistently reported for all listeners and for all listening distances. Conventional loudspeaker design was perceived warmer for both listening distances.

The reduced masking within the LMF band improves clarity, transient reproduction and reduces sound colourations. Some of the improved properties may be due changes in the in–situ magnitude response. Increasing LMF directivity reduces early reflection energy by 3–4 dB (Section 6.3.3). This agrees with Ferekidis & Kempe (2005), reporting similar front–to–back radiation energy difference. Increasing LMF directivity also narrows and focuses the perceived sound image at one-meter listening distance. The wider stereo image in the conventional loudspeaker design at this distance can be explained with the higher levels of lateral reflections in the room. Similar sound image widening was not detected for the three-meter listening distance. Decrease of direct-to-reverberant energy ratio DRR for the three-meter listening distance, due to the early reflection level increase relative to direct sound, may remain too small to cause statistically significant difference in the reproduced audio quality.

ANOVA, PCA and factor analysis indicate that the single most important factor explaining the variance was the loudspeaker directivity (Sections 7.2.3 and 7.2.4). An extensive analysis was carried out to examine data integrity, variable correlations and the source of variance. ANOVA was carried out to find the confidence limits of variance caused by the loudspeaker and to select data for the factor analysis and PCA (Section 7.2.3). In the listening tests, data variance was mainly caused by the loudspeaker directivity difference. The loudspeaker was the single most important factor explaining the variance (see Section 7.2.4). The use of evaluation scales varied between subjects, yielding outliers which presented as negative variance values in the factor analysis. The correlation between the perceived sound quality and the loudspeaker remains the only strong factor (Section 7.2.4).

The working hypothesis (Section 4) was satisfied: LMF directivity increases audio quality after the on-axis sound colour was eliminated as an influence by equalizing both A and B systems to have the same magnitude responses on the axis. The on–axis response is known to be the single most important determinant for the perceived sound quality (Section 2). Some test subjects reported that the audible difference between the conventional and increased directivity loudspeakers systems was very small but the directive loudspeaker system was always ranked better in sound quality over the conventional loudspeaker system.

Small room size limits the amount of acoustical treatment that can be deployed.

At low frequencies, for example in the LMF band, more and heavier absorbers are needed to control room acoustics. This is not feasible in for example crowded and small studio rooms. Physically large loudspeakers typically show less change in directivity across the audible frequency range. Active methods to generate LMF directivity enable characteristics similar to a large loudspeaker to be implemented in a physically small loudspeaker. Increasing acoustic directivity at low frequencies enables more accurate sound reproduction in spaces with less acoustical treatment. More directive loudspeakers also reduce acoustic reflections by console top and floor.

The currently available compact loudspeakers typically show increasing directivity with frequency as their directivity is created only by passive characteristics (transducer sizes, front baffle dimensions, enclosure shapes). At the same time,ITU recommends ITU (2015) that "the directivity index should increase smoothly with frequency". The ITU Recommendation does not clarify the reasoning behind this requirement. More uniform directivity across frequency, achieved by increasing the LMF directivity, did not corrupt the perceived sound quality. It is possible that the ITU Recommendation may originate from the fact that loudspeakers typically demonstrate less directivity towards low frequencies rather than psychoacoustical reasons, and this has simply been adopted in the Recommendation.

9 Conclusion

This Thesis has studied the audible effects of increasing the loudspeaker directivity at low-mid frequencies (LMF). A functional prototype of a compact two–way loudspeaker having an increased LMF directivity has been designed and implemented. An AB listening test has been conducted with two loudspeakers A and B, where the LMF directivity was the only varying characteristic between the loudspeakers A and B. The listening test results were statistically analysed, showing significant improvements in clarity, absence of sound colourations, transient reproduction and accuracy of the virtual sound images improved with increasing LMF directivity. Sound quality impairments were not detected when the LMF directivity increased. The LMF directivity was the main factor contributing to the subjective assessment of the sound quality.

Present recommendations (ITU 2015) require increasing the loudspeaker directivity with increasing frequency. The results of this Thesis do not support such requirement, and may offer evidence to reconsider the recommendations in this area. It is possible that the present recommendations stem mainly from the practical characteristics of existing loudspeaker designs.

References

Allison, R. F. (1974), *The Influence of Room Boundaries on Loudspeaker Power Output.* J. Audio Eng. Soc., vol. 22, no. 5, 314–320.

