THE EFFECT OF STIMULI AND ANGLES ON THE PERCEPTION OF ECHO THRESHOLDS

LEE DAVIS

A thesis submitted to the University of Huddersfield
in fulfilment of the requirements for
the degree of Masters by Research

The University of Huddersfield

November 2016

Abstract

This paper looks into comparisons of time differences recorded for echo thresholds under differing stimuli, angles and listener instructions. Previous research has focused on echo thresholds primarily with regards to level difference or a limited combination of these variables. Contrasting listener instructions between research such as "echo barely audible" and "echo clearly audible" has been shown to produce different thresholds. The former instruction resulted in summing localisation being considered and lower thresholds.

Listeners manipulated sliders on a GUI to reduce the time difference between two randomly selected loudspeakers. Two tests were undertaken to grade the beginning of separation with fusion still evident and complete separation. Orchestral, pink noise burst and speech stimuli were used as continuous, transient and familiar sources respectively. 17 loudspeakers angles were available in total, however a single angle per side of the median plane was chosen randomly by the GUI which produced 10 angles per test. The listener sat in the centre of the room with speakers radiating around them at 0°, ±30°, ±60°, ±90°, ±120°, ±150° and 180° azimuthal intervals at 0° elevation and 0°, ±30° and ±110° at 30° elevation, replicating common multichannel surround setups. The lead sound was presented from the speaker directly in front of the listener at 0° azimuth and 0° elevation.

A Paired-Samples Sign Test was used for significance testing of median differences between graded echo thresholds. There were clear median differences between tests when the marking criteria was different. The orchestral stimulus was overall significantly different to the pink noise and speech stimuli in the fusion test. There were significant differences for half of the angles (those within the median plane or relatively behind the listener position) for the orchestral and pink noise comparison in the separation test. Significant differences were apparent for the majority of angles in the separation test between the orchestral and speech stimuli. For both tests, the pink noise and speech comparison showed no significant differences. Limited significant differences were noted between angles. Median plane angles for the lag sound showed increased echo thresholds.

Table of Contents

Word count: 14,839

List of Tables

List of Figures

Acknowledgements

I would like to offer my greatest appreciation to my supervisor Dr. Hyunkook Lee who has consistently provided guidance and assistance during the difficult moments and always been available to provide critical input. I would also like to thank all of the participants in my listening test for the constructive criticism, enthusiasm and time. Finally, I would like to thank my mother, father and brother for all of their support and understanding.

0 Introduction

An echo is the occurrence of a single or multiple replications of a sound apparent due to a time difference in the arrival at a receiver, such as the human ears, of the same sound from different directions. As sound energy is emitted from a source, such as a loudspeaker, it meets various barriers and obstacles which vary the time of arrival of the sound at a listener's ears. This is processed by the listener and in turn portrays a perception of an acoustic space. Although some reflections may not necessarily be apparent to a listener, they may still have an effect on the overall quality or perception of sound. The primary interest for this research is the perception of an echo rather than its true presence. It deals with psychoacoustics which is the study of the perception of sound. Efficient reproduction of sound fields benefits from dismissal of less important information while retaining that which a human listener can perceive. Dealing with reflections which are only perceptually relevant can in turn save time and provide a better understanding of important audio cues.

Echoes are a difficult physical property of sound to definitively understand and therefore define for every context. In part due to the complexity of it's properties in an environment and our own psychological interpretation. Subjective analysis of a sound field can vary greatly to the actual properties recorded. Even between subject results may be highly disagreeable as familiarity, expectations and hearing inconsistencies can drastically alter the perceptions. We perceive multiple sound changes with varying time differences between sound sources. As the difference increases from zero, the apparent location of the sound source will gradually shift to the earlier sound source location. As it increases further, the later sound source will then not be perceived at all. Beyond this, an echo will begin to be apparent. It becomes difficult for researchers to agree upon when an echo truly begins as there is not a direct time difference where an echo immediately occurs, instead there is typically a gradual transition. The research into the seemingly simple question of the point of separation is undeniably vast, as evidenced by the multiple papers researching different separation descriptors. These are outlined in full in Chapter 1.2. Reflections in context to various objective measures such as those described in this thesis and their ultimate impact on spatial impression have been widely researched, but to reach a simple model which brings these definitions together would be a tricky task. It is interesting to consider whether reflections in objective measures such as those for ASW and LEV are in fact perceivable echoes.

This research intends to combine and expand on the independent variables used in previous research to provide a better understanding of echo thresholds. The time difference is dependent on many factors which include different sound sources (stimuli) and incident sound positions around the acoustic space (angle). An investigation will be performed into the variability of recorded echo thresholds based on these independent variables and instructions provided to listeners. It is hoped that any large discrepancies between echo thresholds with under-researched variables or methods will be highlighted which should encourage further investigative research. There however may be connections with previous research and the data acquired. The research presented has implications for 3D reproduction. Although limited angles are utilised in the top ring at 30° elevation, good estimations are available for intermediate azimuthal angles at this elevation.

Chapter 1 begins with the literature review, outlining the precedence effect and the differences that separate sound images feature as the delay increases. The definition of echo thresholds and researchers' methods of recording it are explored. Chapter 2 continues with a discussion on the general perception of space via the term spatial impression and the key objective measures. Further information is provided regarding these concepts but along the median plane. Our perception of distance and depth is described. Chapter 3 outlines the experiment, beginning with the methodology and intentions for the research. Listeners will be presented with multiple sound scenes based on a direct source from the front of the listener and an echo at different angles. Orchestral, pink noise and speech stimuli variables will be additionally tested. The listener will be asked to reduce the time difference of the echo to the last delay value where an echo can still be heard. There will be two tests each with its own grading criteria for an echo threshold. Test one involves marking where two sound images can be just heard yet there is fusion still present. Test two involves marking the beginning of complete separation. The results are presented next, followed by the critical analysis of the findings. The limitations of the research and future intentions are laid out towards the end of the chapter. Finally in Chapter 4, a summary of the research and the knowledge gained is outlined.

Reflections are a fundamental component of many objective measures for spatial impression. To achieve a better understanding, the research gaps in this basic component need to be tackled. The following motivation and research questions are therefore posited:

- Previous research on echo thresholds (Boerger, 1965a, 1965b; Haas, 1951; Masayuki Morimoto, 2002; Rakerd, Hartmann, & Hsu, 2000) typically focused on reducing the level difference between the lead and lag sound rather than the time difference, therefore the level difference will remain at equal intensity and the results will be compared. When time differences are used as an independent variable, often they are at set values which suggests the research is less interested in the accuracy of this variable compared to others such as sound intensity. Would the ability to finely pick the time difference at more numerous linear angles reveal any patterns?

- Source image depth is considered to extend in front of and behind the listener, however subjective depth cannot be confirmed to reside solely within the median plane although it is certainly an important plane. The research in this thesis will be useful for depth perception research as echo thresholds need to be known first to create depth stimuli. There are only two notable studies which look into echo thresholds in the median plane (Agaeva & Al'tman, 2008; Rakerd et al., 2000). The differences between median plane and horizontal plane angles are curious for varying echo thresholds. Should the angles be treated differently or is it possible to rely on similar delay values?

- Home multichannel systems are becoming more popular due to the increasing interest in 3D sound configurations such as Auro 3D and Dolby Atmos. Therefore testing varying angles at set increments and echo thresholds with height channels would be advantageous. It may be possible to limit the necessity of certain loudspeaker configurations to arrive at an equivalent perception of sound space. This would save time and money for consumers and conclusively encourage surround setups more actively without it appearing to be an overwhelming process. Many precedence effect studies used a dichotic (unique sounds in the left and right ear) headphone presentation (Blauert, 1983; Litovsky, Colburn, Yost, & Guzman, 1999). Furthermore, only a single precedence study is noted using more than two speakers in the azimuthal plane (Litovsky, Rakerd, Yin, & Hartmann, 1997). Presentation of multiple sounds through loudspeakers may produce different results in comparison to a headphone setup due to effects such as interchannel crosstalk and pinna filtering. Will the results directly relate to those from headphone studies?

- Stimuli are believed to be the primary determinant of echo threshold, yet few studies used multiple stimuli in a single study (Litovsky, Colburn, et al., 1999). As an example, Roy and Gimbott (1993) looked into anechoic speech and music. Clicks are by far the most common stimulus in echo threshold related research (Babkoff & Sutton, 1966; Ebata, Sone, & Nimura, 1968; Freyman, Clifton, & Litovsky, 1991; Guttman, 1962; Klemm, 1920; Litovsky, Dizon, & Colburn, 1999; Rosenzweig & Rosenblith, 1950; Yang & Grantham, 1997). Speech is the next most common stimulus tested (Cherry & Taylor, 1954; Haas, 1951; Lochner & Burger, 1958; Meyer & Schodder, 1952). It would be good to avoid researching clicks further. However, which is the most important stimulus and is it a crucial independent variable? It is more trusted to test variables in the same research rather than try to cross-reference research which would risk complications arising due to largely contrasting methods.

1 Localisation in an Enclosed Space

1.1 Precedence Effect

As described briefly in the introduction, as the delay or sound energy between two sounds increases, there are many perceivable effects until an echo is produced. When there is no delay or level difference, the sound image is perceived as a phantom image between the two sound locations. Just noticeable differences between the direct sound and single reflections are noticed between 630 μs and 1 ms (Blauert, 1983) as the delay increases. This is where 'summing localisation' occurs and it is the lower boundary of the range of delays where the law of the first wavefront remains applicable (described below). Summing localisation indicates a single sound is still perceived, however the direction of the sound shifts towards the source. It is dependent on the intensity and time differences between the direct source and reflection in addition to the direction of incidence for both. It is possible to approximate any direction by utilising varied differences — time differences can be substituted for intensity differences and vice versa. In this delay range, the lag sound contributes to this perception. The research in this paper concerns beyond summing localisation as we are not looking for merely a change in direction but at least a suggestion that there are separate images audible. The masked threshold is also important to note here as it is related closely to the beginning of the precedence effect. It is described by Blauert (1983) as "...the level difference between the primary sound and the reflection at which the reflection becomes 'absolutely' audible", however once again this is a definition based on sound level rather than temporal differences. Listeners may use colouration or movement in the sound image to acquire this perception. Beyond approximately 1 ms, as the delay time increases further, a gradual shift of sound image occurs in the direction of the sound source which precedes in time as the lag sound appears to disappear. This is known as the 'law of the first wavefront' or 'precedence effect' which was originally termed by Wallach, Newman, and Rosenzweig (1949). Between 1 – 5 ms (for clicks) the sound image remains perceptively fused and is located towards the direction of the leading speaker (Litovsky, Colburn, et al., 1999). The lag sound no longer has an immediate role in the impression of the direction, yet it understood to still be present. The precedence effect typically disappears at delay values greater than 50 ms for speech. The Haas effect is a specific instance of this researched by Haas (1949) for sounds arriving between 25 – 35 ms who noted that a reflection can be up to 10 dB louder than the lead sound and still not be perceivable. Masayuki Morimoto (2002) tested the upper limit of the precedence effect which was defined as the 'image split'. A music motif stimulus was used and the level of the reflections at angles of $\pm 135°$ were changed in comparison to the direct sound position in front of the listener. It was confirmed that reflections beyond the precedence effect do not always create a sensation of separation but may contribute to listener envelopment (LEV). It has been found for the precedence effect that when the direct sound appears from the same direction as the listener faces, there are no significant differences between reflections (Litovsky, Colburn, et al., 1999). The upper limit of the precedence effect is known as the "echo threshold" and is described in the next chapter.

The research presented by the author aims to relate the different stages of the precedence effect to the listener instruction criteria presented. It is expected that the experiments will provide further information as to the specific cut-off point for multiple stimuli where they begin to feature a breakdown in the precedence effect. The definitions regarding this breakdown however are confusing. For example, once the law of the first wavefront ceases to be evident this may not necessarily mean that the lead and lag sound are immediately perceived as separate images. It is reasonable to imagine a transition period begins with those limits defined by the two tests described in this research.

1.2 Echo Threshold

Beyond the upper limit of the precedence effect is where echoes begin to appear (Blauert, 1983). As outlined in the previous chapter, it is possible that there is a crossover point towards the upper limit of the precedence effect where separation is gradually perceived. The echo threshold has been confirmed as being more difficult to define than summing localisation. This confusion is certainly confirmed by the research of others who have used multiple definitions and methods which vary greatly such as "primary auditory event and equally loud", "echo annoying" and "threshold of indistinction" (Blauert, 1983) which points to different researchers beliefs of the cut-off point. Listener instructions furthermore create different results. Meyer and Schodder (1952) used the criterion "echo barely audible" whereas Lochner and Burger (1958) instead used "echo clearly audible". This resulted in the thresholds for the latter research being naturally larger (Blauert, 1983). These two criteria are similar to the listener instructions provided for the present research however time delay will be considered in contrast. This variability between instructions highlights the importance of an accurate specification for grading echo thresholds.

Blauert defines the setup for echo thresholds as incorporating a base angle between speakers of $\alpha = 80°$ with the median plane intersecting the angle equally. It is questionable whether the primary sound should be at such an angle, although the direct sound may appear from any direction, it would be curious to consider the effect a direct ahead approach has on the results as this is the most common experience (Haas, 1949). The echo threshold was also described as being the shortest delay time at which the second auditory event (reflection) becomes audible which may prove confusing when compared to the masked threshold as it requires the minimum threshold (but for level) that a second auditory event is audible. The echo threshold is not necessarily the time that two sounds split, but merely the perception based on the overall sound quality (Blauert, 1983), therefore this indeed points towards each individual using their own preference in grading a sound image as featuring an echo. Their grading preference may be based on sound level intensity, tone, sound quality or even placebo occurrences. Previously, much of the research has focused on changing the level difference between multiple sound sources. Set angles are usually chosen and they are typically away from the median plane. Furthermore, limited stimuli are tested per research, therefore it is difficult to relate different research if the methods used contrast greatly. The echo threshold has been found to vary immensely (2 – 50 ms) dependent on stimuli (Litovsky, Colburn, et al., 1999), therefore it is important to look into different stimuli in a single test for confirmation. It is worth noting that complex sounds have been found to have an increased echo threshold.

The limit of echo perception has been researched by Rakerd et al. (2000), however this is not enough to provide a full model. Continuous speech was the stimulus of choice. Source intensity of the lagging sound was reduced by the listener until the echo was almost indistinguishable (as faint as possible but still audible). It was discovered that echo thresholds for speech in the vertical plane were higher than the horizontal plane. The definition given was "the level at which a delayed copy of the speech was just barely audible as an image distinct from the direct sound".

Haas (1949) himself in part of his seminal work, set out an experiment investigating "disturbance" as the angle of the reflected sound was modified. The exact definition for this is not described in the literature, however the understanding of the author is that it is when listeners begin to feel uneasiness for specific stimuli as the delay increases. That is, it becomes difficult to decipher exactly what is happening between sources as the overall sound becomes unintelligible. Therefore it is expected the disturbance values would not be identical to the specific test results set in this research. Where the delay values reside in regards to the results however will be interesting. The lead and lag sound intensities were kept identical and constant. A 2 m distance from the source to the listener was maintained and a speech stimulus was used with a speed of 5.3 syllables per second. The equivalent of significantly important findings were specified in that the "critical difference" was where 10 –

20% of listeners felt "disturbed" and the "critical ratio of the delay differences" was where 50% of listeners felt "disturbed". It was discovered that there was little influence on direction of the echo when the direct sound was presented incident from the front, with critical delay differences of 44 ms for 0° and 52 ms for 45°.

Stumpp (1936), as described in Haas (1949), researched into the critical delay time difference with speech as did Haas, however with the primary sound emitted from a lateral direction. The delayed sound was either present from the same location or the opposite lateral angle resulting in a critical delay time difference of 80 ms and 50 ms respectively, indicating a raised echo threshold for same location sources. These comparisons to the research results would be beneficial.

In the research by Damaske (1971), the direct sound was emitted from the front of the listener with a single reflection presented from the side with noise of various pulse widths. The recorded echo thresholds decreased below 15 ms when the pulse width was greater. Beyond 15 ms there was found to be an increase, however pulse width will not be considered in this research.

Schubert and Wernick (1969) performed experiments with high and low pass filtered noise of different durations (20, 50, 100 ms) with a cut-off frequency at 1 kHz and triangular envelope. This resulted in increasing threshold values in relation to the lengthier stimulus with the high pass filtered noise featuring a marginally lower echo threshold than the low pass filtered. The echo thresholds for the 50 ms noise were at 8 ms (HP) and 12 ms (LP) so it would be good to compare this with the 50 ms pink noise used in this research's experiments.

Despite little research into echo thresholds at different angles, Boerger (1965a, 1965b) did provide valuable input. The primary sound was directly in front of the listener and at 25° azimuth with an identical intensity to the reflection. Gaussian tone bursts were used as the stimuli which were one critical band wide. Multiple angles were tested at delays of 10, 25 and 50 ms. The test procedure functioned differently to related research in that the speakers were moved around the listener until the minimum azimuth angle could be recorded for each specified delay time.

