Auditory Spatial Precision in the Horizontal Plane

Joshua Oscar Stevenson-Hoare

A thesis presented for the degree of Doctor of Philosophy

School of Psychology,
Cardiff University,
United Kingdom,
2020

Statement 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD.

Signed _Joshua Stevenson-Hoare_ Date _11/12/2020___________

Statement 2

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is it being submitted concurrently for any other degree or award (outside of any formal collaboration agreement between the University and a partner organisation).

Signed _Joshua Stevenson-Hoare_ Date _11/12/2020___________

Statement 3

I hereby give consent for my thesis, if accepted, to be available in the University's Open Access repository (or, where approved, to be available in the University's library and for inter-library loan), and for the title and summary to be made available to outside organisations, subject to the expiry of a University-approved bar on access if applicable.

Signed _Joshua Stevenson-Hoare_ Date _11/12/2020___________

Declaration

This thesis is the result of my own independent work, except where otherwise stated, and the views expressed are my own. Other sources are acknowledged by explicit references. The thesis has not been edited by a third party beyond what is permitted by Cardiff University's Use of Third Party Editors by Research Degree Students Procedure.

Signed _Joshua Stevenson-Hoare_ Date _11/12/2020___________

Word Count: 32,313

Summary

Our ability to localise sources of sound provides positional information about distal stimuli across the full sphere of possible directions. However, studies of auditory localisation have largely examined only the front hemifield and have typically used methods which are poor at capturing spatial precision. In this thesis a direct measure of spatial precision was performed for positions all around the horizontal plane. It was found that spatial precision shows an asymmetry between front and rear hemifields, such that precision was worse at oblique positions to the rear of a listener than at oblique positions to their front. This was found to be the case for tasks involving relative localisations of static sounds and judgements about the movement of sounds with and without head movements. Pinna cues were implicated as a source of this precision asymmetry. The pinnae were previously thought to only operate for discrimination between front and rear hemifields and judgements of elevation, so this is a novel finding for auditory spatial perception. A simple method of bypassing the outer ear was developed. It was found that the asymmetry was removed when the pinnae were bypassed. The pinnae were shown to provide additional information for location on the horizontal plane for sources whose origin is in the front hemifield and provide no additional information in the rear hemifield. It is this additional information which increases precision in the front hemifield on the horizontal plane, relative to the rear hemifield.

Table of Contents

List of Figures

List of Tables

Acknowledgements

Of course, it goes without saying that I thank my supervisors Tom Freeman and John Culling, without whose expert guidance this project would never have even gotten off the ground. In particular I would like to express my gratitude for the understanding and patience that they showed during a period in which I was struggling more than normal. Without this bedrock I doubt that I would have completed this programme, so for that I am immensely grateful.

To everyone who helped with data collection and analysis, laboratory construction, batch processing, and supplying food and drink when I was in the lab all day – thank you so much!

I would also like to thank all my friends and colleagues who volunteered their ears to be tormented by white noise in the name of scientific progress. I know first-hand how lengthy these experiments can be, so believe me when I saw that I am grateful for you subjecting yourself to all that!

It is here that I single out two people in particular, whose reassuring presence and unwavering support has been a source of comfort and inspiration during the hardest times of these past few years. Спасибо большое Катя и vielen Danke Ellen für deine Hilfe и ваша поддержка. You are the best of eggs.

Then of course there is my family and the revolving door of puppies along the way. It is from my parents that I get my love for programming and for running, both of which were increasingly useful skills as this programme went along. To my grandparents and extended family, thank you for your pride, cheer, and unconditional love. And to Aunty Mandy, you have been more help than I think you could realise, and for that I am in your debt.

GNU Terry Pratchett

Chapter 1 – Introduction

The auditory system provides spatial information that can serve multiple purposes. From the detection of potential threats or opportunities, to guidance of motor functions when other senses are diminished – the ability to localise sound accurately and precisely is greatly of benefit to the hearing individual.

Unlike in the visual system, auditory stimuli originating from different directions are not mapped directly onto a location-preserving sensory array. Instead, the auditory system must calculate the location of a sound source using the cues that are available at the two ears. These cues provide information to location due to the way that sound waves interact with the head and pinnae (outer ear) when sounds arrive from different directions. These cues are not uniformly informative under all conditions or for all sound source directions, and performance is further limited by the computational ability of the brain. These factors result in patterns of performance of listeners in sound localisation tasks that vary with sound source content and location.

The first part of the Introduction discusses our current understanding of these cues and the levels of spatial perceptive ability that they support. The primary focus of much of the empirical work reviewed is to understand the accuracy with which listeners can localise sound sources. Less attention, however, has been given to the precision of these spatial judgements. Given the normal emphasis in psychophysics on exploring the limits of performance, this is surprising.

Understanding spatial precision is vital for understanding spatial perceptive abilities. In auditory localisation literature, changing the signal-to-noise ratio has been shown to affect perceived eccentricity (Garcia et al., 2017) and head movement response (Ege et al., 2018). In literature concerning cross-modal perception the reliability of cues to location has been shown to determine how these cues are combined (Alais & Burr, 2004; Ernst & Banks, 2002). For more complex abilities such as motion perception, Bayesian models in vision science suggest that the percept of velocity is also affected by the precision of sensory signals (Freeman et al., 2010; Hürlimann et al., 2002). Therefore, in order to be able to understand and predict how listeners will perform on auditory localisation tasks, it is necessary to know about the precision of the signals that they are using.

Unfortunately, as discussed in the second part of the Introduction, determining how spatial precision varies with factors such as the location of a sound is difficult on the basis of the existing literature, mainly because the dependent measures used are poor. For instance, the most common type of method involves a pointing response (e.g., with the finger or nose), which confounds sensory precision with the variability of the motor response. Following a critique of the methods used, and an analysis of the existing data, I propose that a more direct approach is needed to determine spatial precision, one based on a method that is well understood.

1.1 Cues to Sound Source Location

The direction of a sound source relative to the head can be specified by two angles, as shown in Figure 1. The centre of the coordinate system is a point inside the head, typically midway along the interaural axis (an axis between the two ears). The sound source can then be thought of as a point on the surface of a sphere. The azimuth specifies the location of the sound horizontally on the sphere. The elevation specifies the vertical location of the sound.

This system indicates azimuth and elevation using degrees of arc. Azimuth is measured in degrees relative to the median plane of the head, a vertical plane running through the centre of the head from front to back. Elevation is measured in degrees relative to the horizontal plane. These can be thought of as analogous to longitude and latitude respectively, as used for locations of the surface of the Earth.

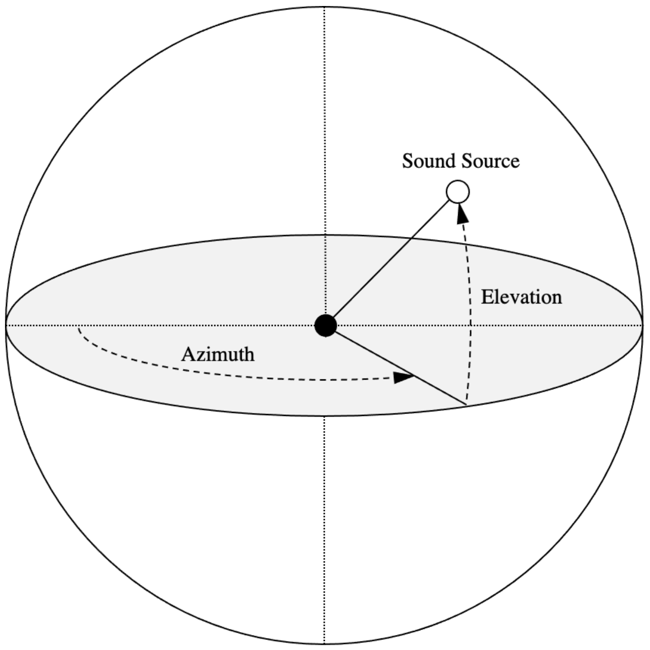

Figure 1. Two-coordinate system showing the azimuth and elevation of a sound source on a surface of a sphere.

Azimuth is typically measured in degrees from 0°, which indicates straight ahead. An azimuth in this scheme must be positive, proceeding clockwise to 359.99…°. Alternatively, azimuth may be specified as positive or negative from 0°, where negative azimuths are to the left of 0° and positive to the right. In this variant, the maximum azimuth is ± 180°.

Elevation is measured as positive or negative degrees from the horizontal plane, with negative elevation indicating a sound below the plane. The maximum displacement from the horizontal plane possible is ± 90°, which corresponds to directly above or below the listener.

Azimuth and elevation are indicated to the listener by various monaural and binaural cues that arise as a result of the direction of the sound source origin relative to the listeners.

1.1.1 **Binaural cues.**

A sound originating from a source on the median plane of the head will generate a proximal stimulus at each ear which is near identical to the stimulus received at the other ear. Any sound source located away from the median plane will produce proximal stimuli at the two ears which differ from each other. These differences are termed interaural differences. There are two primary ways in which these proximal stimuli may differ, and one which is dependent on the stimulus being used. The ear closer to the sound source is referred to as the ipsilateral ear, and the ear further from the stimulus is referred to as the contralateral ear.

1.1.1.1 Interaural time differences.

The difference in time of arrival of the stimuli is referred to as the interaural time difference, or ITD. At room temperature (20°C), the speed of sound in air is approximately 343 ms^{-1}. A sound located off the median plane will reach the ipsilateral ear before it reaches the contralateral one. This can be used to distinguish between sounds left or right of the median plane, as well as the degree of leftness or rightness. The maximal ITD occurs when a sound source is located on the interaural axis as this maximises the difference in time of arrival between the ears. Feddersen, Sandel, Teas, and Jeffress (1957) demonstrated that the maximal ITD for a normal human listener under standard conditions is 690 µs, though different ITDs can be achieved when a different transmitting medium is used, e.g. water (Feinstein, 1973a, 1973b). While the minimum possible ITD is infinitely small, the minimum ITD detectable by the human auditory system is 10 µs (Klumpp & Eady, 1956).

1.1.1.2 Interaural level differences.

The difference in the sound pressure level of a stimulus at the ears is referred to as the interaural level difference, or ILD. Sound waves lose energy as they propagate through air, in accordance with the inverse square law. This states that the energy level of an unimpeded sound wave is inversely proportional to the square of the distance it has travelled from its source. As with the ITD, the ipsilateral ear will experience a greater sound pressure level than the contralateral ear due to the difference in distance from the source. This effect influences all frequencies equally. However, as the ears are not point receivers, but are located on a head, the acoustic ‘shadow’ of the head must also be considered.

Lower frequency sounds have longer wavelengths compared to the size of the head, meaning that the head obstructs the propagation of the sound waves very little, generating a very small ILD. Higher frequency sounds have shorter wavelengths, against which the head is a considerable obstacle. Feddersen et al. (1957) found that ILDs for a source at seven feet could be as much as 20 dB for high frequency sounds.