Ando, Y. (1977), *Subjective Preference in Relation to Objective Parameters of Music Sound Fields with a Single Echo.* J. Acous. Soc. of America, vol. 62, no. 6, 1436–1441.

Angus, J. A. S. (2013), *The Effect of Low Frequency Reflections on Stereo Images.* Proc. Audio Eng. Soc. 134.

Antsalo, P., Karjalainen, M., Mävirta, A. & Välimäki, V. (2004), *Perception of Temporal Decay of Low-frequency Room Modes.* Proc. Audio Eng. Soc. 116.

Backman, J. (1989), *Computation of Diffraction for Loudspeaker Enclosures.* J. Audio Eng. Soc., vol. 37, no. 5, 353–362.

Backman, J. (1995), *A Model of Diffraction from Convex Loudspeaker Enclosure Edges.* Proc. Audio Eng. Soc. 98.

Backman, J. (1999), *Theory of Acoustical Resistance Enclosures.* Proc. Audio Eng. Soc. 106.

Backman, J. (2003), *Low-frequency Polar Pattern Control for Improved In-room Response.* Proc. Audio Eng. Soc. 115.

Ballagh, K. O. (1983), *Optimum Loudspeaker Placement Near Reflecting Planes.* J. Audio Eng. Soc., vol. 31, no. 12, 931–935.

Balogh, G. (1977), *Cardioid Electro–Acoustic Radiator.* US Patent 4,054,748.

Barron, M. (1971), *The Subjective Effects of First Reflections in Concert Halls—the Need for Lateral Reflections.* J. Sound and Vibration, vol. 15, no. 4, 475–494.

Bücklein, R. (1981), *The Audibility of Frequency Response Irregularities.* J. Audio Eng. Soc., vol. 29, no. 3, 126–131.

Bech, S. (1994), *Perception of Timbre of Reproduced Sound in Small Rooms: Influence of Room and Loudspeaker Position.* J. Audio Eng. Soc., vol. 42, no. 12, 999–1007.

Bech, S. (1995), *Timbral Aspects of Reproduced Sound in Small Rooms. I.* J. Acous. Soc. of America, vol. 97, no. 3, 1717–1726.

Benade, A. H. (1985), *From Instrument to Ear in a Room: Direct or via Recording.* J. Audio Eng. Soc., vol. 33, no. 4, 218–233.

Benjamin, E. (2006), *An Experimental Verification of Localization in Two-Channel Stereo.* Proc. Audio Eng. Soc. 121.

Beranek, L. L. (2006), Analysis of sabine and eyring equations and their application to concert hall audience and chair absorption, Vol. 120, ASA, pp. 1399–1410. The J. of the Acous. Soc. of America.

Beranek, L. L. (2012), *Acoustics.* Academic Press, Oxford.

Beranek, L. & Martin, D. W. (1996), *Concert and Opera Halls: How They Sound*, Acous. Soc. of America, New York.

Blauert, J. (1983), *Spatial Hearing: The Psychophysics of Human Sound Localization.* MIT Press, Cambridge.

Boone, M. M. & Ouweltjes, O. (1997), *Design of a Loudspeaker System with a Low-Frequency Cardiodlike Radiation Pattern.* J.Audio Eng. Soc., vol. 45, no. 9, 702–707.

Borenius, J. (1985), *Perceptibility of Direction and Time Delay Errors in Subwoofer Reproduction.* Proc. Audio Eng. Soc. 79.

Boucher, J. M., Debail, B. G., Diquelou, P. Y., Kerneis, Y. & Shaiek, H. (2006), *Enhanced Control of Sound Field Radiated by Co-axial Loudspeaker Systems Using Digital Signal Processing Techniques.* Proc. Audio Eng. Soc. 120.

Bradley, J. S. & Soulodre, G. A. (1995), *The Influence of Late Arriving Energy on Spatial Impression.* J. Acous. Soc. of America, vol. 97, no. 4, 2263–2271.

Bridges, S. (1980), *Effect of Direct Sound on Perceived Frequency Response of a Sound System.* Proc. Audio Eng. Soc. 66.

Case, A. U. (2001), *An Investigation of the Spectral Effect of Multiple Early Reflections.* J. Acous. Soc. of America, vol. 109, no. 5, 2303–2303.