2 Spatial Impression

2.1 ASW and LEV

Spatial impression is understood to be composed of two components which a listener can discriminate against (Bradley & Soulodre, 1995b; Maekawa Morimoto & Maekawa, 1989). The apparent source width (ASW) refers to the perceived width of a sound image and is considered to function with early reflections up to approximately 80 ms. The perception is caused by a fusion of temporal and spatial properties such as the relative level and angle of arrival of early lateral reflections (Bradley & Soulodre, 1995a). There has been confusion over the years regarding the definition of spatial impression, as an example, it was previously referred to as the broadening of the source which would be considered ASW nowadays (Barron, 1971; Barron & Marshall, 1981) The second component, listener envelopment (LEV), relates to the impression a listener has of being surrounded and immersed by the sound irrespective of the ASW sound components. It is considered to be related to the level and angle of late arriving lateral energy. Late reflections are understood to reside beyond approximately 80 ms. However, the cut-off point between ASW and LEV is not strictly 80 ms but is typically at this value for music sources. As an example, speech in contrast typically splits at 50 ms (Masayuki Morimoto, 2002). However, LEV is typically less than 2.2 seconds (Beranek, 2010). Furthermore, Maekawa Morimoto and Maekawa (1989) found that the sensitivity of ASW is affected by LEV, as did Bradley and Soulodre (1995a). The lateral direction of reflections are important for ASW determination (Barron, 1971; Bradley, Soulodre, & Popplewell, 1993; Masayuki Morimoto & Iida, 1995). LEV is affected by level, angular and temporal distributions of late arriving energy (Bradley & Soulodre, 1995a) which implies that it is not truly simple to model accurately. Despite the well grounded history and acceptance of ASW and LEV, Masayuki Morimoto (2002) believes that a division of reflections based on the precedence effect is more essential due to the relation to time and space distribution of reflections which indicates a key motivation for the present research.

2.1.1 Lateral Fraction

The ratio of lateral reflections is an objective measure which is a relatively effective cue for determining spatial impression. A large proportion of lateral reflections increases either the ASW or the LEV dependent on whether early or late reflections are used (Bradley et al., 1993). Subjective judgements of ASW are similar to measures of lateral energy fraction of early arriving sound (LF_0^{80}). Low frequencies are found to be beneficial in providing ASW (Hidaka, Beranek, & Okano, 1995). Equations 2.1 – 2.2 describe how the lateral fraction is calculated. Where $p(t)$ refers to the instantaneous pressure response measured with an omni-directional microphone (Bradley & Soulodre, 1995a). LF can be calculated either for the early reflections (<80 ms) or the late reflections as represented in the differing equations. Each is an effective measure of both ASW and LEV respectively.

$$LF_0^{80} = \left\{ \int_0^{80} p_L^2(t)dt \Big/ \int_0^{80} p^2(t)dt \right\} \tag{2.1}$$

$$LF_{80}^{\infty} = \left\{ \int_{80}^{\infty} p_L^2(t)dt \Big/ \int_{80}^{\infty} p^2(t)dt \right\} \tag{2.2}$$

2.1.2 Late Lateral Strength

The late lateral strength $\mathrm{LG}_{80}^{\infty}$ is an objective measure similar to the lateral fraction. It is a sum of the late lateral energy (> 80 ms). A figure-of-eight microphone records instantaneous sound pressure and compares this to the same source within a free-field environment 10 m from the source. It is averaged over various octave bands ranging from 125 – 1000 Hz which have been found to provide the best results for subjective LEV impressions (Bradley & Soulodre, 1995a). It has also been found an effective predictor for the broadening of a source (Masayuki Morimoto & Iida, 1995) but is not universally agreed upon (Soulodre, Lavoie, & Norcross, 2003). Equation 2.3 sets out how to calculate LG_{80}^{∞}. It is evidently closely related to the LF equation. p_F(t) refers to the instantaneous pressure response of the lateral sound emitted from a figure-of-eight microphone with the directional null pointed towards the source (Bradley & Soulodre, 1995a). p_A(t) refers to the response for the same source but at a distance of 10 m in a free-field.

$$LG_{80}^{\infty} = 10\ log\left\{ \int_{80}^{\infty} p_F^2(t)dt \Big/ \int_{0}^{\infty} p_A^2(t)dt \right\},\ dB \tag{2.3}$$

2.1.3 Interaural Cross-correlation

Interaural cross-correlation is a measure of the similarity of signals received at the ear. The more de-correlated the signals are in comparison to each other, the more the spatial impression widens and head locatedness decreases. IACC values of -1 are out-of-phase in reference to each other, if the IACC is 0 then there is no correlation (they are completely dissimilar), an IACC of +1 indicates the signals are identical. De-correlation of reflective signals — specifically early reflections — has been effective in developing a better depth representation. (Hidaka, Okano, & Beranek, 1992). It is found to be an important control for ASW (Masayuki Morimoto, 2002). To calculate IACC, the maximum absolute of the cross-correlation function for $|\tau| \leq 1$ ms is taken. The function is given in Equation 2.4, where $P_L(t)$ and $P_R(t)$ are equal to the instantaneous pressure response at the left and right ears respectively of a dummy head (Bradley & Soulodre, 1995a). Where $t1$ refers to the earlier time and $t2$ to the later, integrating between different time points is possible. IACC for the total energy can be calculated with $IACC_0^{\infty}$ $(t1 = 0,\ t2 = \infty)$, early sound energy with $IACC_0^{80}$ $(t1 = 0,\ t2 = 0.08\ s)$ and late sound energy with $IACC_{80}^{\infty}$ $(t1 = 0.08\ s,\ t2 = \infty)$.

$$\Phi_{LR} = \frac{1/(t_2 - t_1)\int_{t1}^{t2} P_L(t)P_R(t+\tau)dt}{\sqrt{\Phi_{LL}(0)\Phi_{RR}(0)}} \tag{2.4}$$

IACC itself has been broken down in to early (IACC_0^{80}) and late components ($\text{IACC}_{80}^{\infty}$) and is regarded as an effective measure for ASW and LEV respectively. Judgements of ASW in particular are known to be similar in accuracy to IACCe as described by Okano, Beranek, and Hidaka (1998). Furthermore, an average of the IACC can be taken in octave bands. This was performed by Okano, Hidaka, and Beranek (1994), who averaged the bands centred at 500, 1000 and 2000 Hz for the early reflections and named this IACC_{e3}. They found that smaller values coincided with larger subjective impressions of ASW.

2.1.4 Front-Back Energy Ratio

The front-back energy ratio is another spatial impression objective measure which has been researched extensively. It is a measure of the ratio of sound arriving from the front of the listener to the back. LEV is known to increase as the front-back ratio decreases, which refers to the reflections behind the listener increasing (Masayuki Morimoto & Iida, 1998). Masayuki Morimoto and Iida (1993) found that additionally the energy ratio of reflections from the front of the listener to the back affected envelopment even if the degree of IACC of late reflections was equal. To calculate the ratio, refer to Equation 2.5, where E_f and E_b are the energy of reflections arriving from the front and back of the listener respectively. The angles within the transverse plane are excluded from this equation (Masayuki Morimoto, Iida, & Sakagami, 2001).

$$FBR = 10\ log(E_f/E_b),\ dB \tag{2.5}$$

2.2 Distance & Depth Perception

Distance and depth perception relates to the research due to the questions focused on the median plane. Indeed the initial research direction was involved in looking into depth which is an under-researched area. This research intends to bridge the gap between previous knowledge and provide answers to reflection differences between the vertical and horizontal planes. Depth can be visualised as the range of sound between the front and back of the listener along the median plane but not necessarily confined solely within. The median plane concerns sounds directly in front, behind and above the head. It is notoriously difficult to define depth clearly to subjects. If depth perception is truly confined to the median plane then inter-aural differences would not be evident. The initial time delay gap (ITDG) is an indicator of distance which is essentially pre-delay after the direct sound. It occurs before ASW functions and relates a short delay to a short distance. In the best halls, the ITDG is less than 25 ms, less quality halls it is 25 - 35 ms and beyond this range, an 'arena' sound is experienced (Beranek, 2010). Much of the literature has described depth synonymously with distance. Rumsey (2002) sets out a 'scene-based paradigm' which breaks an environment — such as a concert hall — up into various static scenes based on their descriptive function. The scenes can be applied to width and depth and potentially height. Distance can also be included with this concept. The concept can be applied to individual sound sources, ensembles (groups of sound sources) and the environment. Figure 2.1 illustrates how this would function with regards to depth. As can be seen, depth is envisioned as a range. The advantage of these scenes is clarity in descriptions between researchers.

2.2.1 Sound Intensity

Intensity has long been considered a primary cue for distance perception (Thompson, 1882). It is known to attenuate linearly at 6 dB with every doubling of distance in a free-field environment but this calculation is much less accurate in a reverberant

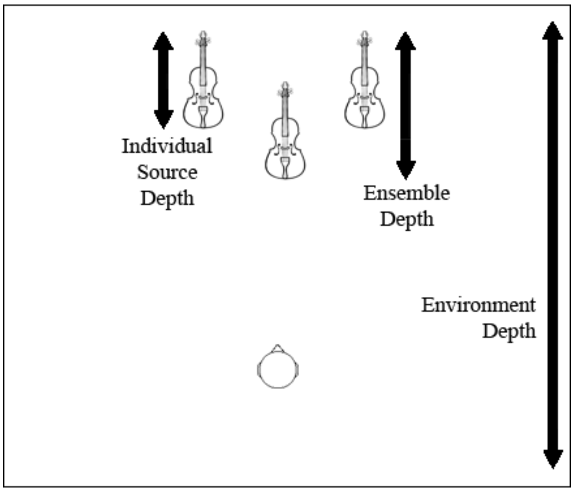

Figure 2.1: Different scenes are presented in relation to depth perception.

environment where the attenuation isn't as noticeable. Various research has however noted that it is a relative cue which means that distance can only be estimated with a reference distance in mind. Additionally, if a sound source is at a large distance, listeners greatly underestimate it (Cochran, Throop, & Simpson, 1968).

2.2.2 Direct-to-Reverberant Ratio

The direct-to-reverberant energy ratio is another primary and absolute cue (Mershon & King, 1975) for distance perception, but specifically in a reverberant space. The ratio of the direct sound energy to the reverberant energy is an effective objective measure for distance. Decreases in the ratio result in a source appearing more distant (Von Békésy & Wever, 1960). The direct-to-reverberant energy ratio allows listeners to estimate distance more accurately than intensity cues and also without a reference distance required. Anechoic distances were underestimated by a factor of nearly 10 whereas echoic environments were found to be more veridical (Mershon & King, 1975; Nielsen, 1993; Von Békésy & Wever, 1960). Distance and spaciousness was found to increase as reverberation did (Steinberg & Snow, 1934).

2.2.3 Critical Distance

As the distance from a source in a reflective space increases, the late reflections build up resulting in a reverberant sound level which remains relatively constant. The critical distance is the point in space where the sound pressure level of the direct sound is the same as the reverberant level (Rumsey, 2001). Beyond this distance is where localisation becomes tricky but immersion increases.

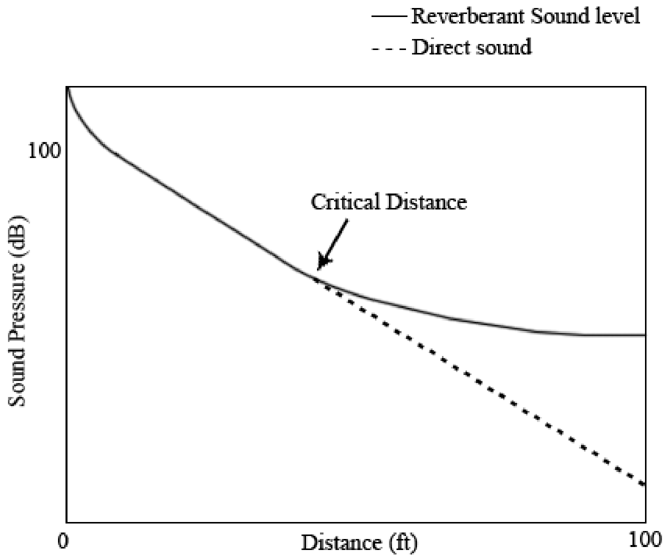

Figure 2.2: The critical distance illustrates how the direct sound energy decrease as the distance from a source increases yet the reverberant energy remains constant. After Rossing (2007)

$$d_c \approx 0.057\sqrt{\frac{V}{RT60} \cdot Q} \tag{2.6}$$

2.2.4 Frequency Bands

Further research on our perception of frequencies by Blauert (1983) resulted in bands of frequencies which feature a certain probability of being represented either in front, behind or above the listener. This study presented various noise stimuli randomly over loudspeakers. With approximately 90% confidence, 1 kHz was located behind the listener. With approximately 75% confidence, 3 kHz was located in front of the listener and 8 kHz directly above.

2.2.5 Familiarity

If a listener is familiar with a sound source then this can reduce inaccuracies in distance estimation, however it is not necessarily 100% effective (Zahorik, 1996). A listener may recognise differences in the intensity or spectral content of a source to that which they have memorised and therefore be able to estimate distance from this. Coleman (1962) showed that when a noise burst stimulus was not familiar to the listeners, distance judgements were inaccurate. It took 10 trials until judgements were more veridical. This was proven by McGregor, Horn, and Todd (1985) but was inconclusive in tests by Nielsen (1991).

3 Experiment

3.1 Methodology

There were two independent variables (IV) to be used in the tests — angle (lag direction) and stimulus. The dependent variable (DV) in the tests was delay (time difference) which was set by the listeners in the interface. The gaps in the research — as alluded to in the final points in Chapter 0 — suggest various options for testing. It was decided that maintaining an equal sound level between two sound source locations was the best approach and emphasis would be placed on the variations of echo thresholds between the independent variables and listener instructions. The following sections outline the process.

3.1.1 Setup

There were 17 possible loudspeakers (Genelec 8040a) used in the experiment. 12 of the available loudspeakers were in the horizontal plane separated evenly in a circle by 30° increments. The remaining five speakers were again in the horizontal plane but at 30° elevation. Their azimuth angles were 0°, ±30° and ±110°. There were limits with the speaker array racking which prevented a ±120° position, therefore ±110° was chosen. These particular angles were chosen due to their future benefit of reproduction on multichannel loudspeaker systems. A single loudspeaker per pair was chosen by a randomisation process to eliminate loudspeakers with the same inter-aural differences being used in the tests. Therefore only 10 unique loudspeaker conditions were used per test. Listeners sat in the centre of the array of speakers facing the 0° azimuth, 0° elevation speaker directly. The original sound was always emitted from this loudspeaker. The distance of the listener from the loudspeakers was two metres for the lower array as was typical for previous research (Haas, 1949) and approximately 2.3 metres for the upper. The report continues with the format

'#° #°' when describing the precise speaker angles, where the first number is the azimuth angle of the speaker and the second is the elevation. For example, 110° 30° is the speaker at 110 azimuth and 30 elevation.

The stimuli used consisted of 3 conditions — The first stimulus was an orchestral clip which was 14.5 seconds in length. It featured string instruments and was an example of a continuous source. The second stimulus was pink noise which was processed in Sound Forge Pro (10.0) to be transient. The pulses were 50 ms in length and featured a one second gap between pulses to prevent overlapping. A 1 ms fade-in was applied to prevent clicks. The third stimuli was continuous German speech acquired from a sound library known as the 'Sound Quality Assessment Material recordings for subjective tests' (SQAM). This was an example of a familiar and continuous source and was found to feature 5 syllables per second. The presentation of the stimuli and tests was randomised to prevent subject bias. All stimuli matched with 48 kHz sampling frequencies to maintain high fidelity. To maintain even sound energy, the stimuli in the patch were calibrated using the Casella CEL-450 Sound Level Meter (Figure 3.1). L_{Aeq} was chosen on the device which measured the equivalent continuous level with an A weighting. The fast time weighting was used for the pink noise and the impulse weighting was used for the orchestral and speech. The SPL meter was positioned directly towards the speaker at 0° 30° and 70 dBa was maintained between stimuli. Level offsets were applied to the lower channels of -0.3 dB to ensure equal sound energy was received at the listener position and a 1 ms delay was introduced into the signal path of the lower speakers to match signal arrival time at the position.

