8-2017

Perception Of Spatial Impression Due To Varying Positions Of The Source On Stage

SungBeen Cho
University of Nebraska-Lincoln, sungbeen@huskers.unl.edu

PERCEPTION OF SPATIAL IMPRESSION DUE TO VARYING POSITIONS OF THE SOURCE ON STAGE

by

Sungbeen Cho

A THESIS

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Master of Science

Major: Architectural Engineering

Under the Supervision of Professor Lily M. Wang

Lincoln, Nebraska

August, 2017

PERCEPTION OF SPATIAL IMPRESSION DUE TO VARYING POSITIONS OF THE SOURCE ON STAGE

Sungbeen Cho, M.S.

University of Nebraska, 2017

Advisor: Lily M. Wang

The listener's surrounding environment affects the perception of sound. The environment includes physical factors such as space size, shape, and finish materials, and psychological factors such as individual difference of impression, and vision. Acoustic spatial impression can be defined as the concept of the type and size of space at which a listener arrives spontaneously when he/she is exposed to an appropriate sound field. This thesis analyzes how physical factors in enclosed spaces affect the acoustic spatial impression, and how sound sources at different positions on stage are perceived in different shapes of spaces.

First, existing spatial impression parameters, IACC (Interaural Cross-Correlation Coefficient) and LF (Lateral Energy Fraction), are analyzed to see how they vary across different source positions. In addition, how acoustic energy changes according to the sound source location is measured and observed in actual spaces, and a new metric named ILD-Correlation Range (ILD-CR) is suggested to understand spatial impression across varying source positions. This metric is based on the Interaural Level Difference (ILD), an important factor in localization perception.

Next, the thesis explores how the shape of the space and the positions of the sound source on stage influence human perception through subjective testing using

auralizations, across rooms with different reverberation times and at two distances between sound source and listener. The ILD is found to vary according to the shape of the space, and localization perception demonstrates significant similarity to the ILD. Other factors besides the shape of the room, such as reverberation time, the distance between sound source and listener, and frequencies are also found to have a significant effect on the listener's spatial perception.

Acknowledgments

I would like to acknowledge my advisor Dr. Lily M. Wang. She helped me to get my research direction right. I would also like to thank Dr. Mahbubul Majumder. Without his passionate participation and input, the statistical analysis could not have been successfully conducted. I would also like to acknowledge Dr. Erica Ryherd. She was not only a committee member of my thesis but also helped me to understand my acoustics more deeply in manner.

I would like to thank the Nebraska Acoustics Group. Special thanks go to Laura Brill; always my colleague and advisor, Dr. Matt Blevins; providing equipment for my experiments, and Dr. Joonhee Lee; always giving comments and advice.

Finally, I must express my very profound gratitude to my parents, parents - in - law, and my wife for providing me with unfailing support and continuous encouragement throughout my years of study.

Table of Contents

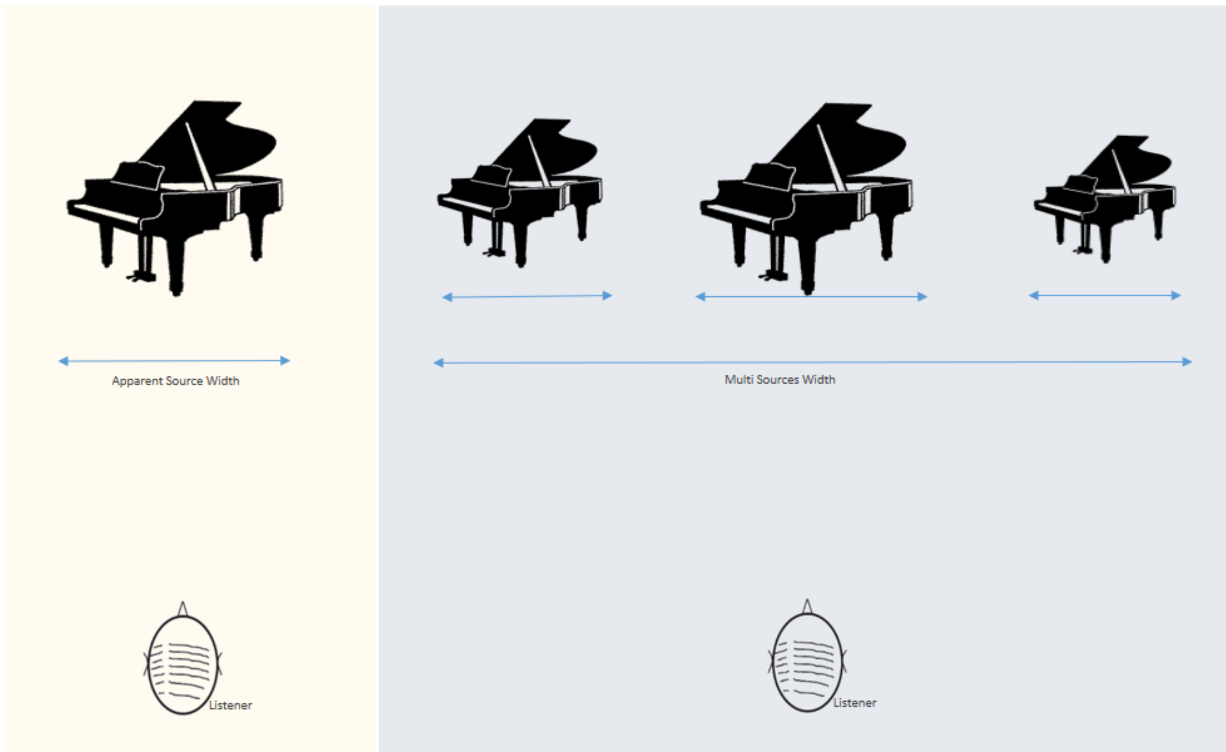
List of Figures

List of Tables

Chapter 1

Introduction

Spatial impression is defined as the concept of the type and size of an actual or simulated space to which a listener arrives spontaneously when he/she is exposed to an appropriate sound field [1]. In the method of expressing the spatial impression, the Apparent Source Width (ASW, figure 1.1 (a)) has been proposed to explain the width of the sound source [2]. In general, the evaluation of the ASW uses the ratio of the lateral energy delivered initially to a listener. A set source location measurement is useful for evaluating ASW. However, since multiple sound source locations are used on stage rather than a single sole sound source position in a performance, the measurement of ASW using a single sound source is not enough to describe the characteristics of sources anywhere on the entire stage (figure 1.1 (b)). For instance, a performance such as a piano recital is played by a single instrument, so the ASW evaluation method using a single sound source position is suitable. However, if a number of people, such as triple, quartet, or chamber music, play at the same time, one should evaluate the entire group together. Therefore, it is important to evaluate how the listener perceives the sound source positions all across the stage and how widely spread the sound source appears on the stage. For these reasons, a way of evaluating how spatial impression changes as sound source position on stage varies is needed.



(a) Single source ASW

(b) Multi source ASW

Figure 1.1: Apparent Source Width(ASW)

This thesis discusses how the existing spatial impression metrics (IACC, LF) reflect the change of acoustic energy from varying sound source locations, and also suggests a new evaluation method, ILD-CR to evaluate spatial impressions according to the sound energy variation from various sound source locations.

The newly proposed metric, ILD-CR is based on the way humans perceive location. To perceive sound location, humans use the difference in time that sound arrives from source to each of the listener's two ears and the level difference between the left and right ears. The time difference between two ears is called Interaural Time Difference (ITD) and the level difference is Interaural Level Difference (ILD) [3]. If the time difference or level difference between the left and right ears is large enough, the sound will be heard as an echo, but if the difference is below the echo threshold, the listener's position perception will change depend on this difference [4]. Since the arrival time to each of the two ears is proportional to the distance between source and receiver changes to the physical spaces do not directly affect Interaural Time Difference(ITD) if the distance remains the same, but Interaural Level Differences (ILD) can vary in each venue because of the shape of spaces and wall finishing materials. In a free-field, listeners do not receive reflections, but in a bounded field, reflections can affect the localization. The transmissions of sound in an enclosed space reach the listener through various acoustic phenomena such as reflection, absorption, diffusion, and refraction. Even though the distance and positions of the sound source and the listener are the same, the transmission path of acoustic energy in the enclosed space differs depending on the shape of the space. Therefore, the transmitted acoustic energy is not the same as in the free field. Furthermore, the values of ILD can be varied depending on the sound source position in the same space, and change in ILD each source position can affect the distance perception between those positions.

ILD-CR is a method of evaluating the spatial impression across a stage by using the ILD of multiple sound source positions. To verify its effectiveness, this thesis examines whether it has valid sound source location information from studying simulations and physical measurements in assorted spaces. This thesis additionally investigates the correlation between human perception and ILD in enclosed space through auditory experiments. Furthermore, to investigate the change of perception according to the change of the sound environment, analysis of results from the experiment have been conducted according to the reverberation time, the distance between the listener and the central sound source location on stage, and across frequencies. Information on gender, age, and degree of music education was also collected to determine the effects of individual differences between subjects in the experiment.

Chapter 2

Literature Review

Location perception of sound sources is a complex process involving not only physical elements but also psychological parts. This process is often referred to as "localization", which means to judge the direction and distance of a sound source [5].

This chapter provides a basic understanding of terminology used in the thesis, and introduces the previously used metrics for spatial impression. Finally, by reviewing related studies, it explains the background to the experiments conducted in this thesis.

2.1 Fundamentals

"Binaural" refers to situations where the sound is delivered to both ears and "diotic" means the stimulus arriving at both ears is identical. If it is different, it is called "dichotic". Blauert [6] classified spatial hearing as the "sound event" which is the acoustic stimulus and "auditory event" which is perceived auditorily. The "auditory event" includes a perceiving system and a describing system. He said that only the person listening to the "sound event" can observe the perceiving system output.

The direction of the sound source is defined relative to the head (Figure 2.1). The description of the angular perception is "azimuth" and "elevation". Both of them are described in terms of degrees, where zero degrees is often considered to be directly ahead of the listener.

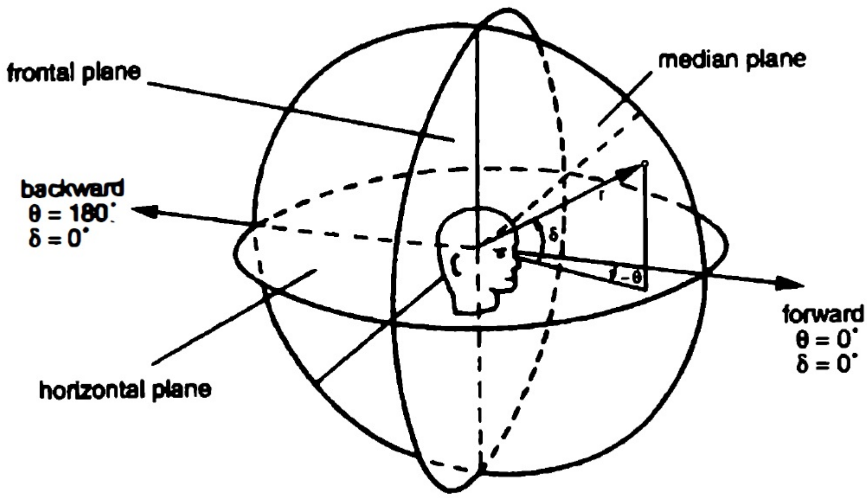


Figure 2.1: Illustration of the coordinate system to define the position of sounds relative to the head (Blauert) [6]

The azimuth is given by the angle θ (Turn clockwise), elevation is given by the angle δ (positive for upward), and r is the distance between a source and listener.

2.2 Localization Cues

2.2.1 Interaural Time Difference (ITD)

Interaural Time Difference (ITD) is a cue for determining the azimuthal position of sounds. It refers to the difference in arrival time between the two ears. This is caused by the distance between the two ears. If the distance to each ear from the sound source is the same, there is no time difference (ITD is zero), but if not, a gap occurs the if azimuth of arriving sound is 90 degrees which are directly opposite one ear, the ITD is typically 690 μs [7].

The ITD can be calculated from the arriving time difference of arrival times between the two ears (figure 2.2, Eq. 2.2.1).

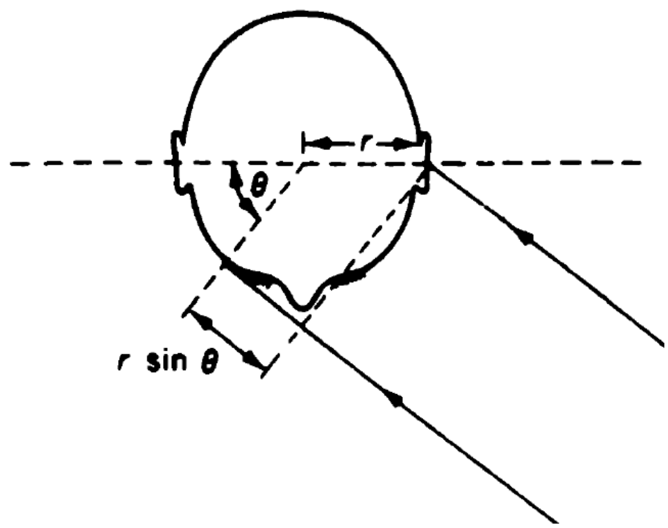


Figure 2.2: Interaural Time Difference (Moore) [7]

$$ITD = \frac{r\theta + r sin\theta}{c} \ldots \qquad \text{(Eq. 2.2.1)}$$

where $-90^\circ \leq \theta \leq +90^\circ$, θ is azimuth, r is the radius of the head, and c is speed of sound [8].

Figure 2.3 plots ITD as a function of azimuth [9].

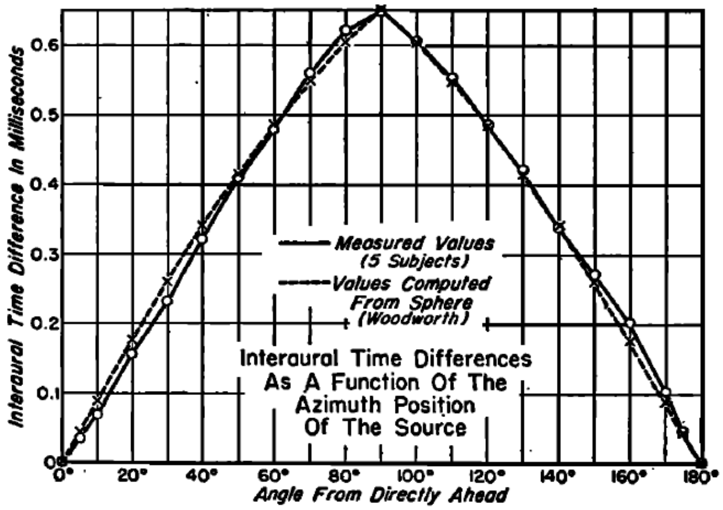


Figure 2.3: ITDs plot by azimuth degrees (Feddersen) [9]

2.2.2 Interaural Level Difference (ILD)

Many scientists in the past have found that the Interaural Level Difference is the only or most important signal parameter in the lateral direction [6].

ILD compares the sound levels at the two ears. At high frequencies, the right ear has a higher level than the left ear when the sound comes from the right side because the head makes a shadow zone. But at low frequencies, the difference is small [10]. ILD consequently is related to the signal frequency and the arrival angle of sound source θ (Figure 2.4).

The following Eq. 3.2.1 is the basic formula for ILD.

$$ILD(r,\theta,\phi,f) = 20\log_{10}\left|\frac{P_R(r,\theta,\phi,f)}{P_L(r,\theta,\phi,f)}\right|(dB)\ldots \qquad \text{(Eq. 2.2.2)}$$

where $P_R(r,\theta,\phi,f)$ and $P_L(r,\theta,\phi,f)$ are the frequency-domain sound pressures at the left and right ears, respectively, generated by a sound source at (r,θ,ϕ) where r is distance, ϕ is elevation, θ is azimuth, and f is frequency [11].

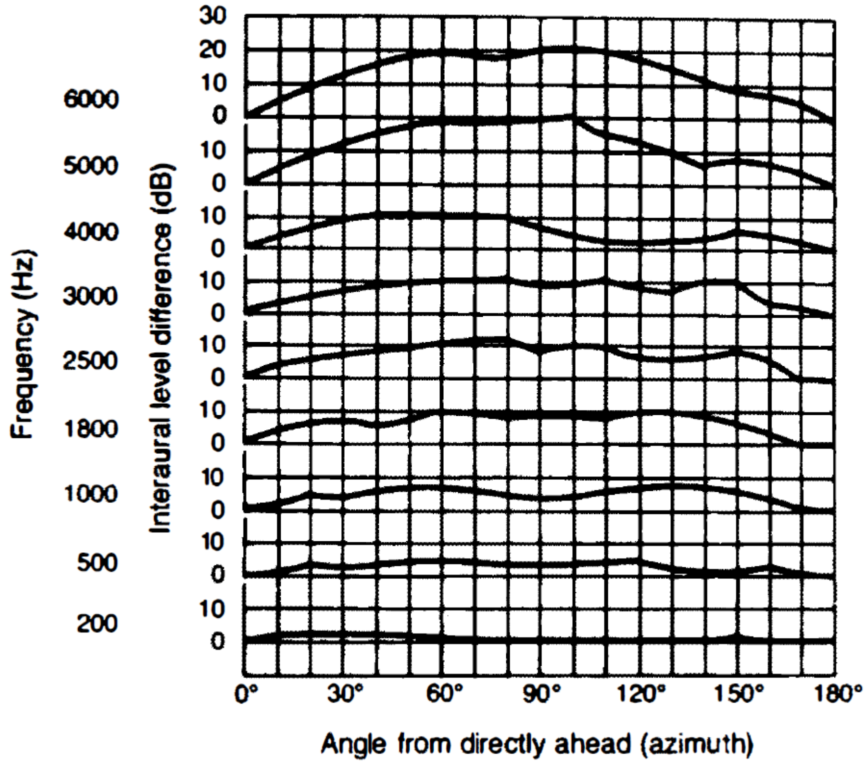


Figure 2.4: ILDs for sinusoidal stimuli plot by azimuth degrees (Feddersen) [9]

ILD may not be similar even though sound sources come from the same direction because of the human body (eg, torso, pinnae, and body etc.) [12], [13].

2.2.3 Head-Related Transfer Function (HRTF)

The pinna, torso, and head affect sound transmission. The degree of the influence varies depending on the frequencies. The Head-Related Transfer Function (HRTF) is an acoustic transfer function that has information about the filtering effect of the anatomical structures for the human and is very individual dependent. It is defined by The difference between the intensities of the measured far-field frequency response when a small microphone is placed in the individual's left- or right-ear, and the intensities of the sound source which measured with a microphone at the center of the head but the head absent. Accordingly, HRTF is dependent on the direction of sound incidence [14]. Figure 2.5 shows the magnitude of the HRTF at various azimuths and elevations.

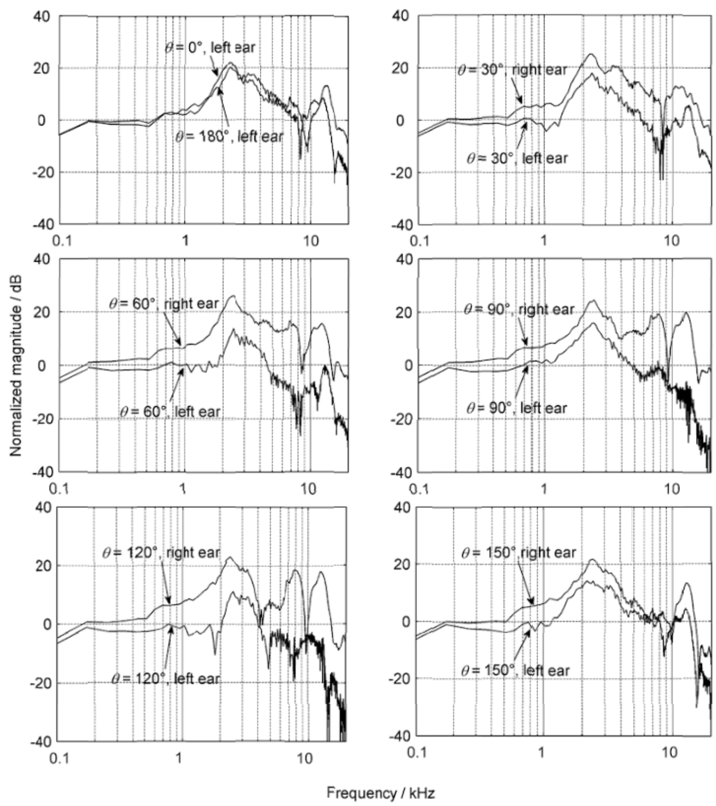


Figure 2.5: Magnitudes of KEMAR HRTFs at various azimuths in the horizontal plane (Xie) [11]

An artificial head simulates the human's anatomical structures. HRTFs are measured by a microphone mounted in the fixed radius to the left or right ears of the artificial head or a human. Measured binaural sound includes the spatial information which is the directional localization cues (ITD, ILD and etc.) for sound sources.

The standard artificial head, KEMAR, was introduced by Burkhard and Sachs in 1972 [15]. It is designed with an average head that represents an "average listener". It is most commonly used because it looks like a human being and easily accessible to the data base.

2.3 Spatial Impression Parameters

One of the aspects of Spatial Impression, "Spaciousness", is widely considered to encompass both Apparent Source Width (ASW) and Listener Envelopment (LEV) [16]. ASW describes the perceived width of the sound source image, while LEV is the subjective feeling that the listener is surrounded by the sound field. The effect of early reflections and later reflections on spatial impression was studied by several people [17], [18], [19], [20]. The ASW is determined by the sound energy received by the listener up to 80 ms after the direct sound in an enclosed space [21], while LEV depends on late arriving sound energy after 80ms.

2.3.1 Lateral Fraction (LF)

After Marshall discovered that early reflections arriving from lateral directions are significant for spaciousness [22], Barron and Marshall derived Lateral Fraction (LF), a linear measure of spatial impression [2]. From subjective listening tests using simulation systems, they found that when the sound arrived at the listener, totally from a lateral direction, spatial impression is maximized, while when the sound arrived from the frontal direction, it was minimized.

LF is the ratio of lateral energy to total initial energy ($0 \sim 80$ms).

$$LF = \frac{\int_{0.005}^{0.08} p^2(t) cos^2\theta dt}{\int_{0}^{0.08} p^2(t)\theta dt} \ldots \qquad \text{(Eq. 2.3.3)}$$

2.3.2 Interaural Cross - Correlation (IACC)

Interaural Cross-Correlation (IACC) is commonly used as a measure of ”spatiality”. Yoichi Ando proposed IACC in his book ”Architectural Acoustics” in 1988 [23].

The following Eq. 2.3.4 is an IACC formula.

$$IACF_{\tau} = \frac{\int_{t_1}^{t_2} P_L(t)P_R(t+\tau)dt}{\left|\int_{t_1}^{t_2} {P_L}^2(t)dt \int_{t_1}^{t_2} {P_R}^2(t)dt)\right|^{1/2}} \ldots \qquad \text{(Eq. 2.3.4)}$$

$$IACC_t(t) = |IACF_t(\tau)|_{max}$$

$IACC_E$= measure of the Apparent Source Width (0 ∼ 80ms)

$IACC_L$= measure of Listener Envelopment (80 ∼ 1,000ms)

In IACC, ASW and LEV are separated by cutoff time. When the cutoff time(Eq. 2.3.4) of IACC is 0 ∼ 80ms, it indicates ASW, while using 80 ∼ 1000ms, it indicates LEV. ASW increases when the correlation of signals reaching both ears decreases, or when IACC values are close to zero [24].

2.4 Related Studies

Research on the relationship between ASW and the spatial impression metrics has been going on for a long time. Barron and Marshall discussed the effect of early lateral reflections on the spatial impression [2]. They found that sound energy in the 125 to 1 kHz octave bands are important to ASW, especially low frequencies. Morimoto and

Posselt [25] carried out a test to study the relationship between early lateral energy and reverberant energy. As a result of the experiment, reverberation contributes to the spaciousness as many early reflections are added. Morimoto and Maekawa [26] have tested the effect of low frequency on ASW at 100 to 5.3kHz, and found that the effect on ASW is great when frequency lower than 510Hz is removed. Bradley found that there was a significant change in all measurements considered as a 30 cm displacement in enclosed space [27].

Toshiyuki Okano et al. [21] showed how IACC and LF are related to ASW. They examined the effect of the low-frequency strength of the source signal on ASW for symphony music.

In the experiment, 12 speakers were installed in the anechoic chamber in a hemisphere arrangement, and then the music signal was played and listeners were asked to respond on the ASW experienced. From comparing the ASW response to the LF and IACC values, LF was found to be a poor representative of ASW at high frequencies (above 500 Hz), while $|1 - IACC|$ was proportional to the ASW response (figure 2.6). They also found that the frequency components lower than 355 Hz affect ASW much more than those higher than 355 Hz.

The effect of high frequency on ASW has been studied by Morimoto and Iida [28]. They studied the effect on ASW by changing the cutoff frequency to 200Hz to 8kHz wide-band noise. The results show that frequency components above 1 kHz do not affect ASW. Signals higher than 1 kHz are consequently not used in the current thesis as the research is interested in ASW and source localization.