Celestinos, A. & Nielsen, S. B. (2008), *Low-Frequency Loudspeaker–Room Simulation Using Finite Differences in the Time Domain—Part 1: Analysis.* J. Audio Eng. Soc., vol. 56, no. 10, 772–786.

Chapelle, P. H. (1973), *The Frequency Response of Loudspeakers: On-Axis Response or Power Response?* Proc. Audio Eng. Soc. 44.

Clark, D. (2010), *Listening Tests for the Effect of Loudspeaker Directivity and Positioning on Auditory Scene Perception.* Proc. Audio Eng. Soc. 129.

d & b Audiotechnik (2015), ‘*ArrayProcessing Feature, Technical White Paper*’. *https://www.dbaudio.com/global/en/service-and-support/downloads/* Accessed February 18 2020.

D’Appolito, J. A. (1983), *A Geometric Approach to Eliminating Lobing Error in Multiway Loudspeakers.* Proc. Audio Eng. Soc. 74.

Davis, M. F. (1987), *Loudspeaker Systems with Optimized Wide-Listening-Area Imaging.* J. Audio Eng. Soc., vol. 35, no. 11, 888–896.

Devantier, A. (2002), *Characterizing the Amplitude Response of Loudspeaker Systems.* Proc. Audio Eng. Soc. 113.

Di Cola, M., Cinanni, D., Manzini, A., Nizzoli, T. & Ponteggia, D. (2009), *Design and Optimization of High Directivity Waveguide for Vertical Array.* Proc. Audio Eng. Soc. 127.

Dodd, M. & Oclee-Brown, J. (2007), *A New Methodology for the Acoustic Design of Compression Driver Phase-Plugs with Concentric Annular Channels.* Proc. Audio Eng. Soc. 123.

EBU (1998), *3276 Listening Conditions for the Assessment of Sound Program Material: Monophonic and Two–Channel Stereophonic.* European Broadcast Union, Geneve.

Evans, W., Dyreby, J., Bech, S., Zielinski, S. & Rumsey, F. (2009), *Effects of Loudspeaker Directivity on Perceived Sound Quality - A Review of Existing Studies.* Proc. Audio Eng. Soc. 126.

Fahy, F. J. & Salmon, V. (1990), *Sound Intensity.* Acous. Soc. America, New York.

Fazenda, B. (2004), *Perception of Room Modes in Critical Listening Spaces.* PhD Thesis, University of Salford.

Fazenda, B. & Wankling, M. (2008), *Optimal Modal Spacing and Density for Critical Listening.* Proc. Audio Eng. Soc. 125.

Fazi, F. M., Shin, M., Olivieri, F. & Fontana, S. (2015), *Low Frequency Performance of Circular Loudspeaker Arrays.* Proc. Audio Eng. Soc. 138.

Feddersen, W., Sandel, T., Teas, D. & Jeffress, L. (1957), *Localization of High-Frequency Tones.* J. Acous. Soc. of America, vol. 29, no. 9, 988–991.

Ferekidis, C. & Kempe, U. (1996), *Room Mode Excitation of Dipolar and Monopolar Low-Freqency Sources.* Proc. Audio Eng. Soc. 100.

Ferekidis, L. & Kempe, U. (2004), *The Beneficial Coupling of Cardioid Low Frequency Sources to the Acoustics of Small Rooms.* Proc. Audio Eng. Soc. 116.

Ferekidis, L. & Kempe, U. (2005), *Controlling the Mode Excitation of Rooms by Using Multiple Low Frequency Cardioids in Multichannel Systems.* Proc. Audio Eng. Soc. 118.

Flindell, I. H., McKenzie, A. R., Negishi, H., Jewett, M. & Word, P. (1991), *Subjective Evaluations of Preferred Loudspeaker Directivity.* Proc. Audio Eng. Soc. 90.

Geddes, E. R. (1986), *Source Radiation Characteristics.* J. Audio Eng. Soc., vol. 34, no. 6, 464–478.

Geddes, E. R. (2006), '*Acoustic waveguide for controlled sound radiation*'. US Patent 7,068,805.

Genelec Oy (2014). Genelec 1238A datasheet, *https://www.genelec.com/1238a/section-downloads*, Accessed Februacy 10, 2020.