When a sound source is located on the interaural axis, the sound pressure level at the contralateral ear is greater than if the sound were located close to, but not on, the interaural axis. This is known as the ‘acoustic bright spot’ (Macaulay et al., 2010). The bright spot arises because of the way sound waves propagate around a near-spherical obstruction. As the waves strike the head at the side of the ipsilateral ear, they travel the same or similar distance around the head to the contralateral ear. Due to the near identical path length, these waves are in-phase, resulting in constructive interference of the waves at the contralateral ear, increasing the sound pressure level. For this reason, the maximal ILD occurs not for sounds on the interaural axis, but for sounds whose origin is at around 50-70 degrees azimuth (Macaulay et al., 2010).

1.1.2 **Monaural cues.**

A sound on the median plane generates little or no interaural differences as it is the same distance from both ears. Using binaural cues alone it would be impossible to tell where a sound was on this plane, creating a loop extending vertically around the head where any location is equally probable. A sound source not located on the median plane does generate interaural differences, but these do not provide an unambiguous cue to location. The ‘cone of confusion’ (Wallach, 1939) is an imaginary cone, where a sound source originating from any location on the cone's surface will create the same interaural differences.

Fortunately, monaural cues can be used to determine both the elevation and whether the sound is in the front or rear hemifield, resolving the cone of confusion and so aiding spatial localisation. These monaural cues arise due to the shape of the outer ear, called the pinna. The shape of the pinna is unique to every individual, like a fingerprint. When a sound wave reaches the outer ear, the various corrugations and convolutions of the pinna diffract and reflect the sound wave before it reaches the ear canal (Blauert, 1997).

By changing the elevation of a sound source, the angle of incidence between the sound wave and the pinnae changes, resulting in different proximal stimuli at the ear canal. The way in which the sound is affected by the pinnae with source elevation is called the head-related transfer function (HRTF). Each individual's HRTF is unique, and so the way in which their own pinna reflects sound with source elevation must be learned. However, there are general trends to which most pinnae adhere, meaning that a generalised HRTF can be used for virtual sound sources to generate the perception of elevation (Wenzel et al., 1993).

Front-back discrimination can also be achieved through the use of pinna cues. The pinnae point forward (Shaw, 1974) and so affect sounds differently depending on the hemifield from which they originate. Indeed, when one ear is occluded so that only monaural cues remain, listeners are still able to distinguish between front and rear locations (Musicant & Butler, 1984).

Monaural cues from the pinna are not equally effective at all frequencies. Lower frequency sounds have a longer wavelength and so are less affected by the comparatively small anatomy of the outer ear. Higher frequency sounds, above approximately 4 kHz, are much more affected by the shape of the pinnae. Consequently, monaural localisation of high-frequency sound is much better than low-frequency sound (Butler & Humanski, 1992).

1.1.3 **Head movement.**

The ‘cone of confusion’ can be also be relieved through the use of head movements, provided that the sound has sufficient duration. There are three axes around which the head may rotate: yaw, pitch, and roll. Yaw is rotation around the vertical axis, looking to the left and to the right while keeping the head upright. Roll is rotation around a horizontal axis parallel to the median plane, tilting the head so that one ear is more elevated than the other. Finally, pitch is rotation around a horizontal axis parallel to the interaural axis. Here, both ears would remain in a similar area of space while the head tilts up and down.

For a sound on the horizontal plane, changing the yaw of the head will alter the ITD and ILD. As the head rotates, one ear will be brought closer to the sound and one brought further away. This can be used to discriminate between front and rear hemifields (Wightman & Kistler, 1999). Following the same principle, a sound not on the horizontal plane may be localised using roll. Similar to yaw, as the roll angle of the head is changed the ears will be brought closer to and further away from the source of sound. This would be more useful when a sound is located near to or on the intersection of the median and horizontal planes (i.e., above or below the listener). This can be used to discriminate between raised or lowered sounds relative to the horizontal plane. The third axis of rotation, pitch, may also be used to discriminate elevation by changing the spectral filtering from the pinna as the angle of coincidence with the stimuli will be changed with pitch angle.

1.2 The Role of Spectral Content in Spatial Localisation

The spectral content of a sound affects the binaural and monaural cues that result from it, causing variation in localisation performance.

Perhaps the earliest efforts to measure human spatial localisation ability were made by Lord Rayleigh (1876). He found that a listener was able to localise the source of a sound from a brief utterance, a sound with a broad spectrum, with relative ease. This, Rayleigh reported, was the case for sounds located in all directions from the listener. However, when that sound was a ‘very forced squeak or a mere grunt’ (p76) then localisation became trickier. He also reported that a ‘low whistle’ was discriminable as to whether it originated from the left or right hemifields but was often confused between front and rear positions. A replication with tuning forks with frequencies as low as 128 Hz revealed that left-right discrimination was robust at frequencies where ILDs would be inaudible, the same difficulty with front-rear discrimination as was found for tones. In a further experiment, Rayleigh (1907) also found that adding ‘little reflecting flaps’ (p231) to a worn headpiece, oriented towards the rear hemifield instead of the front, caused his assistant to make ‘frequent mistakes’ whereas orientation towards the front hemifield did not.

Rayleigh (1907) confirmed that both the intensity of the sound and the phase are important for the discrimination of whether a sound originates from the left or right and suggested that low frequency tones are lateralised using ITDs, while those of high frequency are lateralised using ILDs. This was expanded upon by Hartley and Fry (1921), who calculated interaural level and phase differences for stimuli which varied in azimuthal location and frequency. They demonstrated that the resultant interaural differences were dependent on the frequency of the tone used. The ‘intensity-ratios’ they calculated showed more pronounced curves as the frequency increased. This concurs with what is now known about interaural differences – that interaural level differences become larger as the frequency of a sound is increased.

Stevens and Newman (1936) famously performed a behavioural examination of localisation ability on the roof of the Harvard University Biological Laboratories. They found that localisation ability was poorest for tones around 3 kHz, with performance improving for sounds with frequencies further from this peak. They explained this as being a region where neither interaural level differences nor interaural phase differences were discriminable. A paucity of reliable information therefore led to increased rates of errors of localisation.

Carlile, Delaney, and Corderoy (1999) used full-spectrum noise and noise high-passed or low-passed at 2 kHz to examine the contribution of areas of the frequency spectrum to localisation. They found that removing frequencies above 2 kHz caused an increase in front-back confusions. Removing frequencies below 2 kHz had comparatively little effect on localisation accuracy, though it did increase in the variability in errors in the rear hemifield. Similar findings were made by Musicant and Butler (1984) using noise high- or low-passed at 4 kHz. Note however that these two studies placed the difficult-to-localise 3 kHz region from Stevens and Newman (1936) into different stimuli. This is suggestive that frequencies in this region, while not informative in themselves for tonal stimuli, are not inherently detrimental to localisation ability.

1.3 Spatial Localisation as a Function of Source Azimuth

Localisation performance is not just affected by the spectral content of sounds, but also by the way in which binaural cues change with sound source azimuth.

1.3.1 How localisation accuracy changes with source azimuth.

Stevens and Newman (1936) observed that spatial localisation ability was superior if the origin of the source of sound was directly in front or behind of the listener, as opposed to being towards their side. Based on the nature of binaural cues this is expected as interaural differences change most dramatically for sounds located closer to the median plane and less so towards the interaural axis.

Oldfield and Parker (1984a, 1984b, 1986) conducted a comprehensive examination of spatial localisation ability across positions in a series of three papers. Listeners were blindfolded and asked to point towards the source of sound (a speaker mounted on a swivelling arm) using a hand-held ‘gun’ with antennae mounted to the end. The position of these antennae at the time of response was recorded with three cameras positioned within the laboratory that were able to capture the exact position and orientation of the ‘gun’ so as to calculate the direction indicated by the listener.

Oldfield and Parker (1984a) found that the azimuthal error (measured as degrees from true position) increased as sound sources were moved further towards the interaural axis in the front hemifield. This concurs with the general findings reported by Stevens and Newman (1936), and with the predictions made from binaural cues. They also found that errors continued to increase as sounds were located beyond the interaural axis in the rear hemifield. This is likely to be at least partly attributable to their response method, which will be discussed in section 1.4.5. They also found a largely constant elevation error, though for sound source locations near the median plane in the rear hemifield this was increased.

In their second set of experiments, Oldfield and Parker (1984b) demonstrated that, when the pinnae were prevented from providing a source of information, then localisation ability for azimuth and elevation were not uniformly affected. Listeners were prevented from using information from the pinnae by having casts made of their outer ear, and then moulds applied which flattened out the pinnae, leaving just the ear canal open to the air.

When listeners wore these moulds errors in elevation judgement were greater compared to those found when the pinnae were not occluded. The number of front-back reversals, confusions between front and rear hemifield origins, also increased. Azimuthal localisation (disregarding front-back reversals) was unaffected though. This demonstrated the importance of the pinnae for localising sound not on the horizontal plane, and for resolving the cone of confusion (Wallach, 1939). As azimuthal localisation is reliant on binaural cues, which are unimpeded by occlusion of the pinnae, localisation in this dimension was not affected.

The third set of experiments conducted by Oldfield and Parker (1986) examined how localisation ability is affected by removing cues from one ear entirely, creating a monaural listening environment. With information from one ear only, binaural cues are no longer available. This should worsen performance for azimuth, but leave elevation intact, which was the broad finding of Oldfield and Parker. However, they did find that there was some spared azimuthal localisation ability. They attribute this to the pinna of the unoccluded ear, which does provide some information about azimuth from the spectral filtering it provides. Under normal hearing conditions it was unclear whether this would be impactful on azimuthal localisation ability, or whether it only becomes effective when all other sources of information are removed.

Oldfield and Parker found some reduction in elevation localisation ability with monaural listening, which they believed was due to the removal of information from the pinna on the opposing side of the head. Overall, however, both elevation and front-rear discrimination abilities were largely unaffected by the removal of information from one ear.

Musicant and Butler (1984) examined spatial localisation with and without pinna occlusion, using broadband, high-pass or low-pass stimuli. They replicated the general findings of Oldfield and Parker (1984a), that localisation was best in front of and directly behind the listener, and worse to the side. They also found that localisation was worse in the rear hemifield than in the front hemifield, though see the discussion in section 1.4.5. Importantly, they demonstrated an interaction between whether the pinnae were occluded, and the type of noise used. When the noise contained high-frequencies (broadband or high-pass noise) then occluding the pinnae caused performance to worsen. When the noise contained little to no high frequency noise (low-pass noise) then occluding the pinnae did not significantly affect localisation performance. This is behavioural evidence of the relevance of the pinnae for high frequencies but not for low frequencies.

As pointed out at the start of the Introduction, the primary focus of the studies described above is to understand how accurately listeners can estimate the location of a sound source, typically as a function of azimuth and/or elevation. Some spatial localisation studies also report measures of precision. In the next section I provide a numerical summary across a set of studies that provide recoverable and comparable data on precision. While these data show some consistent changes in precision as a function of position, especially in the comparison between the front and rear fields in the horizontal plane, the methods and dependent measures used are open to question. Consequently, these are reviewed, and this critique is used to motivate the experiments in this thesis.