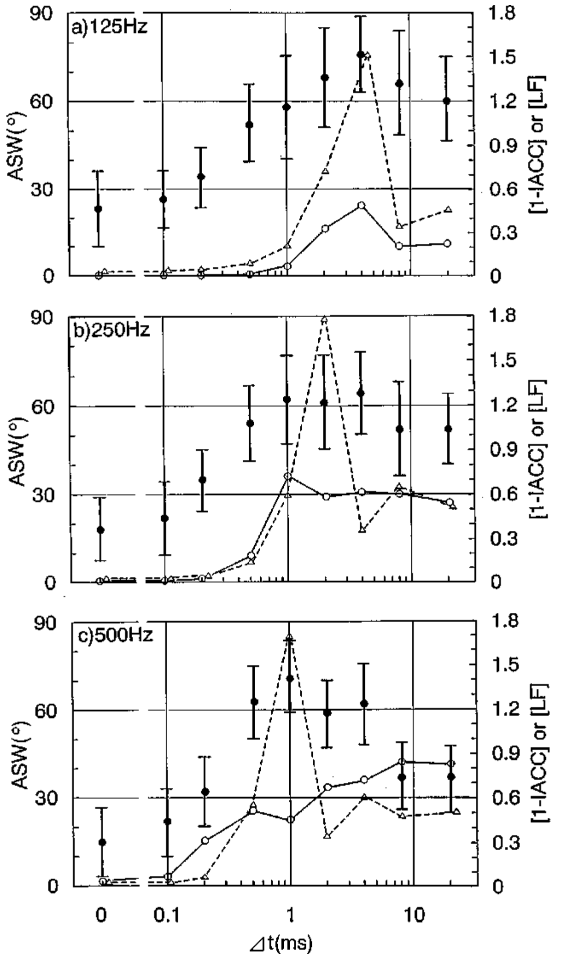


Figure 2.6: ASW (solid circles), $|1 - IACCE|$ (open circles), and LF (open triangles) versus the delay difference between two 90 degree lateral reflections (Toshiyuki) [21]

Other research on the perception of ASW includes, Pollack et al. [29]; in their study of Interaural Noise Cross Correlation and Binaural listening, they reported that the just noticeable difference (JND) increases when the correlation is low between energy delivered to both side ears. In an ASW's study using musical signals, Reichardt et al. [30] found the JND for LF is from 0.06 to 0.09 when LF is in the range of 0.2 to 0.4. Cox et al. [31] found the JND to be about 0.6 for a simulated sound field in an anechoic room convolved with music. Vries [32] found that a small microphone position change in the measurement caused the fluctuation of the results.

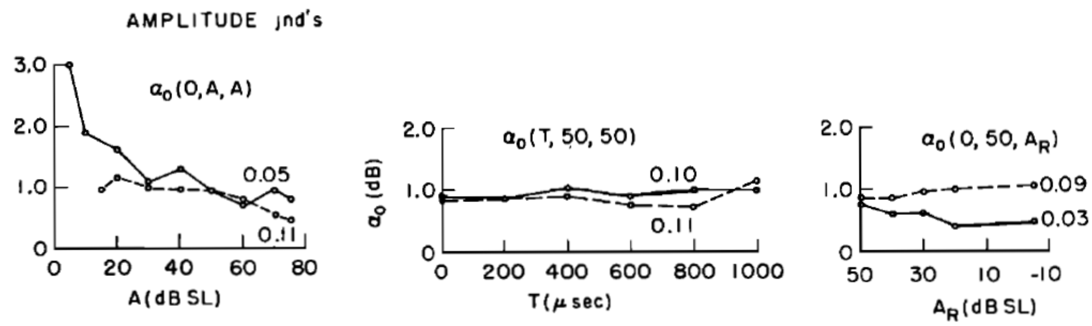


Figure 2.7: JND for the interaural level differences at 500Hz (Hershkowitz) [33]

Auditory research on Interaural Level Difference (ILD) has also been carried out in various ways. Mill studied that the Just Noticeable Difference (JND) for ILD is 0.5 to 1 dB [34]. After then, Hershkowitz performed more JND experiments on ITD and ILD at 500 Hz (Figure 2.7) [33], McFadden tested 250Hz [35], and Wesley conducted at 1000Hz [36]. Yost and Dye [37] measured JNDs for ILD in the range of 200 $\sim$ 5000Hz. The auditory test used the 2AFC method. As a result, JNDs were found to be approximately 0.70 at 200Hz, 0.80 at 500Hz, 1.10 at 1000Hz, 0.65 at 2000Hz, and 0.73 at 5000Hz (reference at ILD=0dB, Figure 2.8). They used the pure tone and the duration was 250ms, gated simultaneously in the left and right with a 10ms $cosine^2$ rise/decay.

However, it is not clear if these JND values are applicable to sound in enclosed spaces which have diffuse reflections because the former experiments were conducted in a free-field. JND and localization in the reverberant environment have been studied by many people [38], [39], [40]. Klockgether [41] studied how ITD and ILD perception

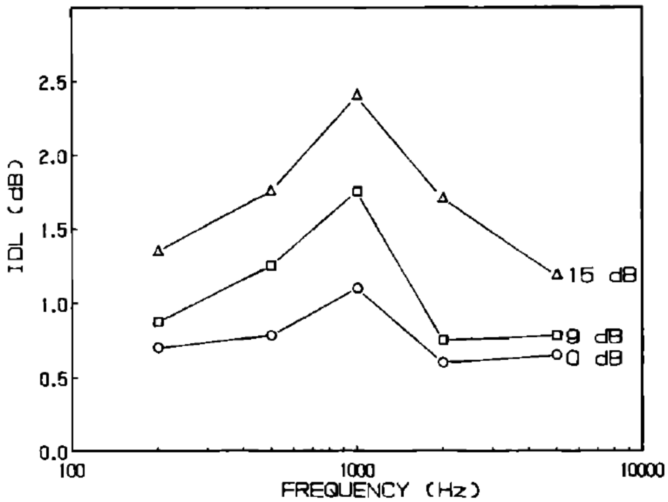


Figure 2.8: JND for ILD in decibels (Yost) [37]

are affected by reverberant spaces. The author investigated the JND of ITD and ILD using the alternative forced choice(AFC) measurement method using Binaural Room Impulse Response (BRIR) manipulation. Experimental results showed that ITD and ILD have increased JND in a reverberant space than in an anechoic space (Figure 2.9). According to the characteristics of the sound source, ITD showed a JND of 2 to 8 times higher. It was analyzed that reverberation had about half the effect on ILD than on ITD.

2.5 Summary

This chapter has covered background regarding fundamentals of spatial impression, the localization cues, and spatial impression parameters, and previous research on ASW and JND for spatial metrics. This thesis reviews existing spatial metrics and suggests a new evaluation metric for the spatial impression of multiple sources in

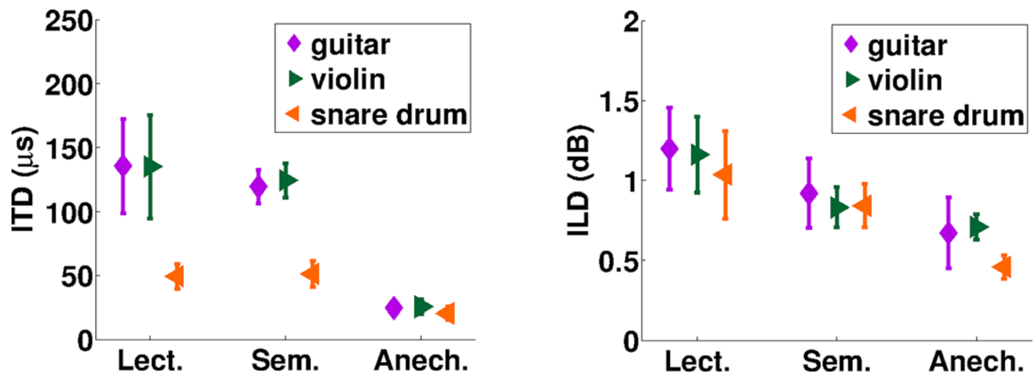


Figure 2.9: Just noticeable difference for the interaural differences (ITD, left panel; ILD, right panel) measured for three different rooms (lecture hall:1.7s, seminar room:0.9s, anechoic) and three different instruments (Stefan) [41]

enclosed spaces.

For the spatial perception evaluation of multiple sound sources, the newly proposed metric is based on the results of previous studies and studies the relationship between the Apparent Source Width (ASW) of a single sound source and the localization of the sound source. To this research, ILD, one of the localization cues, is used as the basis of the new metric. It also conducts auditory experiments to verify the relationship of human perception in reverberant spaces.

Chapter 3

A study of the relationship between ILD and varying sound source location on stage

3.1 Introduction

Since existing methods of evaluating spatial impression (ex. IACC, LF), are based on the measurement of a single sound source, it is focused on changes in spatial impression according to the change of listener position rather than sound source position.

This chapter explores the changes in the value of existing spatial evaluation metrics according to the location of sound sources. How much the existing spatial evaluation metrics (IACC, LF) reflect the change of sound energy due to the varying sound source positions using acoustic simulation is presented first from an acoustical simulation study. The auralization was generated by installing listeners and sound sources in fixed positions on two different spaces.

This chapter proposes subsequently a new metric named ILD-Correlation Range (ILD-CR) to evaluate spatial impression according to the sound energy variation from varying sound source position. To verify whether the new metric (ILD-CR) reflect the positional variation of the sound source, how the values differ depending on the source location using acoustic data measurement in three different types of space is observed and the relationship between the shape of the space and the changing of ILD is analyzed.

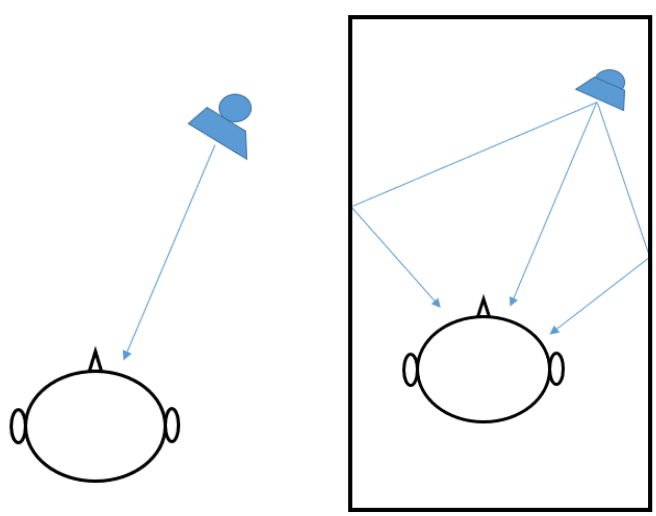
3.2 Methodology

3.2.1 Spatial impression metrics

Typically spatial impression metrics have focused on the relationship between the single source and the listener, quantifying how a stationary source is perceived at the listener's location. This is insufficient to explain the perceived change in spatial impression due to a varying sound source. To represent this, an acoustic metric that can show the change of the energy of the position due to the sound source is needed.

To implement the metric is needed information on energy variation due to source position. Interaural Level Difference(ILD) as the difference between the energy delivered to the left and right ears of a defined receiver provides the information to determine the location perception of sound sources [42]. In this experiment, to compare the energy of the sound source delivered to the listener, the method, cumulating the left and right energy for a certain period of time ($0 \sim 80$ms), is used. In the free field, the sound energy from the same sound source is transmitted in one direction, but in the diffuse field, the direction of the transmitted sound varies due to the reflection sound (Figure 3.1). This method reflects the effect of the early reflections on the listener's perception because the direction of the signal transmitted over time varies depending on the type of space.

The r correlation coefficient between ILD values and sound source position, represented by energy variation [43]. The range of the ILD values across a number of source positions provides a measure of the degree of energy variation. The proposed metric (ILD-CR) is taken as a product these two values.



(a) (b)

Figure 3.1: Sound transmission path : (a) Free field, (b) Diffuse field

The following is the formula for each parameter.

(a) Interaural Level Difference (ILD)

- The Interaural Level Difference (ILD) represents the difference in acoustic energy received at the two ears (left & right).

$$ILD = 10 \log \frac{\int_0^t P_R^2}{\int_0^t P_L^2} \qquad \text{(Eq. 3.2.1)}$$

where P_R is presented at right ear sound energy, and P_L is at the left

(b) r (Linear coefficient)

- r is the degree of linear relationship between ILD and the source position on the stage.

$$r = \frac{\sum(s_i - \bar{s})(ILD_i - \overline{ILD})}{\sqrt{\sum(s_i - \bar{s})^2}\sqrt{\sum(ILD_i - \overline{ILD})^2}} \qquad \text{(Eq. 3.2.2)}$$

where s_i is distance between the stage center and source location

$\bar{s}$ = mean of all s_i, $\overline{ILD}$ =mean of ILD, -1 < r < 1

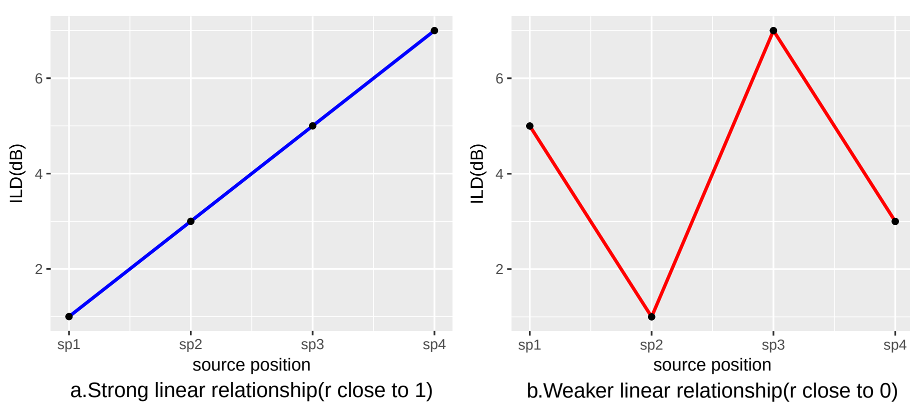


Figure 3.2: Linear relationship example

(c) Range of ILD

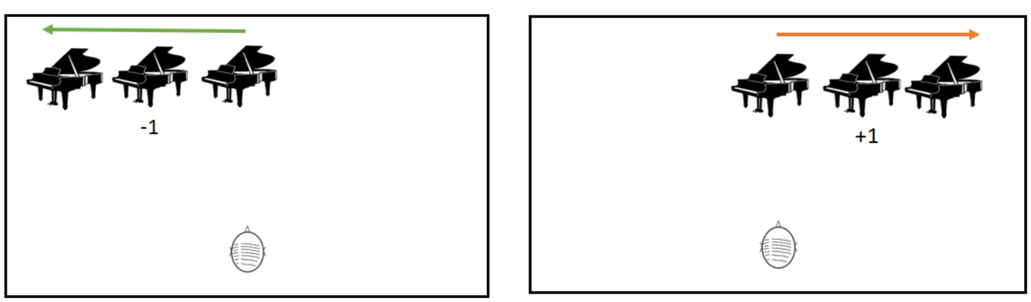
- It indicates the range of ILD across a number of source position on the stage that are being considered

$$\text{Range Of ILD} = \text{Max Of ILD - Min Of ILD} \qquad \text{(Eq. 3.2.3)}$$

(d) Direction of Source Movement (DSM)

- If the direction of the sound source movement is left bound from reference the source, the value is -1, If right, the value is +1 (Figure 3.3)

$$\text{Direction of Source Movement (DSM)} = \text{Left: -1, Right: 1} \qquad \text{(Eq. 3.2.4)}$$



(a) (b)

Figure 3.3: (a) DSM '-1', (b) DSM '+1'

(e) ILD Correlation Range (ILD-CR)

- The ILD-CR is calculated by multiplying the three metrics : r, the Range of ILD, and the DSM. If the ILD-CR value is negative, it means that the sound source movement and the increasing direction of the ILD do not coincide. when the correlation between the ILD value and the movement of the sound source is low (close to zero), the value of ILD-CR decreases by the value of r, even if the range of the ILD is large.

$$\text{ILD-CR} = \text{ILD x } r \text{ x Range Of ILD x DSM} \qquad \text{(Eq. 3.2.5)}$$

3.2.2 Acoustic simulation

Since stage performances are often used multiple sound sources rather than single sound source, it is important to understand how audiences perceive multiple sound sources in spatial impression on a stage.

The acoustic simulation experiment was examined how much the existing evaluation method contains information about the energy variation transmitted to the listener for multiple sound sources. The existing evaluation method used $IACC_E$ and LF, which are representative measurement criteria for spatial impression evaluation, and compared with the proposed ILD-CR.

A sound source set was generated by acoustic simulation program which is ODEON (ver. 11, Denmark) and Subject21 HRTF, based on the 'standard' Kemar head with blocked ear canal and normal sized pinnae. The spaces used in the signal generation had two different shapes which were rectangular and fan.

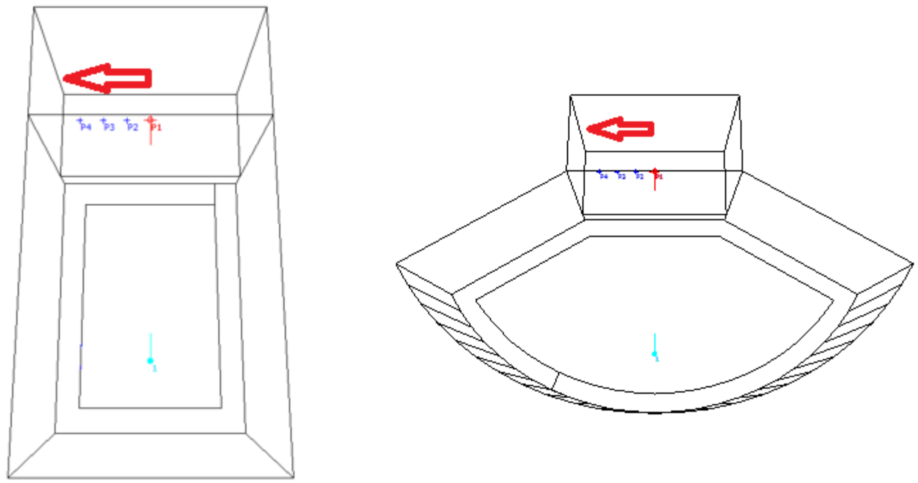
The size of rectangular space is 15 m long, 8 m wide and 7 m high, the volume is 840 m^3. The stage size is 3m long and 8m wide. The average sound absorption (NRC α) of the surface is 0.13. The fan-shaped space is 15 m long, 22 m wide, 7 m high, and the volume is 1380 m^3. The stage size is same as a rectangle model (3m x 8m). The average sound absorption (NRC α) is 0.16. Table 3.1 is the size of spaces used for modeling.

Shape	Size	Volume	Surface area
Rectangular shape	15m x 8m x 7m (L x W x H)	840 m^3	554m^2
Fan shape	15m x 22m x 7m (L x W x H)	1380 m^3	823m^2

Table 3.1: Room size (stage size is same)

The four sound sources were located in across the stage and 1m interval from the center towards stage right (DSM '-1'). In order to obtain more accurate results, the installation interval of sound source was set to the minimum installation interval (1m) allowed by the simulation program. In addition, it considered the difference of the sound level by distance. Receivers were centrally located at 10 m distance from the center sound source as shown in Figure 3.4 (a),(b).

The signals were analyzed by the Matlab program to obtain the LF, $IACC_E$, and ILD-CR values.



(a) Rectangular Shape (b) Fan Shape

Figure 3.4: Spaces used for modeling (Not drawn to same scales)

Table 3.2 shows the set ups applied to the simulation (Odeon).

Impulse Response Length	2000ms
Number of late rays	100,000
Transition Order	2
Number of early scatter rays (per image source)	100
Apply NC curve	NC15

Table 3.2: Simulation program(Odeon) setup

3.2.3 Measurement in Existing Spaces

In acoustic simulation experiment, the proposed ILD-CR was compared with the existing acoustic measurement method. However, since the data generated by the simulation does not apply all the conditions of the actual space, it is necessary to analyze what differences exist in the actual space by measurement. In this experiment, the value of ILD-CR was calculated in the other three shape of spaces and analyzed the relationship between ILD-CR value and space shapes.

3.2.3.1 Microphone Technique

A dummy head microphone is typically used to measure ILD. In this experiment, two cardioid microphones placed 180 ° apart were used to measure the sound energy arriving laterally at the listener position. This method does not include head related transfer functions, that is, the influence of perceiving variations between individuals, and allows comparison of variations in each space by measuring the lateral energy before being received to humans. The measured data from the two cardioid microphones were used in the analyzed using the polar pattern to minimize superposition of sound and to separate the sound from the left and right (figure3.5).

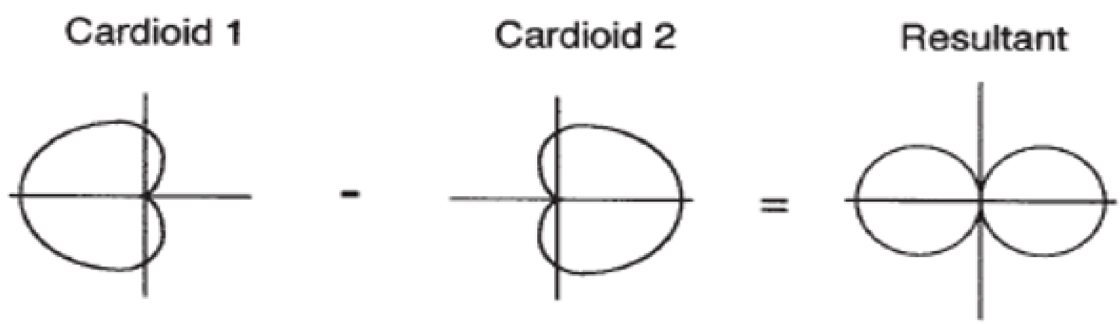


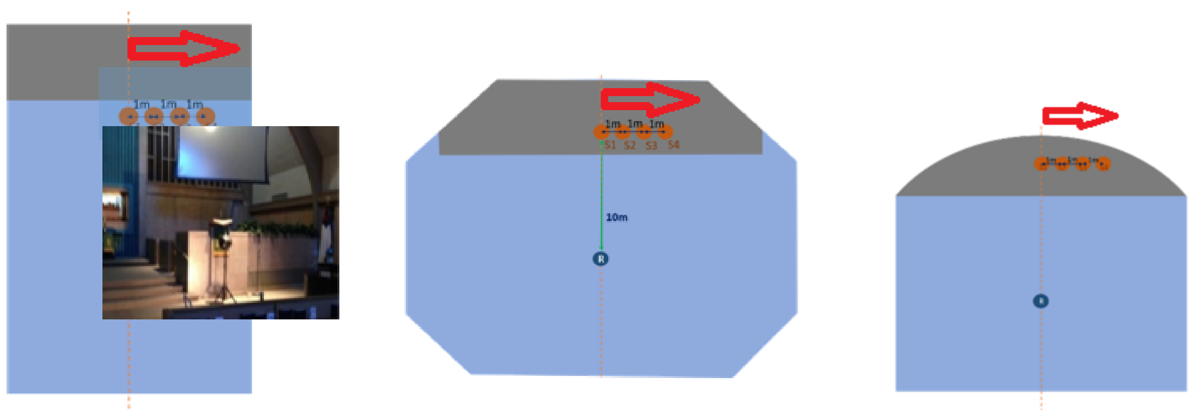
Figure 3.5: Results from Combining dual back to back cardioid microphones (John Eargle) [44]

3.2.3.2 Measurement Procedure

Acoustic measurements were carried out in three spaces of different shapes. Four sound source locations were used at intervals of 1m from center towards stage left

(DSM +1), as opposed to being used in the simulation presented in section 3.2.2, and the distance between the center sound sources and the receiver position was 10m as shown in Figure 3.6.

The swept sine method was used to measure impulse responses at the two cardioid microphones, using EASERA software on a laptop. The speaker used for the sound generation was BAS001 Omnidirectional speaker by Larson Davis accompanies by the amplifier. The microphones were Behringer's C-2 cardioid condenser microphones, and the audio interface used RME's babyface connected to the notebook.



(a) (b) (c)

Figure 3.6: Three types of measurement places : (a) Church1, (b) Church2, (c) Lecture room (Not drawn to same scales)

The first space was a rectangular shaped church (Church1, 3.6 (a)). The reverberation time was 1.4 seconds(mid range Figure 3.7), and there are protruding side walls on both sides of the stage(Figure 3.6 (a)).

The second space was a symmetrical elongated octagon shaped church (Church2, 3.6 (b)). The reverberation time was 0.8 seconds (mid range Figure 3.7).

And the third space which was lecture room (Lecture room, 3.6 (c)) had 0.9

seconds (mid range Figure 3.7) reverberation time, and it had a rectangular shaped audience area and a hemispherical shaped stage.

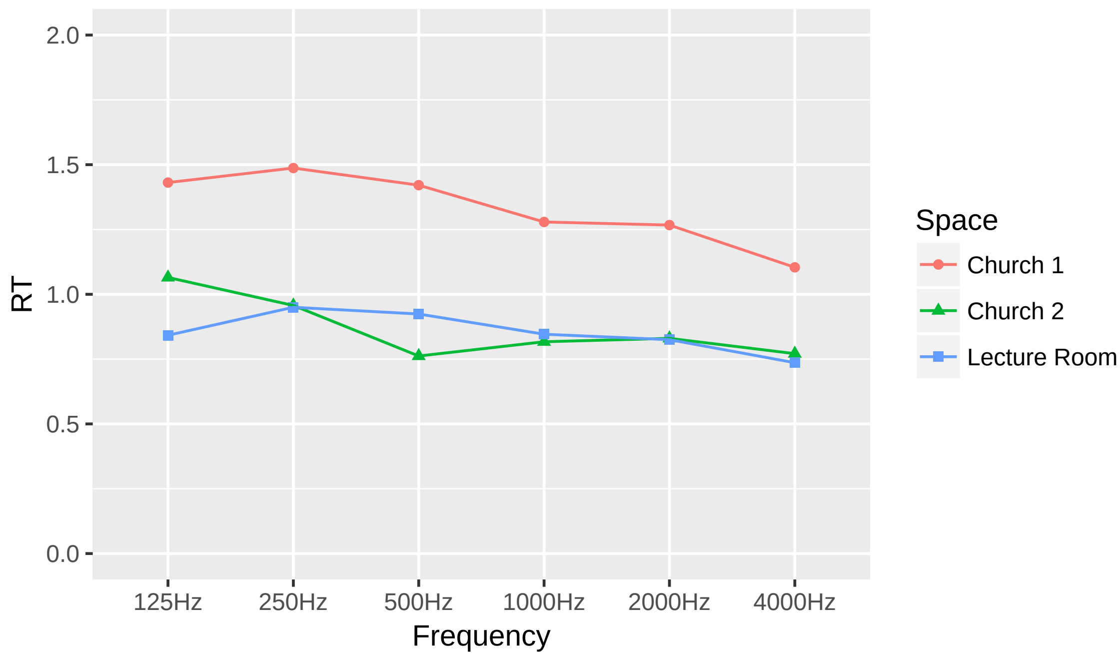


Figure 3.7: Measurement space Reverberation Time

3.3 Results

3.3.1 Acoustic Simulation Result

The following is an analysis of the acoustic simulation results. Figure 3.8 compares the Lateral Energy Fraction(LF) values of the two spaces. The X-axis represents the position of the sound source where SPL is center stage, and the others are each 1m to the audience right and the Y-axis represents the LF value.