Genereux, R. P. (1990), *Adaptive Loudspeaker Systems: Correcting for the Acoustic Rnvironment.* Proc. Audio Eng. Soc. 8, The Sound of Audio.

Gentner, K., Braasch, J. & Calamia, P. (2007), *A Perforated Desk Surface to Diminish Coloration in Desktop Audio-Production Environments.* Proc. Audio Eng. Soc. 123.

Gómez, J., Paiva, R. C. D., Saksela, K., Svedström, T. & Pulkki, V. (2011), *Perception of Spatial Distribution of Room Response Reproduced Using Different Loudspeaker Setups.* Proc. Audio Eng. Soc. 130.

Günel, B., Hacihabiboglu, H. & Kondoz, A. M. (2007), *Source Directivity Simulation in Digital Waveguide Mesh-Based Room Acoustics Models.* Proc. Audio Eng. Soc. 30.

Goldberg, A., Mäkivirta, A. & Varla, A. (2006), *Compensating the Response of Near-Field Loudspeaker Monitors to Minimize the Effects of Desktop Acoustic Loading.* J. Audio Eng. Soc., vol. 54, no. 5, 401–411.

Griesinger, D. (1997), *Spatial Impression and Envelopment in Small Rooms.* Proc. Audio Eng. Soc. 103.

Griesinger, D. (1998), *General Overview of Spatial Impression, Envelopment, Localization, and Externalization.* Proc. Audio Eng. Soc. 15 Audio, Acoustics and Small Spaces.

Griesinger, D. (1999), *Objective Measures of Spaciousness and Envelopment.* Proc. Audio Eng. Soc. 16 Spatial Sound Reproduction.

Griesinger, D. (2009), *The Importance of the Direct to Reverberant Ratio in the Perception of Distance, Localization, Clarity, and Envelopment.* Proc. Audio Eng. Soc. 126.

Griesinger, D. (2012), *Pitch, Timbre, Source Separation, and the Myths of Loudspeaker Imaging.* Proc. Audio Eng. Soc. 132.

Griesinger, D. (2018), *Reproducing Low Frequency Spaciousness and Envelopment in Listening Rooms.* Proc. Audio Eng. Soc. 145.

Groh, A. R. (1974), *High-Fidelity Sound System Equalization by Analysis of Standing Waves.* J. Audio Eng. Soc., vol. 22, no. 10, 795–799.

Gudvangen, S. (2014), *Properties of Gradient Loudspeakers.* Proc. Audio Eng. Soc. 136.

Gunness, D. W. (2018), *Passive Cardioid Speaker.* US Patent 10,123,111.

Haas, H. (1972), *The Influence of a Single Echo on the Audibility of Speech.* J. Audio Eng. Soc., vol. 20, no. 2, 146–159.

Hartman, W. (1983), *Localization of Sound in Rooms.* J Acous. Soc. America, vol. 74, 1380–1391.

Hayashi, S., Mochimaru, A. & Fidlin, P. F. (2012), *The Radiation Characteristics of a Horizontally Asymmetrical Waveguide that Utilizes a Continuous Arc Diffraction Slot.* Proc. Audio Eng. Soc. 133.

Hermes, K. (2019), *Factors Influencing the Spectral Clarity of Vocals in Music Mixes.* Proc. Audio Eng. Soc. 146.

Hill, A. J. (2012), *Analysis, Modeling and Wide-Area Spatiotemporal Control of Low-Frequency Sound Reproduction.* PhD thesis, University of Essex.

Holm, J. (2010), *Applying the Finite Element Method for Modelling Loudspeaker Waveguide Directivity.* Master Thesis, Aalto University.

Holm, J. & Mäkivirta, A. (2013), *Quantifying Diffraction in Time Domain with Finite Element Method.* Proc. Audio Eng. Soc. 5.

Holmes, T. J. (1986), *The Acoustic Resistance Box–A Fresh Look at an Old Principle.* J. Audio Eng. Soc., vol. 34, no. 12, 981–989.

Hongisto, V. (2015), Meluntorjunta. Aalto-yliopiston teknillinen korkeakoulu, signaalinkäsittelyn ja akustiikan laitos, Espoo.

Howze, B. & Henrickson, C. (1983), '*Loudspeaker horn*'. US Patent 4,390,078.

Iding, W. H. (1971), *Unidirectionally Radiating Loudspeakers.* Proc. Audio Eng. Soc. 1ce.