1.4 Precision as a Function of Azimuth

1.4.1 Accuracy versus precision.

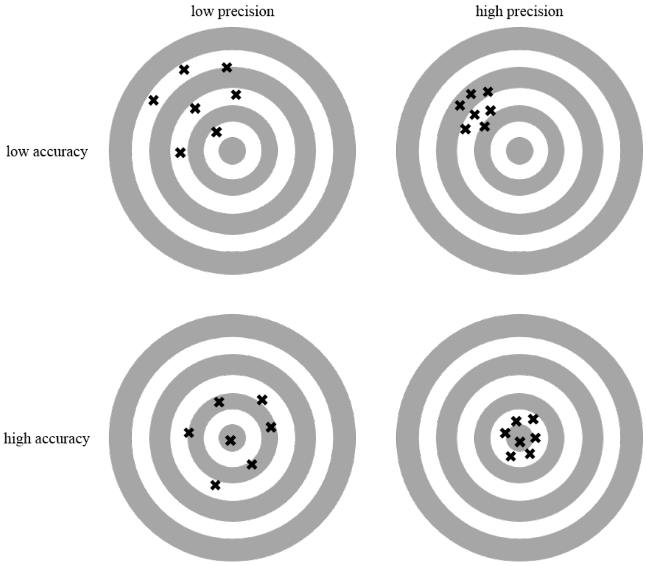

Figure 2. Demonstration of accuracy and precision, using a ‘bullseye’ analogy.

Localisation ability can be quantified in two dimensions: accuracy and precision. A common analogy that is applied when explaining the difference between accuracy and precision is that of throwing darts to hit a bullseye (Figure 2). Throws which are clustered are high in precision. Throws whose average position is over the bullseye are accurate. It is possible for a group of throws to be clustered but consistently far from the bullseye, in which case they would be precise but inaccurate. Alternatively throws which are highly variable but on average hit the bullseye are accurate but imprecise.

Precision is sometimes also referred to as acuity. However, acuity is also taken to mean accuracy in some research (Heffner & Heffner, 2005), so care must be taken in interpreting 'acuity' when it is not explicitly defined. This is discussed in more detail in section 1.4.4.

1.4.2 **Response methods used.**

There are two primary methods of participant response in studies of spatial localisation accuracy: pointing and naming. Pointing methods involve some motor response to gesture towards the origin of the source of sound. Naming methods use a coordinate, degree, or similar system to verbally indicate the sound source origin. These methods purport to examine the same underlying ability, auditory spatial localisation, however their methodology contain different biases that may affect the quality of the data produced.

Pointing methods typically use either a pointing of the hand (Majdak et al., 2010) or some handheld device (Oldfield & Parker, 1984a, 1984b, 1986), or pointing of the head (Carlile et al., 1997, 1999; Makous & Middlebrooks, 1990) or nose (Haber et al., 1993). Naming methods may be in the form of coordinates on the surface of a sphere (Gilkey et al., 1995), degrees of azimuth and elevation (Wenzel et al., 1993), or by the number of the origin speaker (Oldfield & Parker, 1984a, 1984b, 1986). A third set of methods, which are less commonly used in perceptual psychology but often used in audio engineering, are the use of instruments to indicate the location or direction of a sound source. These may be dials (Haber et al., 1993), clock faces (Evans, 1998) or even a sphere (Gilkey et al., 1995; Good & Gilkey, 1996).

1.4.3 **Indirect measurements of precision.**

Some studies have recoverable data on the precision of the responses made by their listeners. I present below a quantitative analysis of the literature on spatial localisation accuracy on the basis of as many studies that I could find that have recoverable precision data. A number of exclusion criteria were implemented before a study's data were included in the analysis, primarily to ensure that adequate comparison could be made between studies.

1. Listener's heads must be fixed during sound presentation.
2. Sound presentation must use real sources, not virtual locations.
3. Examination must be conducted under normal listening conditions (i.e., unimpeded binaural hearing).
4. There must be recoverable precision data including 0° azimuth to permit normalisation (see below).

Table 1

Auditory localisation experiments fitting the exclusion criteria.

Authors	Year	N	Stimulus	Response Method	Precision presented as
Stevens & Newman	1936	2	Tones of varied frequency, clicks	Degrees of azimuth	Mean unsigned error
Oldfield & Parker	1984a	8	White noise	Gun held in hand	Mean unsigned error
Musicant & Butler	1984	8	Broadband noise	Speaker numbering	Error Score[a]
Butler et al.	1990	8	White noise, 30 ms repeating bursts	Speaker numbering	Error Score[a]
Makous & Middlebrooks	1990	6	1.8-16 kHz, 50 kHz sampling rate	Head pointing	Mean unsigned error
Gilkey et al.	1995	3	0.53 - 11.0 kHz, 100 Hz train of 25 µsec pulses	GELP sphere[b]	Mean signed error[c]
Carlile et al.	1997	19	1.0-16 kHz, 80kHz sampling rate	Head pointing	Standard deviation of error
Carlile et al.	1999	6	0.4-16 kHz, 80 kHz sampling rate	Head pointing	Angular extent of major axis of distribution

[a] *Mean number of speakers from truth*

[b] *"God's Eye Location Pointing", touching stylus to sphere*

[c] *Averaged over +5°, -5° elevation*

In total, 26 experiments across 18 studies were considered. Of these, 8 experiments fit all four criteria. The details of these experiments are presented in Table 1. The remaining experiments are tabulated in the Appendix.

In order to present the precision data across studies on a comparable scale, the data have been normalised to the precision reported at 0° azimuth, σ_0,. The precision metric reported at each other azimuth, σ_x, is then scaled by dividing by σ_0 to create σ_x':

$$\sigma_x' = \frac{\sigma_x}{\sigma_0} \tag{1}$$

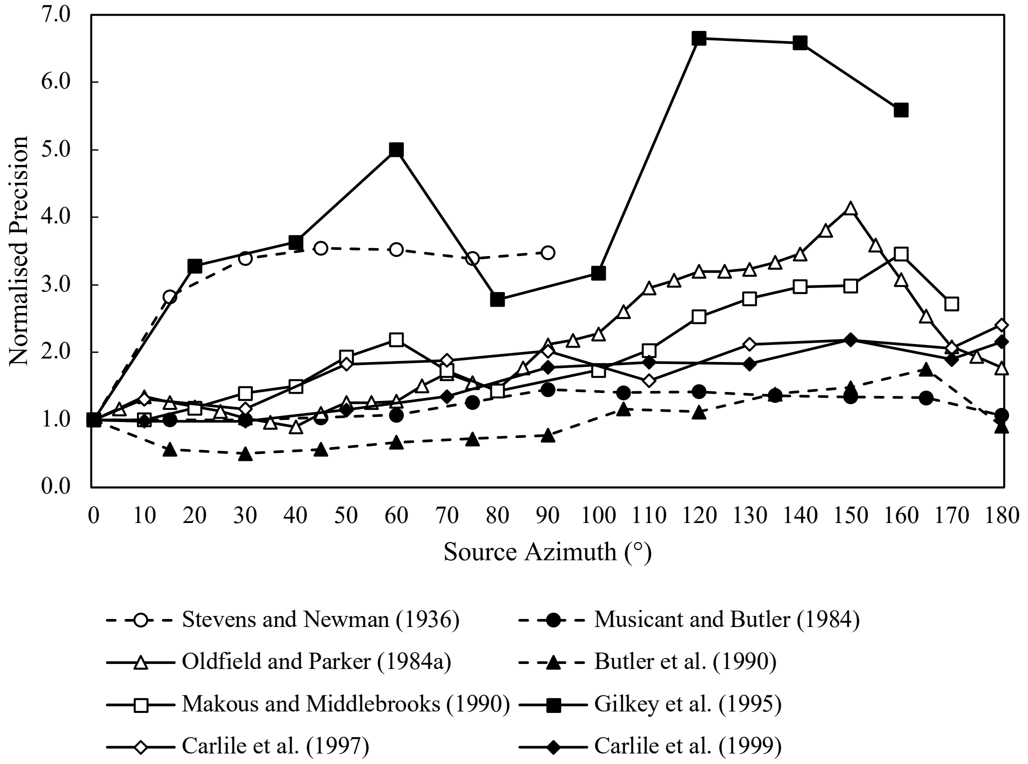

Figure 3. Normalised precision data from sound localisation experiments fitting the exclusion criteria, for sound source locations from 0° to 180° azimuth. Solid lines indicate studies with motor responses. Dashed lines are studies with verbal responses.

Figure 3 plots σ_x' as a function of azimuth. With two exceptions (Gilkey et al., 1995; Stevens & Newman, 1936), all the experiments plotted in Figure 3 show the same pattern of localisation error over azimuth. The lowest error, indicating the highest precision, is found when sound sources are placed at or near the median plane in front. This error increases towards the interaural axis and continues to increase in the rear hemifield. There is finally a reduction in error as sounds are placed close to the median plane behind the head.

The aberrant patterns in Figure 3 are explicable if the details of these studies are examined. In Stevens and Newman (1936) the localisation error reported is averaged over all experiment series in the study. Given that Stevens and Newman used a range of tones, clicks, and hisses, it is not surprising that they should see inflated estimates of localisation error compared to studies which used consistent stimuli. Indeed, as Stevens and Newman only examined positions between 0° and 90° it may be that their error pattern is similar to the others presented, just greater in magnitude for non-0° positions.

Gilkey et al. (1995) can be explained by considering the response method in their study. In this and another study into localisation (Good & Gilkey, 1996), Gilkey and colleagues used a unique method of indicating the location of sound. Listeners used a 'GELP sphere' ("God's Eye Location Pointing"), a sphere with indented lines indicating the major planes. They were instructed to touch a stylus to the surface of the sphere to indicate the perceived location of the sound source. Notably they did so while their head was fixed with a bite bar. This raises issues of both visual input and motor response precision. Listeners were only able to see the 'rear hemifield' of the sphere due to their seating position, so responses to this region of space would benefit from visual guidance as to the location of the stylus. The hand is also not equally dextrous in all orientations, so certain angles of incidence between the stylus and the sphere may have proved harder for listeners to complete than others. While Gilkey et al. report that their listeners were reasonably acute in localising sounds from known locations on the sphere, the stark difference in the pattern of localisation error between this and the other experiments reported in Figure 3 makes it hard to recommend this particular method of reporting sound location.

It must be remembered that these studies only provide indirect measures of spatial localisation precision. Moore, Tollin, and Yin (2008) examined the localisation performance of cats and found that indirect measures of precision were often a poor match for direct measurements of acuity generated from ROC (receiver-operator curve) analysis. They suggest that this is in part due to spatial biases being a source of error in localisation accuracy. A spatial bias, such as a propensity to perceive sounds as being further from the median plane than they really are (Oldfield & Parker, 1984a), would increase the error in localisation accuracy. It would not, however, impact upon direct measures of precision.

1.4.4 **The problem of error metrics.**

There is a general issue within research into auditory spatial localisation in that there is often confusion, if not outright conflation, of precision and accuracy. These are often combined into a single term, ‘acuity’. Although in a true form this would refer to precision alone, this is often not how it is used.