LF means that the Apparent Source Width (ASW) is narrow when the value is smaller, and wide when the value is larger. In Figure 3.8, The difference of the LF values between the spaces is clear, but the difference of the LF values between the positions of the sound sources is not large. This means that even if the position of the sound source changes, the ASW change is not large. Since it does not have the information of the change of the direction or position of the sound source, LF is not sufficient to show the spatial impression of the entire stage.

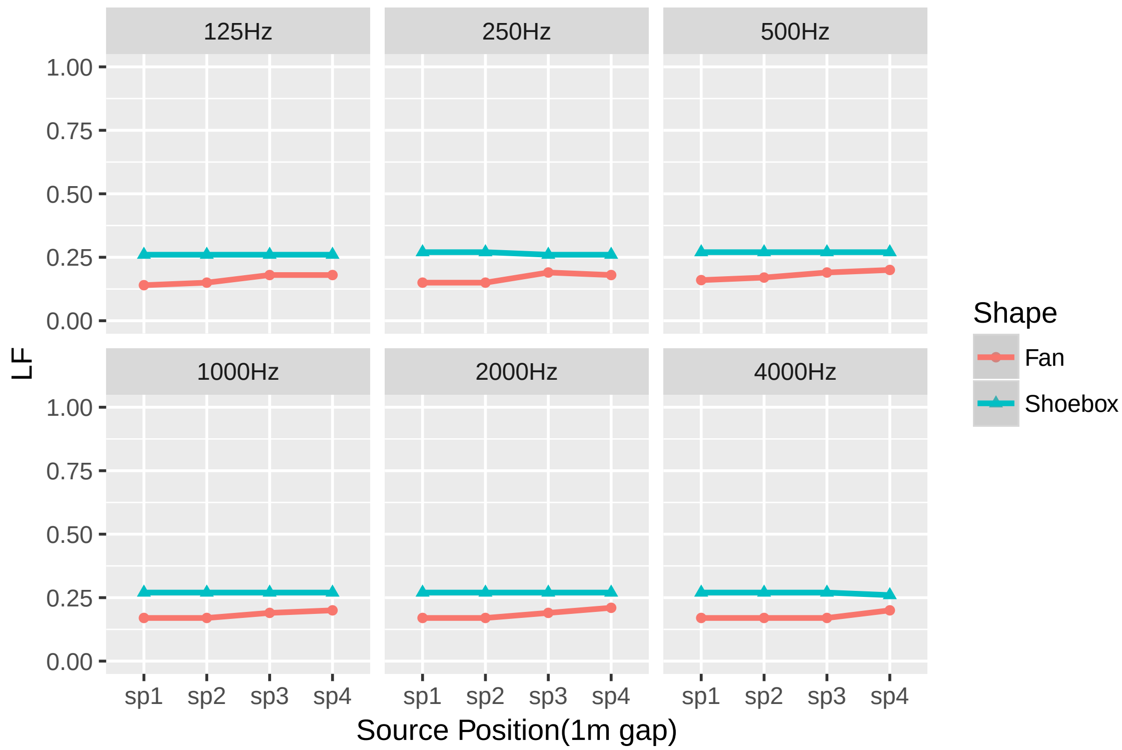


Figure 3.8: LF values from acoustic simulation data across octave bands

Figure 3.9 compares $IACC_E$ values. The X-axis represents the position of the sound source, and the Y axis represents the $IACC_E$ value. $IACC_E$ shows different values according to the location of the sound source. Like ILD, $IACC_E$ uses the difference between sound energy received from left and the right.

$IACC_E$ means that the Apparent Source Width (ASW) is narrow when the value is close to 1, and when the ASW is close to 0, the ASW is wide. In Figure 3.9, the Shoebox shape, 250Hz plot, it shows $IACC_E$ value decreases from sp1 to sp3, and increases in sp4. This means that when the sound source moves from sp1 to sp3, the ASW becomes wider and from sp3 to sp4, the ASW is narrower. It shows the change in ASW. However, Since a decrease or increase in the $IACC_E$ value does not indicate the direction of the sound source, the information about the location of the source is unclear.

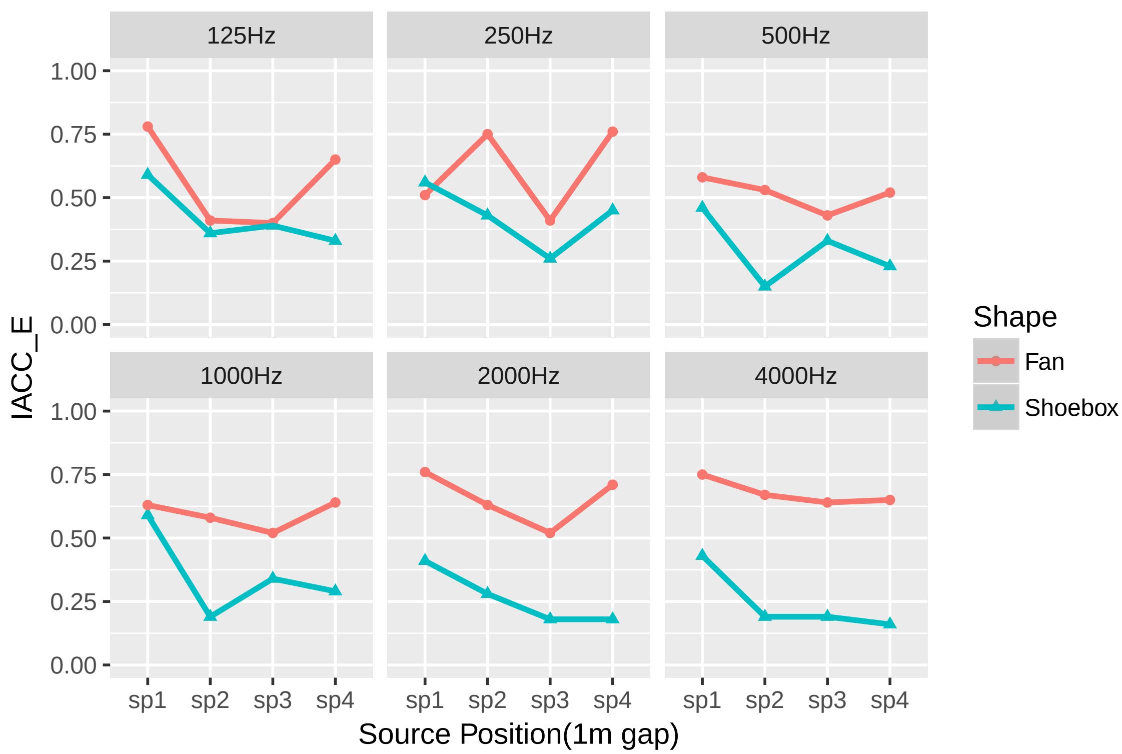


Figure 3.9: $IACC_E$ values from acoustic simulation data across octave bands

The perception of sound source location is influenced by the distance between the listener and source, and the direction of the sound source. $IACC_E$ provides information about the size of the sound source using the ratio of energy delivered to the left and right ears, but it is difficult to determine the relative position between the source and the listener since the direction of sound energy delivery is not clear.

Figure 3.10 shows the ILD values. The X-axis represents the position of the sound source, and the Y axis represents the ILD value(dB). The ILD plot shows the different values according to the position of the sound source. The sound source position sp1 is a case where the sound source is at the center of the stage. In this case, the value of the ILD is close to zero.

The ILD plots give information about the amount and direction of the sound energy delivered from the left and right. If the value of ILD is negative, the sound energy delivered from the left side to the listener is greater than the right, and if positive, the right side energy is greater.

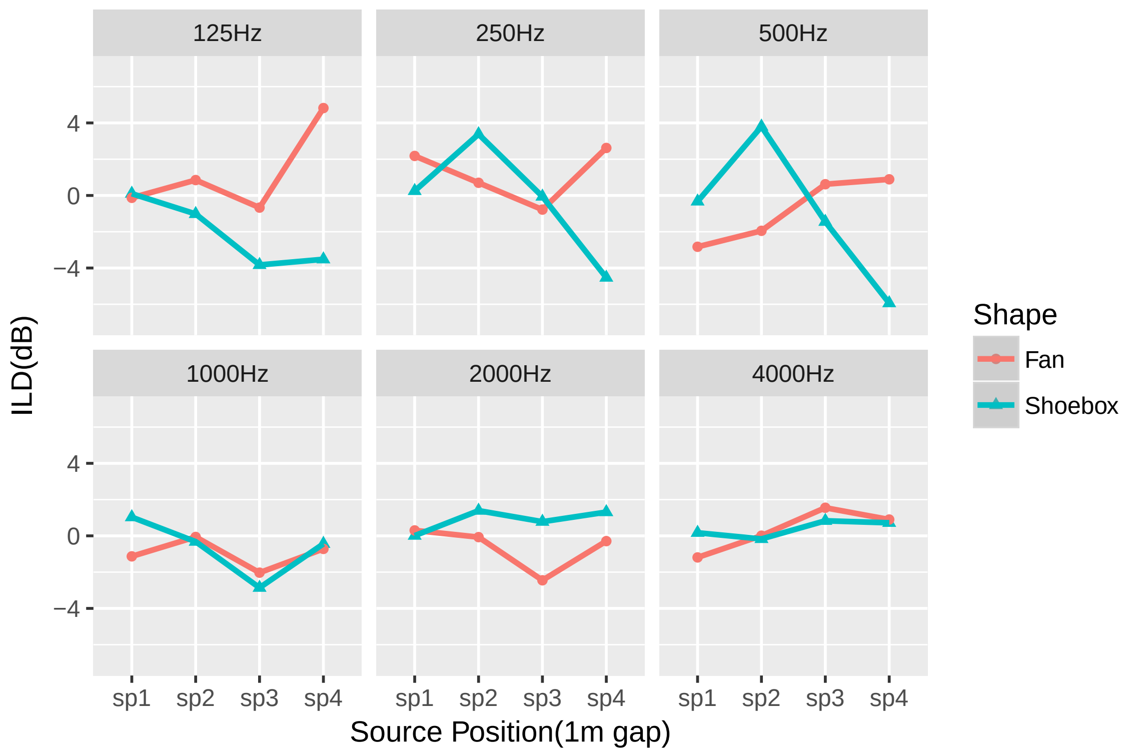


Figure 3.10: ILD values from acoustic simulation data across octave bands

The actual sound source is moving from the center to the left side on stage (Figure 3.4). In this case, the left side of the ILD value is expected to become larger than the right side, but the slopes of the ILD in Figure 3.10 do not always demonstrate a negative slope. In the fan shape and 125Hz plot (Figure 3.10), the ILD value of sp4 is larger than the sp3, and its value is about + 5dB. The sound source located at the extreme left but The sound energy delivered from the right to the listener is greater than the left. This means that the value of ILD at the receiver position is not always proportional to the direction of the source movement and does not match the relative position of the sound source in enclosed space.

In the free field, the ILD has higher values in high-frequency range than low-frequency, but simulation results show that the ILD is close to zero at the frequency range above 2000Hz. For example, in the case, sp4 and shoebox, it has -4 dB value in the low-frequency ranges (125 and 500 Hz), but it is close to 0 dB above 2000 Hz. This is interpreted as having different characteristics in the diffused field.

Next, r, Range of ILD and ILD-CR were calculated using ILD values. The Direction of Source Movement (DSM) value is set to '-1' because the sound source has moved to stage right.

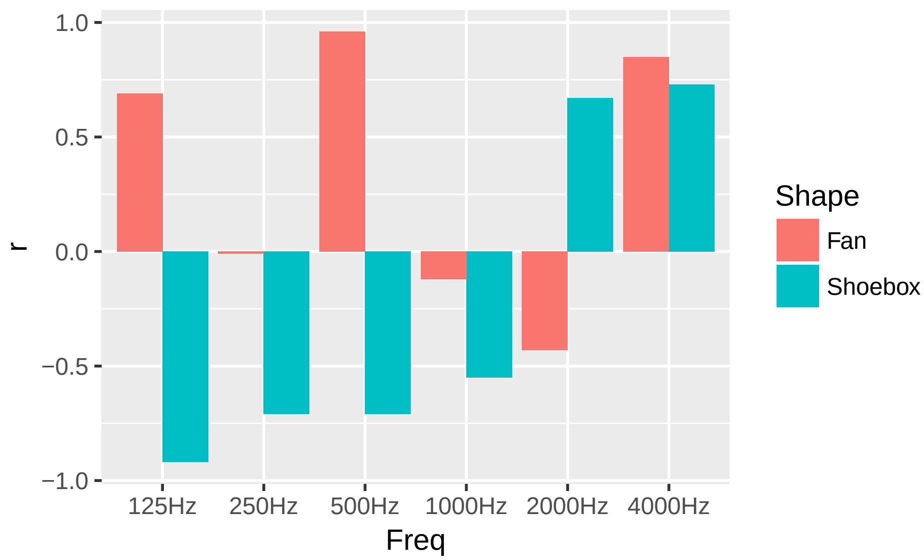


Figure 3.11: r values from acoustic simulation data across octave bands

Since the direction of the sound source is to stage right, the value of r has negative when the direction of the increasing ILD value and the direction of the sound source movement are the same. If the directions are different, the value of r has a positive value. The r which indicates how constant the change of the energy is across source position, and Figure 3.11 shows that the shoebox shape is higher than the fan shape in most of the frequency bands. The fan shape shows a very low correlation in the middle range (250Hz). The r is close to 1 at 125 Hz and 500 Hz, but because it is positive, the direction of the actual sound source and the direction of ILD value do not coincide. At 4000 Hz, both of spaces are in the opposite direction.

The Range (or variation) of the ILD values as the source moves from center to stage right was high in the middle-frequency range for the shoebox shape, and low in the low-frequency and high frequency ranges as shown in Figure 3.12.

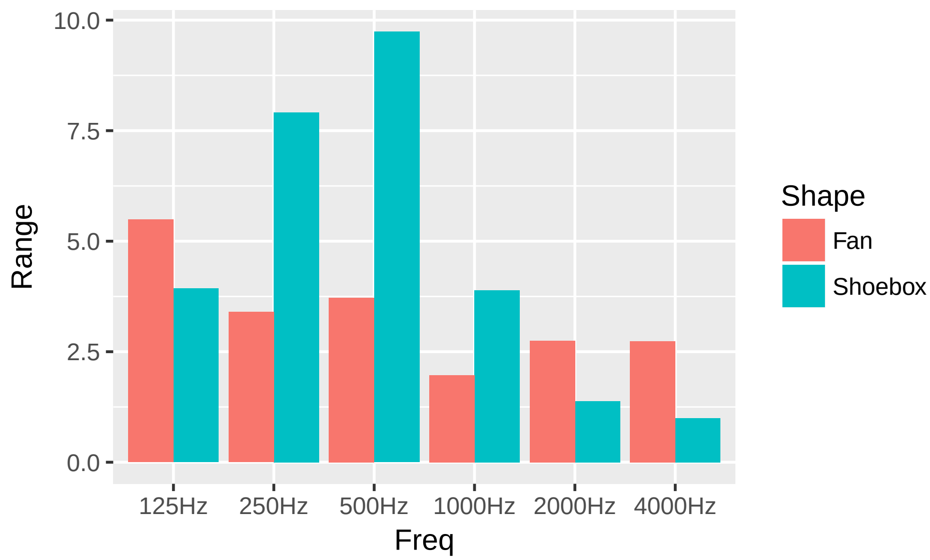


Figure 3.12: Range of ILD values from acoustic simulation data across octave bands

The ILD-CR means if the value has high, it is higher correlated with the change in the position of the sound source and the change in the ratio of the left and right sound energy delivered to the listener and means that the stage has wider spatial impress.

When the ILD-CR value is analyzed figure 3.13, the ILD value, and the sound source movement are proportional to each other, except at the high-frequency range (2000-4000 Hz) in the shoebox shape. However, the fan shape does not show correspondence between the ILD value and the moving direction of the sound sources (negative value), and and the correlation is low. Also, The energy variations transmit to the listeners between the source positions was also small. Therefore, the ILD-CR value was evaluated to be low. Overall, the shoebox shape space was evaluated to have a larger spatial impression than the fan shape without confusion of the sound source positions.

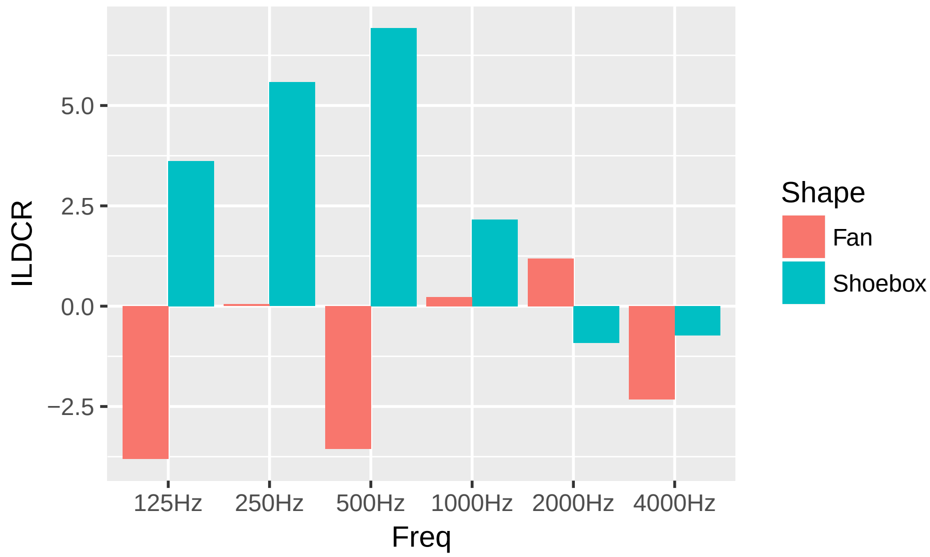


Figure 3.13: ILD-CR values from acoustic simulation data across octave bands

3.3.2 Measurement Results

3.3.2.1 ILD

The measured data were from taken existing spaces using Matlab program. The ILD value was calculated from 0 to 80ms, as shown in Figure 3.14.

(a) Church 1

Church 1 has a rectangular shape and protruding side walls at both sides towards the back of the stage. In figure 3.14, the ILD values are proportional to the sound source positions at sp1 to sp3 because ILD slope has positive when the direction of the sound source is to stage left. However, in sp3 and sp4, the ILD values decrease. It is analyzed that the protruding side walls at the rear stage wall have affected the ILD value. The ILD showed to be sensitive to the shape of the space.

(b) Church 2

Church 2 is a longish octagon shaped space. Above 500 Hz octave band, the moving direction of the sound source movement and the change in the ILD value start to be inconsistent. Above 2000Hz, there is not much change in the left-right energy difference (ILD) value due to the sound source position. The concave shape corner of the space is analyzed as affecting the value. It showed the effect of the space shape is varied in each frequency band.

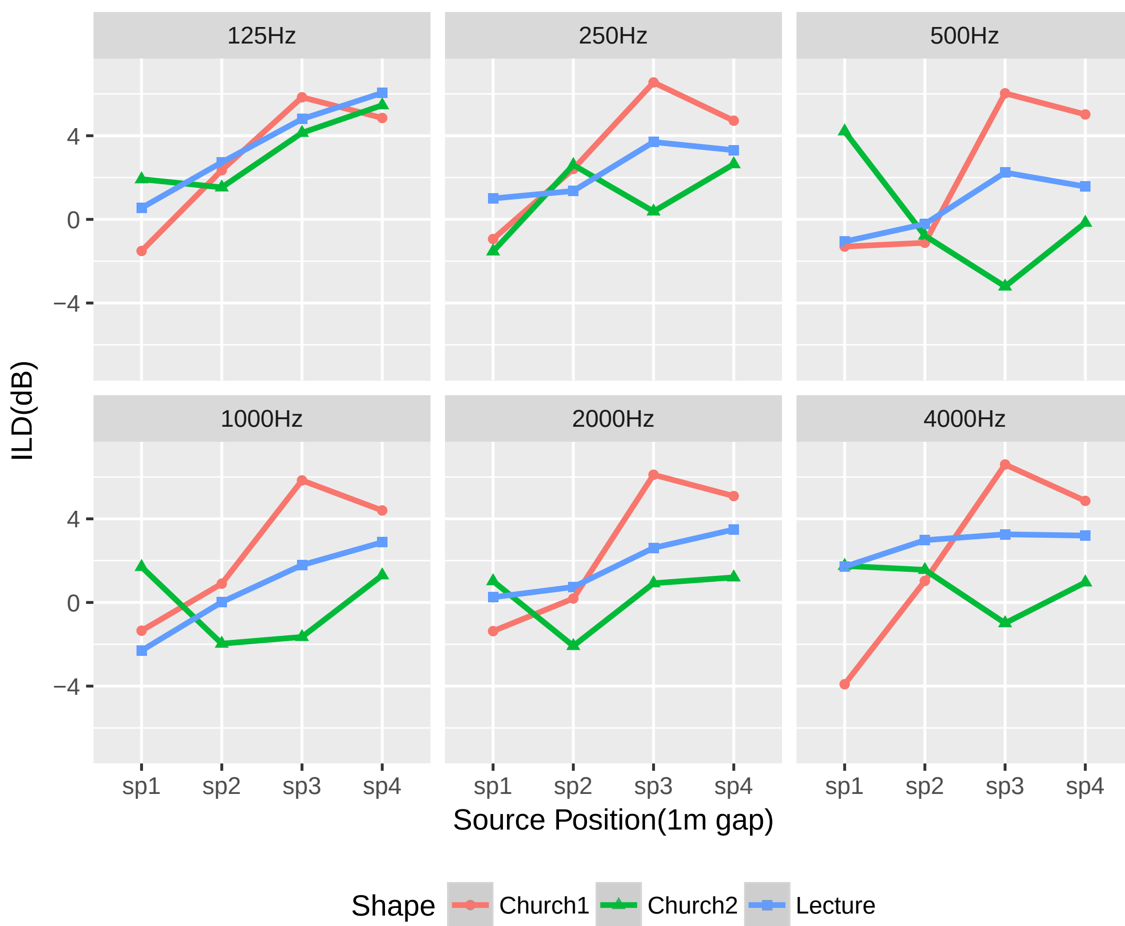


Figure 3.14: ILD values from measurements in existing spaces

(c) Lecture Room

The Lecture Room has hemispherically shaped rear wall to the stage and has a characteristic of getting closer to the wall as the sound source moves outward. The change in the ILD value is proportional to the location of the sound source on the stage as it moves stage left. The slope was analyzed to be gentler compared to the other two spaces. The difference between the ILD values at sound source position source position 1 and source position 4 is not large.

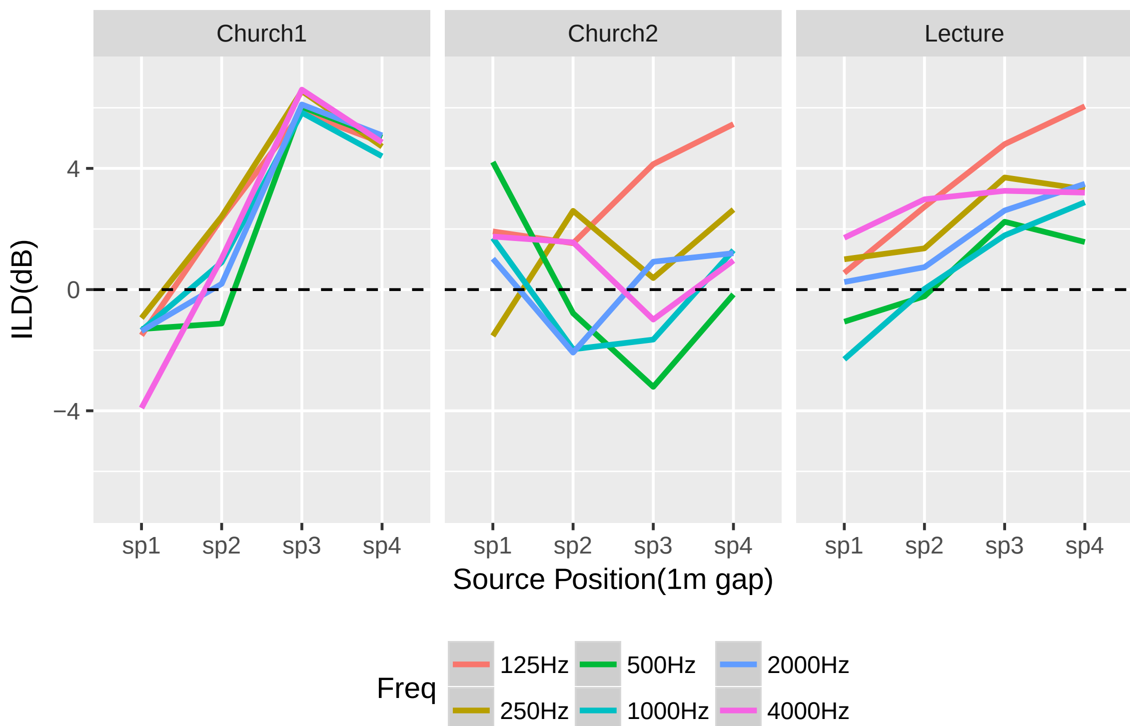


Figure 3.15: ILD value plotted according to the frequency band in each space from measurements in existing spaces

The Figure 3.15 shows the ILD value plotted according to the frequency band in each space. ILD in the free field, higher frequency range value is larger than low frequency. But in this analysis, it does not show the correlation between the frequency and ILD. In the church 1 case, the ILD values at the sp4 location have similar values over the entire frequency bands. In the church 2 case, the ILD value of sp4's position is higher than the 125 Hz and 250 Hz low-frequency band high-frequency band. Int the church 2 case, the ILD values in low-frequency band (125 Hz and 250 Hz) at the sp4 position have higher values than the high-frequency band.