11 experienced listeners, one of which included the researcher, took part in both tests. Therefore this experiment was an example of a repeated measure test. The tests were carried out in an ITU-R BS.1116-compliant critical listening room at the University of Huddersfield. A Logitech G5 wired mouse was used by the subject to perform actions in the Max 7 interface. An iPad was provided with listener instructions as set out in this chapter and was available to use as a mouse mat. The

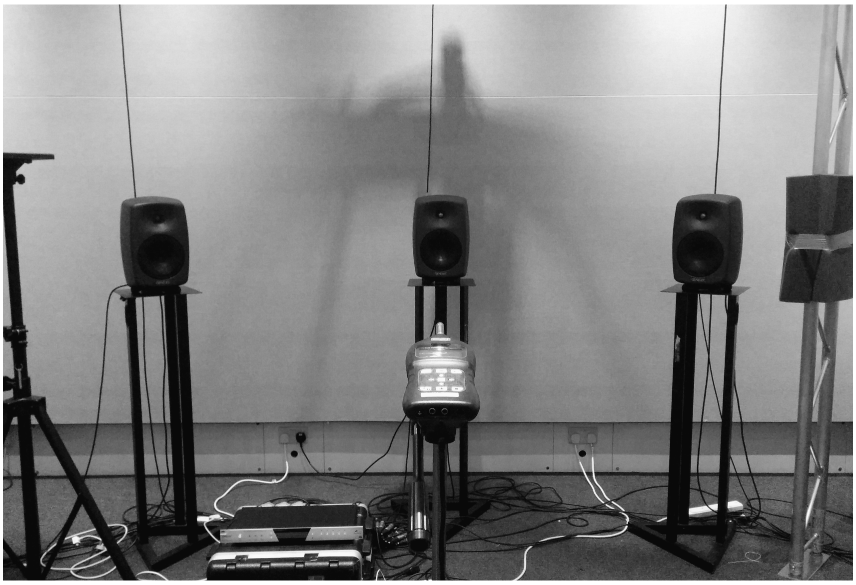

Figure 3.1: The setup of the Casella CEL-450 SPL meter.

Antelope Audio Orion[32] multi-channel AD/DA converter and USB interface was used to map the channels from an Apple Macbook Pro which hosted the listening test interface. The DA volume control in the Orion[32] was set to -17 which applied a level reduction to all channels to approach the 70 dBa SPL required. This value was a typical SPL and close to that of previous researchers of reflective sound testing (72 dBa (Bradley & Soulodre, 1995b), 77 dBa (Barron & Marshall, 1981)) without uncomfortable hearing occurring. The subjects were informed to keep their head located directly forward at the centre speaker and refrain from any head movement once a trial had begun. A microphone stand was placed behind a subject's head to provide guidance as to how far forward the head and body should be situated (Figure 3.2). Head movement was additionally monitored.

3.1.2 Criteria

A set of instructions were given per test to the listeners. There were expectations of significant differences based on the previous research which showed vastly different echo thresholds when unique listener instructions were provided. In test one, the listener graded the time difference where the beginning of separation was apparent between the original and delayed signal. Fusion of the sounds may still be apparent.

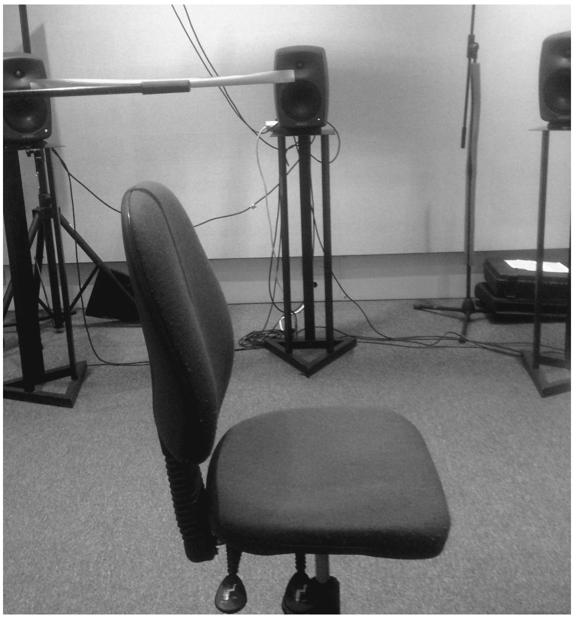

Figure 3.2: The seating position throughout all testing.

In test two, the listener graded the time difference where they could distinguish clear separation between the original and delayed signals. The echo threshold marking criteria provided in the listener instructions to the subjects are outlined below:

> Test 1: "The minimum delay value where sound begins to appear to alternate or move between two locations. Except in the case where a single location is apparent, separation should be evaluated. There may be vast sound energy spread out between locations.".
>
> Test 2: "The minimum delay value where *all* sound begins to clearly appear to alternate or move between 2 locations. Except in the case where a single location is apparent, complete separation should be evaluated. There may be an apparent 'hole' of energy between loudspeakers.".

The movement described in the marking criteria refers to the change in balance of sound energy in the sound field. As a delay value changes, the apparent sound energy from each source may not be consistent to the listener. The summation of these energies will have an effect of perceived direction, sound image width and stereo balance (panning).

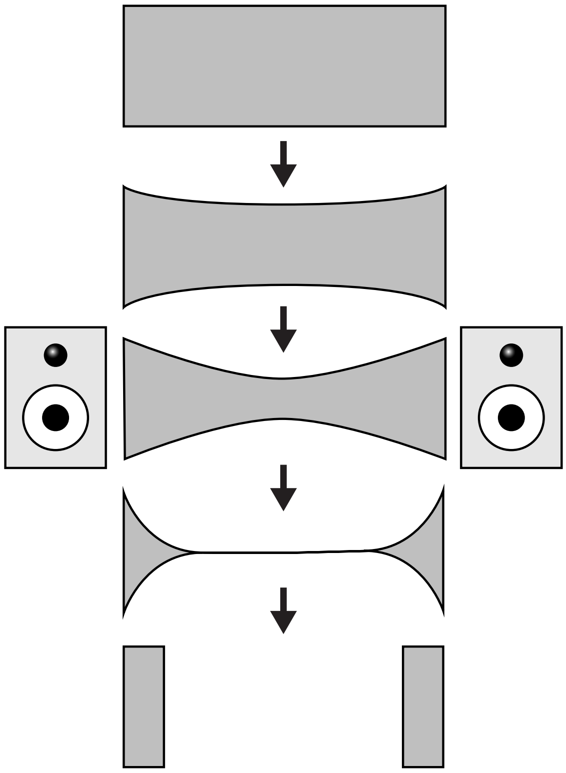

Figure 3.3: Various stages of spread and separation. As the time difference between two loudspeaker sources increases (as illustrated by the progressive images), the sound energy appears weaker between sources until eventually two individual images are achieved. This was a simplification of the process utilised to aid listeners in visualising differences during the listening tests.

To aid in the visualisation of the two echo threshold criteria, a rudimentary graphic indicating different stages in the separation process was provided for the subjects (Figure 3.3). It represents the space between two sound positions and how this space is filled with sound energy. When the energy was equal between channels, the space was filled equally. As they begin to separate, the energy becomes weaker in the middle until (in an ideal world) the two sounds can be heard at the separate distinct locations. The graphic is a simplification of this process as the literature review shows this is not a complete definition, but it did aid listener judgements. For the first test, subjects were grading when the space between the loudspeakers was becoming weaker which indicated energy movement. The second test is exemplified by the graphic with separate boxes.

3.1.3 Listener Interface

The interface for the listening test was created in Cycling '75 Max 7 (7.0.5) (Figure 3.4) which is a visual programming software. It allowed for on-the-fly audio programming changes and conceptualising. It features a presentation mode which hides the visual wires behind the scenes. The interface devised featured four individual discrete sliders. The leftmost slider (stage 1) defined the beginning delay value to implement between the two sound sources in discrete increments of 120 ms ranging from 120 – 600 ms which at the maximum value was beyond recorded echo thresholds. Each subsequent slider to the right subtracted a specific delay value from the previous slider. This particular method is based loosely on a modification of the method of adjustment and adaptive psychophysics methods by Wallis and Lee (2016) from the University of Huddersfield. The stage two slider featured increments of 20 ms ranging from 20 – 120 ms. Stage three featured increments of 4 ms ranging from 4 – 20 ms. Finally, stage four featured increments of 1 ms ranging from 0 – 4 ms. The maximum delay of 600 ms was set per trial. A replication of the original sound was sent after the default delay to a chosen loudspeaker. It was suggested to the subjects to use the buttons next to the sliders rather than the slider itself which were set to the discrete values mentioned previously. The subjects were instructed to examine every slider from left to right and to select on each one the top most value where the marking criteria could still be confirmed. By selecting this value, it was possible to fine-tune grading judgement between discrete values of the next highest slider with the remaining sliders to a 1 ms level of accuracy. An example of how any delay value can be obtained is shown in Table 3.1. It was suggested to the subjects to decrease the sliders as it was logical considering the sliders subtract from each other. Additionally, the GUI was programmed to reset latter sliders if an earlier slider was moved, this was simply to ensure the correct delay was maintained but the benefit was it forced a subject to re-evaluate the test again due to their uncertainty exhibited.

The subjects were able to temporarily mute the audio with the 'sound on/off' button. The restart test button completely restarted the whole set of tests and was rarely used. If a subject did fail to move any sliders this would be recorded. This was beneficial as it could be set as a missing value. Subjects were told to proceed to the next trial when they were confident with their choice. There were two listening

Delay (ms)	Stage 1	Stage 2	Stage 3	Stage 4
41	120 (1)	-60 (3)	-16 (1)	-3 (2)
98	120 (1)	-20 (5)	0 (5)	-2 (3)
333	360 (3)	-20 (5)	-4 (4)	-3 (2)

Table 3.1: An example is provided on how to obtain three random delay values with the four sliders in the GUI. Stage 1 specifies the delay value to begin subtracting from. Subsequent sliders subtract from the previous slider. The values in brackets indicate the button next to the slider which was selected, where 1 was the bottom button to 5 or 6 at the top.

tests to be undertaken by the subject, each which lasted approximately an hour and required a 15 minute mandatory rest period half way through to provide a rest to the ears. The two tests were not to be carried out simultaneously, at least 30 minutes was maintained to again prevent ear fatigue. A ringing tone at approximately 1 kHz was evident occasionally from the speaker racking when loud transients sounded, therefore subjects were informed of this and requested to disregard it. Ear height was set between the tweeter and woofer of the loudspeakers. There were 30 trials to sit through per test which consisted of 10 trials per stimulus which refer to all 10 angles which were accounted for per stimulus as described below. The audio looped continuously unless paused by the listener.

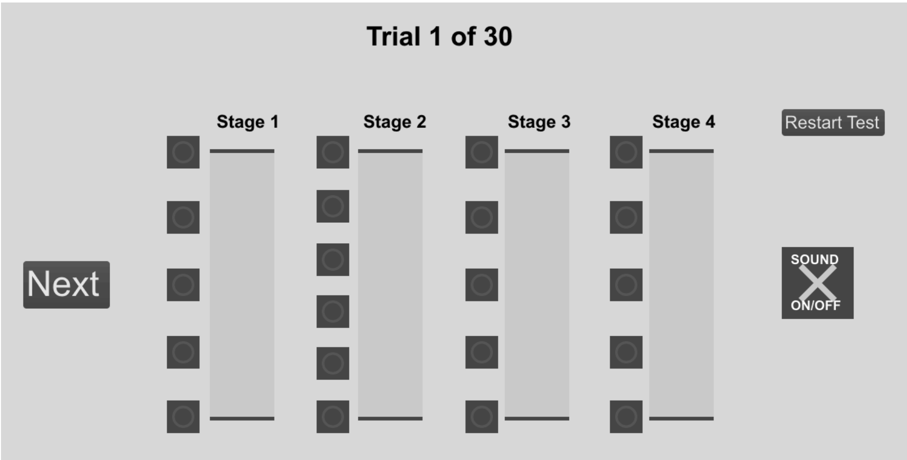

Figure 3.4: The listening test graphical user interface created in Cycling '75 Max (7.0.5).

3.1.4 Familiarisation Testing

Prior to the main testing, a familiarisation Max 7 patch was presented to the subject. The researcher was present as the subject proceeded through the test and pointed out the criteria clearly. Questions were asked to the subject such as "Can you hear the beginning of separation?" (for the test one criteria), "Is a hole in the middle of the sound image apparent?" and "Is there clear separation between the two sound positions?". If the subject's answer to the questions was puzzling or out of the ordinary in comparison to other subjects observations, audible differences were made clear and clarified. The familiarisation test itself was a cut down version of the final test with 6 set varied angles to describe the differences clearly to the subject. These angles were also made known to the subject if required. The familiarisation process took between 20 – 60 minutes dependent on the subject's understanding. The angles presented are given in Table 3.2.

Trial	Angle
1	0° 0°
2	+90° 0°
3	-30° 0°
4	180° 0°
5	-60° 0°
6	-120° 0°

Table 3.2: The angles presented to the listeners per trial in the familiarisation test.

3.1.5 Comments

There were a few observations detailed by the listeners after testing. It was suggested that the orchestral stimulus was difficult to distinguish precise delays with. One subject described difficulty judging the delay value for the orchestral and speech stimuli in test two despite admitting the instructions were completely understandable. The orchestral stimulus would occasionally appear to sound as a stereo image due to frequent fluctuations of the transients between speakers. Many listeners found the 'tail end' of the delayed sound to be the best cue to arrive at their chosen indicator of separation.

3.2 Results

The Max 7 patch produced individual text documents per listener and test (22 total), therefore the data was firstly collected together in a Google Docs spreadsheet. Although there appeared to be noticeable agreement between listeners, the relevance and significance was unclear at this stage. There appeared to be outliers and the subjects occasionally repeated the same delay value for different angles within the same test and stimulus. Two reasons for this repetition are suggested due to the nature of the test — the subject was unclear whether they heard a change and chose a comfortable and familiar value or that there truly is a strong relationship between angles.

The data was imported into the IBM SPSS Statistics (20.0.0) software for analysis. Different layouts provided access to multiple analysis tests, however the structure of the test was repeated measure, therefore the primary file incorporated both independent variables (angle and stimulus) together with the dependent variable (delay). To determine whether parametric or non-parametric testing would be the correct direction, a test of normality was undertaken. The observation of outliers in the spreadsheet highlights the importance of confirmation.

Due to the low sample size (11), the Shapiro-Wilk method was trusted. The orchestral and speech results in test one were found to be normally distributed ($p > .05$). Six angles for the pink noise in test one were not found to be normally distributed — 30° 0° ($p = .020$), 60° 0° ($p = .009$), 90° 0° ($p = .022$), 120° 0° ($p = .002$), 30° 30° ($p = .015$) and 110° 30° ($p = .002$). In test two, the orchestral stimulus was not normally distributed at 90° 0° ($p = .036$), 150° 0° ($p = .049$). Pink noise was not normally distributed at 0° 0° ($p = .004$), 180° 0° ($p = .033$) and 0° 30° ($p = .005$). Speech was not normally distributed at 30° 0° ($p = .047$), 60° 0° ($p = .010$), 90° 0° ($p = .037$) and 110° 30° ($p = .028$).

Ideally a one-way repeated measures ANOVA would have been the preferred test method and it is rather “robust to violations of normality” (A. Lund and Lund, 2013). However, to confirm it was the right test, another assumption was the requirement for no significant outliers. There were evident outliers in every set of data as assessed by inspection of individual boxplots per stimulus and test. A delay value of 0 ms was chosen by one subject for the pink noise at two angles (0° 30° and 150° 0°) in test one which is unexpected, yet there were multiple results where subjects did suggest 1 ms as the optimal delay value. The presence of the outliers may suggest grading difficulty but with a low sample size it was not preferable to discard these results. Therefore, the outliers were kept in the analysis process. Corrective transformations of the data proved unsatisfactory due to the multiple intended analysis iterations required and the necessity for all data to be transformed equally.

Further scrutiny of the boxplots as presented in the next sections indicated patterns in the data. As the test utilised a small sample size and the median was considered, the boxes did feature a lack of symmetry. The outliers at choice angles often matched with the maximum or minimum value of the box of other angles. Please refer to Figures 3.5 – 3.10 for visual representation of the following descriptions. Mean and median values are provided in Tables 3.3 and 3.4 respectively.

Angle	Orchestral	Pink Noise	Speech
0° 0°	70	22	22
30° 0°	55	19	16
60° 0°	46	14	14
90° 0°	50	16	16
120° 0°	47	15	15
150° 0°	53	16	15
180° 0°	70	21	22
0° 30°	67	26	20
30° 30°	50	19	20
110° 30°	45	15	18

Angle	Orchestral	Pink Noise	Speech
0° 0°	162	56	57
30° 0°	101	50	53
60° 0°	106	51	57
90° 0°	97	52	48
120° 0°	103	49	51
150° 0°	115	52	48
180° 0°	149	55	57
0° 30°	147	55	59
30° 30°	112	49	60
110° 30°	104	53	51

Table 3.3: Test one (left) and test two (right) mean values rounded to the nearest millisecond.

3.2.1 Test One Boxplots

The orchestral stimulus in test one didn't exhibit clear differences between angles. The interquartile range (IQR) of the boxes overlapped and the total range of the boxes including whiskers was relatively large at 155 ms ranging from 3 – 158 ms,

Angle	Orchestral	Pink Noise	Speech
0° 0°	63	28	19
30° 0°	55	11	15
60° 0°	45	6	14
90° 0°	36	7	12
120° 0°	50	6	15
150° 0°	51	13	15
180° 0°	60	23	25
0° 30°	56	25	25
30° 30°	48	9	17
110° 30°	51	7	20

Angle	Orchestral	Pink Noise	Speech
0° 0°	175	54	55
30° 0°	80	51	51
60° 0°	86	52	49
90° 0°	86	48	40
120° 0°	68	41	45
150° 0°	85	54	46
180° 0°	116	54	48
0° 30°	105	54	52
30° 30°	96	53	50
110° 30°	90	56	45

Table 3.4: Test one (left) and test two (right) median values.

with angles 0° 0° and 180° 0° most notably occupying this range. The data was the most symmetrical at 60° 0° and 30° 30°, although it was not perfect. The median value of the angles ranged between 36 – 63 ms.