A measurement of localisation acuity is commonly reported using some metric of the ‘error of localisation’ (e.g. Stevens & Newman, 1936). However, this measure is inclusive of both spatial bias (accuracy) and variability of response (precision). Using the dartboard analogy (Figure 2), a dart thrower who hits the same point on the outermost ring of the board every time they throw would have a high ‘error’. The distance from the bullseye is large every time (poor accuracy) but the spread of points is very small (good precision). A different thrower who hits a small distance away from the bullseye (moderate accuracy) but has a spread of points (moderate precision) might be calculated to have the same ‘error’. These two individuals have quite disparate throwing habits (localisation responses) but may be calculated to have the same ‘error of localisation’.

If localisation metrics are not provided as absolute values (i.e., magnitudes rather than vectors), then the collation of multiple results can be misleading. When results are reported as average performances over a set of responses there may be cases, as discussed by Tollins, Populin, and Moore (in Heffner & Heffner, 2005), where results from listeners with opposing spatial biases are combined to produce an average result with no spatial bias. For example, a listener with a bias of +10° and a listener with a bias of –10° would be averaged to a group-level localisation with a bias of 0°. This is clearly unrepresentative of the variability (precision) in the localisations of the listeners and is misleading as to the spatial bias (accuracy) for this sound source.

1.4.5 **The problem of response methods.**

In addition to the error metrics reported in many studies, neither the pointing nor naming methods used are free from problems that may impact on their ability to describe the true precision of localisation ability.

The issue in pointing methods is that the arm and/or head are not capable of pointing to all locations with equal ability. The head, for instance, can only rotate up to 80° in either direction (Swinkels & Swinkels-Meewisse, 2014). The arm is less limited, with a normal range of motion up to -65° on the humeral (horizontal) plane (negative values indicate rearwards motion, with 0° being directly out to the side) (Gates et al., 2016). In both cases, if a sound is located outside of the accessible range, additional body parts must be recruited in order to make a successful response. In most cases this would be a turn of the torso. This additional motor requirement introduces another source of variability which may affect both precision and accuracy. A sound located in the rear hemifield would require some additional turning to identify with a pointing method, whereas a sound located in the front hemifield would be less likely to need this. This may create a confound that artificially inflates estimates of localisation imprecision or inaccuracy in the rear hemifield.

Naming methods do not share the same issue of motor response to the rear hemifield as the pointing response. However, they do have the issue of conceptualising auditory space and producing a verbal response. One issue with these responses is the coarseness of the possible responses. Nielsen (1991) showed that participants have a tendency when presented with a continuous scale to quantise their responses to neat intervals. If, for example, a clock face was used then the maximum level of precision that could be used would be 30 degrees (each number on a clock face being separated by this angle). Similarly, speaker numbering methods are limited by the density of speaker distribution. If sounds only originate from a single speaker at a time this is less disadvantageous, but this would still prevent a listener from indicating a percept of sound location that fell between speakers.

The third, less common method – instrument responses – are also not free from problems. Haber et al. (1993) conducted a set of tests examining nine instrumental methods of indicating sound source origin. Three involved pointing with a body part (nose, chest, finger), two with a rod (cane, 'short stick'), two with a dial (table-mounted, waist-mounted), one drawing on a blank dial, and finally a clock-face verbal response. Haber et al. found that the pointing methods produced fewer errors than the dial methods or the drawing or verbal response methods. However, as Mason (2000) points out, Haber et al.'s subjects were all blind adults with varying durations of loss-of-vision. Visually impaired individuals may have a very different relationship with auditory localisation than normally sighted individuals (e.g. Gori et al., 2014; Zwiers et al., 2001), making comparisons from this study to studies with normally seeing listeners quite hard. Indeed, the drawing response and verbal clock-face response are odd measures to use as many of the 'early-blind' participants (loss before 2 years of age) may have never used a clock or had practice with drawing.

Evans (1998) suggests that abstract methods such as a 'two-step' method may be more promising. This would involve the listener identifying the general region of space from which a sound originates, then using pre-defined responses to specify the location of the sound source relative to a reference sound played from that region. This still imposes a limit on the precision of the response that can be made, though. Using the specifications from Evans, the maximum precision of this measure in a single dimension would be 22.5°, which is far coarser than most pointing method studies.

Whichever method of responding is used, there are issues that can create confounds or limitations within the produced data. It would appear that pointing methods are the least problematic for sounds located in the front hemifield but naming methods may be more fruitful for the rear hemifield. As it would introduce additional issues to switch between methods for different areas, it is unfortunately the case that there is no clear solution to this problem when measuring localisation accuracy is the primary aim of studies.

1.5 Direct Measurements of Spatial Precision for Static Sources

In the studies reviewed above, the measures of precision are necessarily secondary to the primary aim of the research, localisation accuracy. As discussed, this led to issues that stymied the conclusions that could be made about spatial precision. Better measures of precision can be achieved by using methods whose primary aim is to measure the limits of performance. In psychoacoustics the limit of spatial performance is the minimum separation between sound sources required for discrimination of location. This is known as the minimum audible angle (MAA).

For clarity, sound source locations placed within an area of space in which the human hearing system is precise should require only a small angle of separation to be discriminated, that is, heard as coming from spatially separate locations. Sound sources placed within less precise regions of space should require larger angles of separation for the same discrimination performance to be achieved. The MAA therefore represents the precision of the region of space in which the sound sources are located.

The first study of the MAA is most often attributed to Mills (1958). However, they were in fact first systematically studied more than half a century earlier by Daniel Starch (1905). Starch examined the “ability to discriminate between directions of sounds” (p1) in both horizontal and vertical planes, though only the horizontal findings will be reported here. Using an apparatus consisting of three arms mounted on pivots, an attached speaker or speakers could be positioned anywhere on the surface of a sphere (except directly beneath the listener). These speakers could play clicks using a simple circuit or could be made to produce tones by transmitting signals from electrical tuning forks located in a separate room.

Starch reports that his listeners had the smallest MAA (though he did not term it as such) for sound sources located directly in front. The largest MAAs were found for sounds sources located at the side of the head. His listeners, he also reports, were almost but not quite as precise directly behind the head as they were in front. One reason he gives for this is that the pinnae are oriented towards the front such that sounds originating in the front hemifield "are received more easily" (p16), a suggestion that foreshadows some of the conclusions that will be reached in this thesis. Note that Lord Rayleigh (1907) found that listeners made more frequent localisation errors when artificial pinnae ('little reflecting flaps', p231) were oriented towards the rear hemifield compared to the front hemifield, supporting Starch's hypothesis. Starch additionally suggests that we are accustomed to attending to objects in front and so find it easier to attend in this region than behind the head.

Mills (1958) examined not only how the MAA changes with azimuth, but also with other characteristics of the stimulus used. Unlike Starch, Mills only examined the MAA in azimuths from 0° to 90°. The overall pattern of MAA changing with azimuth from Starch was replicated. MAAs were smallest (most precise) when sound sources were located in front of the listener and increased as sounds were placed towards the side of the head, where they were largest. This increase was non-linear, with MAAs changing slowly nearer the median plane and more rapidly as sounds were moved further towards the interaural axis.

The stimuli used by Mills were tone pulses of frequencies between 250 Hz and 10,000 Hz. Mills found that the MAA showed a bimodal response across frequencies. The maximum MAAs for sounds presented at 0° were those for tones with a frequency around 2000 Hz and 8000 Hz. Precision improved for sounds between these frequencies, or lower than 1000 Hz. This was found to be the case across most azimuths. When sounds were presented at 90°, regardless of the frequency used the MAA was more than 40° – the upper limit of measurement in Mills' experiment.

Compared to the studies reviewed earlier that focussed on localisation accuracy, there have been comparatively few studies that have used the MAA to measure precision across the entire horizontal plane, including front and back. This may be in part driven by the idea that localisation in the horizontal plane is driven primarily by binaural cues which are assumed not to be identical between front and rear hemifields (e.g. Mills, 1958; Wallach, 1939). Indeed the only experiment other than Starch (1905) which examines the MAA in the rear hemifield is that of Saberi, Dostal, Sadralobai, and Perrott (1991). They considered how MAAs changed on different planes from horizontal through oblique to vertical. Saberi et al. measured the MAA on the horizontal plane directly behind the head using white noise and found that it did not differ between 180° azimuth and 0°. Consequently, other than Starch (1905), we have no direct measurements of auditory spatial localisation precision in the rear hemifield for positions not on the median plane. But given that precision appears to differ between front and rear hemifields according to those studies that use indirect measures of precision (see Figure 3 above), it seems important to map MAA across the entire horizontal plane.

It should be noted here that the usage of MAA in this thesis exclusively refers to sequential presentations of stimulus pairs. An alternative method is the Concurrent Minimum Audible Angle (CMAA) (Perrott, 1984), in which stimulus pairs are presented simultaneously. Although this method shows a similar increase from the midline towards the interaural axis (Divenyi & Oliver, 1989), CMAAs have always been shown to be worse than MAAs under the same conditions (e.g. source azimuth, spectral content) (Perrott, 1984). This is proposed to be due to the percept of close, simultaneous auditory stimuli being that of a single wide stimulus rather than two separate stimuli (Best, 2004). Evidently the CMAA is not the most precise method of assessing spatial precision as it limits the minimum threshold of performance that can be achieved. Hence the (sequential) MAA is the variant which will be used in this thesis in order to explore the MAA in across the horizontal plane.

1.6 Direct Measurements of Spatial Precision for Moving Sources

A direct measurement of precision using dynamic stimuli is the minimum audible movement angle, or MAMA. Measurement of the MAMA can take two forms. In one, listeners are required to discriminate a moving sound source, usually delivered by a boom-mounted loudspeaker, from a stationary one. In the other, listeners are required to report the direction of source-motion from a single sound presentation.

1.6.1 Detection of motion.

The smallest arc of movement needed to detect motion is the minimum audible movement angle, or MAMA. This was first studied by Harris and Sergeant (1971), though they referred to it as an MAA for a moving sound source. They found that the MAMA was always larger than the MAA for a static sound source. Perrott and Musicant (1977) showed that the minimum arc of movement required for the discrimination of a moving source from a static source varied with the velocity of the stimulus used. Stimuli which moved with a greater velocity resulted in a larger discrimination threshold. This is suggested to be due to a 'minimum integration time' needed for the perception of motion. However, in these experiments the displacement angle was necessarily confounded with duration as this was varied to generate displacement angles with a constant velocity.

Other experiments looking at MAMAs have used a different method, direction discrimination, where listeners are required to judge the direction in which a sound is displaced (Perrott & Marlborough, 1989; Perrott & Tucker, 1988). This method appears to provide very similar minimum thresholds to the motion detection method. Initially this may appear inconsistent, as the discrimination of motion direction involves the collection of more information. However, if detection of motion occurs only due to a detection of change of position then direction discrimination need not have a higher threshold than motion detection. It should be noted that these studies using these two methods are rarely directly comparable due to differences in stimulus spectral content, velocity, and duration, among other factors.

Grantham (1986) examined how the MAMA changed as sounds were presented from different azimuths from 0° to 90°. MAMAs increased as sounds were located further from the median plane towards the interaural axis on the horizontal plane. This is the same pattern of performance as has been observed with MAAs (Mills, 1958). This was found using simulated spatial locations however, with sounds being presented from two loudspeakers which generated a 'moving' stimulus by changing the sound level at each speaker. Subsequently, elements of this study were repeated by Chandler and Grantham (1992) using real sound sources – a loudspeaker on a boom arm in a darkened anechoic chamber. They found that MAMAs for sounds presented at 60° azimuth were still larger than for sounds presented at 0° azimuth. They also corroborated the finding of Perrott and Musicant (1977) that MAMAs increase with sound source velocity. There has been no systematic study of how MAMAs change for sound source azimuths beyond 90°.