3.3.2.2 r, Range of ILD and ILD-CR

This section shows the result of comparing the ILD range and ILD-CR value of each measured space. The Direction of Source Movement (DSM) value is '+1' because the sound source has moved to stage left.

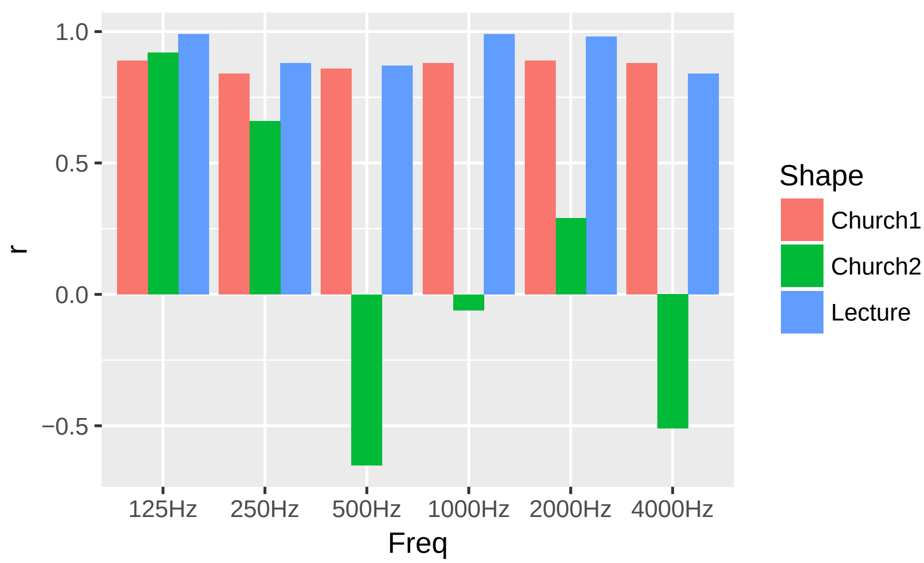


Figure 3.16: r values from measurements in existing spaces

r is higher in the Lecture room, then Church 1, and finally Church 2 (figure 3.16). In the case of the Lecture room, the shifting location of the sound source and the change of the ILD value are proportional and highly correlated with each other across the whole frequency range. In Church 2, the correlation is not high at 500 Hz, and higher the direction of sound source movement and ILD value are inversely proportional. Church 1 shows high r values overall.

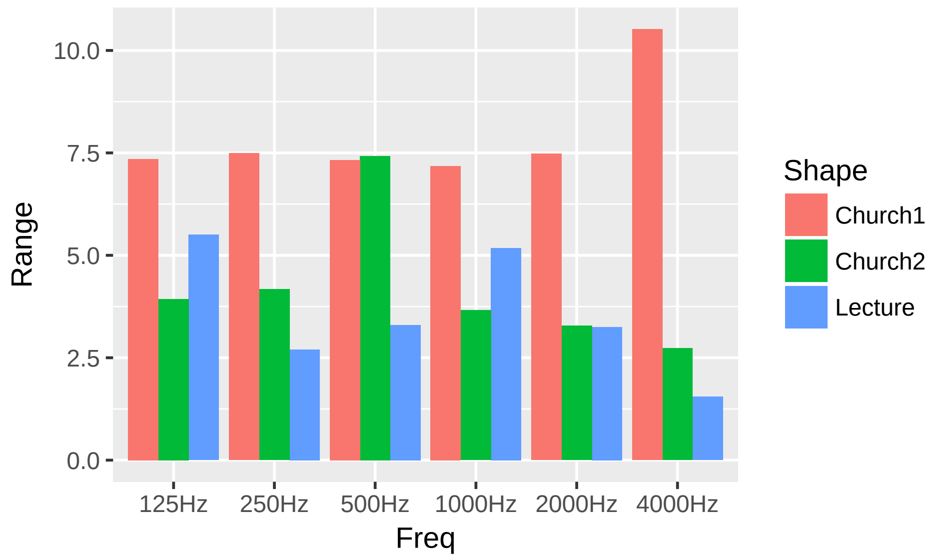


Figure 3.17: Range of ILD values from measurements in existing spaces

The Range of ILD is plotted in Figure 3.17, and is found to be the largest in Church 1 and the lowest in the lecture room. This meant that in the case of Church 1, the difference in ILD values between sp1 and sp4 is largest and in the lecture room is smallest.

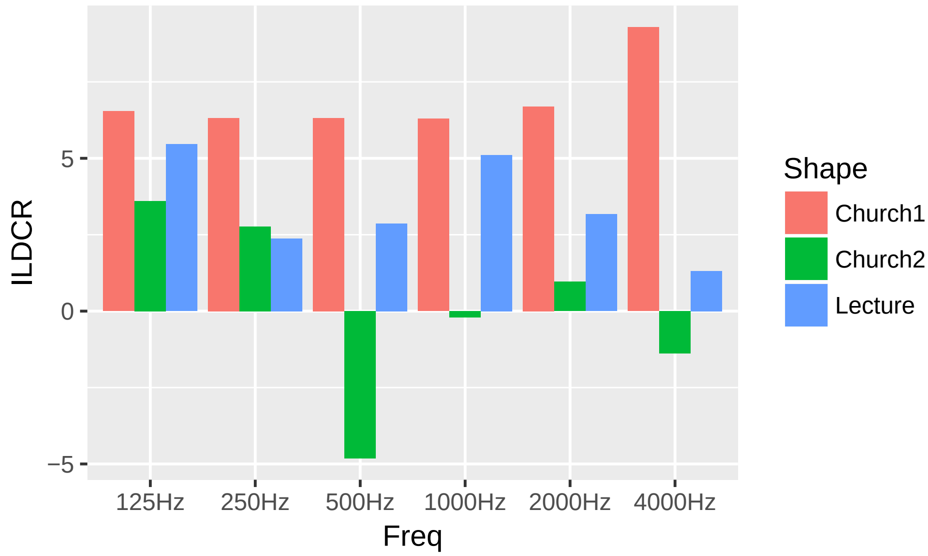


Figure 3.18: ILD-CR values from measurements in existing space

From the analysis of the ILD-CR (figure 3.18), Church1 showed the highest value. Church2 has a low value due to the low correlation between sound source position and ILD value in the mid to high-frequency band, and r has a high correlation in the Lecture room. In the lecture room, r was analyzed as having a high correlation, but the range was not large. So, the ILD-CR value was not high.

In analyzing the ILD-CR values, Church 1 in shoebox form was evaluated to have a larger spatial impression as in the experiment using simulations. However, in the high frequency range above 2000 Hz, the variation of the ILD value due to the sound source movement is not significant in the acoustic simulation, but there was variation in the measurement result. And it showed that the shape of the rear and side walls affects the ILD-CR value. This means that if the ILD is related to the location perception, the ILD-CR can represent a stage spatial impression for the varying sound source location.

3.4 Summary

The objective results from acoustic simulations from two room shapes from the measurement made in three existing spaces demonstrate how well-assorted metrics are at capturing the change in spatial impression as sound source position is moved across a stage.

Lateral Energy Fraction(LF) did not show a significant change due to varying source position across the stage. LF which uses the ratio of the lateral energy to the total sound energy (not distinguishing between the left and right) is not considered to be a good metric fro quantifying how spatial impression may change due to position of the sound source.

$IACC_E$ uses the difference between left and right arriving sound energy. It shows the value changes due to source positions. The $IACC_E$ value can compare the size of the ASW, but does not show the relative position between the sources. So, It is difficult to determine the direction between the source and the listener then.

The acoustic simulation and measurement results show that the ILD values in the diffusion field have different characteristics compared to the free field. In the free field, ILD has higher values in the high-frequency range than in the low-frequency range due to the sound shadow effect of the human head, but this characteristic has not been observed in the experiments conducted in the diffusion field. It is analyzed that the reflections caused by the shape of the space have an influence on the ILD value. The newly proposed ILD-CR metric is based on the ILD value, and thus reflects changes in acoustic energy and position of the sound source depending on the shape of the space.

As the result of measurements and the simulation show the existing acoustic parameters for evaluating spatial impression have limitations in expressing the change in received sound energy due to the change of the position of the sound source. Consequently, a new evaluation method capable of expressing this is needed.

The newly proposed ILD-CR, which describes how linearly correlates ILD values are across different stage position, and how greatly ILD varies through of the range of ILD Values, describe the spatial impression when varying sound source position across the stage.

Chapter 4

Experiments on perception of sound source locations on stage

4.1 Introduction

Since the spatial impression of multiple sound sources on the stage is related to the perceived positional arrangement of each sound source, it is important not only to know how much matches the position of the actual sound source and the position perceived by the listener, but also to know the width of the perceived overall sound source. Since the existing metrics for evaluating spatial impression have limitations in documenting this change a metric called ILD-CR is proposed to effectively express the change in received sound energy according to changing position of the sound source. The ILD-CR includes information about the width of the total sound source perceived by the listener and the difference between the position and the perception of the actual sound source using the ILD, which is an important factor in the location of the sound source. Previous experiments have shown that the ILD used in the ILD-CR at the listener's position varies with the location of the sources in the space. This chapter takes on further step due to source position study how the variation of the ILD due to source position according to the shape of the space affects the perception of spatial impression. Subjective auditory experiment was used to investigate the relationship between change of ILD and the perception of sound source location and examine the effect of other factors besides ILD.

4.2 Methodology

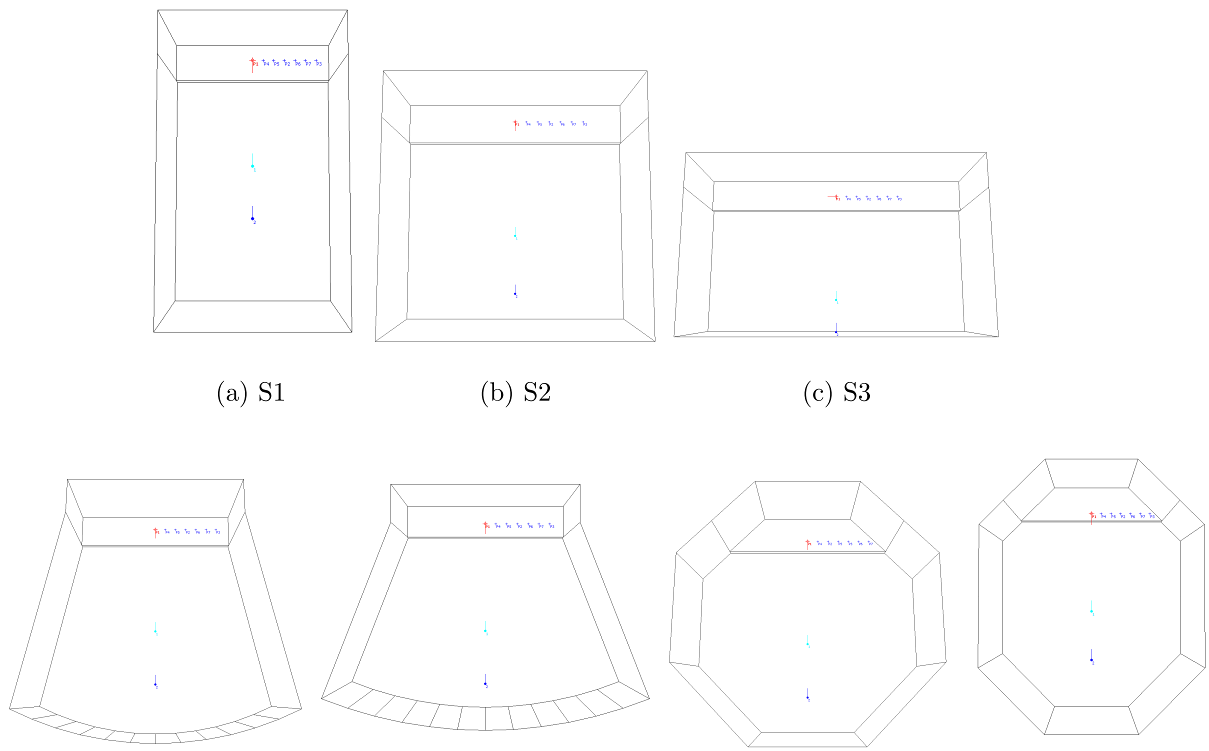
4.2.1 Signal generation

A sound source set was generated by acoustic simulation program which is ODEON (ver.11, Denmark). For auralization an anechoic recording of seven shapes was convoluted with the impulse response and Subject21 HRTF, based on the 'standard' KEMAR head with blocked ear canal and normal sized pinnae, was applied. Table 4.1 lists the parameters used to generate impulse responses in ODEON.

Impulse Response Length	2000ms
Number of late rays	100,000
Transition Order	4
Number of early scatter rays (per image source)	100
Apply NC curve	NC15

Table 4.1: Simulation program(Odeon) setup

Seven different spaces were constructed from one of three shapes: rectangular, fan, and octagon. The rectangular version was varied to be narrow, square, and wide; the fan shape was varied to have two different side wall angles; and the octagon was varied to be regular or elongated. Table 4.2 summarizes the dimensions of each and Figure 4.1 shows the floor plans.



(a) S1 (b) S2 (c) S3

(d) F1 (e) F2 (f) O1 (g) O2

Figure 4.1: 1st row : Rectangular version (S1, S2, S3)
2nd row : Fanshape version (F1, F2), Octagon version (O1, O2)
(Not drawn to same scales; see Table 4.2 for dimension information.)

Shape	Size	Volume	Total surface area
Rectangular version	Narrow (S1, 25m x 15m x 8m)	2972 m^3	1387 m^2
	Square (S2, 19.4m x 19.4m x 8m)	2975 m^3	1366 m^2
	Wide (S3,15m x 25m x 8m)	2920 m^3	1387 m^2
Fan shape version	Narrow(F1, 21.5m x 22m x 8m)	3003 m^3	1364 m^2
	Wide(F2, 19.5m x 26m x 8m)	2981 m^3	1369 m^2
Octagon version	Regular (O1, 21m x 21m x 8m)	2944 m^3	1312 m^2
	Elongated (O2, 23m x 19m x 8m)	2998 m^3	1317 m^2

Table 4.2: Room shape classification

To minimize the influence of physical elements other than the shape of space, the volume of each space and the size of the internal surface area are designed to be the same, and the sound absorption coefficients ($\alpha = 0.2$ or 0.3 depending on the reverberation time conditions) were uniformly assigned on all surfaces.

The sound source consists of seven impulse response, and they located on the stage at 1m interval starting from the center towards stage left, and receivers are centrally located at 10 m and 15 m distance from the center sound source. Impulse responses were simulated and then filtered in Matlab into the 500Hz, and 1000Hz octave bands. Each space was used to simulate auralization with two different reverberant times: 0.7 and 1.6s averaged across the mid-frequencies.

To obtain more accurate results for r(correlation), Range of ILD, and ILD-CR, the interval between sound source positions was set to the minimum allowed by the simulation program of 1m. In the auditory experiment, only auralization from three of the seven sound source positions were used: at 0m (sp1), 3m (sp2), and 6m (sp3) to stage left.

S3, center S-R:15m case which is longish Rectangle shape was excluded from the test because the distance between the source and the listener did not meet 15m. After generating the impulse response, the ILD from certain center-source-receiver-room shape combinations exceeded 3 dB, rather than being zero as is expected with omnidirectional source and receiver on the same center line in a symmetrical room. These samples would likely be perceived as not coming from the center which may affect the experiment, so these combinations were excluded from the experiment. The excluded files include outcome of the S3 (Center S-R dist 10 m, 500 Hz, RT 1.6 sec), and F2 (S-R dist 15 m, 500 Hz, RT 1.6 sec).

A total of 150 auralizations used. One set consists of one reference and one comparison. A total of 100 test sets were used in the experiment because the each space of the same condition had 3 auralizations in total ; Possible experiments under the same conditions example :set1- reference :sp1, comparison :sp2, set2- reference :sp1, comparison :sp3.

.

4.2.1.1 ILD of test signal

Each generated impulse response was filtered into frequency bands using Matlab and then the program calculated the ILD, Using a cutoff time of 80ms. The cut-off time is set to 80ms, which is known to have a large influence on the ASW in the previous study. ILD is the ratio of left to right energy reaching the receiver during the initial

80ms. Since the shape of the space affects the path and quantity of the initial reflection energy, the value of ILD varies depending on each space.

As shown in Table 4.3 and Figure 4.2 to 4.9, the change of ILD was not always proportional to the distance of source location from center stage. Since the sound source moves from the center of the stage to the stage left in all the spaces, the sound energy from the right side of the receivers is expected to be larger than from the left, and the ILD values plotted against source position are expected to exhibit positive slopes, but in some cases negative slopes are found (Example -Figure 4.2 s3, f1). Also, even in the same space and the same location, the ratio of the left and right energy was different according to the frequency band(Example -Figure 4.2 f1 and Figure 4.3 f1).

Shape	Dist S-R	Freq	RT(500Hz)	ILD(dB)			RT(500Hz)	ILD(dB)		
				sp1	sp2	sp3		sp1	sp2	sp3
s1	10m	500Hz	0.7s	0.4	5.96	7.94	1.6s	1.49	5.64	7.16
s2				-0.51	-1.31	-1.58		0.4	0.17	0.46
s3				0.74	-2.3	-4.51				
f1				1.33	-5.47	-4.98		-0.49	-5.87	-7.04
f2				1.53	1.57	-1.88		1.3	3.38	-1.03
o1				0.78	-1.77	-2.22		-1.32	-4.17	-0.61
o2				0.58	3.79	9.97		0.02	4.12	6.59
s1	10m	1000Hz	0.7s	-0.23	3.49	4.72	1.6s	-0.76	3.83	4.23
s2				-0.89	4.03	2.19		-0.54	3.03	2.64
s3				-0.01	2.54	-0.61		-0.68	0.95	-0.27
f1				0.63	-0.61	1.27		-0.48	-0.94	0.97
f2				0.19	1.07	3.18		0.56	-1.99	3.11
o1				-0.68	-4.35	-1.04		-0.35	-2.92	0.68
o2				0.48	4.82	4.67		-2.8	3.66	4.85
s1	15m	500Hz	0.7s	-2.29	4.52	8.86	1.6s	0.27	4.68	9.33
s2				-1.77	1.53	-0.19		-2.53	2.77	2.71
f1				1.78	3.73	0.87		1.54	-3.3	-1.76
f2				0.53	2.45	0.49				
o1				0.15	0.56	1.85		0.24	1.09	-3.45
o2				-2.13	2.42	0.62		1.03	0.98	-1.25
s1	15m	1000Hz	0.7s	-0.56	3.57	4.44	1.6s	-0.33	1.48	4.04
s2				1.48	3.04	4.91		1.43	2.42	3.03
f1				-0.16	1.28	1.46		0.07	-0.87	-2.7
f2				-0.38	-0.38	0.24		1.09	0.04	2.04
o1				-0.05	0.81	0.58		-0.18	-0.21	-2.22
o2				-0.04	0.6	-0.67		-2.49	-0.07	0.55

Table 4.3: ILD values for the test cases. Refer to Table 4.1 for shape. sp1, sp2 and sp3 refer to source positions. Cases where sp1 values were greater than ±3dB are excluded.

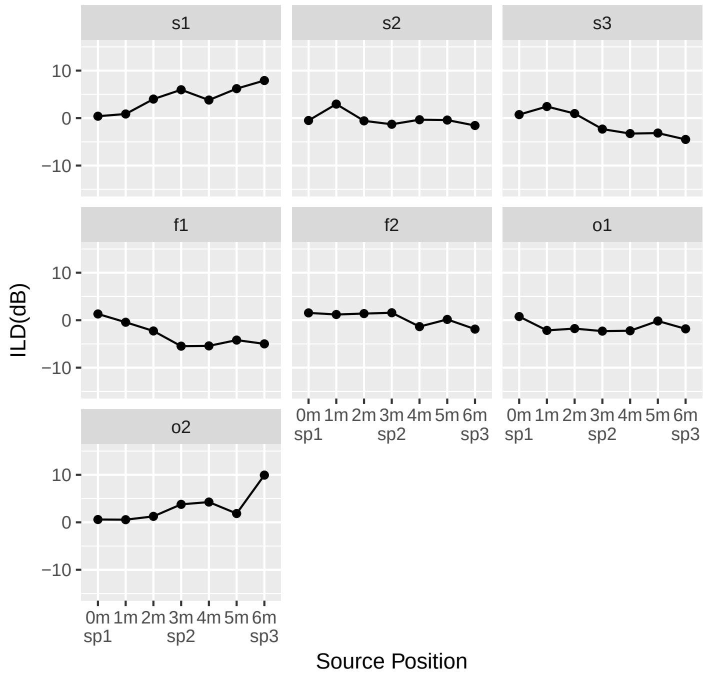


Figure 4.2: ILD calculated for RT=0.7sec, and centerS-R distance=10m for the 500Hz octave band, for the various shaped spaces, marked as shown in Table 4.1, sp1,sp2,sp3 refer to source position

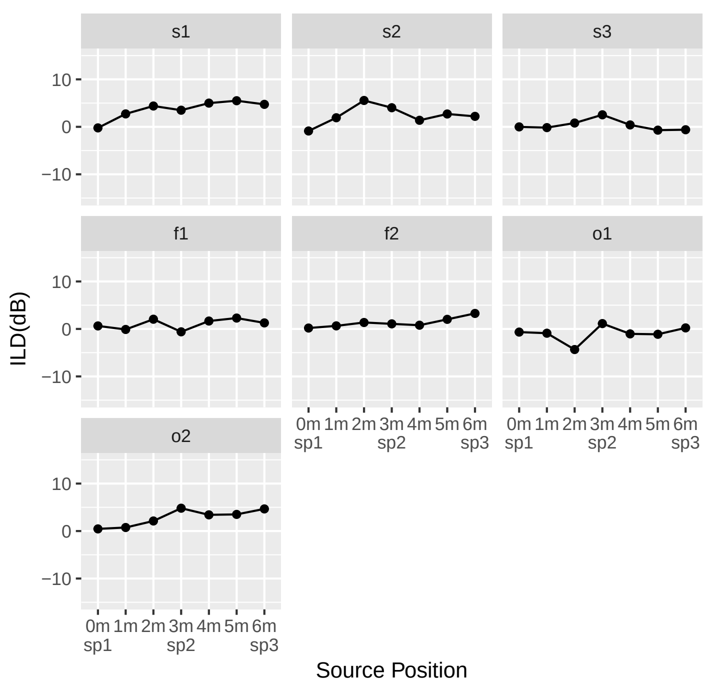


Figure 4.3: ILD calculated for RT=0.7sec, and centerS-R distance=10m for the 1000Hz octave band, for the various shaped spaces, marked as shown in Table 4.1, sp1,sp2,sp3 refer to source position

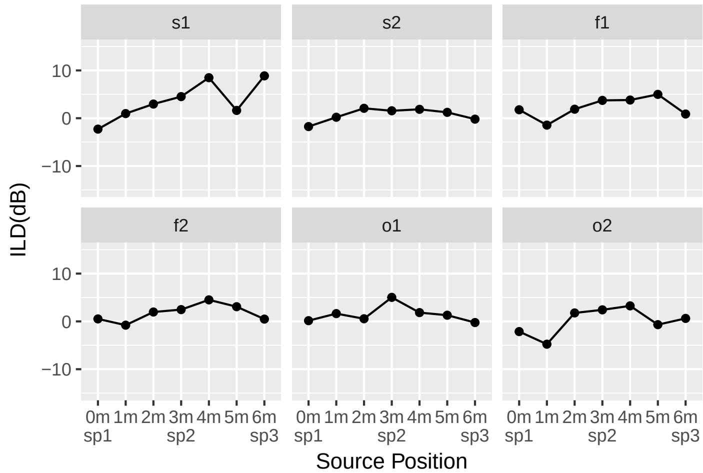


Figure 4.4: ILD calculated for RT=0.7sec, and centerS-R distance=15m for the 500Hz octave band, for the various shaped spaces, marked as shown in Table 4.1, sp1,sp2,sp3 refer to source position

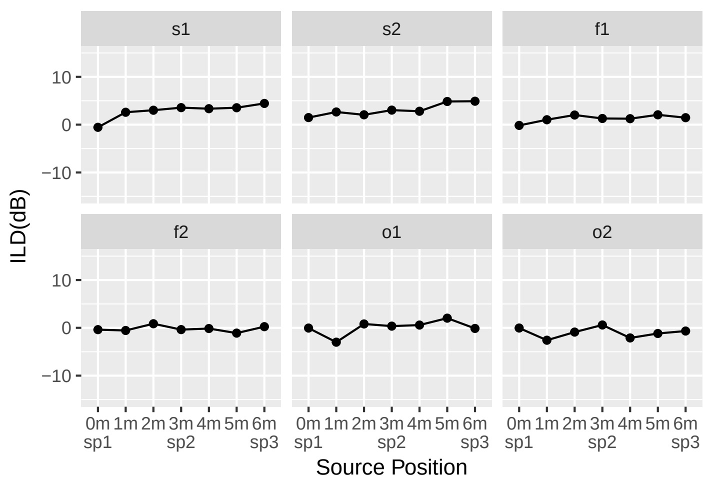


Figure 4.5: ILD calculated for RT=0.7sec, and centerS-R distance=15m for the 1000Hz octave band, for the various shaped spaces, marked as shown in Table 4.1, sp1,sp2,sp3 refer to source position

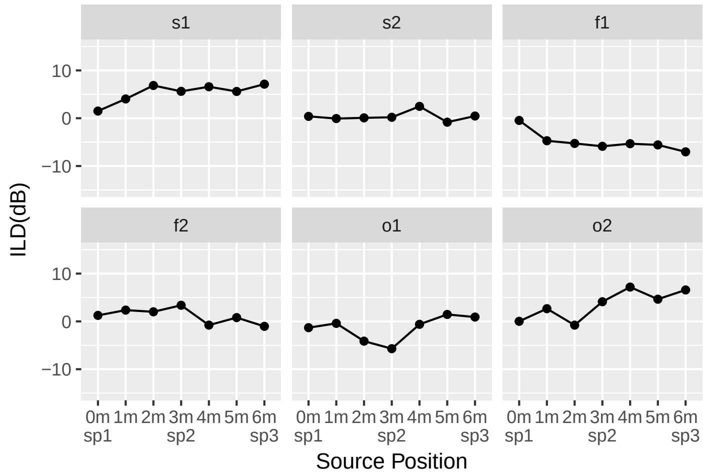


Figure 4.6: ILD calculated for RT=1.6sec, and centerS-R distance=10m for the 500Hz octave band, for the various shaped spaces, marked as shown in Table 4.1, sp1,sp2,sp3 refer to source position