Imamura, H., Marui, A., Kamekawa, T. & Nakahara, M. (2014), *Influence of Directional Differences of First Reflections in Small Spaces on Perceived Clarity.* Proc. Audio Eng. Soc. 136.

ITU (2015), *ITU-R BS.1116-3. Methods for the Subjective Assessment of Small Impairments in audio systems including multichannel sound systems.* International Telecommunications Union, Geneva.

Jensen, R. E. & Welti, T. (2003), *The Importance of Reflections in a Binaural Room Impulse Response.* Proc. Audio Eng. Soc. 114.

Kamaris, G. & Mourjopoulos, J. (2018), *Stereo Image Localization Maps for Loudspeaker Reproduction in Rooms.* Proc. Audio Eng. Soc. 144.

Kang, K. O., Nelson, P. & Park, M. (2008), *Evaluation of Stereophonic Images with Listening Tests and Model Simulations.* Proc. Audio Eng. Soc. 124.

Kantor, K. L. & de Koster, A. P. (1985), *A Psychoacoustically Optimized Loudspeaker.* Proc. Audio Eng. Soc. 77.

Kaplanis, N., Bech, S., Jensen, S. H. & van Waterschoot, T. (2014), *Perception of Reverberation in Small Rooms: a Literature Study.* Proc. Audio Eng. Soc. 55.

Karjalainen, M., Antsalo, P. & Mäkivirta, A. (2003), *Modal Equalization by Temporal Shaping of Room Response.* Proc. Audio Eng. Soc. 23.

Kashani, R. & Wischmeyer, J. (2004), *Electronic Bass Trap.* Proc. Audio Eng. Soc. 117.

Kates, J. M. (2002), *Dipole Loudspeaker Response in Listening Rooms.* J. Audio Eng. Soc., vol. 50, no. 5, 363–374.

Kelloniemi, A., Ahonen, J., Paajanen, O. & Pulkki, V. (2005), *Detection of Subwoofer Depending on Crossover Frequency and Spatial Angle Between Subwoofer and Main Speaker.* Proc. Audio Eng. Soc. 118.

King, R., Leonard, B. & Sikora, G. (2011), *The Practical Effects of Lateral Energy in Critical Listening Environments.* Proc. Audio Eng. Soc. 131.

Klippel, W. (1990*a*), *Assessing the Subjectively Perceived Loudspeaker Quality on the Basis of Objective Parameters.* Proc. Audio Eng. Soc. 88.

Klippel, W. (1990*b*), *Multidimensional Relationship Between Subjective Listening Impression and Objective Loudspeaker Parameters.* J. Acustica, vol. 70, no. 1, 45–54.

Klippel, W. (1995), *Nonlinear System Models for Horn Loudspeakers.* Proc. Audio Eng. Soc. 99.

Kuttruff, H. & Thiele, R. (1954), *Frequency Dependence of the Sound Pressure in Rooms.* J. Acoustica, vol. 4, 614–614.

Laitinen, M.-V., Politis, A., Huhtakallio, I. & Pulkki, V. (2015), 'Controlling the perceived distance of an auditory object by manipulation of loudspeaker directivity'. J. Acous. Soc. America, vol. 137, EL462.

Lehnert, H. (1993), *Auditory Spatial Impression.* Proc. Audio Eng. Soc. 12.

Linkwitz, S. (1992), *Development of a Compact Dipole Loudspeaker.* Proc. Audio Eng. Soc. 93.

Linkwitz, S. (1998), *Investigation of Sound Quality Differences between Monopolar and Dipolar Woofers in Small Rooms.* Proc. Audio Eng. Soc. 105.

Linkwitz, S. (2003), *Why Is Bass Reproduction from a Dipole Woofer in a Living Room Often Subjectively More Accurate Than from a Monopole Woofer?* J. Audio Eng. Soc., vol. 51, no. 11, 1062–1063.

Linkwitz, S. (2007), *Room Reflections Misunderstood?* Proc. Audio Eng. Soc. 123.

Linkwitz, S. (2009), *The Challenge to Find the Optimum Radiation Pattern and Placement of Stereo Loudspeakers in a Room for the Creation of Phantom Sources and Simultaneous Masking of Real Sources.* Proc. Audio Eng. Soc. 127.