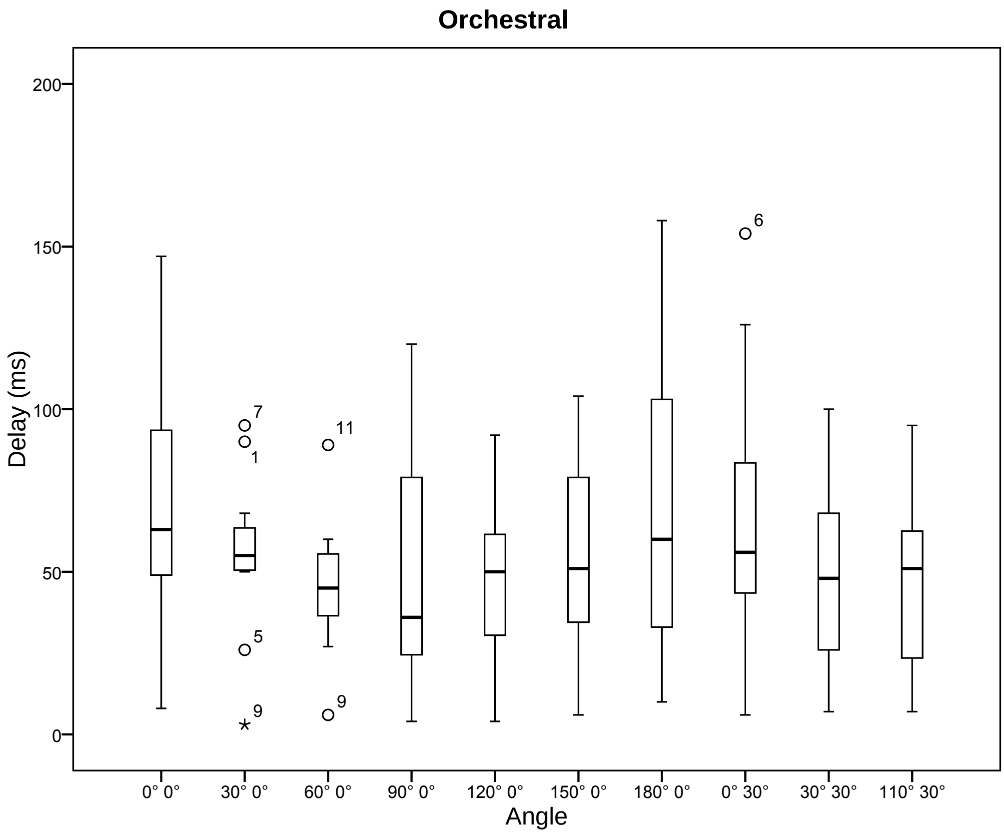

Figure 3.5: Test One Orchestral Boxplot

In comparison to the orchestral stimulus, the pink noise occupied a much smaller range of 0 – 54 ms as illustrated in Figure 3.6. Between angles there appeared to be no clear differences. The median values ranged from 6 – 28 ms and of interest were the angles in the median plane (0° 0°, 180° 0° and 0° 30°) which were at least 10 ms larger than those which weren't. The boxes in the plot were extremely skewed indicating a strong lack of symmetry with only 0° 30° appearing to be moderately symmetrical.

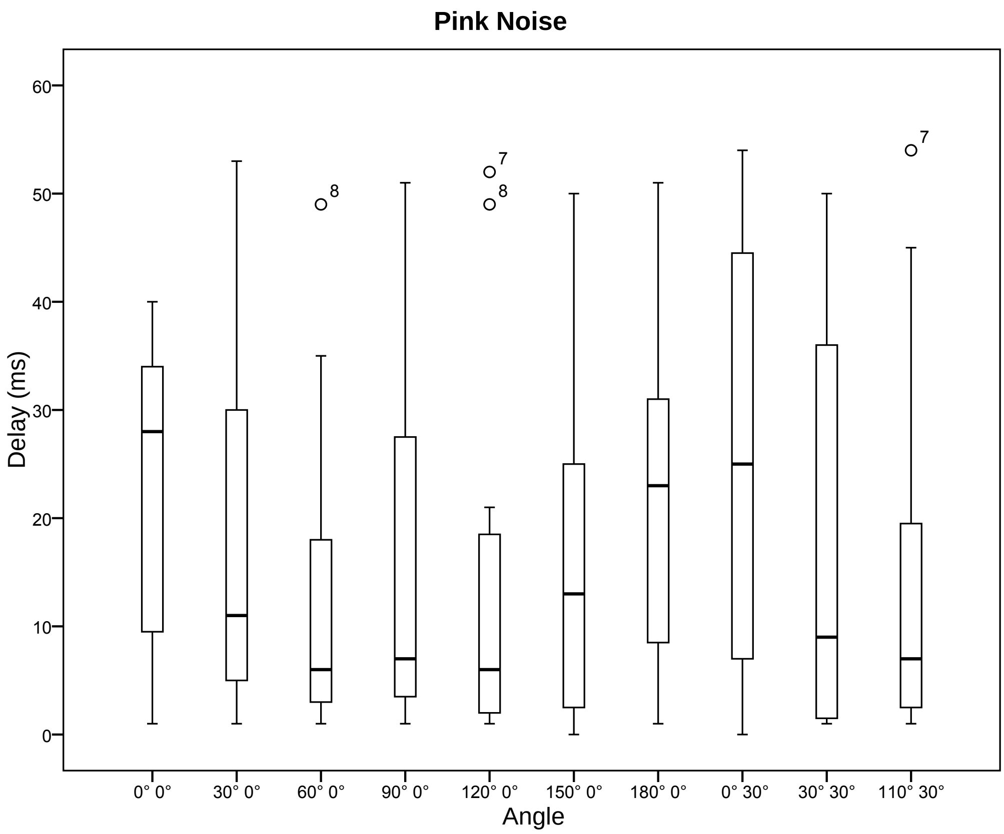

Figure 3.6: Test One Pink Noise Boxplot

The range of the data for the speech stimulus was 3 – 46 ms which was comparable to the pink noise range but with a lower limit 3 ms higher and an upper limit 8 ms lower. No clear differences were evident between angles. The median range was between 12 – 25 ms which is comparable to the median range for the pink noise but with a marginally larger minimum value. The boxplot was not symmetrical.

3.2.2 Test Two Boxplots

The orchestral data in test two ranged from 15 – 305 ms. This extended far beyond the upper limit of the orchestral stimulus in test one by approximately a doubling

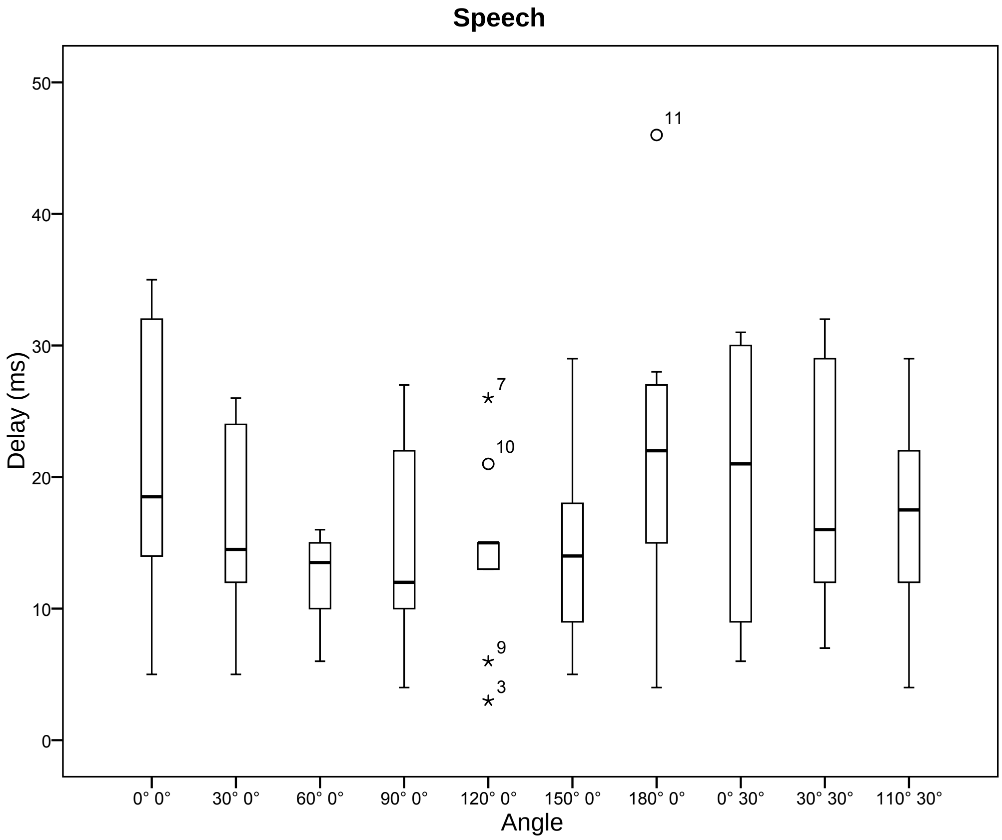

Figure 3.7: Test One Speech Boxplot

although the lower limit was the same. There were no differences to be found between angles. The interquartile range noticeably extended over a large range. The range of median values extended from 68 – 175 ms.

The pink noise data ranged from 26 – 103 ms. In relation to the pink noise data from test one, it was a larger range, although the interquartile ranges were smaller indicating increased agreement between subjects. Between angles there appeared to be no clear differences. The interquartile ranges for angles along the median plane (0° 0°, 180° 0° and 0° 30°) were the smallest. The extreme outliers beyond the maximum sample value originated from a single case. The median values ranged from 41 – 56 ms and there was a lack of symmetry evident.

The speech data ranged from 16 – 146 ms. This was a much larger range in comparison to the speech data in test one. There were no differences between any of the angles. The range of median values was between 40 – 55 ms which was extremely similar to the median values from pink noise within the same test. There

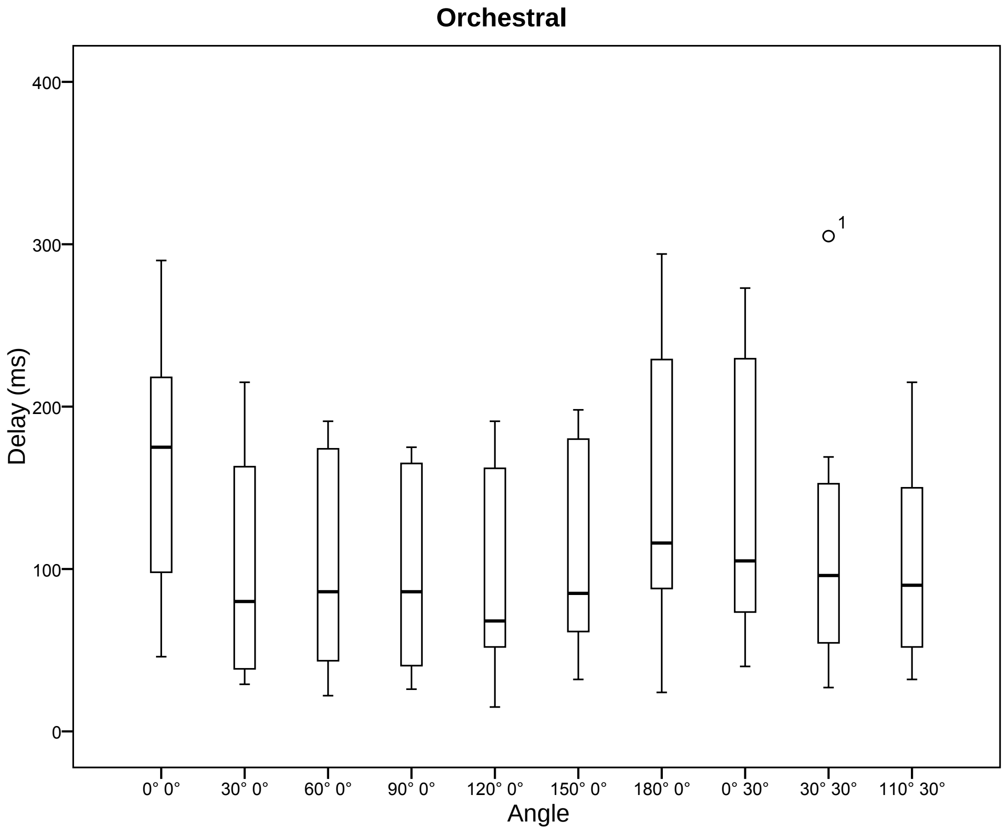

Figure 3.8: Test Two Orchestral Boxplot

was a similar agreement for cases between angles in comparison to the pink noise from the same test as shown by the interquartile ranges. The outliers were from a single case and resided beyond double the median.

3.2.3 Split Boxplots

The final boxplots which are presented in Figures 3.11 & 3.12 split the data into stimuli and test respectively. This created an advantage in focusing on relationships between the data regardless of angle differences.

The orchestral stimulus showed much less agreement between cases in the second test as the data was more spread out. The median of 100 ms in the second test was clearly at a larger delay than in the first test of 52 ms. Despite numerous outliers residing beyond the upper whisker in test one, they did not extend beyond the interquartile range of the second test and certainly did not reach the maximum

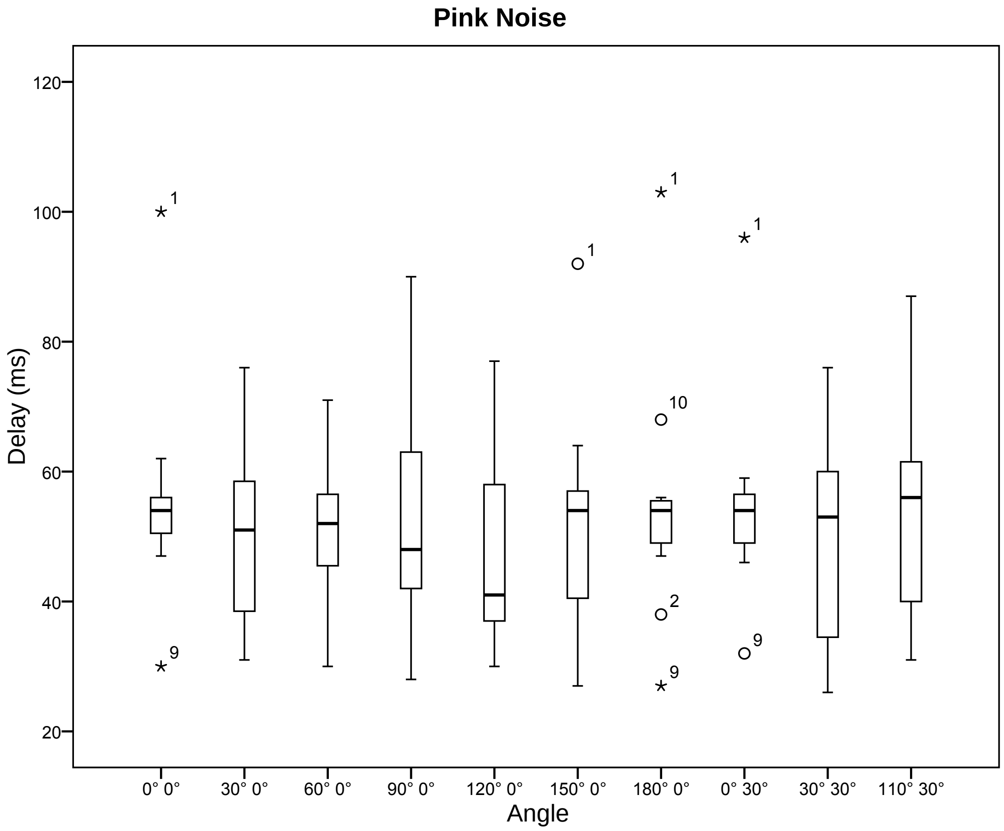

Figure 3.9: Test Two Pink Noise Boxplot

value. In both tests the orchestral stimulus was skewed to the right with test two showing a more significant skew. The outliers for the first test extended to the upper quartile of the second test.

Between tests, the pink noise showed clear differences as the interquartile ranges did not overlap. The IQR of the pink noise in test one did not overlap the IQR of the orchestral stimulus in test one and only marginally for test two. Including the outliers in test two, the pink noise exhibited a similar range between tests but beginning and ending at different and noticeably significant values. Both tests were skewed to the right when considering the outliers. The maximum range extended up to 54 ms in test one which was equal to the median for test two.