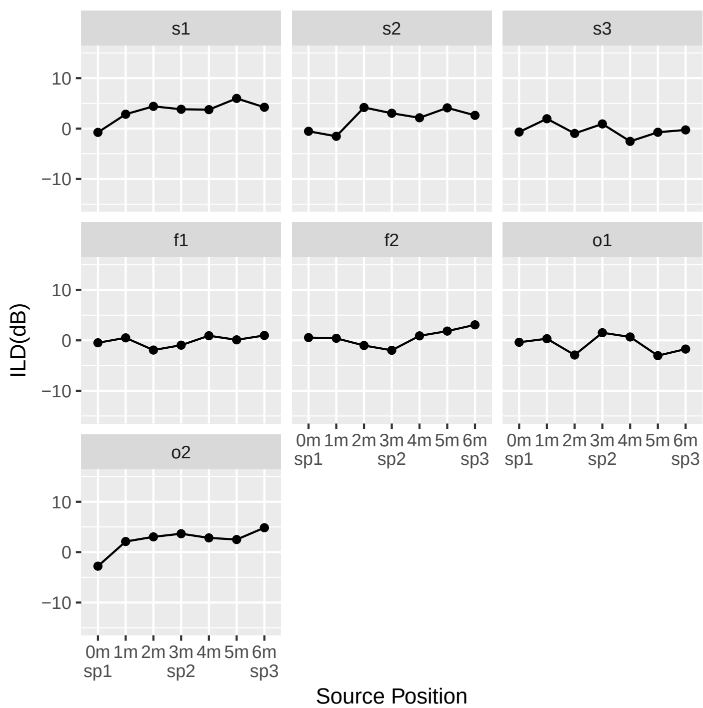


Figure 4.7: ILD calculated for RT=1.6sec, and centerS-R distance=10m for the 1000Hz octave band, for the various shaped spaces, marked as shown in Table 4.1, sp1,sp2,sp3 refer to source position

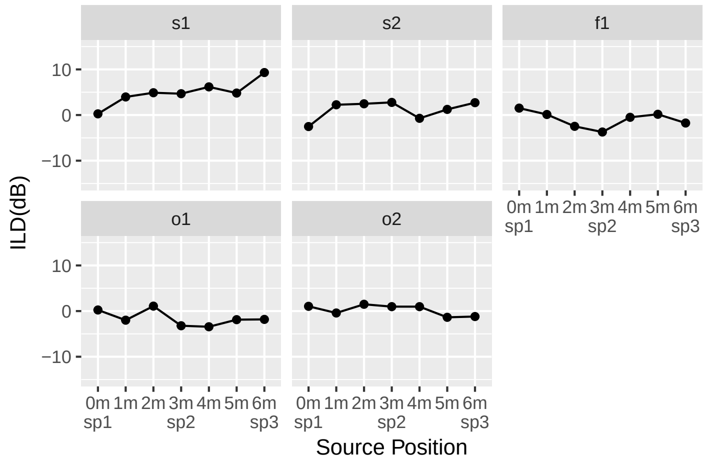


Figure 4.8: ILD calculated for RT=1.6sec, and centerS-R distance=15m for the 500Hz octave band, for the various shaped spaces, marked as shown in Table 4.1, sp1,sp2,sp3 refer to source position

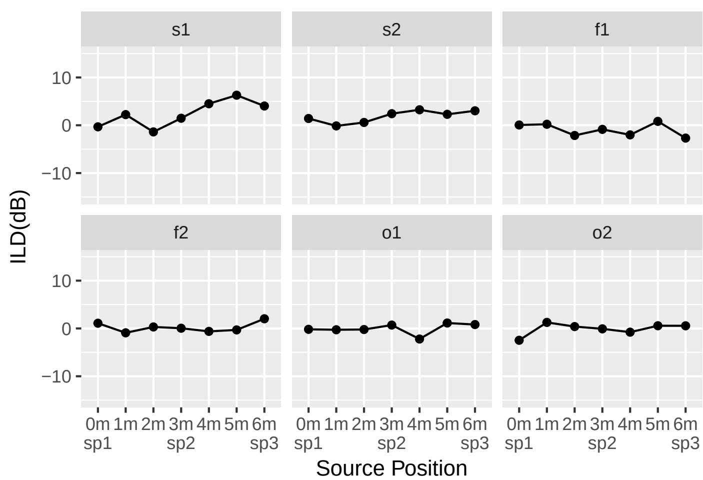


Figure 4.9: ILD calculated for RT=1.6sec, and centerS-R distance=15m for the 1000Hz octave band, for the various shaped spaces, marked as shown in Table 4.1, sp1,sp2,sp3 refer to source position

4.2.1.2 r, Range of ILD and ILD-CR

r represents the degree of linear relationship for ILD across source position on stages. In Table 4.4 and Figure 4.10, since the sound source is being shifted to stage left, if r is negative, it indicates that the ILD values increase inversely to the moving direction of a sound source. This means that the direction of movement of the sound source is not always well-represented by the change in the ILD value.

Shape	Dist S-R(m)	Freq	RT(sec)	r	Range	ILD-CR	Shape	Dist S-R(m)	Freq	RT(sec)	r	Range	ILD-CR
s1	10	500Hz	0.70	0.91	7.52	6.87	o2	15	1000Hz	0.70	-0.02	3.19	-0.07
s2	10	500Hz	0.70	-0.50	4.51	-2.26	s1	10	500Hz	1.60	0.77	5.63	4.32
s3	10	500Hz	0.70	-0.91	6.93	-6.31	s2	10	500Hz	1.60	0.08	3.30	0.28
f1	10	500Hz	0.70	-0.85	6.78	-5.78	f1	10	500Hz	1.60	-0.80	6.57	-5.24
f2	10	500Hz	0.70	-0.80	3.45	-2.77	f2	10	500Hz	1.60	-0.61	4.41	-2.68
o1	10	500Hz	0.70	-0.28	3.07	-0.84	o1	10	500Hz	1.60	0.41	7.17	2.94
o2	10	500Hz	0.70	0.78	9.39	7.33	o2	10	500Hz	1.60	0.80	7.96	6.36
s1	10	1000Hz	0.70	0.83	5.72	4.76	s1	10	1000Hz	1.60	0.76	6.74	5.12
s2	10	1000Hz	0.70	0.25	6.43	1.63	s2	10	1000Hz	1.60	0.65	5.71	3.72
s3	10	1000Hz	0.70	-0.23	3.24	-0.73	s3	10	1000Hz	1.60	-0.31	4.49	-1.37
f1	10	1000Hz	0.70	0.45	2.90	1.29	f1	10	1000Hz	1.60	0.46	2.90	1.35
f2	10	1000Hz	0.70	0.85	3.07	2.63	f2	10	1000Hz	1.60	0.57	5.06	2.86
o1	10	1000Hz	0.70	0.25	5.47	1.37	o1	10	1000Hz	1.60	-0.31	4.56	-1.41
o2	10	1000Hz	0.70	0.85	4.36	3.71	o2	10	1000Hz	1.60	0.75	7.63	5.72
s1	15	500Hz	0.70	0.77	11.14	8.56	s1	15	500Hz	1.60	0.86	9.05	7.81
s2	15	500Hz	0.70	0.37	3.82	1.40	s2	15	500Hz	1.60	0.40	5.29	2.10
f1	15	500Hz	0.70	0.43	6.43	2.78	f1	15	500Hz	1.60	-0.33	5.23	-1.74
f2	15	500Hz	0.70	0.43	5.30	2.30	o1	15	500Hz	1.60	-0.48	4.52	-2.18
o1	15	500Hz	0.70	-0.02	5.26	-0.13	o2	15	500Hz	1.60	-0.60	2.86	-1.70
o2	15	500Hz	0.70	0.49	8.02	3.93	s1	15	1000Hz	1.60	0.76	7.68	5.86
s1	15	1000Hz	0.70	0.83	5.00	4.14	s2	15	1000Hz	1.60	0.75	3.38	2.54
s2	15	1000Hz	0.70	0.91	3.42	3.11	f1	15	1000Hz	1.60	-0.39	3.51	-1.37
f1	15	1000Hz	0.70	0.64	2.22	1.41	f2	15	1000Hz	1.60	0.23	2.94	0.68
f2	15	1000Hz	0.70	-0.03	1.96	-0.06	o1	15	1000Hz	1.60	0.26	3.35	0.89
o1	15	1000Hz	0.70	0.48	5.03	2.42	o2	15	1000Hz	1.60	0.41	3.74	1.53

Table 4.4: r, Range of ILD, and ILD-CR value for all test cases. Refer to Table 4.1 for shape. sp1, sp2 and sp3 refer to source positions.Cases where sp1 values were greater than $\pm$3dB are excluded.

For example, when the reverberation time is 0.7 seconds, and the distance between the center source and the receiver is 10 m, the r value is found to be negative at the 500 Hz band in the room shapes of s2, s3, f1, f2, and o1. Also, according to the shape of space, the degree of correlation shows a great difference. For example, in the same condition above (RT=0.7, S-R=10m, 500Hz), the change of the ILD value in shape s1 is highly correlated with the sound source position, but shape o1 has low correlation. And, in the RT is 0.7 seconds, S-R=15m, 500Hz case, a positive correlation between the source position and ILD is sound in most shapes. This is quite different from the previous case where the S-R distance was 10m. It demonstrates that difference in energy between left and right ears at the receiver depends heavily on the position of the listener and the shape condition of the room.

The result of r, Range of ILD and ILD-CR are visualized in Figure 4.10, Figure 4.11 and Figure 4.12. Refer to Table 4.2 for classification of space type.

Figure 4.10: r (the degree of linear relationship)

By making the average sound absorption coefficient (α 0.2 or 0.3), volume (3000 m^3) and surface areas (1380 m^2) of the seven spaces are made close to each other, it minimized the difference of energy loss due to sound absorption. Figure 4.11 showed the range of the ILD value is significantly different depending on the shape of the space and conditions.

Direction of Source Movement(DSM) is '+1' because the direction of sound source movement is to the right.

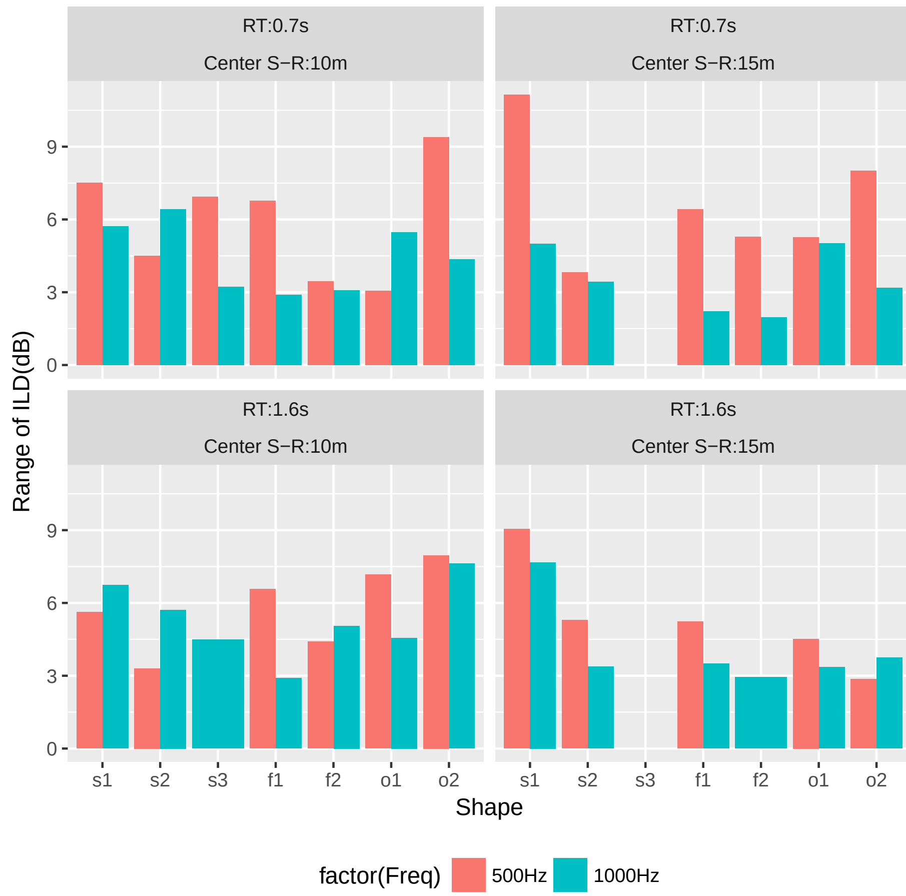


Figure 4.11: Range of ILD

The ILD-CR is determined by r, Range of ILD and DSM. If the value is negative, it means that the direction of shifting sound position is inversely proportional to the ILD value. In the RT=0.7 sec, center S-R distance=10m, and 500Hz case, the direction of the shifting sound position and the change of the ILD values do not coincide in the room shapes of s2, s3, f1, f2 and o2. In the RT=0.7 sec, center S-R distance 15m, and 500Hz case, range of ILD had 3.82 dB, but the ILD-CR is 1.4, much smaller than range of ILD because of $r(0.37)$ is small.

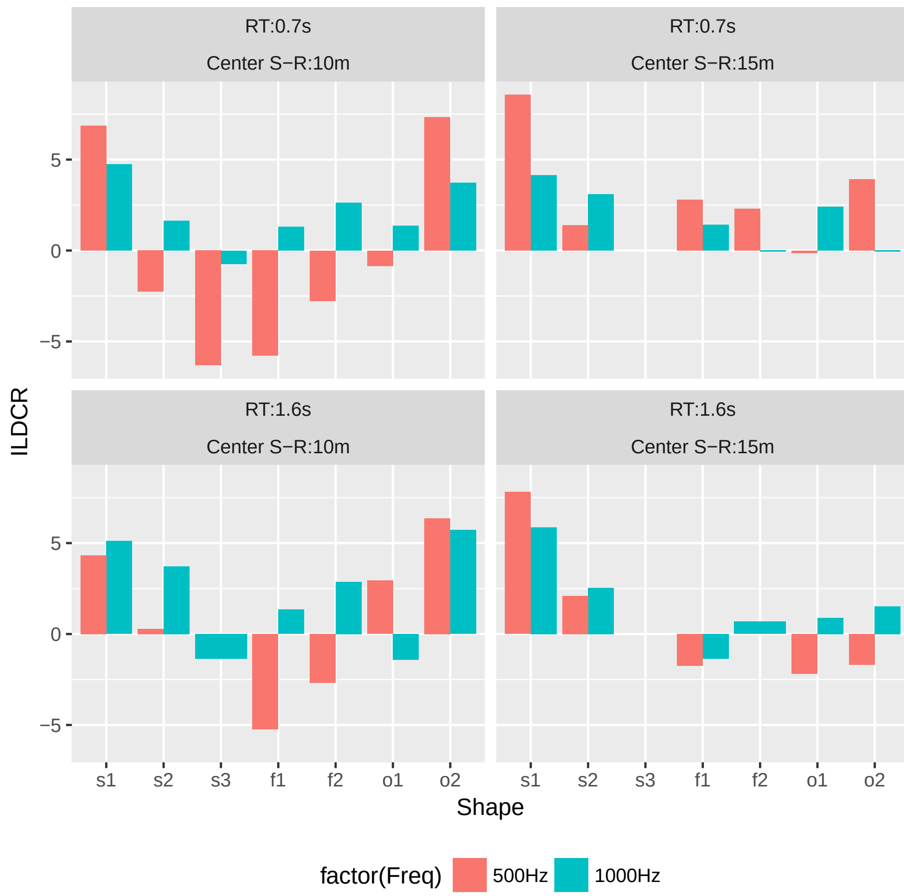


Figure 4.12: ILD-CR

4.2.1.3 Auditory subjective test

This experiment approved by the Institutional Review Board (IRB) for the Protection of Human Subjects in the University of Nebraska - Lincoln.

(a) **Participants**

The subjective assessment was completed by 22 participants. Participants were screened to have a threshold below 25 dB hearing level from the 125 Hz to 8,000Hz octave bands. General information on the participants are given in Table 4.5

Gender	Age	Musical Training
Female : 7 Male : 15	20∼25 :11 26∼30 : 8 31∼35 : 1 36∼40 : 0 41∼45 : 2	None : 10 1∼3 years : 4 More 3 years : 8

Table 4.5: Participant demographics

None of the participants indicated that they had perfect pitch. Participants were paid 15 dollars for their time.

(b) **Procedure**

Basic demographic information (Age, Gender, Musical Experience) was collected during the orientation of this study to correlate with data. Testing was conducted in a dedicated testing room, free from outside interaction, so subjects were maintain anonymity in participation of the study. Before the test, experimenter explained the process and purpose of the test to taskers. To minimize external noise, the test was conducted in the Nebraska sound booth. The experiment period is between August and December 2016. Total experimental time varied depending on the subjects; on average it took about one hour.

The experiment was conducted using a headphone system (Beyer dynamic DT-880 pro) connected to a laptop computer running Meridian Explorer audio interface. The sound samples were adjusted to a maximum of 75dBA to minimize the influence of loudness variance.

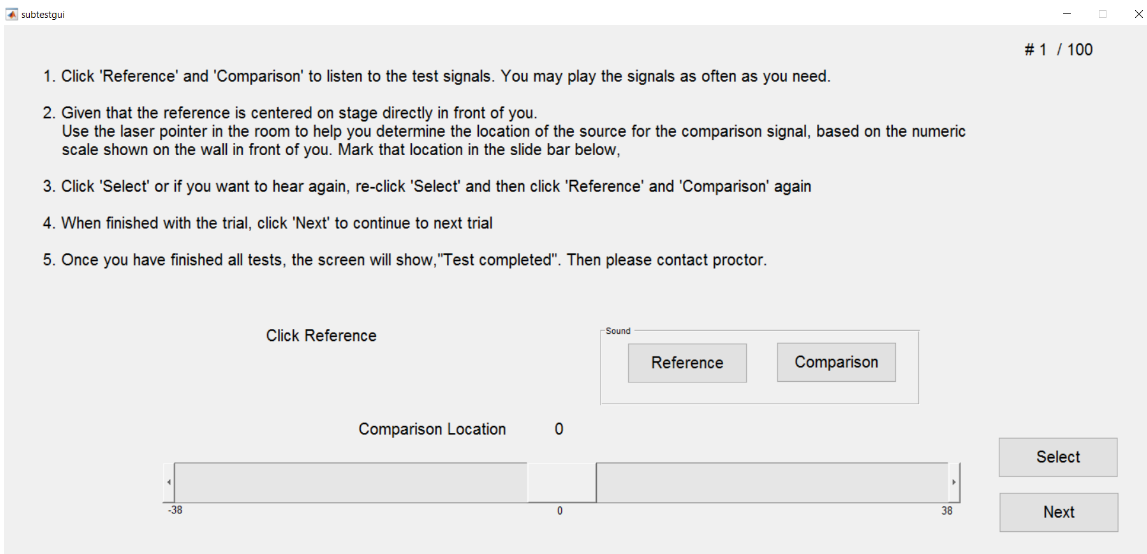
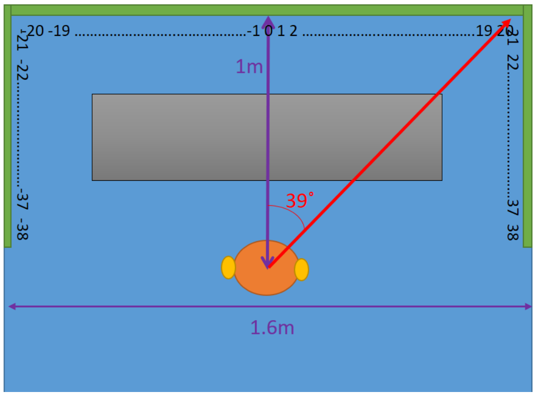


Figure 4.13: GUI screenshot

Experiments were conducted using a test program written in Matlab. As a Graphic User Interface (GUI) consisted of a section describing the experiment sequence, a section for listening to the sound samples, and section for inputting responses.The position of the reference sound source is assumed to be coming from a position at the center of the stage and participant are asked to input relative position of the comparison signal using the slide on the GUI (Figure 4.13).

To help the participants determine the location of the source for the comparison signal, inside the booth, there was a numbered scale attached to the front and sides of the participant's seat that participant used to indicate the perceived location of the sound source (Figure 4.14); laser pointer was provided to assist participants with deciding which number to input. The position of the chair of the experimenter was specified so that the distance (1m) between the ruler and the chair was always constant.



(a) chamber dimension

(b) chamber picture

Figure 4.14: Nebraska sound booth

Participant could listen to the reference and comparison samples as many time as they wish. When moving to the next test set, a pause of 5 seconds was taken to minimize influence from the previous test. The experiment consisted of a total of 100 comparisons using all samples shown in Table 4.3. The test set was administered using Latin square design minimize any bias in the experiment sequence.

(c) **Numerical scale conversion to angle**

The numerical scale attached to the front and side walls ranged from -38 to +38. Zero was located in the center. The distance between the participant and center of the scale was 1m and the interval between each scale marked was 4cm (Figure 4.14 (a)). The numeric value given in the participant's response was used for analysis by converting the number to the angle indicating the direction of the sound source.

The formula used to convert the numerical value on the scale to an angle is: given in Eq. 4.2.1

$$Angle = ATAN(Numericalscalex0.04)x180/\pi$$

$$(when -20<Numericalvalue<20)$$

$$Angle = 90 - ATAN((45 - numericalvalue)x0.04/0.8)x180/\pi$$

$$(whenNumericalvalue>20)$$

$$Angle = -90 - ATAN((-45 - numericalvalue)x0.04/0.8)x180/\pi$$

$$(when -20>Numericalvalue)$$

(Eq. 4.2.1)

The converted angles are shown as Table 4.6.

Numerical.Scale	Angle	Numerical.Scale.1	Angle.1	Numerical.Scale.2	Angle.2	Numerical.Scale.3	Angle.3
-38	-70.71	-18	-35.75	2	4.57	22	41.01
-37	-68.20	-17	-34.22	3	6.84	23	42.27
-36	-65.77	-16	-32.62	4	9.09	24	43.60
-35	-63.43	-15	-30.96	5	11.31	25	45.00
-34	-61.19	-14	-29.25	6	13.50	26	46.47
-33	-59.04	-13	-27.47	7	15.64	27	48.01
-32	-56.98	-12	-25.64	8	17.74	28	49.64
-31	-55.01	-11	-23.75	9	19.80	29	51.34
-30	-53.13	-10	-21.80	10	21.80	30	53.13
-29	-51.34	-9	-19.80	11	23.75	31	55.01
-28	-49.64	-8	-17.74	12	25.64	32	56.98
-27	-48.01	-7	-15.64	13	27.47	33	59.04
-26	-46.47	-6	-13.50	14	29.25	34	61.19
-25	-45.00	-5	-11.31	15	30.96	35	63.43
-24	-43.60	-4	-9.09	16	32.62	36	65.77
-23	-42.27	-3	-6.84	17	34.22	37	68.20
-22	-41.01	-2	-4.57	18	35.75	38	70.71
-21	-39.81	-1	-2.29	19	37.23		
-20	-38.66	0	0.00	20	38.66		
-19	-37.23	1	2.29	21	39.81		

Table 4.6: Conversion of numerical values on scale in sound booth to angle

4.2.2 Linear Mixed-Effects Model

In this experiment, the subject was asked to indicate the direction of the source from sound source emanates, given a reference sample that emanates from a center source tow the relative sound sources were located to stage left 3m from center source and 6m to stage left from the center source in the same room and condition. The samples were classified by the conditions of the distance between the center source and the listener, the octave band frequency, the reverberation time, and the ILD, while participants classified were by gender, music training experience, and age.

The basic form of the multiple linear regression model is,

$$Y_i = \alpha + \beta_1 X_i + \cdots \cdot + \beta_k X_{ik} + \varepsilon_i \cdots \qquad \text{(Eq. 4.2.2)}$$

Where Y_i is the value of response, and β_1, β_2, $\cdots$, β_k are the slope coefficients. X_i, $\cdots$ X_{ik} are the regressors for observation i [45]

The multiple linear regression model formula for this investigation is given in Eq. 4.2.3.

$$response \sim \text{Source} + \text{Shape} + \text{Dist} + \text{Freq} + \text{RT} + \text{ILD} + \text{age} + \varepsilon \ldots \qquad \text{(Eq. 4.2.3)}$$

where X_i, $\cdots$ X_{ik} are Source, Shape, Dist, Freq, RT, and ILD. k is 6 for this model.

However, because this analytical model elicited multiple responses from each participant, those responses that come from the same participant cannot be regarded as independent from each other. This would violate the independence assumption. Since each subject judges the position of the source with different criteria, this will be a unique factor affecting all the responses of the subject, making these various responses inter-dependent rather than independent. To solve this problem, random effects are added to the subject responses. This can resolve the non-independence by assuming a different "baseline" value for each subject. So, each subject can have a different baseline for their own perception of sound source location.