Lipshitz, S. P. & Vanderkooy, J. (1985), *Experiments in Direct/Reverberant Ratio Modification.* Proc. Audio Eng. Soc. 79.

Mäkivirta, A., Antsalo, P., Karjalainen, M. & Välimäki, V. (2003), *Modal Equalization of Loudspeaker-Room Responses at Low Frequencies.* J. Audio Eng. Soc., vol 51, no. 5, 324–343.

Mäkivirta, A., Väisänen, J., Martikainen, I., Lund, T. & Naghian, S. (2017), *The Acoustic Design of Minimum Diffraction Coaxial Loudspeakers with Integrated Waveguides.* Proc. Audio Eng. Soc. 142.

Miller III, R. R. (2005), *Physiological and Content Considerations for a Second Low Frequency Channel for Bass Management, Subwoofers, and LFE.* Proc. Audio Eng. Soc. 119.

Møller, M., Olsen, M., Agerkvist, F. T., Dyreby, J. & Munch, G. K. (2010), *Circular Loudspeaker Arrays with Controllable Directivity.* Proc. Audio Eng. Soc. 128.

Morimoto, M. (1997), *The Role of Rear Loudspeakers in Spatial Impression.* Proc. Audio Eng. Soc. 103.

Mourjopoulos, J. & Paraskevas, M. (1991), *Pole and Zero Modeling of Room Transfer Functions.* J. Sound and Vibration, vol. 146, no. 2, 281–302.

Norcross, S. G., Soulodre, G. A. & Lavoie, M. C. (2003), *Temporal Aspects of Listener Envelopment in Multichannel Surround Systems.* Proc. Audio Eng. Soc. 114.

Noy, D., Hauser, G. & Storyk, J. (2003), *Low Frequency Absorbers - Applications and Comparisons.* Proc. Audio Eng. Soc. 115.

Olive, S. E. (2003), *Differences in Performance and Preference of Trained Versus Untrained Listeners in Loudspeaker Tests: A Case Study.* J. Audio Eng. Soc., vol. 51, no. 9, 806–825.

Olive, S. E. (2004*a*), *A Multiple Regression Model for Predicting Loudspeaker Preference Using Objective Measurements: Part I - Listening Test Results.* Proc. Audio Eng. Soc. 116.

Olive, S. E. (2004*b*), *A Multiple Regression Model for Predicting Loudspeaker Preference Using Objective Measurements: Part II - Development of the Model.* Proc. Audio Eng. Soc. 117.

Olive, S. E. & Martens, W. L. (2007), *Interaction between Loudspeakers and Room Acoustics Influences Loudspeaker Preferences in Multichannel Audio Reproduction.* Proc. Audio Eng. Soc. 123.

Olive, S. E., Schuck, P. L., Sally, S. L. & Bonneville, M. (1995), *The Variability of Loudspeaker Sound Quality Among Four Domestic-Sized Rooms.* Proc. Audio Eng. Soc. 99.

Olive, S. E. & Toole, F. E. (1989), *The Detection of Reflections in Typical Rooms.* J. Audio Eng. Soc., vol. 37, no. 7/8, 539–553.

Olive, S. E., Welti, T. & McMullin, E. (2013), *Listener Preferences for In-Room Loudspeaker and Headphone Target Responses.* Proc. Audio Eng. Soc. 135.

Olivier, L. B. (2010), *Subwoofer Directivity Control.* Thesis, Aalborg University.

Olson, H. F. (1957), *Acoustical Engineering*, van Nostrand, New York.

Olson, H. F. (1973), *Gradient Loudspeakers.* J. Audio Eng. Soc., vol. 21, no. 2, 86–93.

Panzer, J. & Ponteggia, D. (2011), *Inverse Distance Weighting for Extrapolating Balloon-Directivity-Plots.* Proc. Audio Eng. Soc. 131.

Perrott, D. R. & Buell, T. N. (1982), *Judgments of Sound Volume: Effects of Signal Duration, Level, and Interaural Characteristics on the Perceived Extensity of Broadband Noise.* J. Acous. Soc. of America, vol. 72, no. 5, 1413–1417.

Poletti, M. & Betlehem, T. (2013), *Design of a Prototype Variable Directivity Loudspeaker for Improved Surround Sound Reproduction in Rooms.* Proc. Audio Eng. Soc. 52.