The speech in test one featured the smallest range in comparison to the other stimuli and tests which indicated the highest agreement between subjects. Its IQR was closest in size to the IQR of the pink noise in test two. The median for the speech in test one of 15 ms was marginally larger than the median for the pink

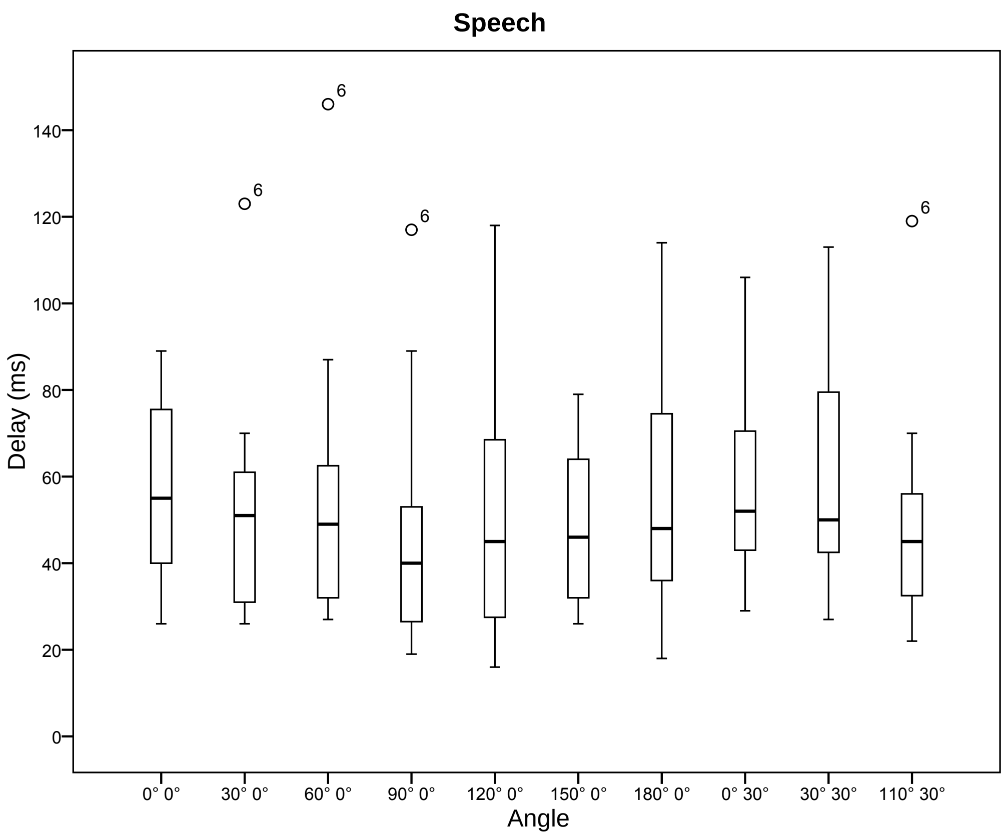

Figure 3.10: Test Two Speech Boxplot

noise of 11 ms in the same test. The opposite case was apparent in the second test, with an overall median of 49 ms for the speech and median of 54 ms for the pink noise. Despite this feature, the difference in the median between these two stimuli was evidently minimal. The symmetry of the speech for both tests indicated a right skew. There were clear differences between the tests. The box for the speech in test two relatively closely resembled the orchestral stimulus in test one, the median values respectively were 49 ms and 52 ms.

3.2.4 Friedman Test

Due to the presence of frequent outliers, it was evident that non-parametric testing would be the most trustworthy, although less powerful method to proceed with. The Related-Samples Friedman's Two-Way Analysis of Variance by Ranks test is the non-parametric equivalent of the repeated measures ANOVA test which allows for

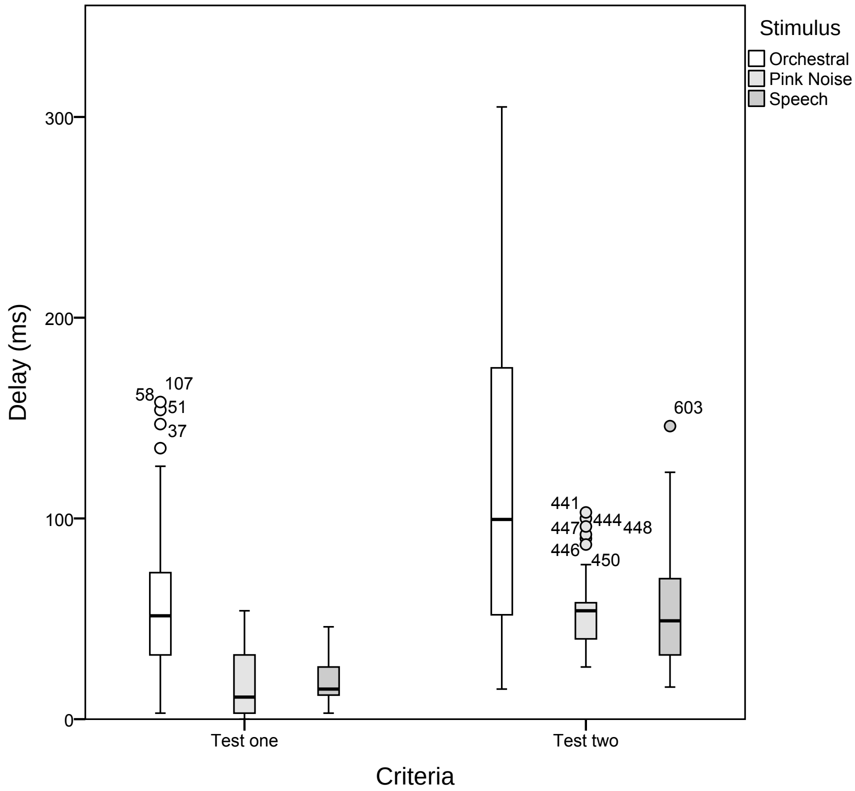

Figure 3.11: Box plot of criteria differences split by stimulus.

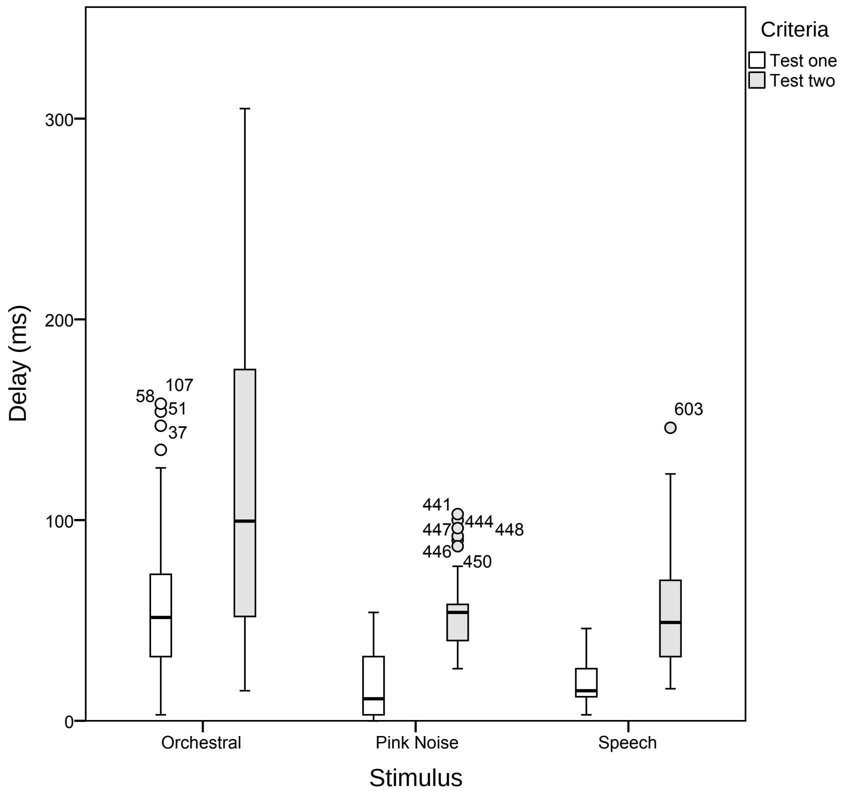

Figure 3.12: Box plot of stimulus differences split by criteria.

the comparison of multiple groups at once. The null hypothesis for a Friedman test (H_0) is where the distribution of scores in each group are the same. The alternative hypothesis (H_A) is where at least two of the groups' distributions differ. Each independent variable was tested separately as the Friedman test functions with a single categorical IV with three or more groups.

Initially, the angle was the independent variable and analysed per stimulus and test. In reference to the significance and test statistic provided by the data in Table 3.5, the results of the test showed statistically significant differences between angles for the orchestral stimulus in test one ($X^2(9) = 21.171$, p = .012), test two ($X^2(9) = 28.890$, p = .001) and the pink noise in test one ($X^2(9) = 23.947$, p = .004). There were not significant differences between angles for the pink noise stimulus in test two ($X^2(9) = 7.180$, p = .618) and the speech stimulus in test one ($X^2(9) = 13.847$, p = .128) and test two ($X^2(9) = 13.545$, p = .139).

Test	Orchestral	Pink Noise	Speech
1	.012 (21.171)	.004 (23.947)	.128 (13.847)
2	.001 (28.890)	.618 (7.180)	.139 (13.545)

Table 3.5: Friedman test results showing statistically significant and not significant differences (p) between angles. The test statistic is provided in brackets.

The stimuli were then chosen as the independent variable and analysed per angle and test. To avoid repetition, due to the vast amount of comparisons, please refer to the tables for the in-depth statistics. Table 3.6 lists the significant differences and test statistics for this test. It can be clearly seen that there were significant differences between stimuli for the majority of angles excluding 30° 0° in test two ($X^2(2) = 5.628$, p = .060) and 60° 0° in test two ($X^2(2) = 3.818$, p = .148).

3.2.5 Paired-Samples Sign Test

To determine which groups from the Friedman Test were statistically significantly different it was required to run pairwise comparisons. Automatically generated post hoc tests used the Related Samples Wilcoxon Signed Ranked Test to create pairwise comparisons with a Bonferroni correction. However, due to the lack of symmetry

Angle	Test 1	Test 2
0° 0°	.005 (10.619)	.002 (13.000)
30° 0°	.002 (12.600)	.060 (5.628)
60° 0°	<.001 (15.488)	.148 (3.818)
90° 0°	.006 (10.093)	.020 (7.818)
120° 0°	.001 (13.818)	.019 (7.953)
150° 0°	.005 (10.619)	.001 (13.476)
180° 0°	.013 (8.727)	.003 (11.455)
0° 30°	.001 (13.190)	.001 (13.636)
30° 30°	.027 (7.190)	.020 (7.818)
110° 30°	.020 (7.818)	.002 (12.047)

Table 3.6: Friedman test results showing statistically significant and not significant (p) differences between stimuli. The test statistic is provided in brackets.

in the distribution of differences, the Wilcoxon test was not truly reliable due to its assumption of a symmetrical distribution. The Paired-Samples Sign Test is less powerful but does not have this requirement and would be trusted for in-depth pairwise comparisons. However, the results of the Wilcoxon Test are provided in Appendix A in Figures A.1 and A.2 for the sake of interest. Needless to say, the results from the Wilcoxon tests were similar to those from the Paired-Samples Sign Test.

To prevent an excess of table data in this chapter, the full results are provided in Appendix B in Tables B.1 – B.24. The essential details remain in the main text. In this test, positive (+) and negative (-) differences occur between pairs of data. The median of paired differences (or frequently mentioned here as median differences for short) refers to the average of the differences acquired per listener for a single pair of data. The null hypothesis for the sign test (H_0) is where the median of paired differences = 0. The alternative hypothesis (H_A) is where the median of paired differences $\neq 0$. In other words, for the null hypothesis to be true, half of the differences should be positive and half should be negative. The number of positive or negative signs can be useful to know and may indicate a trend, however this must be tested for significance. The exact significance value was recorded for every test due to the small sample size. Median values have been reported previously, please see Table 3.4 for comparison with the median differences.

Between Tests

The median of paired differences in this section are in relation to test one subtracted from test two which was expected to result in positive values. There were clear increases in median difference between tests for every angle and stimulus. The orchestral stimulus showed a large increase ranging between 47 – 84 ms in comparison to the pink noise (26 – 37 ms) and speech stimuli (21 – 38 ms). The positive, negative and tied paired differences refer to the difference taken between two groups which then has the median evaluated of this difference. This suggests whether the subject chose delay values which increased or decreased per group. The results confirm that an increase in delay was specified by a strong majority of subjects. The orchestral stimulus showed the most variety with the least positive values being evident for the angle 30° 0° of seven with four negative and zero tied (no change) values. The pink noise and speech stimuli showed many positive increases with the least being the speech stimulus at 90° 0° with a respectable nine positive, one negative and one tied value. The significant difference data shows the orchestral stimulus was not statistically significantly different at angles 30° 0° ($p = .549$), 60° 0° ($p = .065$), 90° 0° ($p = .227$), 120° 0° ($p = .065$) and 30° 30° ($p = .065$). The remaining angles for the orchestral stimulus and all of the pink noise and speech were statistically significantly different.

Figure 3.13 shows the median differences plotted between tests. It was clear that there was large variance for the orchestral stimulus in comparison to the pink noise and speech stimuli which are relatively more even between angles. There were still differences however between the latter two. The orchestral stimulus showed a large decrease in median difference with angles away from the median plane.

Between Stimuli

Each column in the table data represents the value of a stimuli subtracted from another. O = orchestral, P = pink noise and S = speech, this key to the stimuli appears in the table legend. The stimuli were subtracted in this order due to expec-

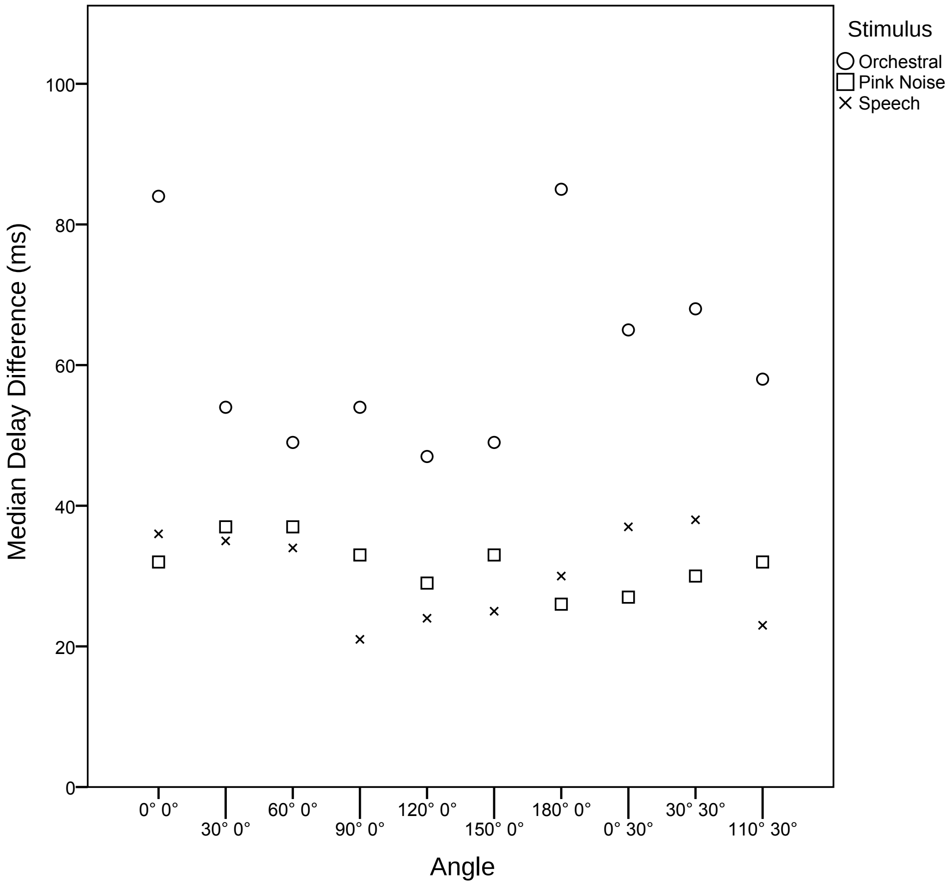

Figure 3.13: The median difference between tests for each stimulus at different angles.

tations of positive values from initial impressions of the data. The results would be correct regardless of which stimulus is subtracted from the other but would have to be reversed. In test one, the largest median differences were between the orchestral and pink noise (O-P) stimuli ranging 26 – 47 ms, and the orchestral and speech (O-S) stimuli ranging 21 ms – 40 ms. Between the speech and pink noise stimuli, the median of paired differences did not range to such an extent (-3 – 9 ms). In test two, there were again the most increases in median difference between angles for O-P (1 – 122 ms) and O-S (18 – 119 ms), these ranges were much larger than for test one with the noticeable outlier in the range occurring at 0° 0°. Between S-P, there was a median difference of only 21 ms in comparison (-11 – 10 ms). In test one, a large majority of positive differences were evident for the O-P relationship with ten or 11 positive differences. The least positive differences occurred at angles 0° 0° and 30° 30° (nine positive, one negative and one tied). O-S showed marginally less positive values with the two speakers at 30° 30° and 110° 30° showing the least but still eight positive differences. The S-P comparison appeared to be much less conclusive with a slight preference to positive values for the speech stimulus. In test two, the conclusion was similar, most notably positive values were observed less often for the O-P and in particular S-P relationship in comparison to O-S. In test one, every angle for the O-P relationship proved statistically significantly different ranging p = .001 – p = .021. Most of the angles for O-S were statistically significantly different, the exceptions were 90° 0° (p = .065), 180° 0° (p = .065), 30° 30° (p = .109) and 110° 30° (p = .227). No angles were statistically significantly different for S-P, with significance values ranging p = .344 – 1. In test two, half of the angles for the O-P relationship were significant, this is a strong contrast to test one. These were angles behind the listener or along the median plane. For O-S, the angles feature statistically significant differences at all angles except 60° 0° (p = .065) and 30° 30° (p = .065). There were no statistically significant differences for any angle with regards to the S-P analysis.