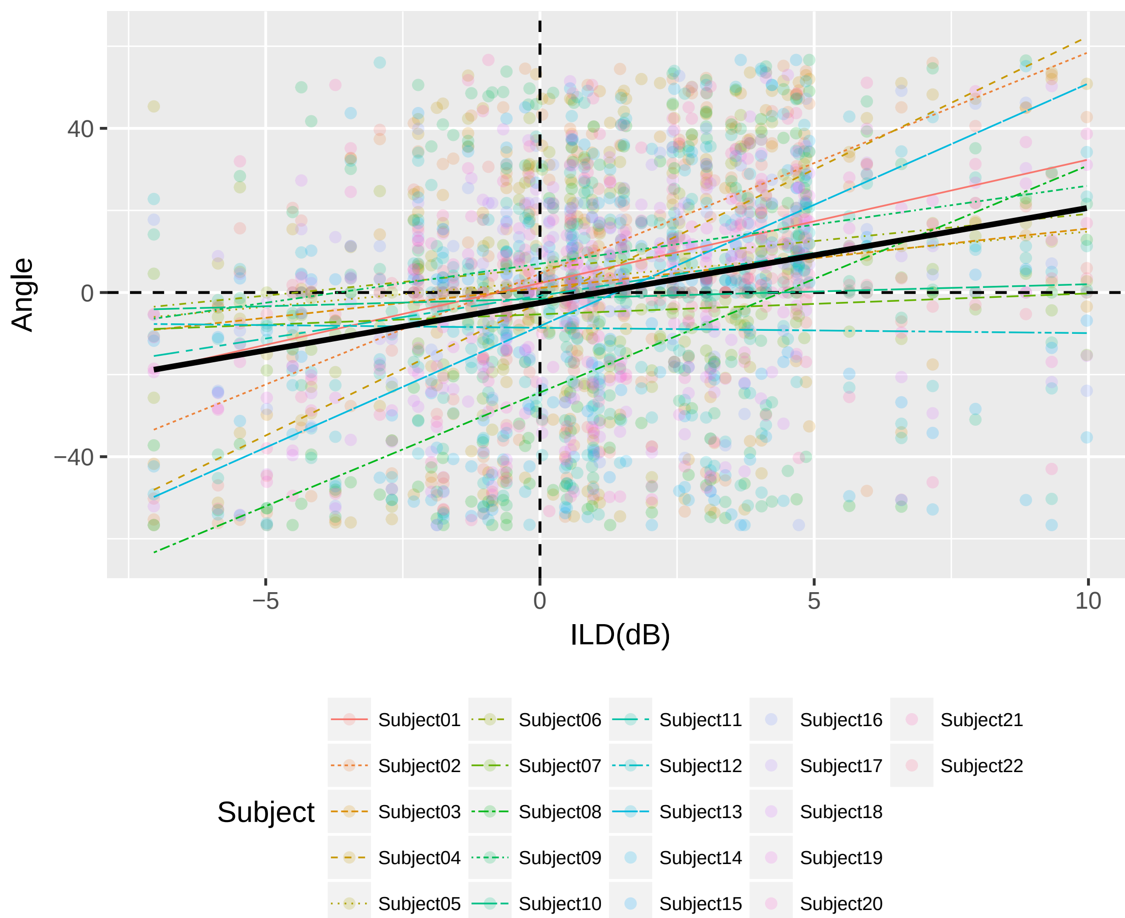


Figure 4.15: Each subject linear regression and the multiple linear regression model plot

In Figure 4.15, the bold black line is the visualization of Eq. 4.2.3 and other lines in the plot show the source location responses and how linear there are for each subject. Each subject's responses linear regression lines demonstrate a different slope and intercept compared to the multiple linear regression model. When the ILD value is zero, the response angle is expected to be close to zero because the ratio of the left and right sound energy is the same. In fact, multiple linear regression model plot (bold black lines) shows similar results. However, each subject's responses linear regression plot shows a significant difference in values depending on the Subject. It means that the subjects have different baselines.

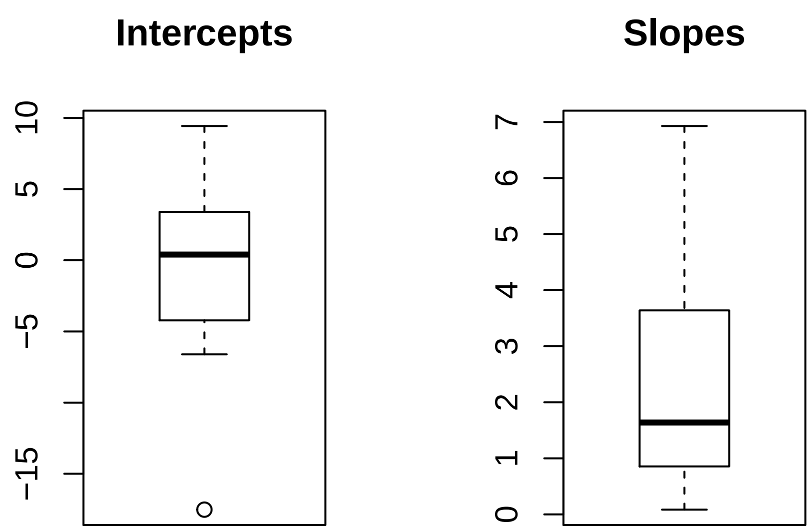


Figure 4.16: Boxplots for subject coefficients for the least-squares regression of location perception on ILD

To analyze this a little more, Figure 4.16 shows the distribution of intercepts and slopes from each subject's responses. The average of total intercepts is spread widely from -4 to 3. The slope also has a value between 0.9 and 3.7 and is spread out. Figure 4.17 shows the ninety-five percent confidence intervals for the least-squares intercepts and slopes for the within subject regressions of location perception. The confidence intervals for each subject are different, and the regression coefficients (intercept and slope) are also widely varying from subject to subject.

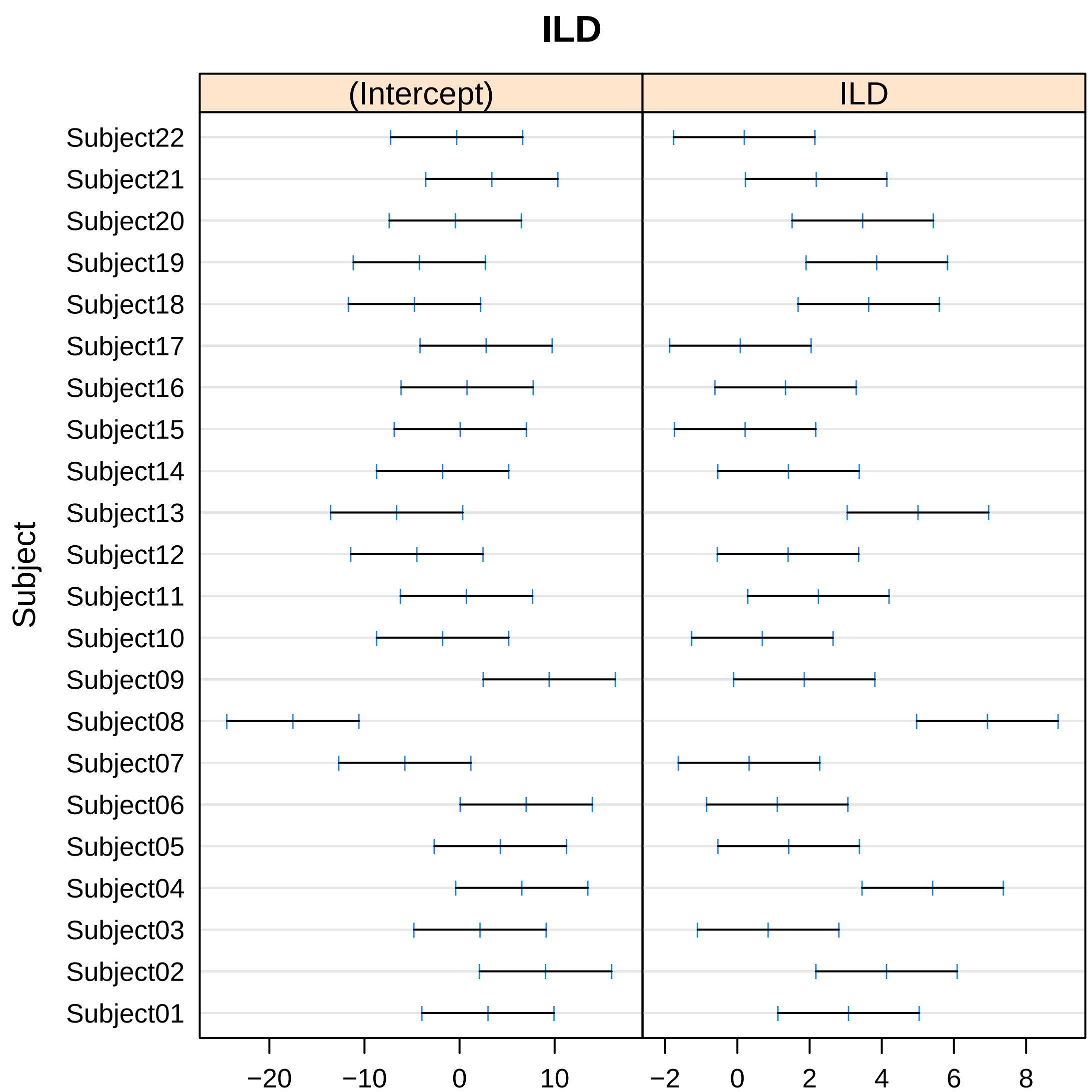


Figure 4.17: Ninety-five percent confidence intervals for least-squares intercepts(left) and slopes(right)

Figure 4.18 visualizes the standard angle of the residual per each subject. It shows different residuals for each subject and the difference is significant, with a lot of individual variation. For example, the residual mean of subject 9 was relatively higher than zero, while subject 8 had a significantly lower.

These data demonstrate why this experiment requires a mixed model. The multiple linear regression model is described with fixed effects and common errors, but in the mixed model, random effects are to the add the fixed effects. In this case, the random effect is ”subject”, as this characterizes the idiosyncratic variation that is due to individual differences.

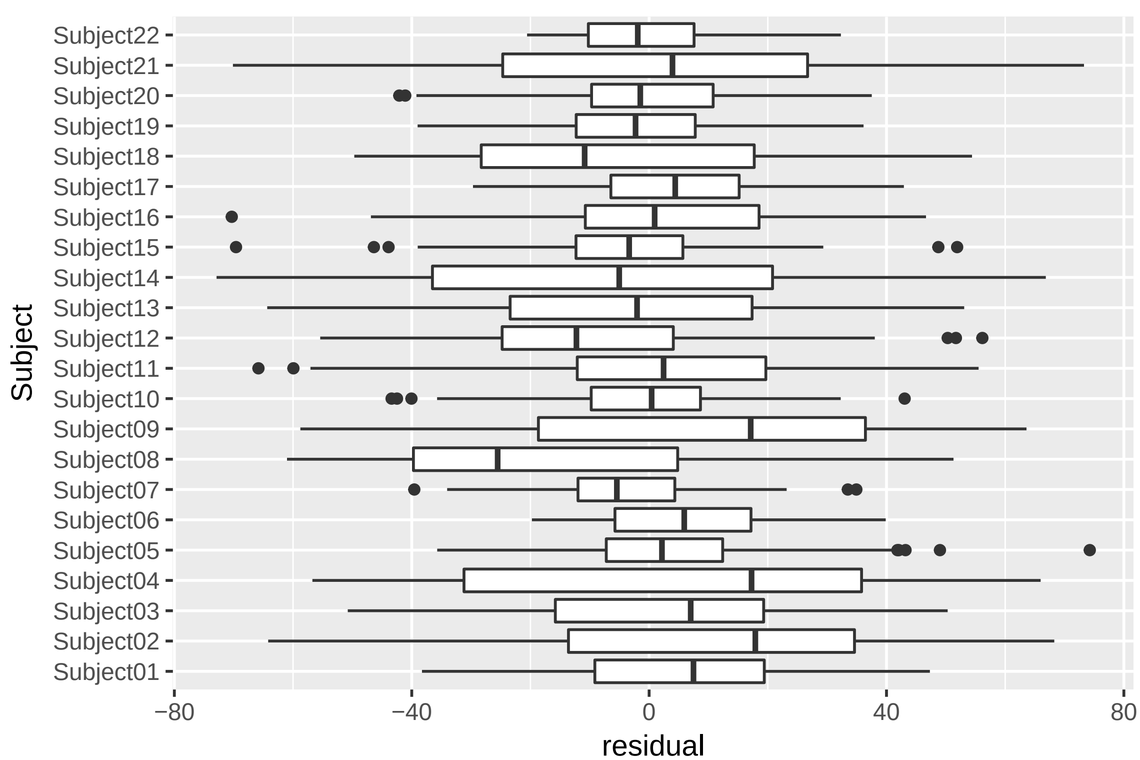


Figure 4.18: residual plot without random-effect

The basic form of the mixed model [46] is as follows.

$$Y_{ij} = \beta_1 + \beta_2 X_{2ij} + \cdots + \beta_p X_{pij} + \delta_{1i} Z_{1ij} + \cdots + \delta_{qi} Z_{qij} + \varepsilon_{ij} \cdots \qquad \text{(Eq. 4.2.4)}$$

Where Y_{ij} is the value of response; β_1, β_2, $\cdots$, β_p are the fixed-effect coefficients, which are identical for all groups; X_{2ij}, $\cdots$ x_{pij} are the fixed-effect regressors for observation j in group i; there is also implicitly a constant regressor, $x_{1ij} = 1$; δ_{1i}, $\cdots$, δ_{qi} are the random-effect coefficients for group i; Z_{1ij}, $\cdots$, Z_{qij} are the random-effect regressors; Group, i, is the number of subjects (22) in this experiment [45].

In Eq. 4.2.4, the first part, $\beta_1 + \beta_2 X_{2ij} + \cdots + \beta_p X_{pij}$, is the fixed part which is the intercept and the slope coefficient of regressors (explanatory variable) times regressors, and the remaining, $\delta_{1i} Z_{1ij} + \cdots + \delta_{qi} Z_{qij}$, is the random part.

The updated random effect linear model for this experiment is then given in Eq. 4.2.5.

$$response \sim \text{Shape} + \text{Source} + \text{Dist} + \text{Freq} + \text{RT} + \text{ILD} + \text{age} + (1 + ILD \mid \text{Subject}) + \varepsilon \ldots \qquad \text{(Eq. 4.2.5)}$$

Where Y_{ij} = Angle response, X_{2ij}, $\cdots$ X_{pij} are Shape (Shape of space); Source (Source position(sp1, sp2, sp3)); Dist (distance between center source and listener(10, 15m));

Freq (frequency octave band (500, 1000Hz)); RT (reverberation time(0.7, 1.6s)); and ILD, Z_{1ij} is Subject.

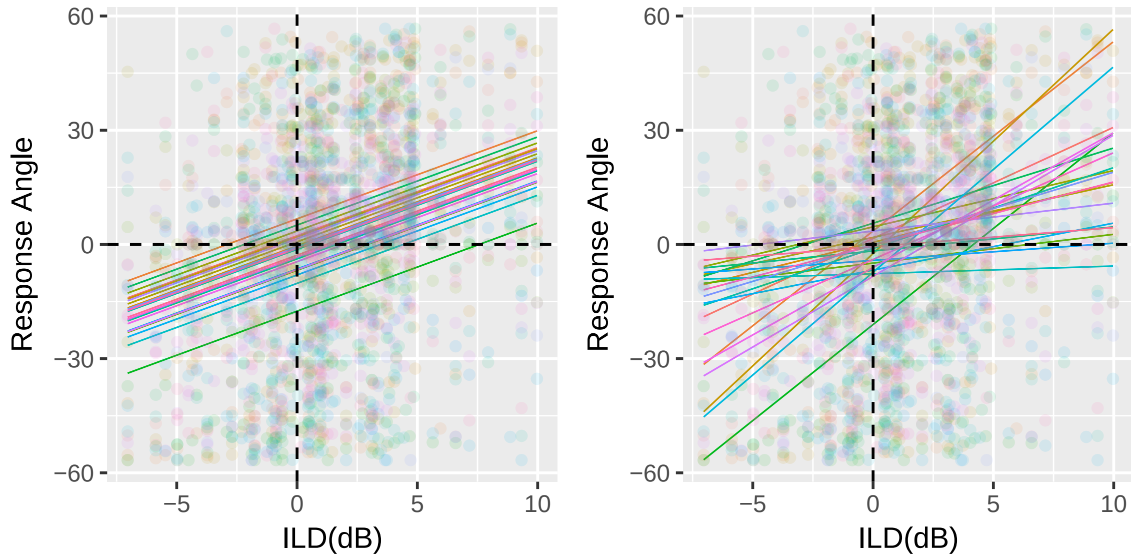


Figure 4.19: Example of Random intercept(left) and slope(right) model

The random effect model can be divided into a random intercept model and a random slope model. The random intercept model is where each group has a different intercept. In this experiment, it means the difference of the location perception reference point of each subject. Figure 4.19(left) visualized the random intercept model. The random slope model is that not only intercepts but also slopes vary group by group. In this experiment, it means the difference of the location perception depending on the ILD change of each subject. Figure 4.19(right) visualized the random slope model.

Figure 4.20 visualizes the standard angle of the residual per each subject using the random-effect model(Eq. 4.2.5). Comparing Figure 4.18 (without random-effect), the mean value of the each subject residual is more close to zero.

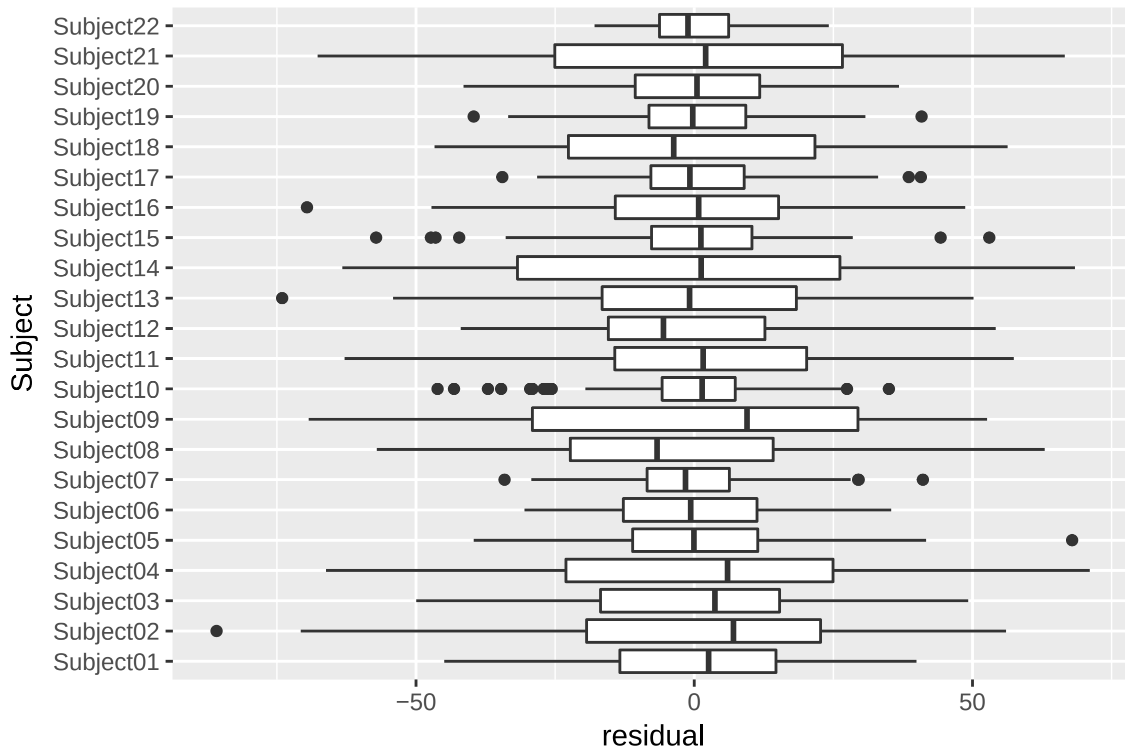


Figure 4.20: residual plot with random-effect

4.3 Result

4.3.1 Relationship between ILD and location perception

The following analysis is the relationship between ILD of the assorted impulse responses and the angle response from the subjective test using a mixed effect model (Eq. 4.2.5). Table 4.7 shows the variance and standard deviations of the subject and residual in random effect. The value of the standard deviation for "Subject" explains the variability of the random effect between subjects. The intercept of the Subject group is the variability of intercept between groups, and the ILD is the slope variability between groups. Residual is the "random" deviations from predicted values that are not due to the subject.

Groups	Name	Variance	Std.Dev
Subject	(Intercept)	41.31	6.43
Subject	ILD	3.62	1.90
Residual		553.02	23.52

Table 4.7: Random Effects Variance and Std.Dev

	Estimate	Std. Error	t value
(Intercept)	1.82	3.91	0.47
Shapes2	0.96	1.92	0.50
Shapes3	-5.69	2.81	-2.03
Shapef1	-17.56	2.31	-7.59
Shapef2	-16.85	2.08	-8.09
Shapeo1	-0.13	2.25	-0.06
Shapeo2	-1.94	1.85	-1.05
Sourcesp3	-6.61	1.01	-6.57
Dist	0.43	0.21	2.09
Freq	0.01	0.00	5.57
RT	-7.18	1.13	-6.35
ILD	1.40	0.46	3.04

Table 4.8: Fixed Effects Estimate, Std.Error, and estimate t-value

Table 4.8 represents the fixed effect values. The values show the relationship between the response and each fixed effect. For ILD, the estimate value is 1.4. This means that if the angle response increases by 1.4, the ILD increases by 1.

The mixed effect model (Eq. 4.2.5) was obtained by using the forward selection method. In the following analysis, a variable was considered for addition or subtraction from the set of explanatory variables based on the criterion. The criterion was evaluated by fitting the P-value using the Likelihood Ratio Test.

To analysis the affect of ILD effect to the angle response, let's calculate the p-value of the ILD for the angle using a random intercept model. To obtain the p-value in the mixed model, the Likelihood Ratio Test was used. To calculate the effect of ILD or response angle, a full-model and null-model are made and then compared using ANOVA in 'R'.

H_0 : ILD have no effect to the angle effect.

$$\begin{aligned} &Full - model : response \sim \text{ ILD} + (1| \text{ Subject}) + \varepsilon \\ &Null - model : response \sim \ (1| \text{ Subject}) + \varepsilon \end{aligned} \quad \ldots \qquad \text{(Eq. 4.3.6)}$$

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
ILD.Null.model	3	20776.08	20793.17	-10385.04	20770.08			
ILD.Full.model	4	20595.90	20618.68	-10293.95	20587.90	182.18	1	0.0000

Table 4.9: Relationship between ILD and response angle

	Estimate	Std. Error	t value
(Intercept)	-2.51	1.37	-1.83
ILD	2.32	0.17	13.78

Table 4.10: ILD Fixed Effects Estimate, Std.Error, t-value in ILD Full-model

ILD affected the response angle (χ^2 (1) = 182.18, p < 0.05)(Table 4.9), lowering it by about 2.32 $\pm$ 0.17(standard error)(Table 4.10). This means that if the ILD value increases 2.32 dB, the angle (or perception of the sound source location) increases by 1 degree.

However, this result assumed that the effect of ILD is the same for all subjects. But, the effect of ILD could be different for each subject; one must check interference using a random slope model. Table 4.11 shows the slope coefficient of ILD. In this table, the coefficient of ILD for each subject is the same, but the random slope model is allowed to have different slopes for the effect of ILD.

	(Intercept)	ILD
Subject01	2.23	2.32
Subject02	6.64	2.32
Subject03	-0.41	2.32
Subject04	1.94	2.32
Subject05	0.64	2.32
Subject06	3.47	2.32
Subject07	-6.75	2.32
Subject08	-17.40	2.32
Subject09	5.01	2.32
Subject10	-3.77	2.32
Subject11	-1.31	2.32
Subject12	-10.16	2.32
Subject13	-3.48	2.32
Subject14	-7.98	2.32
Subject15	-6.42	2.32
Subject16	-0.68	2.32
Subject17	1.48	2.32
Subject18	-6.58	2.32
Subject19	-4.51	2.32
Subject20	-3.33	2.32
Subject21	-1.05	2.32
Subject22	-2.83	2.32

Table 4.11: The coefficients of the intercept model

Eq. 4.3.7 is the ILD random slope model.

$$\text{response} \sim \text{ILD} + (1{+}\text{ILD}|\ \text{Subject}) + \varepsilon \ldots \qquad \text{(Eq. 4.3.7)}$$

In Table 4.12, the ILD coefficient is different from Table 4.11. This means that the effects of ILD on each Subject are different. Let's check whether the difference of each Subject's influence should be taken in the results. In the same method as above,

	(Intercept)	ILD
Subject01	1.53	2.92
Subject02	3.48	4.96
Subject03	0.51	1.58
Subject04	-2.40	5.87
Subject05	1.81	1.39
Subject06	4.55	1.49
Subject07	-4.88	0.77
Subject08	-20.89	5.02
Subject09	5.50	1.98
Subject10	-1.73	0.65
Subject11	-1.06	2.12
Subject12	-7.65	0.22
Subject13	-7.27	5.38
Subject14	-6.71	1.24
Subject15	-4.18	0.47
Subject16	-0.17	1.92
Subject17	3.45	0.75
Subject18	-8.33	3.70
Subject19	-6.03	3.53
Subject20	-3.94	2.80
Subject21	-0.23	1.67
Subject22	-0.61	0.52

Table 4.12: The coefficients of the random slope model

the p-value was obtained from the random slope model of ILD.

H_0 : The ILD has the same effect on the each subject.

$$\begin{aligned} Null-model &= response\sim \text{ILD} + (1|\text{ Subject}) + \varepsilon \\ Full-model &: response\sim \text{ILD} + (1\text{+ILD}|\text{ Subject}) + \varepsilon \end{aligned} \qquad \ldots \qquad \text{(Eq. 4.3.8)}$$

ILD appears to affect the response angle.(χ^2 (1) = 18.87, p < 0.05)(Table 4.13),

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
ILDr.Null.model	5	20536.09	20564.57	-10263.05	20526.09			
ILDr.Full.model	6	20519.23	20553.40	-10253.61	20507.23	18.87	1	0.0000

Table 4.13: Relationship between ILD and response angle according to the random slope model

	Estimate	Std. Error	t value
(Intercept)	-2.51	1.44	-1.75
ILD	2.32	0.42	5.46

Table 4.14: ILD random slope model Estimate, Std.Error, t-value in ILD Full-model

lowering it by about 2.32 $\pm$ 0.42(standard error)(Table 4.14). H_0 hypothesis is rejected, so the difference of the ILD effect on each Subject is statistically significant.

Figure 4.21 visualizes the relationship between ILD and a response angle by subject. First, it shows the ILD and the response angle (location perception) were analyzed in a proportional relationship. However, the baseline of sound source location perception between the subjects is not the same, and the effect on the ILD varies.