Pulkki, V. & Karjalainen, M. (2015), *Communication Acoustics: an Introduction to Speech, Audio and Psychoacoustics*, John Wiley and Sons, Chichester.

Queen, D. (1979), *The Effect of Loudspeaker Radiation Patterns on Stereo Imaging and Clarity.* J. Audio Eng. Soc., vol. 27, no. 5, 368–379.

Rasmussen, S. & Rasmussen, K. B. (1994), *On Loudspeaker Cabinet Diffraction.* J. Audio Eng. Soc vol. 42 no. 3 147–150.

Rodenas, J. A. & Aarts, R. M. (2001), *Sweet Spot Widening for Stereophonic Sound Reproduction.* Proc. 2001 IEEE Workshop, 191–194.

Sabine, W. C. (1927), *Collected Papers on Acoustics.* Harvard University Press, Cambridge.

Salmensaari, O., Borenius, J. & Backman, J. (1993), *Improving the Listening Conditions in Sound Control Rooms by Mounting the Loudspeaker in a Large Baffle.* Proc. Audio Eng. Soc. 94.

Salmi, J. & Weckström, A. (1982), *Listening Room Influence on Loudspeaker Sound Quality and Ways of Minimizing It.* Proc. Audio Eng. Soc. 71.

Sato, K. & Haneda, Y. (2017), *Filter Design of a Circular Loudspeaker Array Considering the Three-Dimensional Directivity Patterns Reproduced by Circular Harmonic Modes.* Proc. Audio Eng. Soc. 142.

Schulein, R. B. (1975), *In Situ Measurement and Equalization of Sound Reproduction Systems.* J. Audio Eng. Soc., vol. 23, no. 3, 178–186.

Schultz, T. J. (1963), *Problems in the Measurement of Reverberation Time.*

Sinclair, R. (1978), *The Dispersion of Re-Entrant Exponential Horns.* Proc. Audio Eng. Soc. 60.

Sridhar, R., Tylka, J. G. & Choueiri, E. (2016), *Metrics for Constant Directivity.* Proc. Audio Eng. Soc. 140.

Staffeldt, H. (1974), *Correlation Between Subjective and Objective Data for Quality Loudspeakers.* J. Audio Eng. Soc., vol. 22, no. 6, 402–415.

Staffeldt, H. (1984), *Measurement and Prediction of the Timbre of Sound Reproduction.* J. Audio Eng. Soc., vol. 32, no. 6, 410–414.

Subkey, A., Cabrera, D. & Ferguson, S. (2005), *Localization and Image Size Effects for Low Frequency Sound.* Proc. Audio Eng. Soc. 118.

Taylor, R. & Keele, Jr., D. B. D. (2017), *Theory of Constant Directivity Circular-Arc Line Arrays.* Proc. Audio Eng. Soc. 143.

Taylor, R., Manke, K. & Keele, Jr., D. B. D. (2017), *Constant Directivity Circular-Arc Arrays of Dipole Elements.* Proc. Audio Eng. Soc. 143.

Theiss, B. & Hawksford, M. J. (1998), *Binaural Model-Based Measurements of Phantom Images.* Proc. Audio Eng. Soc. 105.

Thiele, G. & Kügler, C. (1992), *Loudspeaker Reproduction: Study on the Subwoofer Concept.* Proc. Audio Eng. Soc. 92.

Toole, F. (2015), 'The measurement and calibration of sound reproducing systems'. J. Audio Eng. Soc., vol. 63, no. 7/8, 512–541.

Toole, F. E. (1986*a*), *Loudspeaker Measurements and Their Relationship to Listener Preferences: Part 1.* J. Audio Eng. Soc., vol. 34, no. 4, 227–235.

Toole, F. E. (1986*b*), *Loudspeaker Measurements and Their Relationship to Listener Preferences: Part 2.* J. Acous. Eng. Soc., vol. 34 no. 5 323–348.

Toole, F. E. (1999), *Audio–Science in the Service of Art.* Harman International.

Toole, F. E. (2017), *Sound Reproduction: The Acoustics and Psychoacoustics of Loudspeakers and Rooms.* Focal Press, Oxford.

Toole, F. E. & Olive, S. E. (1988), *The Modification of Timbre by Resonances: Perception and Measurement.* J. Audio Eng. Soc., vol. 36, no. 3, 122–142.