Between Angles

The final set of comparisons were between each individual angle for each test and stimulus, this creates a lot of data, but the results from the Friedman test were unclear, so clarification is important. In every subsequent table, the angle in the column is that which was subtracted from. Due to redundant comparisons being omitted, to acquire the missing interactions, the angle should be found in the row instead of the column and the median difference/positive and negative values should be reversed.

For the orchestral stimulus in test one, the largest median differences occurred with angle combinations which included a speaker along the median plane. For 0° 0°, 13 ms was the largest value when paired with 120° 0°. At 180° 0° the range of increases in median difference was 7 – 19 ms. Increasing the angle between 90° 0° – 150° 0° showed that different combinations of angles resulted in a decrease in median difference for the lesser angle. The smallest median difference was 0 ms (no change) for 0° 0° and 30° 0°; 0° 0° and 0° 30°; 60° 0° and 150° 0°; 120° 0° and 30° 30°. The largest median difference range was 22 ms at 30° 0° (-7 – 15 ms) and the smallest median difference range was 10 ms at 150° 0° (-8 – 2 ms). When observing height channels, there was no difference between 0° 0° and 0° 30°. There was a miniscule median increase for 30° 0° over 30° 30° of 3 ms and median decrease for 120° 0° over 110° 30° of -3 ms. 180° 0° in comparison to 0° 0° had an increase in median difference of 7 ms. Both 0° 0° and 180° 0° showed the most positive values in comparison to the other angle combinations. The only statistically significant increases in median difference were with the angles 0° 0° and 60° 0° ($p = .012$); 0° 0° and 150° 0° ($p = .012$); 0° 0° and 30° 30° ($p = .021$); 0° 0° and 110° 30° ($p = .021$).

In comparison to the orchestral stimulus in test one, the pink noise in test one featured many interactions which showed 0 ms median difference indicating no change. The majority of increases in median difference, including some of the largest, were present for interactions with at least one speaker located in the median plane. For 0° 0° the range was 9 ms (0 – 9 ms); 180° 0° the range was 6 ms (0 – 6 ms); 0° 30° the range was 9 ms (-1 – 8 ms); 110° 30° the range was 9 ms (-9 – 0 ms). The smallest median difference range was 5 ms at 30° 0° (-4 – 1 ms) and 150° 0° (-5 – 0 ms). Regarding the median plane, there was a 1 ms increase in median difference between 0° 0° and 0° 30°; 30° 0° and 30° 30°. There was no median difference between 120° 0° and 110° 30°; 0° 0° and 180° 0°. Angles along the median plane showed the most positive values in comparison to negative and tied. 0° 0° with 150° 0° stood out the most with ten subjects specifying a larger delay with a negative value from one subject. There were again limited pairs which were statistically significant, all of which were confirmed to feature at least one speaker which was located in the median plane. The most numerous were at 180° 0° with 60° 0° ($p = .004$), 90° 0° ($p = .039$), 120° 0° ($p = .021$) and 110° 30° ($p = .021$) altogether. The remaining angles were 0° 0° and 150° 0° ($p = .012$); 60° 0° and 0° 30° ($p = .039$).

For the speech in test one, the largest increases in median differences occurred where one speaker is located in the median plane. At 0° 0° the range is 5.5 ms (0.5 – 6 ms) with the largest difference in comparison to 90° 0°. At 180° 0° the range is 8 ms (-1 – 7 ms) and 0° 30° the range is 9 ms (-2 – 7 ms). The smallest median difference was 0 ms (no change) which was evident with multiple angles for 120° 0° and 150° 0°. No median difference was observed additionally between 30° 0° and 110° 30°. The largest median difference range was 9 ms at 0° 30° (-2 – 7 ms) and the smallest median difference range was 3 ms at 150° 0° (-3 – 0 ms). There were few median differences between lower and upper channels, ranging from 2 – 4 ms. Between 0° 0° and 180° 0° there was only a 1 ms difference. Angles along the median plane once again showed the most positive (increased) values. In decreasing positive quantity order, the median plane angles are as follows: 0° 0°, 180° 0° and 0° 30° There was only the case of a single statistically significant median difference between 90° 0° and 110° 30° ($p = .039$).

For the orchestral stimulus in test two, the range of median differences was exceptionally large as was seen in the boxplots. The largest median difference was 85 ms between 0° 0° and 110° 30°. The smallest median difference was 0 ms (no change) when at least one speaker was a height channel not in the median plane: 60° 0° and 30° 30°; 120° 0° and 110° 30°; 150° 0° and 110° 30°. The largest median difference range was 106 ms at 110° 30° (-85 – 21 ms) and the smallest median difference range was 48 ms at both 90° 0° (-45 – 3 ms) and 30° 30° (-36 – 12 ms). The median difference decreases more as angles become closer together. Between equivalent lower and upper speakers, the median differences are relatively small for 0° 0° and 0° 30° (16 ms); 30° 0° and 30° 30° (-10 ms) and non-existent for 120° 0° and 110° 30°. In comparison to angles out of the median plane, 0° 0° and 180° 0° had a relatively low median difference between them of 15 ms. For 0° 0° the overwhelming quantity of positive differences suggests the subjects agreed that it required a larger delay value. The statistically significant median differences are limited to 0° 0° with 30° 0° ($p = .001$), 60° 0° ($p = .012$) and 90° 0° ($p = .012$); 0° 30° with 30° 0° ($p = .021$) and 60° 0° ($p = .012$); 120° 0° with 150° 0° ($p = .032$).

For the pink noise stimulus in test two, the ranges of the median of paired differences are vastly smaller in comparison to the orchestral stimulus in the same test. The largest median difference was 11 ms between 0° 30° and 120° 0°. The smallest median difference was 0 ms (no change) which occurred at least once at every angle. The largest median difference range was 13 ms at 0° 30° (-2 – 11 ms) and the smallest median difference range was 4 ms at 30° 0° (-4 – 0 ms). Between lower and upper channels there were no median differences except for 0° 0° and 0° 30° (2 ms); 120° 0° and 110° 30° (-5 ms). A 2 ms increase in median difference was likewise observed between 0° 0° and 180° 0°. There was no particular angle which stood out as the positive and negative values appeared to be relatively evenly spread. There were no statistically significant median differences present.

For the speech stimulus in test two, the largest median difference was 16 ms between 0° 30° and 150° 0°. The smallest median difference was 0 ms (no change) between 110° 30° and 0° 0°; 110° 30° and 30° 0°; 90° 0° and 120° 0°. The largest median difference range was 18 ms at 150° 0° (-16 – 2 ms) and the smallest median difference range was 7 ms at 0° 0° (-3 – 4 ms). There were no clear patterns in median differences as the angles increased or decreased. Angles along the median plane and at 30° 30° were found to have the most increases in median difference. Between the lower and upper channels there were no median differences which stood out. There was a 1 ms increase in median difference between 0° 0° and 180° 0°. At 90° 0°, interactions with four angles resulted in positive increases in median difference as specified by 9 subjects. 60° 0° featured two angles with 9 positive values. There were no statistically significant median differences to be reported.

3.3 Discussion

We now run over the main points from the results section and provide analysis beginning with general observations and concluding with comparisons to the literature. The median delay values for the orchestral stimulus were always greater than that for the other stimuli. The pink noise median values in the second test correlate well with the expectations for a truly separated signal as described in the listener instructions for the test. The values provided were within -9 – +6 ms of the length of the pink noise transient (50 ms). The 54 ms maximum value for the stimulus in test one indicates that some listeners did confuse it with the instructions for the second test. In both tests, the stimulus was found to feature the least median differences between angles. In test two, the most agreement between subjects was for the loudspeakers located in the median plane in comparison to any other test, stimulus and angle. In test two, there were less median differences between angles compared to test one. There was the most agreement between listeners for the speech in test one in comparison to test two, however there were more median differences in comparison to the pink noise tests. As signals became more complex, the subjects appeared to find it more difficult to agree on a delay value represented by the orchestral stimulus results in test one. As these complex signals overlapped it became difficult to interpret the beginning point of the sound and the delayed point due to frequent changing transients. Considering the sample size for the tests there were many outliers, a greater sample size may have removed the outliers. However it is believed that understanding the marking criteria and familiarisation is more important to eliminating outliers and reducing the range of delay values. In particular for the orchestral stimulus, but evident for the pink noise and speech, was a smaller median difference between tests as the angle diverged from the median plane. The evidence points towards an increase in median difference at angles situated in the median plane in comparison to those which were not. Stumpp (1936) found that when the reflection came from the same direction as the direct sound, that there was a larger echo threshold in comparison to when the reflection came from a different angle which agrees with the results in this paper. It does appear that the same location is not critical to the increase in comparison to the angle merely being along the median plane although a lateral direct source was used which should be considered. There were no statistically significant differences between angles for the pink noise and speech stimuli in test two and the remaining differences often occurred where there was a single speaker in the median plane. Clear delay differences were found between the orchestral and pink noise stimuli and the orchestral and speech stimuli. Between the speech and pink noise however there were extremely minimal differences. Unanimous statistically significant differences were found between tests for the pink noise and speech stimuli. The orchestral stimulus showed statistically significant differences for half of the angles. Between the upper and lower channels, there was limited difference ranging from 0 – 4 ms. There was minimal difference between 0° 0° and 180° 0°, typically ranging from 0 ms – 2 ms except for the orchestral stimulus in test one which was 7 ms. Blauert (1983) has mentioned how the smallest echo thresholds for speech begin around 20 ms. This agrees closer to the results in test one as expected relatively closer than test two. However, a lower threshold was specified by the listeners in the current research as shown by the median value of 15 ms. Rakerd et al. (2000) found that there were no significant differences between combinations of loudspeakers for the masked threshold in the horizontal plane. This appears to be in agreement with the research presented yet they did compare variations in level difference. The research focused on the lead sound from multiple angles in addition to directly in front of the listener at 0° 0°. Although when the direct sound was in front of the listener the echo thresholds were once again lower for the delayed sound in the left and right speakers compared to the front. Another finding was that echo thresholds in the vertical plane were higher than the horizontal plane as was the case in this research. Furthermore, Yang and Grantham (1997) discovered that the separation of loudspeakers did not have a consistent effect on the echo thresholds in the horizontal plane. However, it was measured with click stimuli. The echo threshold has been found to be significantly reduced when spatial separation is implemented (Litovsky & Colburn, 1998) which may have had an impact on the research. Agaeva and Al'tman (2008) tested echo thresholds within the median plane, the results of which show that the white noise of 5, 10, 20 and 100 ms in length featured echo threshold values ranging from 5 – 14 ms in comparison to the much extended range for the pink noise in the present research. However the median values were at 6 – 28 ms in test one which is certainly comparable. Butler and Humanski (1992) found that localisation in the horizontal plane is more accurate than in the vertical plane. The range of values in the present research for the median angles is not necessarily larger to reflect this, however the large median values may reflect this. In comparison to the tests performed by Schubert and Wernick (1969), the pink noise in test one resembled the results much, whereas in test two they certainly didn't. An 11 ms median was calculated overall for test one and a 54 ms delay in test two. The 50 ms noise results by Schubert and Wernick (1969) were 8 ms (HP) and 12 ms (LP) indicating their thresholds were more of a first impression of separation. Additionally, the noise stimulus in the present research was not identical as it wasn't high or low passed and a triangular envelope was not applied. The results agree with those by Haas (1949) for the implications of different angles. There were negligible differences to be found. The specification of "critical delay differences" and "disturbance" in the Haas test were certainly intriguing and comparisons between the results were important. Critical delay difference values of 44 ms and 55 ms for the speech are much closer to the results in test two indicating disturbance occurs nearer to the separation point than the beginning of separation. It is interesting to note that as the precedence effect is known to disappear above approximately 50 ms, this is approximately where the median for the speech occurred in test two.

3.3.1 Implications of the Results

The results provided in this thesis suggest that echo thresholds should be considered as dynamic and dependent on a multitude of factors. By providing such information it will be simpler in the future to design spatial impression modelling software or objective measures which are both veridical yet also omit superfluous information resulting in an efficient approach. There is the potential for creating more realistic multichannel surround sound recordings with the knowledge that the specific angles themselves are not critical as long as the speakers surround the listener. With the renewed interest in median plane representation of sound in recent years, the research provides important clarification that the these angles should feature a larger delay difference in comparison to those on the horizontal plane to be noticeable as an echo. Previous research did not utilise channels at elevation angles above or below 0° when away from the median plane, this information will therefore be more relevant to multichannel surround systems in comparison. Most importantly, the research may provide agreement between researchers as to how to describe echo thresholds in the future, therefore leading to more accurate and defined research.

3.3.2 Limitations & Future Work

Although the research has not touched the surface of the knowledge required for a full model of echo thresholds, it does provide suggestions as to the correct path. The extreme outliers and large whiskers in the boxplots may show misunderstanding or difficulty judging the criteria and therefore a greater sample size and deep familiarity with the test procedure should be considered for future work. There is still much work to be done in considering the best listener instructions. Due to the vast changes which happen as the delay between two sounds increases, it is imperative that each listener is marking the same criteria. A simple and contrasting explanation for each criteria would work best but may be difficult to acquire. There should be a finer grading available in the GUI between 0 and 1 ms as it appeared that a few subjects tried to approach a smaller value than 1 ms, although this would be classed as summing localisation. It may be wise to perform a set of tests where the listeners have to grade when a sound is beginning to separate or is clearly separated. This is effectively a backwards version of the test in this research but would not have been easy to set up prior to the results gathered here. The benefit of this additional research would be discovering each listeners individual perception of the instructions and it would prevent the bias from this research when performing the second test (whether that be test one or two) — they would not prejudge that the delayed sound should be either higher (than test one) or lower (than test two.) It would be informative to re-sit the test with different lengths and band-limited pink noise to see if the patterns continue. Further complex signals need to be examined as the orchestral stimulus in this research became confusing for many listeners which may have been the nature of longer complex signals but it should be queried nonetheless. It would be interesting to repeat the same parameters as this research but with the direct sound away from the front of the listener as was the case in previous research (Rakerd et al., 2000). This would provide a further model. An obvious limitation in the research was maintaining the reflection level at the same level as the direct sound, although this was intentional, it is not realistic by any means and should be considered.

4 Conclusion

To fill in the gaps that were left by previous research we have looked into multiple stimuli, angles and listener instruction. In contrast, the sound level intensity remained constant between lead and lag sounds unlike a large majority of prior research. Despite these differences, the general setup remained equivalent. It was found that between criteria for all stimuli there were clear significant differences which implies that echo thresholds should not be considered only for a single stimulus. The orchestral stimulus featured significant differences in comparison to the pink noise and speech stimuli, however the pink noise and speech were not generally significantly different from each other. The disturbance as described by Haas and many others previously is understandable with complex signal such as the orchestral and speech stimuli as the ears of the listeners find it difficult to comprehend what is being heard. Disturbance is not directly an echo threshold and should be considered as more intermediary phenomenon between the beginning of separation and the eventual split. It is possible that this disturbance caused uncertainties in delay grading and therefore suggests the lack of significant differences. Between angles, there were overall no clear differences to reported which is agreeable with the past research. Angles within the median plane were generally found to require larger delay values than those which weren't. This agreed again with past research which cited larger delay values along the median plane. There were no clear differences between the upper channels and their equivalent lower channels which has been as of yet to the knowledge of the author not been researched for echo thresholds. Between 0° 0° and 180° 0° there were additionally no differences to be found. The echo threshold limit is in conclusion found to be much more flexible than being stated as the upper limit of the precedence effect. It is suggested that the beginning of separation is closer to that of summing localisation and for complex and "disturbing" sources the delay values required may largely exceed previously believed upper limits of the precedence effect. It is hoped that this research helps to bridge the gaps left by previous research in addition to bringing it together in a more meaningful manner.

Bibliography

Agaeva, M. Y. & Al'tman, Y. A. (2008). Echo thresholds measured in the vertical and horizontal planes. *Human Physiology*, *34*(6), 678–684.

Babkoff, H. & Sutton, S. (1966). End point of lateralization for dichotic clicks. *The Journal of the Acoustical Society of America*, *39*(1), 87–102.