4.3.2 Relationship between other explanatory variables and location perception

To investigate how the shape of the space affects the angle, the shape explanatory variable is added to the linear model (Eq. 4.3.9) and is plotted by shape in Figure 4.22.

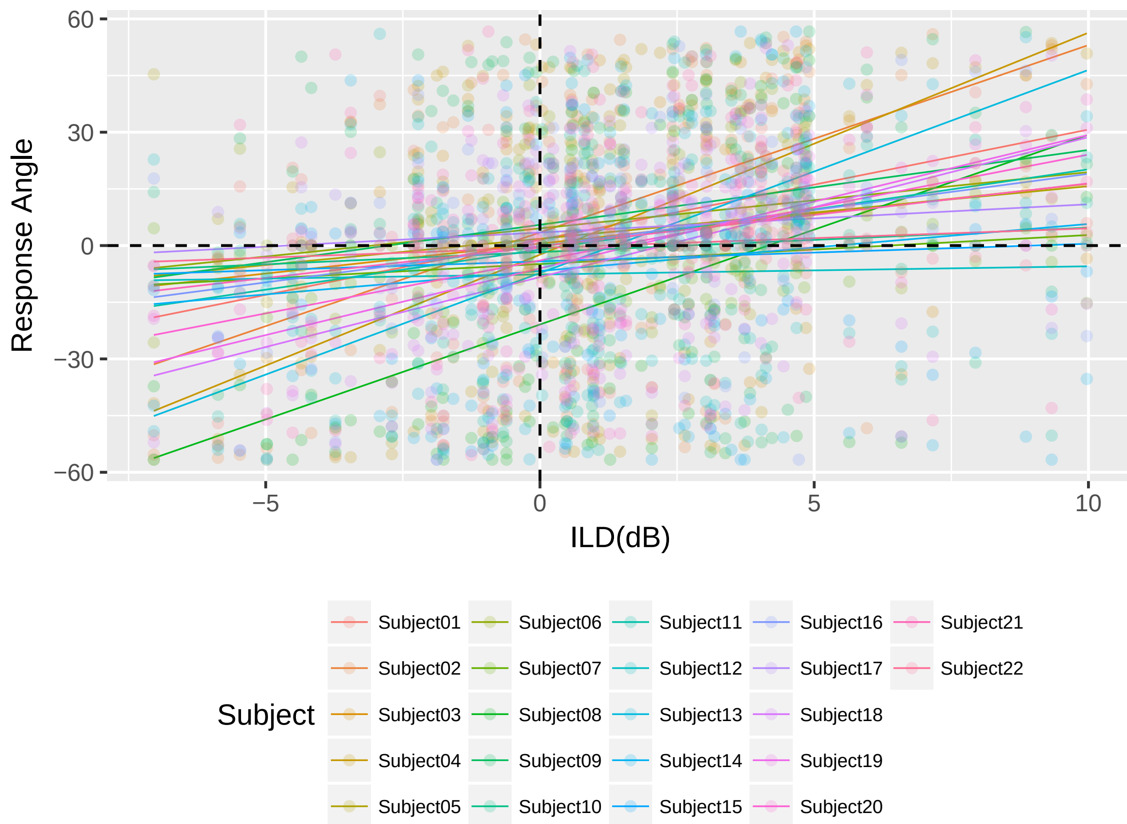


Figure 4.21: ILD full random slope model (Eq. 4.3.8) predict plot by subject

H_0 : Shapes have no effect on the angle responses.

$$\begin{aligned} Null-model &: response\sim \text{ILD} + (1\text{+ILD}|\ \text{Subject}) + \varepsilon \\ Full-model &: response\sim \text{ILD} + \text{Shape} + (1\text{+ILD}|\ \text{Subject}) + \varepsilon \end{aligned} \quad \ldots \quad \text{(Eq. 4.3.9)}$$

As shown in Figure 4.22, the range of ILD values varies depending on each shape (Figure 4.11), and also since angle is proportional to the ILD value, perception of the source position is affected by hall shape. Shape also appears to have a statistically significant effect on angle.(x^2 (1) = 176.61, p < 0.05)(Table 4.16).

	Estimate	Std. Error	t value
(Intercept)	3.28	2.18	1.50
ILD	1.57	0.45	3.48
Shapes2	1.53	1.96	0.78
Shapes3	-3.75	2.79	-1.35
Shapef1	-16.42	2.36	-6.97
Shapef2	-15.40	2.11	-7.28
Shapeo1	0.95	2.30	0.41
Shapeo2	-1.53	1.89	-0.81

Table 4.15: ILD, Shape random slope model Estimate, Std.Error, t-value in ILD Full-model

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
Shape.Null.model	6	20519.23	20553.40	-10253.61	20507.23			
Shape.Full.model	12	20354.61	20422.97	-10165.31	20330.61	176.61	6	0.0000

Table 4.16: Relationship between ILD and response angle according to Shape in random slope model

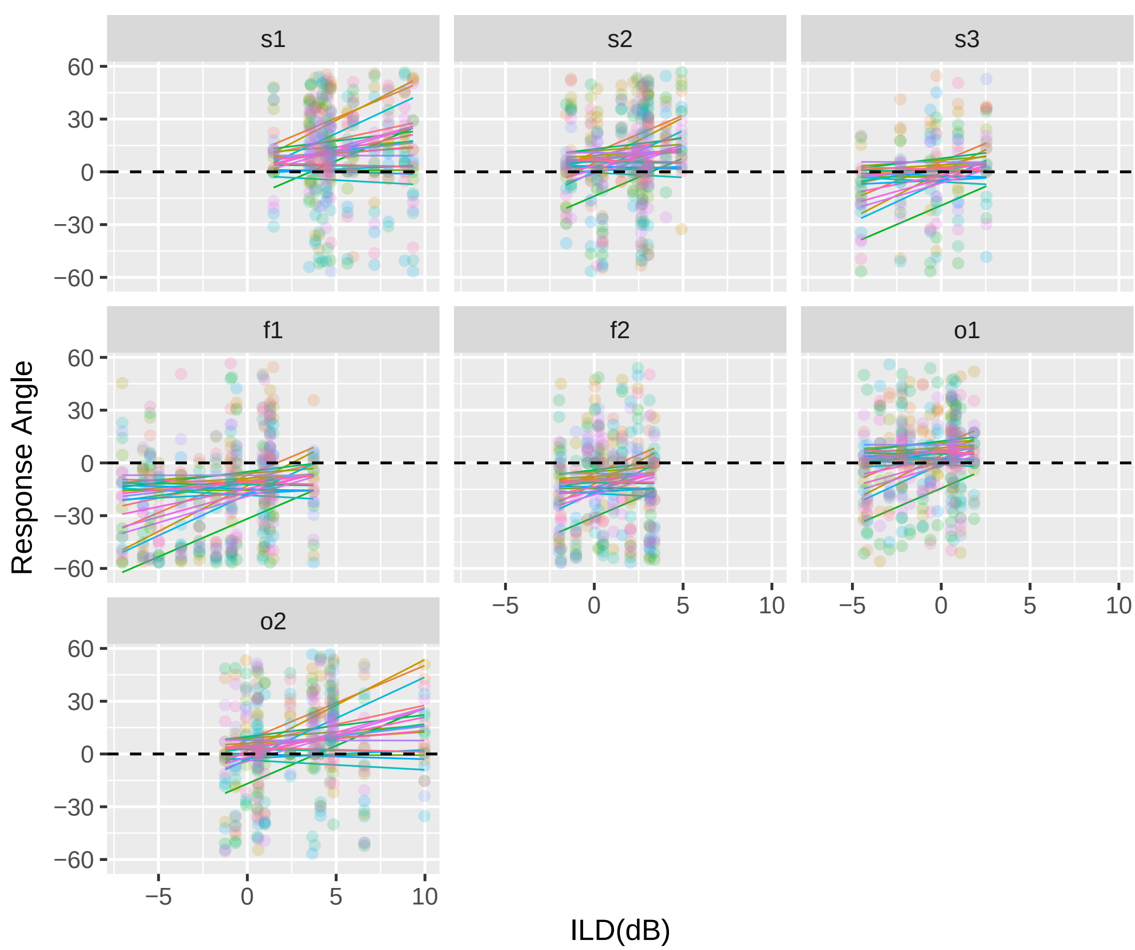


Figure 4.22: ILD random slope model by Shapes

Next, let's check whether the reverberation time affects the response angle. The method is the same as the above. First, a full model and a null model are made.

H_0 : Reverberation time has no effect on the angle responses.

$$RT.Null - model : response \sim \text{ ILD + Shape +(1+ILD| Subject)} + \varepsilon \quad \cdots$$

$$RT.Full - model : response \sim \text{ RT + ILD + Shape + (1+ILD| Subject)} + \varepsilon \quad \text{(Eq. 4.3.10)}$$

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
RT.Null.model	12	20354.61	20422.97	-10165.31	20330.61			
RT.Full.model	13	20320.52	20394.57	-10147.26	20294.52	36.09	1	0.0000

Table 4.17: Relationship between ILD and response angle according to reverberation time

	Estimate	Std. Error	t value
(Intercept)	12.02	2.61	4.60
RT	-6.91	1.14	-6.03
ILD	1.42	0.45	3.14
Shapes2	1.02	1.95	0.52
Shapes3	-5.70	2.79	-2.04
Shapef1	-17.45	2.34	-7.45
Shapef2	-16.51	2.11	-7.84
Shapeo1	-0.02	2.28	-0.01
Shapeo2	-1.90	1.88	-1.01

Table 4.18: ILD and RT Fixed Effects Estimate, Std.Error, t-value in ILD Full-model

Reverberation time (RT) affects the response angle (χ^2 (1) = 36.09, p < 0.05), lowering it by about 1.42 ± 0.45(standard error)(Table 4.17). In Table 4.18, reverberation time and angle response were inversely proportional.

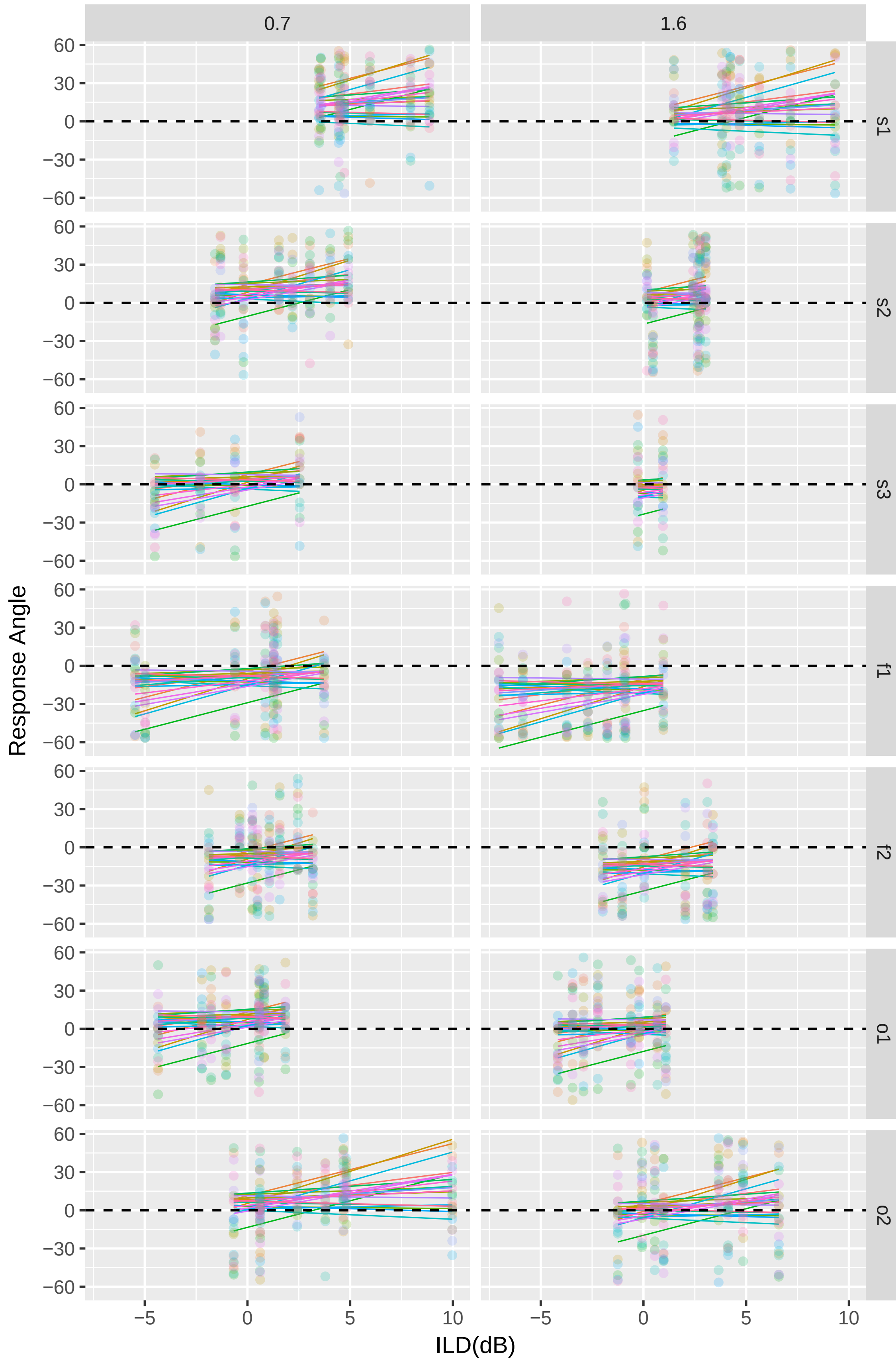


Figure 4.23: RT linear random slope model

$$RT.ILD.model : RT \sim \text{ILD} + \varepsilon \ \cdots \qquad \text{(Eq. 4.3.11)}$$

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.1434	0.0102	111.5517	0.0000
ILD	-0.0091	0.0029	-3.0996	0.0020

Table 4.19: ILD and Reverberation time Estimate, Std.Error, t-value in ILD RT model

Figure 4.23 visualizes the relationship between RT and response angle by subject. Although the shape of the space is the same, the range of the ILD varied when the reverberation time differs. Since the reverberation time is inversely proportional to the ILD value (p-value < 0.05, Eq. 4.3.11, Table 4.19), the response angle is found to increase when the reverberation time decrease.

In the same way, The p-values for distance(distance between center sound source and listener:10, 15m), octave band frequency(500, 1000Hz) are calculated and presented in Table 4.21 to 4.32.

1. Distance: 10m, 15m

   H_0 : Distance between center sound source and listener has no effect on the angle responses.

$$Dist.Null - model : response \sim \text{ILD} + \text{Shape} + \text{RT} + (1 + ILD|\ \text{Subject}) + \varepsilon$$
$$Dist.Full - model : response \sim \text{Dist} + \text{ILD} + \text{Shape} + \text{RT} + (1 + ILD|\ \text{Subject}) + \varepsilon \tag{Eq. 4.3.12}$$

	Estimate	Std. Error	t value
(Intercept)	6.22	3.71	1.68
Dist	0.46	0.21	2.20
ILD	1.41	0.45	3.12
Shapes2	0.98	1.95	0.51
Shapes3	-4.58	2.83	-1.62
Shapef1	-17.51	2.34	-7.48
Shapef2	-16.38	2.10	-7.79
Shapeo1	-0.08	2.28	-0.03
Shapeo2	-1.92	1.88	-1.02
RT	-6.85	1.14	-5.99

Table 4.20: ILD and distance random slope model Estimate, Std.Error, t-value in ILD Full-model

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
Dist.Null.model	13	20320.52	20394.57	-10147.26	20294.52			
Dist.Full.model	14	20317.68	20397.43	-10144.84	20289.68	4.84	1	0.0278

Table 4.21: Relationship between ILD and response angle according to distance

The distance between the center sound source and the listener affected the angular response (χ^2 (1) = 4.84, p(0.03)<0.05), lowering it by about 0.46 $\pm$ 0.21(standard error)(Table 4.21, Table 4.20).

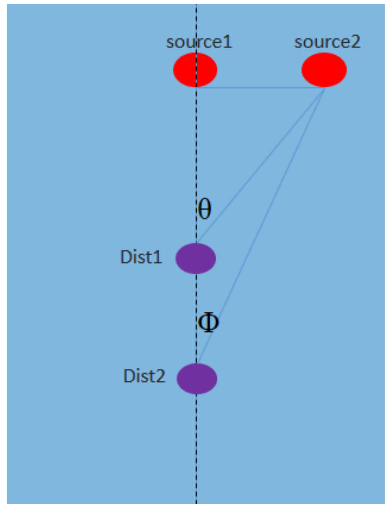


Figure 4.24: Angle of listener and sound source along distance

$$Dist.ILD.model : ILD \sim \text{Dist} + \varepsilon \ \cdots \qquad \text{(Eq. 4.3.13)}$$

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.3991	0.3515	1.1354	0.2563
Dist	0.0698	0.0280	2.4903	0.0128

Table 4.22: ILD and Distance from center source Estimate, Std.Error, t-value in ILD RT model

When the distance from the center source is farther away, the physical angle becomes smaller (Figure 4.24: $\theta > \Phi$). However, it is analyzed that the angle response increases when the distance from the center source is longer. (Eq. 4.3.13, Table 4.22). This analysis indicates that ILD can affect the angle response.

2. Frequency: 500Hz, 1000Hz

   H_0 : Frequency has no effect on the angle responses.

$$Freq.Null - model : response \sim \text{Dist + ILD + Shape + RT +}$$

$$(1 + ILD|\text{ Subject}) + \varepsilon$$

$$Freq.Full - model : response \sim \text{Freq + Dist + ILD + Shape + RT +}$$

$$(1 + ILD|\text{ Subject}) + \varepsilon$$

(Eq. 4.3.14)

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
Freq.Null.model	14	20317.68	20397.43	-10144.84	20289.68			
Freq.Full.model	15	20288.72	20374.16	-10129.36	20258.72	30.96	1	0.0000

Table 4.23: Relationship between ILD and response angle according to Frequency

	Estimate	Std. Error	t value
(Intercept)	-0.86	3.90	-0.22
Freq	0.01	0.00	5.58
Dist	0.44	0.21	2.09
ILD	1.28	0.45	2.83
Shapes2	0.54	1.94	0.28
Shapes3	-6.45	2.83	-2.28
Shapef1	-18.40	2.33	-7.90
Shapef2	-17.40	2.10	-8.30
Shapeo1	-0.91	2.27	-0.40
Shapeo2	-2.23	1.86	-1.20
RT	-7.25	1.14	-6.37

Table 4.24: ILD and Frequency random slope model Estimate, Std.Error, t-value in ILD Full-model

Octave band frequency affected the response (χ^2 (1) = 30.96, $p < 0.05$), lowering it by about 0.01 $\pm$ 0(standard error)(Table 4.23, Table 4.24).

3. Source position: sp2(3m to stage left from sp1), sp3(6m to stage left from sp1)

H_0 : Source position has no effect on the angle responses.

$$Source.Null-model : response\sim \text{Freq + Dist + ILD + Shape + RT +}$$
$$(1 + ILD|\text{ Subject}) + \varepsilon$$
$$Source.Full-model : response\sim \text{Source + Freq + Dist + ILD + Shape + RT +}$$
$$(1 + ILD|\text{ Subject}) + \varepsilon$$

(Eq. 4.3.15)

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
Source.Null.model	15	20288.72	20374.16	-10129.36	20258.72			
Source.Full.model	16	20247.79	20338.93	-10107.90	20215.79	42.93	1	0.0000

Table 4.25: Relationship between ILD and response angle according to source position

	Estimate	Std. Error	t value
(Intercept)	1.82	3.89	0.47
Sourcesp3	-6.61	1.00	-6.58
Freq	0.01	0.00	5.58
Dist	0.43	0.21	2.10
ILD	1.40	0.45	3.11
Shapes2	0.96	1.92	0.50
Shapes3	-5.69	2.80	-2.03
Shapef1	-17.56	2.31	-7.60
Shapef2	-16.85	2.08	-8.11
Shapeo1	-0.13	2.25	-0.06
Shapeo2	-1.94	1.84	-1.05
RT	-7.18	1.13	-6.37

Table 4.26: ILD and source position random slope model Estimate, Std.Error, t-value in ILD Full-model

Source position affected the response (χ^2 (1) = 42.93, p < 0.05), lowering it by about -6.61 ± 1(standard error)(Table 4.25). The effect of the sound source position on the angle response is statistically significant. However, it does not mean that the moving direction of the sound source position is proportional to the angle response.

4. Gender: male, female

H_0 : Gender has no effect on the angle responses.

$$Gender.Null-model : response\sim \text{Source} + \text{Freq} + \text{Dist} + \text{ILD} + \text{Shape} + RT + (1 + ILD|\ \text{Subject}) + \varepsilon$$

$$Gender.Full-model : response\sim \text{gender} + \text{Source} + \text{Freq} + \text{Dist} + \text{ILD} + \text{Shape} + RT + (1 + ILD|\ \text{Subject}) + \varepsilon$$

(Eq. 4.3.16)

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
gender.Null.model	16	20247.79	20338.93	-10107.90	20215.79			
gender.Full.model	17	20248.31	20345.14	-10107.15	20214.31	1.49	1	0.2227

Table 4.27: Relationship table between ILD and response Angle according to gender

Gender does not have a statistically significant effect on the response (χ^2 (1) = 1.49, p > 0.05)(Table 4.27). This is as expected.

	Estimate	Std. Error	t value
(Intercept)	4.25	4.31	0.99
genderM	-3.57	2.82	-1.27
Sourcesp3	-6.61	1.00	-6.58
Freq	0.01	0.00	5.58
Dist	0.43	0.21	2.10
ILD	1.40	0.45	3.11
Shapes2	0.96	1.92	0.50
Shapes3	-5.69	2.80	-2.03
Shapef1	-17.56	2.31	-7.60
Shapef2	-16.85	2.08	-8.11
Shapeo1	-0.13	2.25	-0.06
Shapeo2	-1.94	1.84	-1.05
RT	-7.18	1.13	-6.37

Table 4.28: ILD and gender random slope model Estimate, Std.Error, t-value in ILD Full-model

5. Age

Participants were divided into five age groups. See table 4.5 for Participant demographics.

H_0 : Age has no effect on the angle responses.

$$Age.Null - model : response\sim \text{Source + Freq + Dist + ILD + Shape +}$$

$$RT + (1 + ILD|\text{ Subject}) + \varepsilon$$

$$Age.Full - model : response\sim \text{age + Source + Freq + Dist + ILD + Shape +}$$

$$RT + (1 + ILD|\text{ Subject}) + \varepsilon$$

(Eq. 4.3.17)

Age affected on the response (χ^2 (1) = 6.58, p < 0.05)(Table 4.29), lowering it

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
Age.Null.model	16	20247.79	20338.93	-10107.90	20215.79			
Age.Full.model	17	20243.22	20340.05	-10104.61	20209.22	6.58	1	0.0103

Table 4.29: Relationship between ILD and response angle according to age

	Estimate	Std. Error	t value
(Intercept)	-12.95	6.56	-1.97
age	0.55	0.20	2.77
Sourcesp3	-6.61	1.00	-6.58
Freq	0.01	0.00	5.58
Dist	0.43	0.21	2.10
ILD	1.40	0.45	3.11
Shapes2	0.96	1.92	0.50
Shapes3	-5.69	2.80	-2.03
Shapef1	-17.56	2.31	-7.60
Shapef2	-16.85	2.08	-8.11
Shapeo1	-0.13	2.25	-0.06
Shapeo2	-1.94	1.84	-1.05
RT	-7.18	1.13	-6.37

Table 4.30: ILD and age random slope model Estimate, Std.Error, t-value in ILD Full-model

by about 0.55 ± 0.2(standard error)(Table 4.30).

6. Musical training

Participants were divided into three training groups. See table 4.5 for Participant demographics.

H_0 : Musical training has no effect on the angle responses.

$$Freq.Null - model : response \sim \text{age} + \text{Source} + \text{Freq} + \text{Dist} + \text{ILD} +$$

$$Shape + RT + (1 + ILD|\ \text{Subject}) + \varepsilon$$

$$Freq.Full - model : response \sim \text{training} + \text{age} + \text{Source} + \text{Freq} + \text{Dist} + \text{ILD} +$$

$$Shape + RT + (1 + ILD|\ \text{Subject}) + \varepsilon$$

(Eq. 4.3.18)

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
training.Null.model	17	20243.22	20340.05	-10104.61	20209.22			
training.Full.model	19	20246.94	20355.17	-10104.47	20208.94	0.27	2	0.8720

Table 4.31: Relationship table between ILD and response Angle according to musical training

	Estimate	Std. Error	t value
(Intercept)	-12.46	6.90	-1.81
trainingM	0.72	3.35	0.21
trainingN	1.58	2.81	0.56
age	0.50	0.22	2.25
Sourcesp3	-6.61	1.00	-6.58
Freq	0.01	0.00	5.58
Dist	0.43	0.21	2.10
ILD	1.40	0.45	3.11
Shapes2	0.96	1.92	0.50
Shapes3	-5.69	2.80	-2.03
Shapef1	-17.56	2.31	-7.60
Shapef2	-16.85	2.08	-8.11
Shapeo1	-0.13	2.25	-0.06
Shapeo2	-1.94	1.84	-1.05
RT	-7.18	1.13	-6.37

Table 4.32: ILD and musical training random slope model Estimate, Std.Error, t-value in ILD Full-model

Musical training does not have a statistically significant effect on the response (χ^2 (2) = 0.27, p > 0.05)(Table 4.32). This is as expected.

4.3.3 Interaction

The following result shows if the two factors are inter-dependence (interaction). Likelihood ratio test was used for verification. Age, Shape, Freq, Source, ILD, Shape, Dist, and RT combinations were used in the test. Full model and Reduced model were created in the following manner.

Ex)

$$\begin{aligned} Full.model &: response\sim \text{age} + \text{RT} + (1\text{+ILD}|\ \text{Subject}) + \varepsilon \\ Reduced.model &: response\sim \text{age} * \text{RT} + (1\text{+ILD}|\ \text{Subject}) + \varepsilon \end{aligned} \tag{Eq. 4.3.19}$$

Table 4.33 shows the significant interaction factor combinations.