The frequency of the stimulus used to measure the MAMA affects the detection and discrimination thresholds found. Perrott and Tucker (1988) found that MAMAs were best for tones with a frequency below 1000 Hz, although they did not find any significant differences between MAMAs for frequencies above this. They do remark on the visual similarity between the pattern of performance they found across frequencies for the MAMA and that found for MAAs by Mills (1958). Chandler and Grantham (1992) directly compared MAAs and MAMAs across frequencies and found that the patterns were highly similar. This, they suggest, shows that MAMAs are reliant on the same underlying cues as MAAs and it is consistent with the idea that auditory motion is essentially an inference of position change based on two successive "snapshots" of the auditory scene.

Perrott and Marlborough (1989) examined whether the MAMA task could be performed with just the start and end points of the movement trajectory and found that listeners were capable of performing this task. They compared this to a standard MAMA with constant motion throughout and found that performance was better when listeners heard a moving stimulus throughout rather than just the start and end points. This, they inferred, is evidence against a simplistic view of motion detection as a comparison of localisations. However, proponents of snapshot theory suggest that it consists of a series of successive localisations at far shorter intervals than that between the static stimuli used in Perrott and Marlborough's experiment. Grantham (1985, in Grantham, 1997) found that MAMAs showed consistent thresholds as the inter-stimulus interval was reduced down to 100 ms, and this task could still be performed with shorter intervals albeit with poorer performance. Furthermore, the pairs of static sounds used in Perrott and Marlborough (1989) resemble an MAA task more than a judgement of motion displacement. Though, this would seem inconsistent with their findings as an MAA threshold is normally smaller than a MAMA threshold. Grantham (1997) explains that this may be because the stimuli used were short bursts of only 10 ms, which would have provided poor spatial information (Chandler & Grantham, 1992).

1.6.2 **Discrimination of velocity.**

Carlile and Best (2002) examined the ability of the auditory system to discriminate velocity using a task with varying displacement cues. In the first condition the start and end point of the reference and test sounds were randomly varied within a range, preventing these being used as cues to displacement. In the second condition the stimuli were given a standard duration with the trajectory centred. This provided displacement cues from the location of the start and end points, as these covaried with the velocity of the stimuli. In a third condition the start point was fixed in space and the end point covaried with stimulus velocity. This also provided a displacement cue, but the difference between the end points of the reference and test stimuli was doubled. Carlile and Best showed that the threshold for velocity discrimination was lower when there were reliable displacement cues available (conditions 2 and 3). The best performance was in condition 3 when the displacement cue from the endpoint of the trajectory was doubled. Listeners were still able to discriminate velocity when displacement cues from the start and end points were disrupted, though (condition 1).

This demonstrates the informativeness of displacement cues to the discrimination of velocity, but also that the auditory system can still function when they are impaired. Freeman et al. (2014) showed that the auditory system prefers to use displacement and duration cues but will use velocity cues if forced to do so by the other cues being made uninformative. This may initially appear to provide some evidence for direct-motion theory as the auditory system appears to process velocity itself. However, Freeman et al. point out that in vision – where there are known to be low-level detectors tuned to motion – velocity dominates over other cues. Given that velocity is not dominant in auditory motion perception it seems unlikely that similar low-level detectors tuned to velocity exist in the auditory system.

Velocity perception and discrimination is thus likely to be a high-level ability generated from inferences made by low-level displacement and duration detectors. For example, Zakarauskas and Cynader (1991) demonstrate that the auditory system could use the first derivative of sound level changes at the ear to detect the velocity of sound sources. However, this only allows for a relative judgement as source intensity will affect the perceptible level change (Warren, 1982). For tasks where a reference velocity is provided (e.g. Carlile & Best, 2002), this is adequate. If an absolute measure of velocity is required, such as estimating time-to-contact (Shaw et al., 1991), this method of assessing velocity is no longer sufficient (Guski, 1992). Consequently, the auditory system is a poor source of auditory velocity information.

1.7 Thesis Outline

The second chapter of this thesis lays out the primary method of assessment of spatial precision that will be used in the experiments presented herein. This is the minimum audible angle (MAA) task. The methods of generating and presenting stimuli are given, and the rationale behind the spatial characteristics of these methods is provided and shown empirically.

In the third chapter the MAA task is applied to the full horizontal plane to directly examine how spatial precision changes with the source azimuth of stimuli. It is found that spatial precision is worse at oblique positions in the rear hemifield than at their counterparts in the front hemifield. A control experiment is also performed to rule out any effects of the acoustics of the laboratory in which the previous experiment was performed. A third experiment extends the examination of spatial precision on the horizontal plane to tasks involving judgements about moving stimuli, by using a minimum audible movement angle task, which tests the threshold for direction discrimination.

The fourth chapter tests spatial precision changes with azimuth when listeners perform a more complex task, in which they are required to integrate information about their own head movements to make accurate assessments of the movement of presented stimuli.

The fifth chapter of this thesis attempts to examine why differences in spatial precision are found between hemifields. The spectral content of stimuli is varied to test how stimulus frequency affects this asymmetry, and then inference is made about the role of the outer ear. A method is generated to examine this possibility and shown in an acoustic mannikin to be effective. This is then tested in human listeners and the effects of removing the contributions of the outer ear on spatial precision are shown.

The final chapter summarises the previous chapters and discusses the implications of the findings of this thesis for methodology, research into spatial localisation, and some applications of the findings. Limitations and future research directions are also discussed.

Chapter 2 – Defining a Direct Measure for Spatial Precision

2.1 Chapter Overview

In section 1.3 it was discussed how indirect measures of spatial localisation precision can often include other sources of variability that become confounds. Direct measures, such as the minimum audible angle (MAA), do not have this issue as they use the same response for all sound source locations. However, there are other important considerations that must be made when designing this task. This chapter describes the primary task used throughout this thesis, the MAA. Hypotheses about the effects of stimulus presentation method and spatial characteristics are made, based on mathematical predictions. These are then tested in an experiment presented using the task. The results from this experiment serve to establish the optimal method of generating and presenting stimuli using an array of loudspeakers.

2.2 The MAA Task

2.2.1 **Stimulus generation and presentation.**

Broadband white noise stimuli (200 Hz – 20 kHz) were used, with a sampling rate of 44,100Hz. The stimulus duration was 500 ms.

Stimuli were localised in space by generating independent noise for each speaker and applying a Gaussian-shaped spatial weighting function centred on the location at which the sound source was intended to appear. The standard deviation of the Gaussian weighting function was 0.7 times the speaker separation, or 5.25°. The rationale behind this number, and the method used to obtain it is explained in section 2.4.

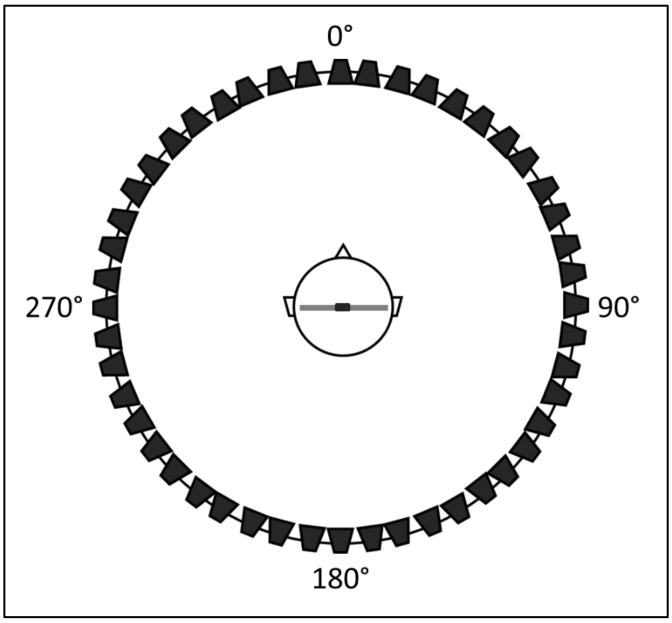

Figure 4. Schematic representation of the ring-shaped array of speakers used in the laboratory, with cardinal positions indicated. The listener is depicted (not to scale) in the centre of the ring, oriented towards the 0° location. A head tracker is also shown being worn by the listener.

A circular array of speakers was used to present the stimuli, shown in Figure 4. This contained 48 speakers separated by 7.5° of arc. The radius of the ring was 1.2m, and speakers were mounted on a ring suspended from poles. The ring was placed at a height from the floor equivalent to the height of the ears of an average seated listener, 1.4 metres. The listener was seated in the centre of the speaker array.

Sound levels were calibrated for each speaker so that the measured mean sound level at the centre of the array was 70 dB. This is slightly louder than usual speech, but quieter than speech in noisy environments such as a train or aeroplane (Olsen, 1998). This level was chosen to be sufficiently loud as to prevent strain but not so loud as to be uncomfortable for extended periods of listening.

2.2.2 **Centring the listener.**

2.2.2.1 Calculation of the effects of mis-centring the listener.

If the listener was not correctly centred within the circular array of speakers, then the perceived location of a presented stimulus (effective azimuth, θ') will not correspond to the actual location of the stimulus on the speaker array (source azimuth, θ). A mathematical analysis was used to assess the tolerance of participant positioning. The effective azimuth can be calculated using the source azimuth, the ring radius R, and the amount by which the head is offset within the array, Δ (2-(6). These measurements are shown in diagram form in Figure 5.

$$L_1 = R\sin\theta \tag{2}$$

$$L_2 = R\cos\theta \tag{3}$$

$$L_3 = L_2 + \Delta \tag{4}$$

$$\tan\theta' = \frac{L_1}{L_3} = \frac{R\sin\theta}{R\cos\theta + \Delta} \tag{5}$$

$$\theta' = \operatorname{atan}\left(\frac{R\sin\theta}{R\cos\theta + \Delta}\right) \tag{6}$$

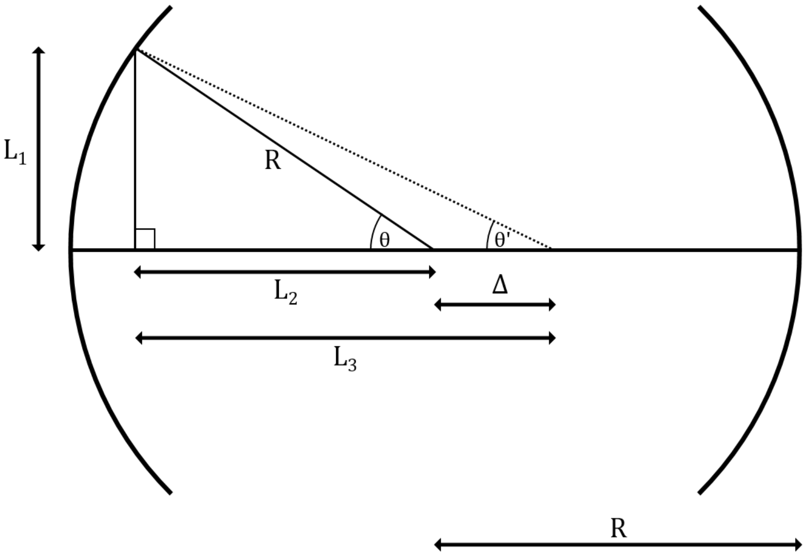

Figure 5. Diagram demonstrating effective azimuth relative to source azimuth.