Barron, M. (1971). The subjective effects of first reflections in concert halls—the need for lateral reflections. *Journal of sound and vibration*, *15*(4), 475–494.

Barron, M. & Marshall, A. H. (1981). Spatial impression due to early lateral reflections in concert halls: the derivation of a physical measure. *Journal of Sound and Vibration*, *77*(2), 211–232.

Beranek, L. L. (2010). Listener envelopment lev, strength g and reverberation time rt in concert halls. In *Proceedings of 20th international congress on acoustics, ica.*

Blauert, J. (1983). *Spatial hearing: the psychophysics of human sound localization.* MIT Press.

Boerger, G. (1965a). *Die lokalisation von gausstönen.*

Boerger, G. (1965b). Über die trägheit des gehörs bei der richtungsempfindung. *Proc 5th Int Congr Acoustics, Liège.*

Bradley, J. S. & Soulodre, G. A. (1995a). Objective measures of listener envelopment. *The Journal of the Acoustical Society of America*, *98*(5), 2590–2597.

Bradley, J. S. & Soulodre, G. A. (1995b). The influence of late arriving energy on spatial impression. *The Journal of the Acoustical Society of America*, *97*(4), 2263–2271.

Bradley, J. S., Soulodre, G. A., & Popplewell, N. (1993). Pilot study of simulated spaciousness. *The Journal of the Acoustical Society of America*, *93*(4), 2283. doi:10.1121/1.406555

Butler, R. A. & Humanski, R. A. (1992). Localization of sound in the vertical plane with and without high-frequency spectral cues. *Perception & psychophysics*, *51*(2), 182–186.

Cherry, E. C. & Taylor, W. K. (1954). Some further experiments upon recognition of speech with one and with two ears. *The Journal of the Acoustical Society of America*, *26*, 554–559.

Cochran, P., Throop, J., & Simpson, W. (1968). Estimation of distance of a source of sound. *The American journal of psychology*, 198–206.

Coleman, P. D. (1962). Failure to localize the source distance of an unfamiliar sound. *The Journal of the Acoustical Society of America*, *34*(3), 345–346.

Damaske, P. (1971). Head-related two-channel stereophony with loudspeaker reproduction. *The Journal of the Acoustical Society of America*, *50*(4B), 1109–1115.

Ebata, M., Sone, T., & Nimura, T. (1968). On the perception of direction of echo. *The Journal of the Acoustical Society of America*, *44*(2), 542–547.

Freyman, R. L., Clifton, R. K., & Litovsky, R. Y. (1991). Dynamic processes in the precedence effect. *The Journal of the Acoustical Society of America*, *90*(2), 874–884.

Guttman, N. (1962). A mapping of binaural click lateralizations. *The Journal of the Acoustical Society of America*, *34*(1), 87–91.

Haas, H. (1949). The influence of a single echo on the audibility of speech. *Journal of the Audio Engineering Society*, *20*(2), 146–159. English Translation (1972).

Haas, H. (1951). Über den einfluss eines einfachechos auf die hörsamkeit von sprache. *Acta Acustica united with Acustica*, *1*(2), 49–58.

Hidaka, T., Beranek, L. L., & Okano, T. (1995). Interaural cross-correlation, lateral fraction, and low-and high-frequency sound levels as measures of acoustical quality in concert halls. *The Journal of the Acoustical Society of America*, *98*(2), 988–1007.

Hidaka, T., Okano, T., & Beranek, L. (1992). Interaural cross correlation (iacc) as a measure of spaciousness and envelopment in concert halls. *The Journal of the Acoustical Society of America*, *92*(4), 2469–2469.

Klemm, O. (1920). Untersuchungen über die lokalisation von schallreizen iv: über den einfluss des binauralen zeitunterschieds auf die lokalisation. *Arch ges Psychol*, *40*, 117–145.

Litovsky, R. Y. & Colburn, H. S. [H S]. (1998). Precedence effects in the azimuthal and sagittal planes, 53.

Litovsky, R. Y., Colburn, H. S. [H Steven], Yost, W. A., & Guzman, S. J. (1999). The precedence effect. *The Journal of the Acoustical Society of America*, *106*(4), 1633–1654.

Litovsky, R. Y., Dizon, R. M., & Colburn, H. (1999). Studies of the precedence effect in the median-sagittal and azimuthal planes in a virtual acoustic space. *submitted to The Journal of the Acoustical Society of America (unpublished).*

Litovsky, R. Y., Rakerd, B., Yin, T. C., & Hartmann, W. M. (1997). Psychophysical and physiological evidence for a precedence effect in the median sagittal plane. *Journal of neurophysiology*, *77*(4), 2223–2226.

Lochner, J. & Burger, J. (1958). The subjective masking of short time delayed echoes by their primary sounds and their contribution to the intelligibility of speech. *Acta Acustica united with Acustica*, *8*(1), 1–10.

Lund, A. & Lund, M. (2013). One-way repeated measures anova in spss statistics. Retrieved from https://statistics.laerd.com/premium/rma/repeated-measures-anova-in-spss-4.php

McGregor, P., Horn, A. G., & Todd, M. A. (1985). Are familiar sounds ranged more accurately? *Perceptual and motor skills*, *61*(3f), 1082–1082.

Mershon, D. H. & King, L. E. (1975). Intensity and reverberation as factors in the auditory perception of egocentric distance. *Perception & Psychophysics*, *18*(6), 409–415.

Meyer, E. & Schodder, G. R. (1952). *Über den einfluss von schallrückwürfen auf richtungslokalisation und lautstärke bei sprache, von erwin meyer und georg r. schodder*. Vandenhoeck und Ruprecht.

Morimoto, M. [Maekawa] & Maekawa, Z. (1989). Auditory spaciousness and envelopment. In *Proc. 13th int. congr. on acoustics* (Vol. 2, pp. 215–218).

Morimoto, M. [Masayuki]. (2002). The relation between spatial impression and the precedence effect.

Morimoto, M. [Masayuki] & Iida, K. (1993). A new physical measure for psychological evaluation of a sound field: front/back energy ratio as a measure for envelopment. *The Journal of the Acoustical Society of America*, *93*(4), 2282–2282.

Morimoto, M. [Masayuki] & Iida, K. (1995). A practical evaluation method of auditory source width in concert halls. *Journal of the Acoustical Society of Japan (E)*, *16*(2), 59–69.

Morimoto, M. [Masayuki] & Iida, K. (1998). Effects of front/back energy ratios of early and late reflections on listener envelopment. *The Journal of the Acoustical Society of America*, *103*(5), 2748–2748.

Morimoto, M. [Masayuki], Iida, K., & Sakagami, K. (2001). The role of reflections from behind the listener in spatial impression. *Applied Acoustics*, *62*(2), 109–124.

Nielsen, S. H. (1991). *Distance perception in hearing.*

Nielsen, S. H. (1993). Auditory distance perception in different rooms. *J. Audio Eng. Soc*, *41*(10), 755–770. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=6982

Okano, T., Beranek, L. L., & Hidaka, T. (1998). Relations among interaural cross-correlation coefficient (iacce), lateral fraction (lfe), and apparent source width (asw) in concert halls. *The Journal of the Acoustical Society of America*, *104*(1), 255–265.

Okano, T., Hidaka, T., & Beranek, L. L. (1994). Relations between the apparent source width (asw) of the sound field in a concert hall and its sound pressure level at low frequencies (gl), and its inter-aural cross correlation coefficient (iacc). *The Journal of the Acoustical Society of America*, *96*(5), 3268–3268.

Rakerd, B., Hartmann, W. M., & Hsu, J. (2000). Echo suppression in the horizontal and median sagittal planes. *The Journal of the Acoustical Society of America*, *107*(2), 1061–1064.

Rosenzweig, M. R. & Rosenblith, W. A. (1950). Some electrophysiological correlates of the perception of successive clicks. *The Journal of the Acoustical Society of America*, *22*(6), 878–880.

Rossing, T. D. (2007). *Springer handbook of acoustics*. Springer.

Roy, A. & Gimbott, F. (1993, March). Automatic echo determination in measured room impulse responses. In *Audio engineering society convention 94*. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=6590

Rumsey, F. (2001). Spatial audio (focal, oxford).

Rumsey, F. (2002). Spatial quality evaluation for reproduced sound: terminology, meaning, and a scene-based paradigm. *Journal of the Audio Engineering Society*, *50*(9), 651–666.

Schubert, E. D. & Wernick, J. (1969). Envelope versus microstructure in the fusion of dichotic signals. *The Journal of the Acoustical Society of America*, *45*(6), 1525–1531.

Soulodre, G. A., Lavoie, M. C., & Norcross, S. G. (2003, June). Objective measures of listener envelopment in multichannel surround systems. In *Audio engineering society conference: 24th international conference: multichannel audio, the new reality*. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=12284

Steinberg, J. C. & Snow, W. B. (1934). Physical factors*. *Bell System Technical Journal*, *13*(2), 245–258. doi:10.1002/j.1538-7305.1934.tb00661.x

Stumpp, H. (1936). *Experimentalbeitrag zur raumakustik*. Verlag von R. Oldenbourg.

Thompson, S. P. (1882). On the function of the two ears in the perception of space. *Philosophical Magazine Series 5*, *13*(83), 406–416. doi:10.1080/14786448208627205. eprint: http://dx.doi.org/10.1080/14786448208627205

Von Békésy, G. & Wever, E. G. (1960). *Experiments in hearing*. McGraw-Hill New York.

Wallach, H., Newman, E. B., & Rosenzweig, M. R. (1949). A precedence effect in sound localization. *The Journal of the Acoustical Society of America*, *21*(4), 468–468.

Wallis, R. & Lee, H. (2016). The reduction of vertical interchannel crosstalk: the analysis of localisation thresholds for musical sources. *presented at the 140th AES convention*.

Yang, X. & Grantham, D. W. (1997). Echo suppression and discrimination suppression aspects of the precedence effect. *Perception & psychophysics*, *59*(7), 1108–1117.

Zahorik, P. (1996). Auditory distance perception: a literature review. *Phd Preliminary Examination, University of West Maddison, Department of Psychology*, 19.

Appendix A Wilcoxon Test Results

The Wilcoxon test showed statistically significant differences between angles only for the orchestral stimulus in test two between 0° 0° and 30° 0°, p = .013.

Angle	O-P	O-S	S-P
0° 0°	.043	.009	.
30° 0°	.002	.022	.
60° 0°	.001	.004	.
90° 0°	.006	.	.
120° 0°	.002	.009	.
150° 0°	.009	.043	.
180° 0°	.032	.032	.
0° 30°	.017	.003	.
30° 30°	.043	.	.
110° 30°	.017	.	.

Angle	O-P	O-S	S-P
0° 0°	.012	.004	.
30° 0°	.	.	.
60° 0°	.	.	.
90° 0°	.	.017	.
120° 0°	.	.032	.
150° 0°	.023	.002	.
180° 0°	.032	.004	.
0° 30°	.004	.004	.
30° 30°	.017	.	.
110° 30°	.043	.003	.

Table A.1: Wilcoxon test results showing statistically significant (p) differences between stimuli at different angles for test one (left) and test two (right). O = orchestral, P = pink noise, S = speech.

Angle	Orchestral	Pink Noise	Speech
0° 0°	.003	.003	.006
30° 0°	.041	.008	.005
60° 0°	.026	.003	.003
90° 0°	.018	.003	.007
120° 0°	.013	.003	.003
150° 0°	.006	.003	.003
180° 0°	.026	.003	.009
0° 30°	.006	.007	.003
30° 30°	.016	.009	.003
110° 30°	.017	.003	.003

Table A.2: Wilcoxon test results showing statistically significant (p) differences between tests at different angles and stimuli.

Appendix B Paired-Samples Sign Test Results

B.1 Between Tests

Angle	Orchestral	Pink Noise	Speech
0° 0°	84	32	36
30° 0°	54	37	34.5[1]
60° 0°	49	37	34
90° 0°	54	33	21
120° 0°	47	29	24
150° 0°	49	33	25
180° 0°	85	26	30
0° 30°	65	27	37
30° 30°	68	30	38
110° 30°	58	32	23

Table B.1: Median of paired differences in milliseconds between tests at different angles and stimuli. (Test 2–Test 1).

Angle	Orchestral	Pink Noise	Speech
0° 0°	11.0.0	11.0.0	11.0.0
30° 0°	7.4.0	10.1.0	10.0.0
60° 0°	9.2.0	11.0.0	11.0.0
90° 0°	8.3.0	11.0.0	9.1.1
120° 0°	9.2.0	11.0.0	11.0.0
150° 0°	10.1.0	11.0.0	11.0.0
180° 0°	10.1.0	11.0.0	10.1.0
0° 30°	10.1.0	10.1.0	11.0.0
30° 30°	9.2.0	10.1.0	11.0.0
110° 30°	9.1.1	11.0.0	11.0.0

Table B.2: Positive, negative and tie values respectively between tests at different angles and stimuli. (Test 2–Test 1)

[1]Due to a missing value from a subject during testing, a median is calculated between 10 subjects here.

Angle	Orchestral	Pink Noise	Speech
0° 0°	.001	.001	.012
30° 0°	.549	.012	.002
60° 0°	.065	.001	.001
90° 0°	.227	.001	.021
120° 0°	.065	.001	.001
150° 0°	.012	.001	.001
180° 0°	.012	.001	.012
0° 30°	.012	.012	.001
30° 30°	.065	.012	.001
110° 30°	.021	.001	.001

Table B.3: Sign test results showing statistically significant and not significant (p) differences between tests at different angles and stimuli.

B.2 Between Stimuli

Angle	O-P	O-S	S-P
0° 0°	44	37	-1
30° 0°	43	38	2.5
60° 0°	29	28	5
90° 0°	26	24	5
120° 0°	34	24	5
150° 0°	38	36	2
180° 0°	47	36	-1
0° 30°	33	40	-3
30° 30°	34	36	6
110° 30°	29	21	9

Angle	O-P	O-S	S-P
0° 0°	122	119	-4
30° 0°	29	26	-6
60° 0°	34	33	-2
90° 0°	1	18	-7
120° 0°	25	32	-3
150° 0°	43	38	-3
180° 0°	61	52	-8
0° 30°	51	63	2
30° 30°	43	40	10
110° 30°	34	45	-11

Table B.4: Median of paired differences in milliseconds between stimuli at different angles for test one (left) and test two (right). O = orchestral, P = pink noise, S = speech.

Angle	O-P	O-S	S-P
0° 0°	9.1.1	10.1.0	4.6.1
30° 0°	11.0.0	9.1.0	6.4.0
60° 0°	11.0.0	10.0.1	6.5.0
90° 0°	10.1.0	9.2.0	7.3.1
120° 0°	11.0.0	10.1.0	6.5.0
150° 0°	10.1.0	9.1.1	6.4.1
180° 0°	10.1.0	9.2.0	5.6.0
0° 30°	10.1.0	10.0.1	4.6.1
30° 30°	9.1.1	8.2.1	6.5.0
110° 30°	10.1.0	8.3.0	7.4.0

Angle	O-P	O-S	S-P
0° 0°	10.0.1	10.1.0	4.6.1
30° 0°	7.4.0	9.1.1	4.7.0
60° 0°	7.4.0	9.2.0	5.6.0
90° 0°	7.4.0	10.1.0	3.8.0
120° 0°	9.2.0	9.1.1	5.6.0
150° 0°	9.1.1	11.0.0	4.6.1
180° 0°	10.1.0	10.1.0	4.7.0
0° 30°	10.1.0	11.0.0	6.5.0
30° 30°	9.2.0	9.2.0	8.3.0
110° 30°	10.1.0	10.1.0	3.7.1

Table B.5: Positive, negative and tie values respectively between stimuli at different angles for test one (left) and test two (right). O = orchestral, P = pink noise, S = speech.

Angle	O-P	O-S	S-P
0° 0°	.021	.012	.754
30° 0°	.001	.021	.754
60° 0°	.001	.002	1
90° 0°	.012	.065	.344
120° 0°	.001	.012	1
150° 0°	.012	.021	.754
180° 0°	.012	.065	1
0° 30°	.012	.002	.754
30° 30°	.021	.109	1
110° 30°	.012	.227	.549

Angle	O-P	O-S	S-P
0° 0°	.002	.012	.754
30° 0°	.549	.021	.549
60° 0°	.549	.065	1
90° 0°	.549	.012	.227
120° 0°	.065	.021	1
150° 0°	.021	.001	.754
180° 0°	.012	.012	.549
0° 30°	.012	.001	1
30° 30°	.065	.065	.227
110° 30°	.012	.012	.344

Table B.6: Sign test results showing statistically significant and not significant (p) differences between stimuli at different angles for test one (left) and test two (right). O = orchestral, P = pink noise, S = speech.