Factors	χ DF	χ^2	P-Value
age*Shape	6	13.07	0.042
age*Freq	1	3.93	0.048
age*Source	1	7.02	0.008
Source*Shape	6	40.51	0
Freq*ILD	1	4.59	0.032
Dist*Shape	5	33.77	0
ILD*Shape	6	14.87	0.021

Table 4.33: Interaction P-Values

In Table 4.33, the shape has the interaction with other factors which are age, source position, distance, and ILD (p-value is less than 0.05). It meant that the effect of age, source, distance, ILD on the response angle is varied for shapes. Also, when analyzing the relationship between the distance between center sound source and listener and the shape, and source position and the shape, the effect of shape on the response angle for each source positions is statistically different.

The fitting regression model with interaction is as follows.

response $\sim$ age + Source + Freq + Dist + ILD + Shape + RT + age*Shape + age*Freq +age*Source + Freq*ILD + Source*Shape + Dist*Shape + ILD*Shape + (1+ILD| Subject) + ε

4.4 Summary

This chapter investigates how the ILD in indoor spaces is related to the position of the sound source and impacted by other variables, including space shape, RT, distance S-R, frequency. In the seven types of space used in the experiment, the ILD did not change in regular manner along with the location of the source. It means that the ILD value does not always change in the same direction according to the direction of shifting sound source position.

As shown in the experimental results, the relationship between the ILD value and the source position was statistically highly significant. The reverberation time was inversely proportional to the ILD value, and the ILD value was analyzed to be proportional to the location perception response. This means that if the reverberation time is long, it becomes difficult to perceive the sound source location correctly. In fact, most of the participants who participated indicated in the exit interview that it was easier to find sound source location in the reverberation time of 0.7 seconds than for that of 1.6 seconds.

Other experimental factors related to participant demographics (ex: gender and musical training experience) excluding age were observed to have no statistically significant effects on perception of source location in this study.

Chapter 5

Conclusion and Future Work

5.1 Conclusion

The perception of sound source position in an enclosed space is influenced not only by the sound signal emitted by the source but also by environmental variables. In this thesis, the sound transmission pattern and subsequent position perception change according to the surrounding environment. While previous studies have focused on the perception of a single sound source location, this study focused more on the change of perception as the source position varies and proposes a method of evaluating the changing spatial impression according to how the position of the sound source changes.

Since this new metric should show the change of the energy according to the position of the sound source, the proposed metric is based on Interaural Level Difference(ILD) which is an important factor for human location perception. In the free field, the ILD depends on the difference in acoustic energy delivered from a single direction. However, in the enclosed space, since the sound energy can be transmitted to the listener from additional various directions due to reflection, the influence of physical elements is great. The reflected sound energy can be transmitted from directions that are not coincident with the actual position of the sound source, thereby impacting localization.

Therefore, the thesis first investigated how the value of ILD changes with source location according to the shape of space using acoustic simulation and physical measurement methods. In the simulation experiment, models of a fan shape and shoe box shaped halls, which are typical forms of the theater, were created, and the change of the ILD value was observed according to the location of the sound source. In the measurement, changes in ILD values were analyzed in three different shapes of spaces. In both experiments, the changing of ILD values reflected the shape of the space.

In chapter 3, the simulation and measurement results showed that the acoustic energy delivered to the listener depends on the location of the source and also affects ILD values. The existing metrics, LF and $IACC_E$ for spatial impression were found to be insufficient to compare different source positions.

For evaluating the spatial impression of sound sources at different positions on a stage, this thesis used ILD to propose a new metric. The simulation and measurement results show that the ILD contains the location information of the sound source and shows that the value changes according to the shape of the space. For example, it shows the amount of change in the ILD value is not relative to the degree of the sound source movement. In both the simulation and measurement, the sound source was shifted to positions that were 1m apart, but the variation of the ILD value varied less regularly. Furthermore, in some of the results, the change in ILD was observed to be opposite to the direction of sound source movement. For example, the sound source moved to the right of center but the ILD values did not increase as expected. The proposed ILD-CR metric uses the magnitude of the left-right energy ratio based on the ILD, but also the change in the relative position information of the sound sources.

In chapter 4, an auditory experiment was used to explore whether the position perception changed proportionally with the change of ILD. Seven types of space were modeled with the same volume, and the change in ILD values was observed in different acoustic environments. There was a statistically significant relationship between the ILD value and the location perception although there was a difference in the location perception according to the participant. As the ILD value increased, the perception increased proportionally. In the case of the reverberation time, the longer the time (0.7 sec vs. 1.6 sec), the less the change of perception. This means that the long reverberation time makes it difficult to perceive the position of the sound source. The finding that the change of ILD value in the enclosed space is highly correlated with the sound source location perception of the listener shows that the proposed metric, ILD-CR, has proper information for the spatial impression evaluation metric because it uses sound source's location information.

The proposed ILD-CR metric estimates how similar the sound source movement is with positional perception change with space. According to the results of experiments, the perception of the position of the sound source changes in proportion to the ILD value. That is, the patterns of change and perception of ILD have the similar pattern. Based on these results, ILD-CR using similarity (range) and linearity (r) of ILD is considered to be suitable for evaluating the change of location of sound sources in enclosed spaces.

5.2 Recommendations for Future Work

It is difficult to evaluate the spatial impression from a stage due to multiple sound source positions simply by using ASW of each sound source. In this thesis, a method of evaluating the spatial impression across the stage by using the change of perception according to the position of the sound source is proposed. This section will discuss future research directions based on the findings as well as limitations of the presented work.

5.2.1 Test signals

Narrow band signals (500 and 1000 Hz) used in the experiments have been reported to have a significant impact on ASW in previous studies. However, it is not only important to know how each frequency band affects perception, but it is also important to know if changes may be due to the interaction between each frequency. Therefore, a study on how perceptive changes occur through experiments using wide band signals is suggested.

5.2.2 Auditory Experiment Method

The experiment was conducted in such a way that the listener was asked their perception of the test signal at each location of the stage. However, this experimental method can not tell the effect of the interaction when listeners hear multiple sound sources from different locations at the same time. Experiments are suggested to see how listeners respond when multiple sources simultaneously generate signals.

5.2.3 Measurement

The test signals used in the experiments were generated by the acoustic simulation program Odeon. However, even though the acoustic simulation technology is quite developed, auralization differs from sound in the actual environment. Therefore, to observe the response of listeners more precisely, it is necessary to use sound sources recorded in actual spaces. This is necessary not only for the auditory experiment but also for the acoustic data analysis. Chapter 3 analyzed the relationship between the shape of space and acoustic parameters using physical measurements but only in three different spaces. Gathering additional physical data on how acoustic parameters change depending on the shapes of spaces is encouraged.

Bibliography

[1] U. Lehmann Reichardt W. “Raumeindruck als Oberbegriff von Raumlichkeit und Halligkeit”. In: *Acustica* 40.277-289 (1978).

[2] M. Barron and A. H. Marshall. “Spatial impression due to early lateral reflections in concert halls: The deviation of a physical measure”. In: *J. Sound* Vib. 77.211-232 (1981).

[3] L. Rayleigh. “On our perception of sound direction”. In: *Philosophical magazine* 13.214‘232 (1907).

[4] E. B. Newman Wallach Hans and M. R. Rosenzweig. “A Precedence Effect in Sound Localization”. In: *J.Acoust* 21.4.468 (1949).

[5] Durand R Begault. *3-D Sound for Virtual Reality and Multimedia*. Boston: AP Professional, 1995.

[6] Jens Blauert. *Spatial Hearing*. MIT Press, 1983.

[7] Brian C. J. Moore. *An Introduction to the Psychology of Hearing*. Leiden: Brill, 2013.

[8] Strutt J. “On out perception of sound direction”. In: *Philos. Mag* 13.214-232 (1907).

[9] D. C. Teas W. E. Feddersen T. T. Sandel and L. A. Jeffress. “Localization of high-frequency tones”. In: *Acoust.* Soc. Am., 29.988-991 (1957).

[10] William M. Hartmann and Zachary A. Conbstan. “Interaural level differences and the level-meter model”. In: *Acoust.* Soc. Am., 112.1037-1045 (2002).

[11] Bosun. Xie. *Head-Related Transfer Function and Virtual Auditory Display*. J Ross, 2013.

[12] Vaillancourt M. M Shaw E. A. G. “Transformation of sound-pressure level from the free field to the eardrum presented in numerical form”. In: *Acoust.* Soc. Am., 78.1120-1123 (1985).

[13] W. M. Zhang P. X. Hartmann. “On the ability of human listeners to distinguish between front and back”. In: *Hearing Research* 260.30-46 (2010).

[14] M. Vorlander. “Past, Present and Future of Dummy Heads”. In: *conference of Federation of the iberoamerican acoustical societies* (2004).

[15] R.M. Sachs M.D. Burkhard. “KEMAR the Knowles Electronics Manikin for Acoustic Research”. In: *Report No. 20032-1* iNDUSTRIAL rESEARCH pRODUCTS,inC. (1972).

[16] Raatgever F. Ptter J. and J. Bilsen. “Measures for spaciousness in room acoustics based on a binaural strategy”. In: *Acta Acust.* China.429-443 (1995).

[17] John S. Bradley and Gilbert A. Soulodre. "The Influence of Late Arriving Energy on Spatial Impression". In: *J.Acust. Soc. Am.* 97.4.2263-2271 (1995).

[18] R.D. Reich J.S. Bradley and S.G. Norcross. "On the combined effects of early- and late-arriving sound on spatial impression in concert halls". In: *J.Acust. Soc. Am.* 108.2.651-61 (2000).

[19] M. Barron. "Measured Early Lateral Energy Fractions In Concert Halls And Opera Houses". In: *Journal of Sound and Vibration* 232.1.79-100 (2000).

[20] M. Barron. "Late Lateral Energy Fractions and the Envelopment Question in Concert Halls". In: *Applied Acoustics* 62.2.185-202 (2001).

[21] Takayuki Hidaka Toshiyuki Okano Leo L. Beranek. "Relations among interaural cross-correlation coefficient (IACCE), lateral fraction (LFE), and apparent source width (ASW) in concert halls". In: *Acoustical Society of America* S0001-4966∼98 05907-4.43.55.Fw, 43.55.Gx, 43.55.Hy (1998).

[22] Marshall AH. "A note on the importance of room cross-section in concert halls". In: *J. Sound.Vib.* 5(1).100-112 (1967).

[23] Yoichi. Ando. *Architectural Acoustics: Blending Sound Sources, Sound Fields, and Listeners.* New York: AIP, 1998.

[24] J. Blauert and W. Lindemann. "Spatial Mapping of intracranical auditory events for various degrees of interaural coherence". In: *J.Acoust* Soc.Am.79 (3) (1986).

[25] M. Morimoto and Christoph Posselt. "Contribution of reverberation to auditory spaciousness in concert halls". In: *J. Acoust.* Soc. Jpn. 10.87-92 (1988).

[26] M. Morimoto and Z. Maekawa. "Effects of low frequency components on auditory spaciousness". In: *Acustica* 66.190-196 (1988).

[27] J.S. Bradley and R. E. Halliwell. "Accuracy and reproducibility of auditorium acoustics measure". In: *Proc. INst. of Acoustics* 10.399-406 (1988).

[28] M. Morimoto and K. Iida. "propriate frequency bandwidth in measuring interaural cross-correlation as a physical measure of auditory source width". In: *Acoust. Sci. Tech* 26.179-184 (2005).

[29] Irwin Pollack and W.J. Trittipoe. "Binaural Listening and Interaural Noise Cross Correlation". In: *J. Acoust.* 31.1250-1252 (1959).

[30] W. Reichardt and W. Schmidt. "Die horbaren Stufen des Raumeindruckes bei Musik". In: *Acustica* 17.175-179 (1966).

[31] W.J. Davies T.J. Cos and Y.W. Lam. "The sensitivity of listeners to early sound field changes in auditoria". In: *Acustica* 79.27-41 (1993).

[32] D. de Vries. "Spatial fluctuations in measures for spaciousness". In: *J.Acust. Soc. Am.* 110.947-956 (2001).

[33] R.M. Hershkowitz and N.I. Durlach. "Interaural Time and Amplitude JNDs for a 500Hz Tone". In: *J.Acust. Soc. Am.* 46.1464-1467 (1969).

[34] A. W. Mills. "Laterization of high-frequency tones". In: *J. Acoust.* Soc. Am. 32.132-134 (1960).

[35] Lloyd A. Jeffress Dennis McFadden and Harold L. Ermey. “Difference in Interaural Phase and Level in detection and Lateralization:250Hz”. In: *J.Acust. Soc. Am.* 50.1484-1493 (1971).

[36] D.Wesley Grantham. “Interaural intensity discrimination:insensitivity at 1000Hz”. In: *J.Acust. Soc. Am.* 75.1191-1194 (1984).

[37] William A. Yost and Raymond Dye. “Discrimination of interaural differences of level as a function of frequency”. In: *J.Acust. Soc. Am.* 83.1846-1851 (1988).

[38] B. Shinn-Cunningham and K. Kawakyu. “Neural Representation of Source Direction in Reverberant Space”. In: *2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics* IEEE Cat.No.03TH8684 (2003).

[39] S. Devore and B. Delgutte. “Effects of Reverberation on the Directional Sensitivity of Auditory Neurons across the Tonotopic Axis: Influences of Interaural Time and Level Differences”. In: *Journal of Neuroscience* 30.23.7826-837 (2010).

[40] Antje Ihlefeld and Barbara G. Shinn-Cunningham. “Effect of Source Spectrum on Sound Localization in an Everyday Reverberant Room”. In: *J.Acust. Soc. Am.* 130.1.324-33 (2011).

[41] Stefan Klockgether and Stevenvande Par. “Just noticeable differences of spatial cues in echoic and anechoic acoustical environments”. In: *J. Acoust.* 140,Issue 4 (2016).

[42] G Von Bekesy. “Uber das Richtungshoren bei einer Zeitdifferenz oder Lautstarkeungleichheit der beidseitigen Schalleinwirkungen”. In: *Phys* Z. 31.824-838,857-868 (1930).

[43] Karl Pearson. "Note on Regression and Inheritance in the Case of Two Parents". In: *Proceedings of the Royal Society of London* 58.-1.240-42 (1895).

[44] John Eargle. *The Microphone Book*. Oxford: Focal, 2004.

[45] John Fox. *Applied Regression Analysis and Generalized Linear Models*. Los Angeles: SAGE, 2016.

[46] Nan M. Laird and James H. Ware. "Random-Effects Models for Longitudinal Data". In: *Biometrics* 38.4.963 (1982).

Appendix A

Matlab Code

1 GUI Matlab code

```
function varargout = subtestgui(varargin)
% SUBTESTGUI MATLAB code for subtestgui.fig
%      SUBTESTGUI, by itself, creates a new SUBTESTGUI or raises the existing
%      singleton*.
%
%      H = SUBTESTGUI returns the handle to a new SUBTESTGUI or the handle to
%      the existing singleton*.
%
%      SUBTESTGUI('CALLBACK',hObject,eventData,handles,...) calls the local
%      function named CALLBACK in SUBTESTGUI.M with the given input arguments.
```

```
%
%      SUBTESTGUI('Property','Value',...) creates a new SUBTESTGUI or
    raises the
%      existing singleton*.  Starting from the left, property value
    pairs are
%      applied to the GUI before subtestgui_OpeningFcn gets called.  An
%      unrecognized property name or invalid value makes property
    application
%      stop.  All inputs are passed to subtestgui_OpeningFcn via
    varargin.
%
%      *See GUI Options on GUIDE's Tools menu.  Choose "GUI allows only
    one
%      instance to run (singleton)".
%
% See also: GUIDE, GUIDATA, GUIHANDLES

% Edit the above text to modify the response to help subtestgui

% Last Modified by GUIDE v2.5 09-May-2016 11:33:16

% Begin initialization code - DO NOT EDIT
gui_Singleton = 1;
gui_State = struct('gui_Name',       mfilename, ...
                   'gui_Singleton',  gui_Singleton, ...
                   'gui_OpeningFcn', @subtestgui_OpeningFcn, ...
                   'gui_OutputFcn',  @subtestgui_OutputFcn, ...
```

```
                    'gui_LayoutFcn',  [] , ...
                    'gui_Callback',   []);
if nargin && ischar(varargin{1})
    gui_State.gui_Callback = str2func(varargin{1});
end

if nargout
    [varargout{1:nargout}] = gui_mainfcn(gui_State, varargin{:});
else
    gui_mainfcn(gui_State, varargin{:});
end
% End initialization code - DO NOT EDIT


% --- Executes just before subtestgui is made visible.
function subtestgui_OpeningFcn(hObject, eventdata, handles, varargin)
% This function has no output args, see OutputFcn.
% hObject    handle to figure
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
% varargin   command line arguments to subtestgui (see VARARGIN)

% Choose default command line output for subtestgui
handles.output = hObject;

% Update handles structure
guidata(hObject, handles);
```

```


%slider intial
set(handles.inum,'String',get(handles.slidernum,'Value'));
set(handles.slidernum,'value',0);

% UIWAIT makes subtestgui wait for user response (see UIRESUME)
% uiwait(handles.figure1);
%test data read
global fnames     %wav file name
global numfids    %wav file amount
global s           %wav file
global Fs          %sample rate
global K
global tss        %test set amount(total 52)
global j
global sig
global sel
global testset
global setnum
global sliderVal
global setitnum


%load wav file

sliderVal=0;
```

```
j=1;
sel=0;
fnames = dir('*.wav');
numfids = length(fnames);
s = cell(1,numfids);

for K = 1:numfids
 [s{K},Fs] = audioread(fnames(K).name);     %load wav file to s

end

%test data order read

tss=xlsread('testset.xls','sheet3');   %testset number

testset=xlsread('testset.xls','sheet4');   %testset order latin square

setitnum=xlsread('testset.xls','setitnum');     %testset order number(
    tasker number, row number)

setnum=[testset(setitnum,:)  , testset(setitnum+21,:)];  %testset
    question number



xlswrite('testresult.xls',setitnum,['Result',num2str(setitnum)],'A1:A1')
     %testset order number write
```

```
xlswrite('testresult.xls',setnum,['Result',num2str(setitnum)],'A2:AZ2') %testset question number write

set(handles.setnumid,'String',sprintf('# %d  / 104',j));    % show test question number at screen
%set(handles.setnumid,'String',j);




% --- Outputs from this function are returned to the command line.
function varargout = subtestgui_OutputFcn(hObject, eventdata, handles)
% varargout  cell array for returning output args (see VARARGOUT);
% hObject    handle to figure
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Get default command line output from handles structure
varargout{1} = handles.output;


% --- Executes on button press in reference.
function reference_Callback(hObject, eventdata, handles)
% hObject    handle to reference (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
global fnames
```

```
global numfids
global s
global Fs
global K
global tss
global j
global sig
global setnum
jj=0;
jj=setnum(j)*2-1;

%ss=[s{tss(jj)}; s{tss(jj+1)}];
sound(s{tss(jj)}*1,Fs)                %replay reference
set(handles.infom,'String','Using the slide bar, select the location of
    the comparision, or if you want to hear it again, click Reference
    and Comparison again')

% --- Executes on button press in comp.
function comp_Callback(hObject, eventdata, handles)
% hObject    handle to comp (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
global fnames
global numfids
global s
global Fs
global K
```

```
global tss
global j
global sig
global setnum
jj=0;
jj=setnum(j)*2-1;

%ss=[s{tss(jj)}; s{tss(jj+1)}];
sound(s{tss(jj+1)}*1,Fs)                    %replay conparison
set(handles.infom,'String','Using the slide bar, select the location of
    the comparision, or if you want to hear it again, click Reference
    and Comparison again')


% --- Executes on slider movement.
function slidernum_Callback(hObject, eventdata, handles)
% hObject    handle to slidernum (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)

% Hints: get(hObject,'Value') returns position of slider
%        get(hObject,'Min') and get(hObject,'Max') to determine range of
    slider
global sliderVal
global j

sliderVal(j)=get(hObject,'Value');
```

```
assignin('base','sliderVal',sliderVal(j));      %slider bar value assign
set(handles.inum,'String',num2str(sliderVal(j)));  %show slider bar
    value to screen


% --- Executes during object creation, after setting all properties.
function slidernum_CreateFcn(hObject, eventdata, handles)
% hObject    handle to slidernum (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    empty - handles not created until after all CreateFcns
    called

% Hint: slider controls usually have a light gray background.
if isequal(get(hObject,'BackgroundColor'), get(0,'
    defaultUicontrolBackgroundColor'))
    set(hObject,'BackgroundColor',[.9 .9 .9]);
end


% --- Executes on button press in next.
function next_Callback(hObject, eventdata, handles)
% hObject    handle to next (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
global j
global sel
global sliderVal
```

```
global setnum
global setitnum


if sel==0                                %without 'Next' button'
     set(handles.infom,'String','Please, input number & click the select
    button')

elseif sel==1                            %with 'Next' button'
     set(handles.inum,'String',0);
     set(handles.infom,'String','Please, click Reference and Comparison')
        xlswrite('testresult.xls',sliderVal,['Result',num2str(setitnum)],'
    A3:AZ3')
        xlswrite('testset.xls',setitnum+1,'setitnum')

     set(handles.slidernum,'value',0);

        j=j+1;
       %set(handles.setnumid,'String',j);
         set(handles.setnumid,'String',sprintf('# %d  / 104',j));
     sel=0;

elseif j==52                             % when it is completed
     xlswrite('testresult.xls',sliderVal,['Result',num2str(setitnum)],'A3
    :AZ3')
     set(handles.infom,'String','Thanks, test completed')
     xlswrite('testset.xls',setitnum+1,'setitnum')
```

```
end



% --- Executes on button press in select.
function select_Callback(hObject, eventdata, handles)
% hObject    handle to select (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
global sel
if sel==0
sel=1;
set(handles.infom,'String','Please, click Next button or if you want to
    hear again, re-click Select and then click Reference and Conarison
    again')
elseif sel==1
    sel=0
    set(handles.infom,'String','Please, input number & click Select')
end
```

2 ILD Calculation Matlab code

```
clear
tg=0.08;  %cutoff time
CF=[125, 250, 500, 1000,2000,4000, 8000]; %frequency band
aname=['c1.wav','c2.wav','c3.wav','c4.wav','c5.wav','c6.wav','c7.wav'];
aa=1;
```

```
bb=0;

%oct=1; %one octave band

fnames = dir('*.wav');
numfids = length(fnames);
ss = cell(1,numfids);

for K = 1:numfids
 [ss{K},Fs] = audioread(fnames(K).name);     %load wav file to s


for aa=1:7
%while (aa<8)
     bb=bb+1;

%Left sound energy

s = ss{K}(:,1);

Fc=CF(aa);
oct=1;                            % Octave band select one octave
h=IrFil(s,Fs,Fc,oct);           % Filtering impulse response


hh=abs(hilbert(h)); %make smooth curve
t=[0:1/Fs:(length(hh)-1)/Fs]; %sound source time length cal
```

```
L(length(hh):-1:1)=0;
E(length(hh):-1:1)=0;
L(length(hh):-1:1)=[];

%Energy curve plot
e=20*log10(hh./max(hh));

%Schroeder integration
c=1.5;
tl=1/(c*0.1);
te=[0:1/Fs:round((length(hh)-1)/tl/Fs)];  %Schroeder curve time length
    cal, length pro->round
L(round(length(hh)/tl):-1:1)=(cumsum(hh(round(length(hh)/tl):-1:1).^2)./
    sum(hh(1:round(length(hh)/tl))).^2);
E=10*log10(L./max(L));


%%Acoustic parameter Calculation
for i=round((length(hh)-1)/tl) :-1:1  %%Find the time when the direct
    sound starts
    E0=10*log10(L(i)./max(L));
            if E0>=max(E)-0.001
                t00=i;
            break;
    end
end

```

```
Ec50=cumsum(hh(t00+(tg*Fs):-1:t00).^2);
LLL=max(Ec50);


%Right sound energy

s=0;
te=0;
t=0;

s=ss{K}(:,2);
FcR=Fc;
oct=1;                          % Octave band select one octave
hR=IrFil(s,Fs,FcR,oct);         % Filtering impulse response

hhR=abs(hilbert(hR)); %make smooth curve
t=[0:1/Fs:(length(hhR)-1)/Fs]; %sound source time length cal
LR(length(hhR):-1:1)=0;
E(length(hhR):-1:1)=0;
LR(length(hhR):-1:1)=[];

%Energy curve plot
bb=bb+1;
eR=20*log10(hhR./max(hhR));

%Schroeder integration
```

```
tl=1/(c*0.1);
te=[0:1/Fs:round((length(hhR)-1)/tl/Fs)];  %Schroeder curve time length
    cal, length pro->round
LR(round(length(hhR)/tl):-1:1)=(cumsum(hhR(round(length(hhR)/tl):-1:1)
    .^2)./sum(hhR(1:round(length(hhR)/tl))).^2);
ER=10*log10(LR./max(LR));


 Ec50R=cumsum(hhR(t00+(tg*Fs):-1:t00).^2);
 RRR=max(Ec50R);
 LLL;


ratioLR=(10*log10(RRR./LLL)); %ILDc calculate

%print

Ec50=0;
Ec50R=0;

ToratioLR(K,aa)=ratioLR;
ToLLL(K,aa)=LLL;
ToRRR(K,aa)=RRR;
end

end
%print to excel file
```

```
xlswrite('ILD10.xls',ToratioLR,'ILD')
xlswrite('ILD10.xls',ToLLL,'LLL')
xlswrite('ILD10.xls',ToRRR,'RRR')
```

3 Impulse Response filter Matlab code

```
function [ fil_data ] = IrFil(s,Fs,Fc,oct )
if Fc <= 125
    n=2;
else
    n=3;
end

if oct==1
    a=2^(1/2);
elseif oct==1/3
    a=2^(1/(2*3));
else
    error('Error!You should put 1 or 1/3')
end

ir_limits=[Fc/a,Fc*a]/(Fs/2); %Fs/2-->Nyquist
[coef_b,coef_a]=butter(n,ir_limits);% bandpass filter
fil_data=filter(coef_b,coef_a,s);

end
```