For example, if the head was offset 25 cm towards the rear of the ring, then a sound with a source azimuth of 45° would have an effective azimuth of 41.08°, an mislocation error of nearly 4°. The error between the source and effective azimuths varies with both the offset of the listener from the centre, and the source location of the sounds. Figure 6 shows the differential of the effective and source azimuths for a theoretical listener offset by 25cm towards the rear of the ring.

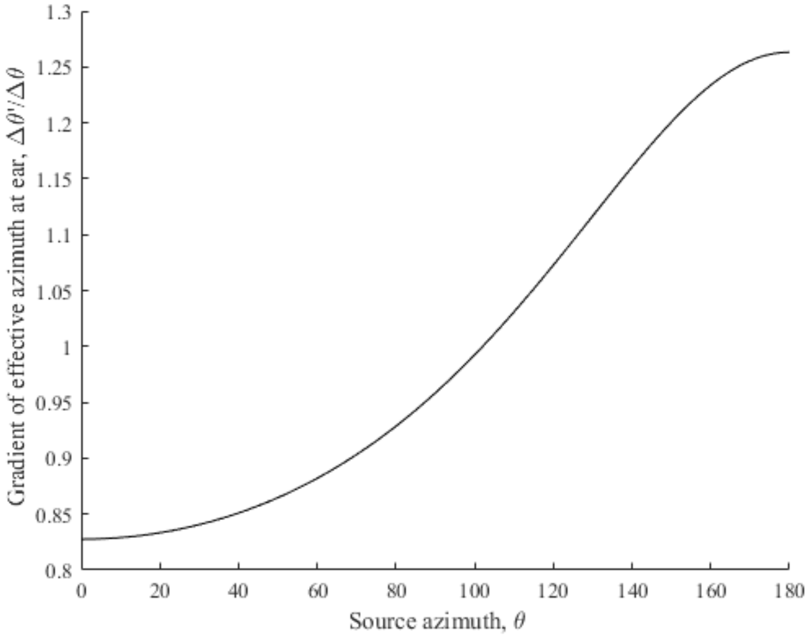

Figure 6. Plotted differential of effective azimuth relative to source azimuth, for a head offset of 25cm.

2.2.2.2 Comparison of centring methods.

To avoid artefactual mislocations of azimuth, it was necessary that listeners be appropriately centred within the speaker array. A method was developed that provided a simple and effective method of ensuring this was the case.

The interaural axis of the listeners was centred in the ring using a taut cord as shown in Figure 7(B). Listeners placed their chin into a chinrest and maintained fixation on the speaker positioned directly in front of them. The experimenter then took a cord attached to the frame at 270° azimuth and pulled it taut in the air above the listener's head. It was then aligned to a point on the frame at 90° azimuth and gently lowered, while maintaining tension, onto the listener's head. This provided a straight line which bisects the ring array horizontally as viewed from above. This line could be sighted down by the experimenter, viewing from 90° azimuth towards 270° azimuth. This allowed the experimenter to determine if the listener's ear canal (and interaural axis) was located on a vertical plane given by this line to within 1 cm. Adjustments forward and backward were made for each listener and the method repeated until their interaural axis was satisfactorily aligned.

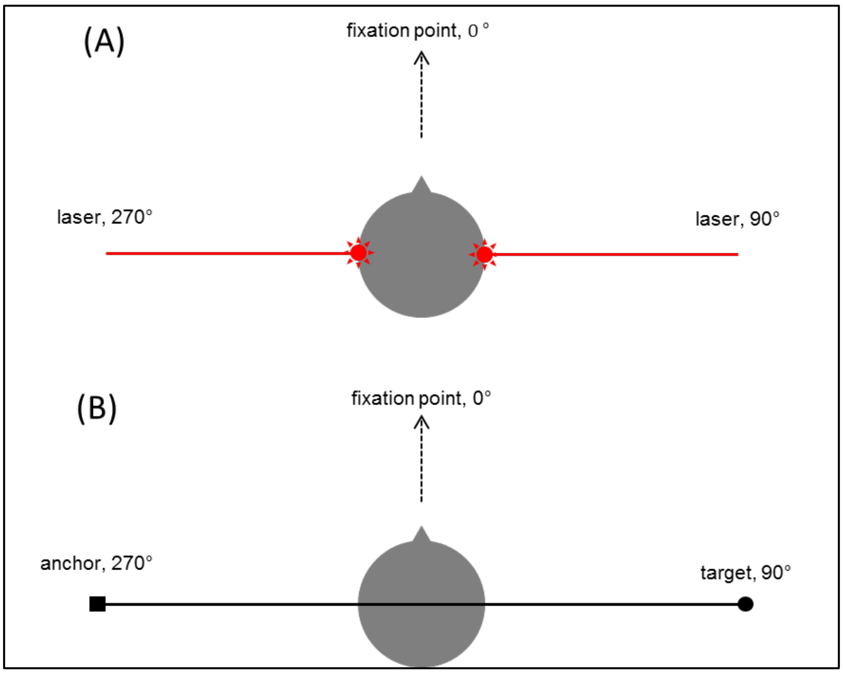

Figure 7. Two methods of centring the listener's head within the speaker ring, using laser alignment (A) and the cord method (B).

Using a head and torso manikin (KEMAR), this method was compared with a method using lasers mounted at 90° and 270°, shown in Figure 7(A), and it was found that the disparity between methods was less than 1 cm when performed correctly. The cord method was chosen as it was simpler to conduct and carried less risk of accidental injury to listeners, even though the lasers used were low power (1 mW) and unlikely to cause retinal damage. Additionally, while both methods require regular alignment checks, the cord method is easily verified by checking the anchor point whereas the laser method requires viewing of both sides of the head in quick succession to reduce incidental movement of the listener's head between checks at each ear. Consequently, the cord method was used in all experiments where the head was required to be stationary within the ring.

During experiments, the maintenance of listeners' head position and orientation was tracked using a magnetic tracker (Polhemus Liberty) with six degrees of freedom, mounted onto a headband. This sampled the listener's head position for 200 ms prior to the start of each trial. From this, six axes of position and orientation data could be recovered.

Listeners completed all blocks using a chinrest, which was adjusted to a comfortable height prior to starting each session. This chinrest was aligned with a line marked on the floor of the laboratory which bisected the ring array from 0° azimuth to 180°. This ensured that the listener's head was aligned left/right within the ring.

Collectively, these measures ensured that errors of source location misperception were removed as experimental confounds. They also enabled for the retroactive analysis of head position and orientation data for outlier removal and for data analysis.

2.2.3 **Head position data analysis.**

To examine listeners' head position, the variability was used rather than the absolute position as it is less affected by constant or long-term offsets, e.g., a listener relaxing into a slightly different position following the start of the experiment. The measurement chosen, median absolute deviation, was non-parametric as these data were highly positively skewed - suggesting that most changes in position were very small (examples of this analysis are provided in section 2.5.4). This was calculated for each block and plotted using the six available degrees of freedom.

2.3 Task Procedure

2.3.1 **Presentation and participant response.**

A practice block was provided prior to the start of the experiment to familiarise listeners with the task and stimuli. Stimuli were then presented in blocks, in a randomised order. Each block presented sounds from one source location only. Although the source locations used could be any azimuth on the array, azimuths were chosen which were a multiple of 7.5°. This corresponded to a centroid located directly on a speaker.

On each trial a pair of sounds was presented successively with an inter-stimulus interval of 100 ms, the minimum interval found by Grantham (1985, in Grantham, 1997) to display consistent performance levels. The listener's task was to identify the location of the second source of each pair relative to the first. Stimuli presented in the front and rear quadrants of the ring were judged as either being to the left or to the right of the first stimulus. Stimuli in the left and right quadrants were judged as being in front or behind the first stimulus. All judgements were relative to the listeners' own reference frame. A stimulus displaced clockwise from the previous stimulus in the front quadrant should therefore be judged to be displaced rightward, whereas that stimulus would have to be displaced anti-clockwise in the rear quadrant to be judged as displaced rightward, and so on.

A 2-down-1-up staircase task was used, with two interleaved staircases. This method converges on 70.7% accuracy (Levitt, 1971). The calculation of the separation angle between presented sound sources, Φ, is given in (7). This separation angle was increased or decreased following either a single incorrect response or two consecutive correct responses respectively. Steps of separation angle were logarithmic, using a value of θ that was increased or decreased by 0.25. Each change from a series of correct responses to incorrect responses or vice versa was classed as a reversal. When ten reversals were reached on a given staircase, that staircase terminated, and when both staircases reached ten reversals the task finished.

$$\Phi = 7.5 \times 2^{\theta} \tag{7}$$

Each pair of sounds was offset by a symmetric amount either side of the centroid depending on the value from the staircase. An offset of up to ± 7.5°, selected at random from a uniform distribution, was added to both centroids in a pair to reduce the predictability of the location of the sound source.

Responses were made using a gamepad controller, with a diamond-shaped arrangement of buttons. The leftmost and rightmost buttons corresponded to ‘leftward’ or ‘rightward’ displacements, respectively. The topmost and bottommost buttons corresponded to displacements ‘forwards’ and ‘backwards’, respectively. This mapping was chosen as it provides an intuitive relationship between the spatial judgement being performed by listeners and the appropriate button to press. No feedback was given following response.

2.3.2 **Threshold calculation and outlier removal.**

For each staircase the first two reversals were discarded to limit the inclusion of button press errors from threshold calculation, as an initial ‘incorrect’ response followed by a ‘correct’ response would count as two reversals. The remaining eight reversals were averaged to generate the threshold performance for that staircase.

Thresholds were averaged across all staircases and repetitions for each condition and for each participant. Outliers were then examined using the modified Z-score method (Iglewicz & Hoaglin, 1993). This method was chosen as it is more robust to outliers than its parametric cousin, the Z-score (Jones, 2016), while still remaining simple to calculate. The modified Z-score calculates outlier status using the median absolute deviation (MAD) within a dataset, rather than the mean and standard deviation. The cut-off recommended for this method is ±3.5 MADs from the median. Cut-off values were generated for each condition independently, as median threshold values varied considerably between conditions.

2.4 Stimulus Spatial Width and Source Angular Separation

Most of the early measurements of auditory localisation accuracy (Rayleigh, 1907; Stevens & Newman, 1934, 1936) and of auditory spatial precision (Mills, 1958; Starch, 1905) used stimulus presentation methods that approximated point sources. For a point source, the spatial extent of the stimulus itself is infinitely small, as the sound waves expand outwards from a single location. Although the apparatus used in each of these experiments was not a true point source, they were very small sources such as a single loudspeaker, horn, or tuning fork. This method is still used by more recent researchers (Butler et al., 1990; Carlile et al., 1997, 1999; Oldfield & Parker, 1984a), but this is not the only method that can be employed.