Tylka, J. G., Sridhar, R. & Choueiri, E. (2015), *A Database of Loudspeaker Polar Radiation Measurements.* Proc. Audio Eng. Soc. 139.

van der Werff, J. (2001), *Electronically Controlled Loudspeaker Arrays Without Side Lobes.* Proc. Audio Eng. Soc. 110.

Vanderkooy, J. (2006), The acoustic center: A new concept for loudspeakers at low frequencies. Proc. Audio Eng. Soc. 121.

Walker, R. (1994), *Early Reflections in Studio Control Rooms: The Results From the First Controlled Image Design Installations.* Proc. Audio Eng. Soc. 96.

Wang, S., Wang, H. & Zhang, C. (2016), *Relative Contribution of Interaural Time and Level Differences to Selectivity for Sound Localization.* Proc. Audio Eng. Soc. 140.

Warusfel, O., Derogis, P. & Causse, R. (1997), *Radiation Synthesis with Digitally Controlled Loudspeakers.* Proc. Audio Eng. Soc. 103.

Weckström, A. (1987), *Directive Loudspeaker Design.* Proc. Audio Eng. Soc. 82.

Zacharov, N. (1997), *On the Loudspeaker Directivity Considerations for 5.1-Channel Audio-Visual Reproduction: A Subjective Appraisal.* Proc. Audio Eng. Soc. 102.

Zacharov, N. (1998), *Subjective Appraisal of Loudspeaker Directivity for Multichannel Reproduction.* J. Audio Eng. Soc. vol. 35 no. 4 288–303.

Zacharov, N. et al. (2000), *Perceptual Studies on Spatial Sound Reproduction Systems*, PhD Thesis, Helsinki University of Technology.

Zwicker, E. & Fastl, H. (1999), *Psychoacoustics: Facts and Models.* Springer, Berlin.

Appendix A: Questionnaire

GENELEC®

Test ID:

SUBJECTIVE EVALUATION

Name: Date: Distance:

Sample#1: Noise burst

<u>Q1: Sound Image Location</u>

DUT 1

Left Right

DUT 2

Comments:...

<u>Q2: Sound Image Width</u>

DUT 1

Narrow Wide

DUT 2

Comments:...

<u>Q3: Colouration, low-mid frequency</u>

DUT 1

cold Warm

DUT 2

Comments:...

1

GENELEC®

Sample #2: Metal

Q4: Clarity, low frequency

DUT 1
Poor, muddy | Good, clear
DUT 2

Comments:...

Q5: Clarity, mid frequency

DUT 1
Poor, muddy | Good, clear
DUT 2

Comments:...

Q6: Sound Image Definition

DUT 1
Unclear | Focused, Precise
DUT 2

Comments:...

Q7: Colouration, low-mid frequency

DUT 1
Cold | Warm
DUT 2

Comments:...

GENELEC®

Sample#3: Hard Rock

Q8: Clarity, low frequency

DUT 1
Poor, muddy — Good, clear
DUT 2

Comments:...

Q9: Clarity, mid frequency

DUT 1
Poor, muddy — Good, clear
DUT 2

Comments:...

Q10: Clarity, vocals

DUT 1
Poor, muddy — Good, clear
DUT 2

Comments:...

Q11: Colouration, low-mid frequency

DUT 1
Cold — Warm
DUT 2

Comments:...

GENELEC®

Sample#4: Fusion Jazz

Q12: Clarity, low frequency

DUT 1

Poor, muddy — Good, clear

DUT 2

Comments:..

Q13: Clarity, mid frequency

DUT 1

Poor, muddy — Good, clear

DUT 2

Comments:..

Q14: Clarity, melodies

DUT 1

Poor, muddy — Good, clear

DUT 2

Comments:..

Q15: Mid frequency transients

DUT 1

Poor, muddy, slow — Good, clear, fast

DUT 2

Comments:..

Appendix B: Comments from Listening Tests

1 m Listening Distance

- Echoiy sound was missing from 1 m listening round compared to the 3m listening round

3 m Listening Distance

- Subjectively different frequency responses between DUTs A and B
- Not much real low frequency in any sample
- Unfamiliar music.
- Q8: DUT 1 (A) has thin bass
- Q2: Sound seems to widen in up-down direction when changed from DUT 2 (B) to 1 (A).

Appendix C: Listening Room Illustration