B.3 Between Angles

B.3.1 Orchestral Test One

Angle	0° 0°	30° 0°	60° 0°	90° 0°	120° 0°	150° 0°	180° 0°	0° 30°	30° 30°
30° 0°	0	.	.	.	.	.	.	.	.
60° 0°	12	10	.	.	.	.	.	.	.
90° 0°	11	5	2	.	.	.	.	.	.
120° 0°	13	9	2	-2	.	.	.	.	.
150° 0°	8	6	0	-2	-2	.	.	.	.
180° 0°	-7	-7	-11	-19	-19	-8	.	.	.
0° 30°	0	-1	-1	-9	-6	-5	10	.	.
30° 30°	8	3	-6	-3	0	-1	13	4	.
110° 30°	8	15	-1	-5	-3	-1	15	5	0

Table B.7: Median of paired differences in milliseconds between angles for the orchestral stimulus in test one.

Angle	0° 0°	30° 0°	60° 0°	90° 0°	120° 0°	150° 0°	180° 0°	0° 30°	30° 30°
30° 0°	4.5.2	.	.	.	.	.	.	.	.
60° 0°	10.1.0	8.3.0	.	.	.	.	.	.	.
90° 0°	8.2.1	6.4.1	6.5.0	.	.	.	.	.	.
120° 0°	9.2.0	7.3.1	8.2.1	4.6.1	.	.	.	.	.
150° 0°	10.1.0	7.4.0	5.5.1	4.7.0	4.7.0	.	.	.	.
180° 0°	2.8.1	3.8.0	3.8.0	2.9.0	3.8.0	2.8.1	.	.	.
0° 30°	5.5.1	4.6.1	3.6.2	3.8.0	2.9.0	3.7.1	8.2.1	.	.
30° 30°	9.1.1	7.4.0	4.7.0	5.6.0	4.5.2	5.6.0	9.2.0	7.4.0	.
110° 30°	9.1.1	8.3.0	5.6.0	4.7.0	4.7.0	4.6.1	9.2.0	7.4.0	4.4.3

Table B.8: Positive, negative and tied values respectively between angles for the orchestral stimulus in test one.

Angle	0° 0°	30° 0°	60° 0°	90° 0°	120° 0°	150° 0°	180° 0°	0° 30°	30° 30°
30° 0°	1	.	.	.	.	.	.	.	.
60° 0°	.012	.227	.	.	.	.	.	.	.
90° 0°	.109	.754	1	.	.	.	.	.	.
120° 0°	.065	.344	.109	.754	.	.	.	.	.
150° 0°	.012	.549	1	.549	.549	.	.	.	.
180° 0°	.109	.227	.227	.065	.227	.109	.	.	.
0° 30°	1	.754	.508	.227	.065	.344	.109	.	.
30° 30°	.021	.549	.549	1	1	1	.065	.549	.
110° 30°	.021	.227	1	.549	.549	.754	.065	.549	1

Table B.9: Sign test results showing statistically significant and not significant (p) differences between angles for the orchestral stimulus in test one.

B.3.2 Pink Noise Test One

Angle	0° 0°	30° 0°	60° 0°	90° 0°	120° 0°	150° 0°	180° 0°	0° 30°	30° 30°
30° 0°	4	.	.	.	.	.	.	.	.
60° 0°	7	0	.	.	.	.	.	.	.
90° 0°	6	0	0	.	.	.	.	.	.
120° 0°	7	0	0	0	.	.	.	.	.
150° 0°	5	0	0	0	1	.	.	.	.
180° 0°	0	-3	-4	-2	-3	-2	.	.	.
0° 30°	1	-1	-4	-7	-1	-4	1	.	.
30° 30°	7	1	0	0	0	0	1	6	.
110° 30°	9	0	0	0	0	0	6	8	1

Table B.10: Median of paired differences in milliseconds between angles for the pink noise stimulus in test one.

Angle	0° 0°	30° 0°	60° 0°	90° 0°	120° 0°	150° 0°	180° 0°	0° 30°	30° 30°
30° 0°	7.3.1	.	.	.	.	.	.	.	.
60° 0°	8.2.1	5.2.4	.	.	.	.	.	.	.
90° 0°	8.2.1	5.2.4	4.3.4	.	.	.	.	.	.
120° 0°	8.2.1	4.4.3	2.3.6	3.3.5	.	.	.	.	.
150° 0°	10.1.0	5.5.1	3.5.3	5.5.1	6.4.1	.	.	.	.
180° 0°	5.4.2	2.8.1	0.9.2	1.8.2	1.9.1	2.8.1	.	.	.
0° 30°	6.5.0	2.7.2	1.8.2	2.8.1	2.7.2	1.7.3	6.5.0	.	.
30° 30°	7.3.1	6.2.3	4.5.2	4.5.2	4.5.2	4.4.3	7.3.1	7.2.2	.
110° 30°	8.2.1	5.2.4	4.3.4	4.4.3	3.5.3	5.5.1	9.1.1	8.2.1	6.2.3

Table B.11: Positive, negative and tied values respectively between angles for the pink noise stimulus in test one.

Angle	0° 0°	30° 0°	60° 0°	90° 0°	120° 0°	150° 0°	180° 0°	0° 30°	30° 30°
30° 0°	.344	.	.	.	.	.	.	.	.
60° 0°	.109	.453	.	.	.	.	.	.	.
90° 0°	.109	.453	1	.	.	.	.	.	.
120° 0°	.109	1	1	1	.	.	.	.	.
150° 0°	.012	1	.727	1	.754	.	.	.	.
180° 0°	1	.109	.004	.039	.021	.109	.	.	.
0° 30°	1	.180	.039	.109	.180	.070	1	.	.
30° 30°	.344	.289	1	1	1	1	.344	.180	.
110° 30°	.109	.453	1	1	.727	1	.021	.109	.289

Table B.12: Sign test results showing statistically significant and not significant (p) differences between angles for the pink noise stimulus in test one.

B.3.3 Speech Test One

Angle	0° 0°	30° 0°	60° 0°	90° 0°	120° 0°	150° 0°	180° 0°	0° 30°	30° 30°
30° 0°	0.5	.	.	.	.	.	.	.	.
60° 0°	5	3.5	.	.	.	.	.	.	.
90° 0°	6	2	-1	.	.	.	.	.	.
120° 0°	3	0	0	-1	.	.	.	.	.
150° 0°	3	2.5	0	0	0	.	.	.	.
180° 0°	1	-4	-7	-4	-5	-3	.	.	.
0° 30°	2	-1	-7	-4	-4	-2	2	.	.
30° 30°	2	-4	-1	-2	-2	-1	4	-1	.
110° 30°	3	0	-5	-3	-2	-3	2	2	2

Table B.13: Median of paired differences in milliseconds between angles for the speech stimulus in test one.

Angle	0° 0°	30° 0°	60° 0°	90° 0°	120° 0°	150° 0°	180° 0°	0° 30°	30° 30°
30° 0°	5.4.1	.	.	.	.	.	.	.	.
60° 0°	8.2.1	7.3.0	.	.	.	.	.	.	.
90° 0°	7.2.2	6.2.2	3.7.1	.	.	.	.	.	.
120° 0°	7.4.0	5.5.0	3.5.3	5.6.0	.	.	.	.	.
150° 0°	8.3.0	7.3.0	4.5.2	5.5.1	4.5.2	.	.	.	.
180° 0°	6.5.0	4.6.0	2.9.0	3.8.0	3.8.0	2.7.2	.	.	.
0° 30°	9.2.0	4.6.0	3.6.2	4.7.0	2.6.3	4.6.1	6.4.1	.	.
30° 30°	7.3.1	3.7.0	3.7.1	4.7.0	2.7.2	4.7.0	7.4.0	2.7.2	.
110° 30°	6.5.0	4.4.2	2.9.0	1.8.2	2.8.1	2.8.1	6.4.1	7.4.0	6.5.0

Table B.14: Positive, negative and tied values respectively between angles for the speech stimulus in test one.

Angle	0° 0°	30° 0°	60° 0°	90° 0°	120° 0°	150° 0°	180° 0°	0° 30°	30° 30°
30° 0°	1	.	.	.	.	.	.	.	.
60° 0°	.109	.344	.	.	.	.	.	.	.
90° 0°	.180	.289	.344	.	.	.	.	.	.
120° 0°	.549	1	.727	1	.	.	.	.	.
150° 0°	.227	.344	1	1	1	.	.	.	.
180° 0°	1	.754	.065	.227	.227	.180	.	.	.
0° 30°	.065	.754	.508	.549	.289	.754	.754	.	.
30° 30°	.344	.344	.344	.549	.180	.549	.549	.180	.
110° 30°	1	1	.065	.039	.109	.109	.754	.549	1

Table B.15: Sign test results showing statistically significant and not significant (p) differences between angles for the speech stimulus in test one.

B.3.4 Orchestral Test Two

Angle	0° 0°	30° 0°	60° 0°	90° 0°	120° 0°	150° 0°	180° 0°	0° 30°	30° 30°
30° 0°	54	.	.	.	.	.	.	.	.
60° 0°	44	-4	.	.	.	.	.	.	.
90° 0°	45	-3	-3	.	.	.	.	.	.
120° 0°	57	-18	5	-1	.	.	.	.	.
150° 0°	40	-18	-10	-16	-15	.	.	.	.
180° 0°	15	-42	-27	-45	-45	-44	.	.	.
0° 30°	16	-40	-36	-40	-41	-32	-1	.	.
30° 30°	36	-10	0	3	-10	5	4	24	.
110° 30°	85	-21	-13	-19	0	0	23	15	-12

Table B.16: Median of paired differences in milliseconds between angles for the orchestral stimulus in test two.

Angle	0° 0°	30° 0°	60° 0°	90° 0°	120° 0°	150° 0°	180° 0°	0° 30°	30° 30°
30° 0°	11.0.0	.	.	.	.	.	.	.	.
60° 0°	10.1.0	4.6.1	.	.	.	.	.	.	.
90° 0°	10.1.0	4.6.1	5.6.0	.	.	.	.	.	.
120° 0°	9.2.0	5.6.0	7.3.1	5.6.0	.	.	.	.	.
150° 0°	9.2.0	2.8.1	3.8.0	3.8.0	1.9.1	.	.	.	.
180° 0°	6.4.1	2.9.0	2.9.0	3.8.0	3.8.0	3.8.0	.	.	.
0° 30°	6.4.1	1.9.1	1.10.0	2.9.0	3.8.0	2.9.0	4.6.1	.	.
30° 30°	9.2.0	4.6.1	5.5.1	6.5.0	5.6.0	6.4.1	8.3.0	9.2.0	.
110° 30°	9.2.0	4.7.0	4.7.0	4.7.0	5.4.2	5.5.1	7.3.1	7.3.1	5.6.0

Table B.17: Positive, negative and tied values respectively between angles for the orchestral stimulus in test two.

Angle	0° 0°	30° 0°	60° 0°	90° 0°	120° 0°	150° 0°	180° 0°	0° 30°	30° 30°
30° 0°	.001	.	.	.	.	.	.	.	.
60° 0°	.012	.754	.	.	.	.	.	.	.
90° 0°	.012	.754	1	.	.	.	.	.	.
120° 0°	.065	1	.344	1	.	.	.	.	.
150° 0°	.065	.109	.227	.227	.021	.	.	.	.
180° 0°	.754	.065	.065	.227	.227	.227	.	.	.
0° 30°	.754	.021	.012	.065	.227	.065	.754	.	.
30° 30°	.065	.754	1	1	1	.754	.227	.065	.
110° 30°	.065	.549	.549	.549	1	1	.344	.344	1

Table B.18: Sign test results showing statistically significant and not significant (p) differences between angles for the orchestral stimulus in test two.

B.3.5 Pink Noise Test Two

Angle	0° 0°	30° 0°	60° 0°	90° 0°	120° 0°	150° 0°	180° 0°	0° 30°	30° 30°
30° 0°	0	.	.	.	.	.	.	.	.
60° 0°	0	0	.	.	.	.	.	.	.
90° 0°	2	-4	-2	.	.	.	.	.	.
120° 0°	10	0	1	2	.	.	.	.	.
150° 0°	3	0	1	4	-4	.	.	.	.
180° 0°	2	-4	1	0	-10	0	.	.	.
0° 30°	2	-3	-2	-1	-11	-4	0	.	.
30° 30°	3	0	-1	2	0	2	5	5	.
110° 30°	0	-4	-4	-4	-5	-2	5	2	-4

Table B.19: Median of paired differences in milliseconds between angles for the pink noise stimulus in test two.

Angle	0° 0°	30° 0°	60° 0°	90° 0°	120° 0°	150° 0°	180° 0°	0° 30°	30° 30°
30° 0°	5.4.2	.	.	.	.	.	.	.	.
60° 0°	5.3.3	5.5.1	.	.	.	.	.	.	.
90° 0°	6.5.0	4.6.1	4.7.0	.	.	.	.	.	.
120° 0°	7.3.1	5.5.1	6.3.2	7.2.2	.	.	.	.	.
150° 0°	7.2.2	5.5.1	6.4.1	6.4.1	4.6.1	.	.	.	.
180° 0°	6.2.3	4.6.1	7.4.0	5.4.2	3.7.1	4.4.3	.	.	.
0° 30°	6.4.1	3.7.1	5.6.0	3.6.2	3.7.1	3.7.1	4.5.2	.	.
30° 30°	7.4.0	5.4.2	4.6.1	6.4.1	5.5.1	7.4.0	7.4.0	8.3.0	.
110° 30°	5.5.1	3.7.1	4.7.0	4.7.0	3.8.0	5.6.0	6.5.0	6.5.0	2.8.1

Table B.20: Positive, negative and tied values respectively between angles for the pink noise stimulus in test two.

Angle	0° 0°	30° 0°	60° 0°	90° 0°	120° 0°	150° 0°	180° 0°	0° 30°	30° 30°
30° 0°	1	.	.	.	.	.	.	.	.
60° 0°	.727	1	.	.	.	.	.	.	.
90° 0°	1	.754	.549	.	.	.	.	.	.
120° 0°	.344	1	.508	.180	.	.	.	.	.
150° 0°	.180	1	.754	.754	.754	.	.	.	.
180° 0°	.289	.754	.549	1	.344	1	.	.	.
0° 30°	.754	.344	1	.508	.344	.344	1	.	.
30° 30°	.549	1	.754	.754	1	.549	.549	.227	.
110° 30°	1	.344	.549	.549	.227	1	1	1	.109

Table B.21: Sign test results showing statistically not significant (p) differences between angles for the pink noise stimulus in test two.

B.3.6 Speech Test Two

Angle	0° 0°	30° 0°	60° 0°	90° 0°	120° 0°	150° 0°	180° 0°	0° 30°	30° 30°
30° 0°	4	.	.	.	.	.	.	.	.
60° 0°	-2	-1	.	.	.	.	.	.	.
90° 0°	4	6	4	.	.	.	.	.	.
120° 0°	-1	5	4	0	.	.	.	.	.
150° 0°	4	1	2	-2	-1	.	.	.	.
180° 0°	1	-5	-3	-9	-2	-10	.	.	.
0° 30°	-3	-2	-5	-15	-8	-16	-3	.	.
30° 30°	-1	-2	-1	-5	-5	-14	-4	2	.
110° 30°	0	0	3	-3	-4	1	5	2	5

Table B.22: Median of paired differences in milliseconds between angles for the speech stimulus in test two.

Angle	0° 0°	30° 0°	60° 0°	90° 0°	120° 0°	150° 0°	180° 0°	0° 30°	30° 30°
30° 0°	7.4.0	.	.	.	.	.	.	.	.
60° 0°	4.6.1	4.6.1	.	.	.	.	.	.	.
90° 0°	6.4.1	9.2.0	9.1.1	.	.	.	.	.	.
120° 0°	5.6.0	8.3.0	9.2.0	5.5.1	.	.	.	.	.
150° 0°	6.5.0	6.5.0	7.4.0	4.7.0	4.6.1	.	.	.	.
180° 0°	6.5.0	4.7.0	4.7.0	4.7.0	5.6.0	3.7.1	.	.	.
0° 30°	4.7.0	3.8.0	4.7.0	3.8.0	3.8.0	2.8.1	3.6.2	.	.
30° 30°	4.7.0	5.6.0	3.7.1	2.9.0	2.8.1	2.8.1	5.6.0	6.5.0	.
110° 30°	5.5.1	5.4.2	8.2.1	2.9.0	2.8.1	6.4.1	7.4.0	6.5.0	7.3.1

Table B.23: Positive, negative and tied values respectively between angles for the speech stimulus in test two.

Angle	0° 0°	30° 0°	60° 0°	90° 0°	120° 0°	150° 0°	180° 0°	0° 30°	30° 30°
30° 0°	.549	.	.	.	.	.	.	.	.
60° 0°	.754	.754	.	.	.	.	.	.	.
90° 0°	.754	.065	.021	.	.	.	.	.	.
120° 0°	1	.227	.065	1	.	.	.	.	.
150° 0°	1	1	.549	.549	.754	.	.	.	.
180° 0°	1	.549	.549	.549	1	.344	.	.	.
0° 30°	.549	.227	.549	.227	.227	.109	.508	.	.
30° 30°	.549	1	.344	.065	.109	.109	1	1	.
110° 30°	1	1	.109	.065	.109	.754	.549	1	.344

Table B.24: Sign test results showing statistically significant and not significant (p) differences between angles for the speech stimulus in test two.