Using multiple sound sources concurrently, it is possible to create a perceived source whose location is not congruent with any of the physical sources of sound. This method is capable of generating stimuli with larger spatial extents than that of the single-source method, allowing for examination of the effect of spatial characteristics of stimuli on localisation ability. However, point sources can still be generated by using a single sound source at a time (Perrott & Saberi, 1990; Saberi et al., 1991).

The use of multiple concurrent sound sources has been largely restricted to studies of auditory motion perception (Brimijoin, 2018; Freeman et al., 2017; Grantham, 1986). Studies which have used laboratory set ups with multiple available sound sources to measure localisation accuracy or precision of static sounds have typically only used a single sound source at a time (Chandler & Grantham, 1992), rather than multiple sources simultaneously. Measuring precision with a fixed point source puts a lower bound on the precision that can be measured because the smallest possible MAA that can be measured is equivalent to the angular separation between speakers.

Limiting the sensitivity of the precision measure employed is not problematic if the angular separation between sound sources is smaller than the lowest theoretical MAA. In

Perrott and Saberi (1990), for example, the angular separation between speakers was 0.46°. This is smaller than the lowest MAA normally found which is around 1° (Mills, 1958). The loudspeaker array employed in Perrott and Saberi's experiment was 7 metres from the participant and only subtended 13.3° of arc, though. If a similar density of speakers was used to cover the full horizontal plane, more than 780 such sound sources would be required. Using a lower number of sound sources, with concurrent sound presentation to generate 'phantom' sources, can offer a more reasonable alternative.

As the spatial extent of a stimulus presented using the multiple-source method can be varied, this raises the question of what the optimal spatial extent of the stimulus is. Theoretically, this should be as small as possible, to approximate a point source. However, as has been described above, this is prohibitive to physically produce. The best source spatial extent should then be the one which is the narrowest spatial extent possible using a given array while avoiding artefactual cues, such as changes in overall sound level with sound location.

Using the Gaussian weighting method described in section 2.2, sounds can be generated down to arbitrarily small spatial widths (extent in one dimension, i.e., the horizontal plane). However, good generation of a single perceived sound using this method is reliant on the power of the stimulus being consistent regardless of the location of the perceived sound source. By calculating the power for a stimulus for which the centroid of the Gaussian weighting function is located directly on a speaker and comparing it to a sound for which the centroid falls between speakers, we can determine the minimum width for the Gaussian weighting function. The minimum width (critical value) of the Gaussian weighting function is the width that provides a consistent stimulus power level, regardless of the location of the centroid. This is illustrated in Figure 8.

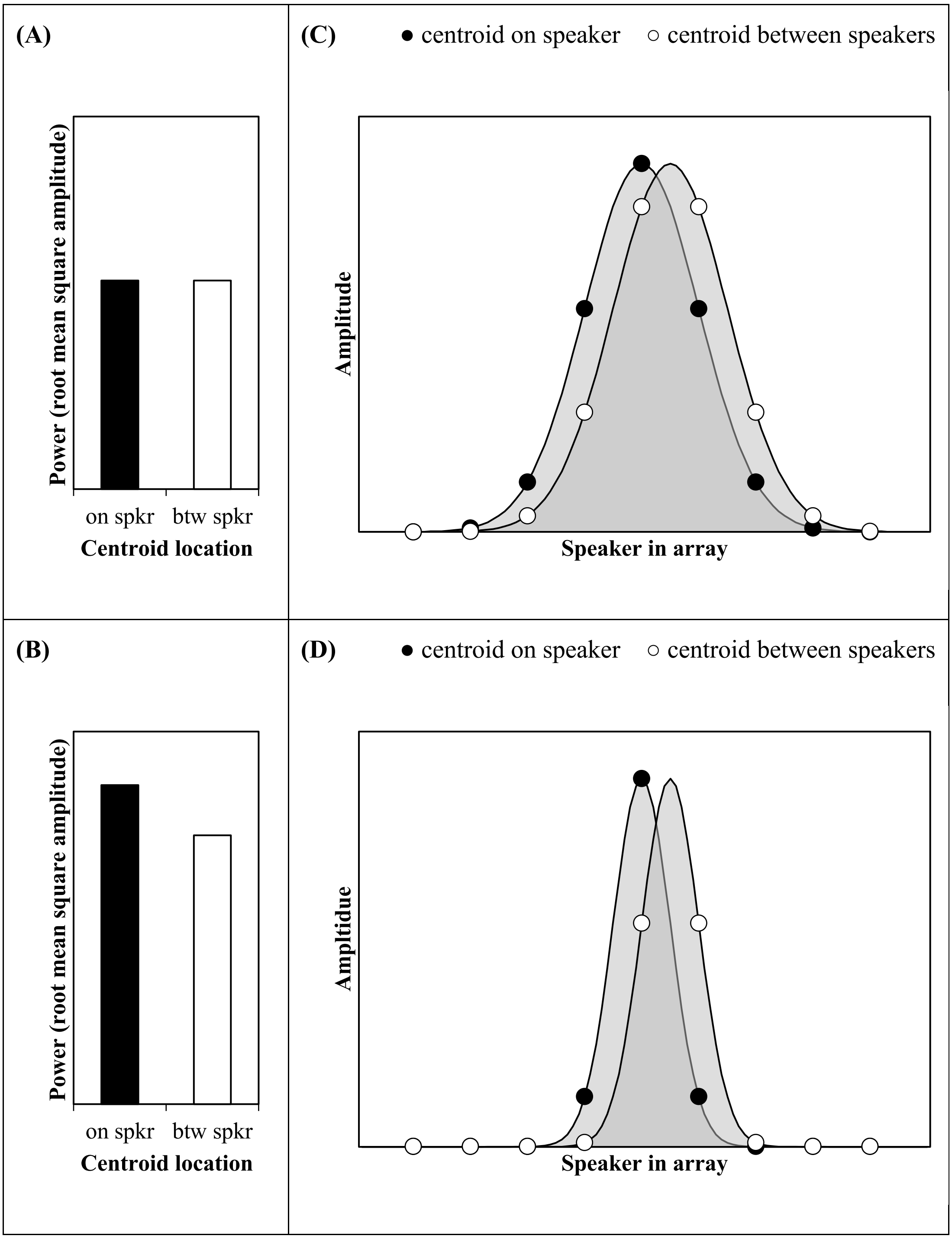

Figure 8. (A) Power level for a stimulus with a Gaussian weighting function standard deviation above critical value; (B) power level for a stimulus with a Gaussian weighting function standard deviation below critical value; (C) amplitudes for an array of sources, circles indicate individual sources, curves indicate underlying distribution of sound (SD above critical value); (D) source amplitudes and distribution for sound (SD below critical value).

As with the single-source method, the angular separation between sound sources in the multiple-source method is also very important. The critical value remains a constant when expressed in speaker units as it is the result of calculation independent of perception. Therefore, the angular breadth that corresponds to this critical value will vary with the density of the loudspeakers. A multiple-source array with a high density of loudspeakers would therefore have a smaller critical value, in degrees of azimuth, than an array with a low density of loudspeakers. If this critical value is insufficiently small, then the measured MAA may instead be limited by the critical value – not the minimum perceptible angle between sources. The relationship between the critical value and the MAA is unknown, so an important empirical query is how densely spaced the loudspeakers must be to measure the limits of perception.

The experiment that follows examines how the auditory spatial precision, measured using the MAA, differs with the standard deviation of the Gaussian weighting function used, and with the density of the sound sources with which these stimuli are presented. The lowest MAAs should be found for stimuli with a spatial width at the critical value.

2.5 Method

2.5.1 **Participants.**

Seven listeners took part in this experiment (1 male, ages 19 – 25). All self-reported normal hearing.

2.5.2 **Task and stimuli.**

All stimuli were presented from a sound source azimuth of 0°. The power levels for these standard deviations, depending on where the centroid is located, are shown in Figure 9. It can be seen that as the standard deviation of the Gaussian weighting function increases, the difference between the power levels decreases.

Table 2

Spatial width of the stimulus per angular separation between sources.

SD in speaker units	7.5° angular separation	15° angular separation
0.35	2.625°	5.25°
0.50	3.75°	7.5°
0.70	5.25°	10.5°
1.00	7.5°	15°
1.40	10.5°	21°
2.00	15°	30°
2.80	21°	42°

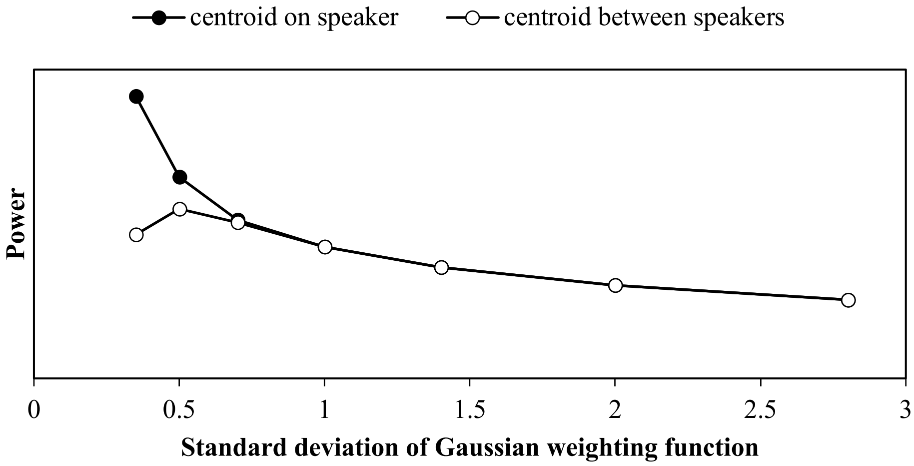

Figure 9. Calculated power for stimuli with a centroid which falls either on a speaker or between speakers.

2.5.3 **Procedure.**

One block was performed per condition, and participants completed all blocks in a random order. This experiment was completed in two sessions, each of approximately one hour, on separate days to prevent fatigue.

2.5.4 **Results.**

2.5.4.1 Head position and orientation.

The head position and orientation of the listeners was tracked using six degrees of freedom. The median absolute deviation (MAD) was calculated for each block and these are shown in Figure 10.

Head position axes (X, Y, Z) show that the MAD in head position was less than 2 cm in all blocks, and less than 0.5 cm in the vast majority of blocks. Head orientation axes (yaw, pitch, roll) show that the MAD for orientation was less than 4° in all blocks, and the modal orientation variability was less than 0.5°. This demonstrates the effectiveness of using a chinrest to keep the participant's head stationary within the ring.

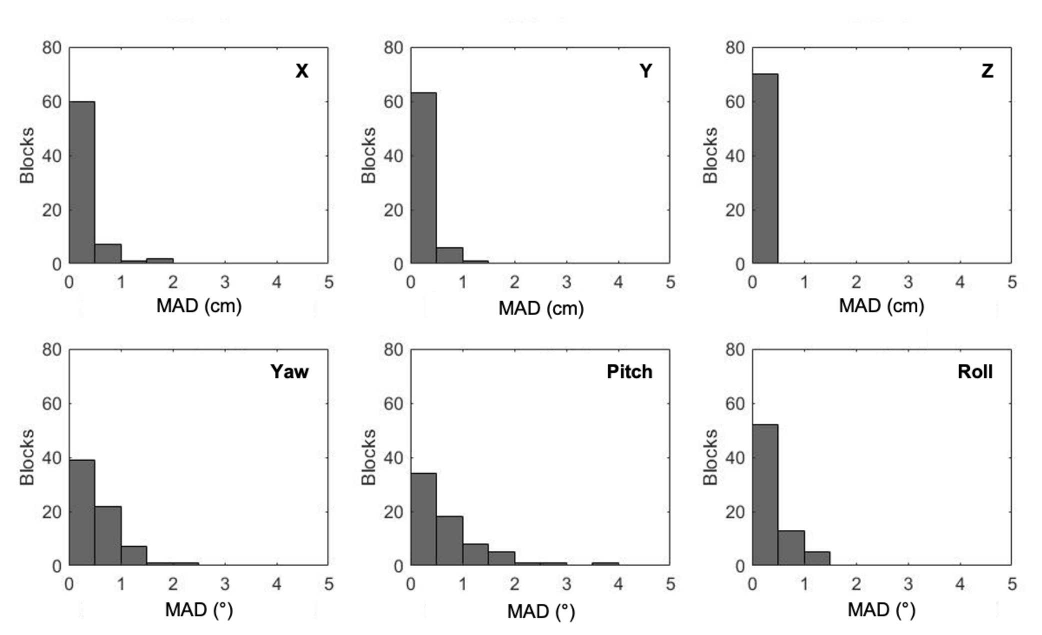

Figure 10. Histograms of median absolute deviation of head movement per block, in positional and rotation axes.

2.5.4.2 MAAs.

These data can be analysed and plotted in two ways. Firstly, by the SD of the weighting function in speaker units (Figure 11(A)), which shows how source density affects measured spatial precision. Secondly, by the SD of the weighting function in degrees

(Figure 11(B)), which shows how the spatial extent of the stimulus affects the measured spatial precision. As the primary issue under investigation is the minimum necessary spatial extent of the stimulus as it is presented to the listener, analysis will be performed on the results in units of degrees.

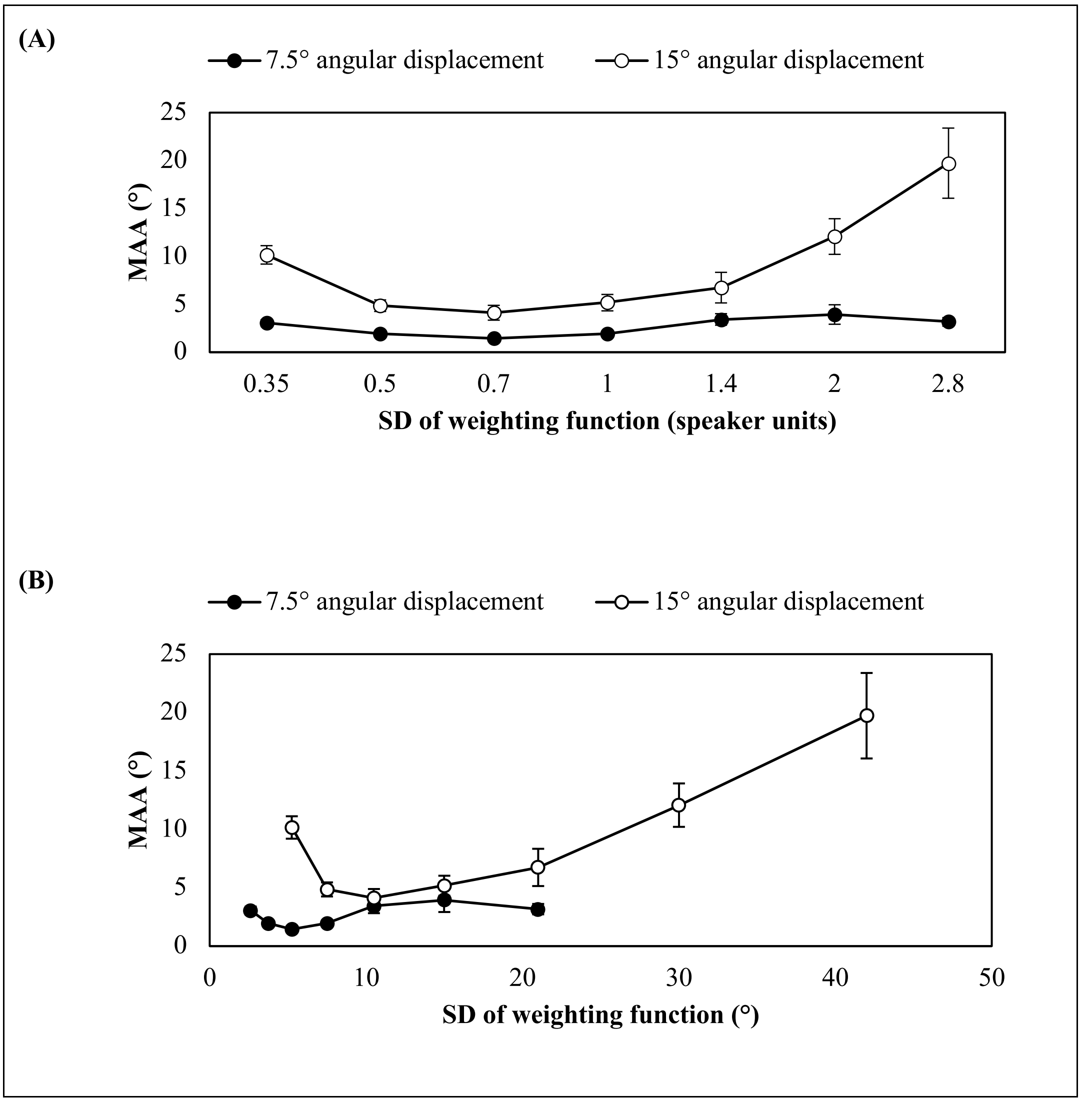

Figure 11. Minimum audible angles for different Gaussian weighting function standard deviations, per speaker separation angle. (A) Data plotted using standard deviation in speaker units and (B) in degrees. Error bars are ± standard error.

Not all spatial width SDs (in degrees) were presented at both speaker separation angles. Therefore, an ANOVA was performed on only those SDs which were present in both conditions. This found that there was a significant interaction between spatial width and speaker separation angle, $F(4,36) = 16.384, p < 0.001$. There were also main effects of spatial width, $F(4,36) = 6.463$, $p < 0.001$, and speaker separation angle, $F(1,9) = 42.794$, $p < 0.001$.

T-tests showed that there was a significant difference in MAA between speaker separation angles at spatial widths of 5.25° ($t(9) = 10.631$, $p < 0.001$), 7.5° ($t(9) = 6.067$, $p < 0.001$), 15° ($t(9) = 2.316$, $p < 0.05$), and 21° ($t(9) = 2.557$, $p < 0.05$). There was no significant difference between speaker separation angles at 10.5°, $t(9) = 0.828$, $p = 0.429$.

These data show that the separation angle between loudspeakers does affect the precision of the measured localisation ability. Wider speaker separation angles increase the measured MAA. The spatial width of the stimulus used also affects the precision with which a listener can localise a sound. Increasing the spatial width also increases the MAA, but there is also an interaction such that MAAs are particularly affected when the spatial width is smaller than the critical value.

2.5.5 **Discussion.**

Minimum audible angles (MAAs) were measured using a range of stimulus spatial widths, and a source angular separation of either 7.5° or 15°. The pattern of MAA performance is broadly similar between angular separation conditions. For both angular separations the lowest MAAs were found for a Gaussian weighting function with a standard deviation of 0.7 speaker units. This is the critical value, the narrowest width of stimulus possible while maintaining power level consistency between centroid locations (Figure 9). This therefore confirms the hypothesis that this would be the condition resulting in the greatest precision.

Below the critical value the MAAs increased rapidly. This is likely due to the inconsistency in sound level produced by stimuli of this width. If the centroid of the stimuli falls ‘between’ speakers, the overall power produced will be lower than if the centroid falls ‘on’ a speaker. This can be seen in Figure 9, where there is a clear separation between power levels for stimuli with a centroid ‘on’ or ‘between’ speakers, below 0.7 speaker unit standard deviations. If the stimulus pairs on a trial have different sound levels, this may lead to confounding perceptions such as change of distance from the listener (Shinn-Cunningham, 2000) which could decrease performance on this task.

Above the critical value MAAs also increased, but the gradient of this increase was shallower. As the stimuli become more spatially diffuse they will become more similar to sound field stimuli (Griffiths & Green, 1999). Sound field stimuli are not perceived as discrete localisable sources, which may explain why precision increased.

It can be seen from Figure 11 that most MAAs measured using a 15° source angular separation were higher than MAAs measured using the condition with a 7.5° source angular separation. Although, only the smallest and largest stimulus spatial widths were significant following correction. This finding confirms the suggestion that the density of sound sources is important for measuring the spatial precision abilities of the auditory system. Using a larger angular separation can obscure the true spatial precision of a listener. As this is true using the MAA task, the lowest level assessment of auditory spatial abilities, it is likely to be true for more higher-order abilities too. Therefore, researchers performing tasks which assess abilities such as compensation for movement or object tracking should consider the density of sources available in order to be sure that they have not obtained an artificially inflated measure.

The demonstration of 0.7 speaker units as the critical value aligns with predictions made from prior calculations, suggesting that this is the best value to be used in all MAA experiments going forward. The 7.5° angular separation is the density of the speakers in the laboratory and therefore cannot be reduced further. The MAA found using the critical value was approximately 1.4°, only slightly larger than the 1° consistently found using point sources (Mills, 1958; Perrott & Saberi, 1990; Saberi et al., 1991). An MAA of 1° also represents the theoretical limit of the human auditory system, as this corresponds to an ITD of 10 μs (Klumpp & Eady, 1956). Consequently, there is very little to be gained from using a higher density of sound sources.

2.6 Chapter Discussion

Using the MAA task overcomes the problems of response confound described in section 1.4.2. It provides a good direct measure of precision, which cannot be easily inferred from studies of spatial localisation. It is still important though to ensure that appropriate precautions have been made when constructing and conducting this task. These considerations and the proper application of this task are described in this chapter.

In section 2.2 the general method of the MAA task employed elsewhere in this thesis was described. A two-alternative forced choice procedure is provided, which forces listeners to make a judgement about the relative location of sound sources. The separation between pairs of sound sources is changed using a staircase, which decreases as listeners perform well and increases if they make an incorrect judgement. This staircase converges to a 70.1% performance threshold, which is taken as the MAA in this task.

It is very important that the listener be aligned correctly within a speaker array, particularly a circular one. If the sources are not located in the correct positions with respect to the listener's head, or the listener is improperly oriented, then the measured precision will not reflect the true precision of the auditory system. The specific methods of aligning the participant (laser and cord) were compared, and the rationale given for the cord method which was used in all experiments where appropriate. The cord method was found to be accurate to within 1 cm of the true centre of the speaker array. A chinrest was also employed to ensure that the listeners stay at the correct position within the ring. Measurements of head position and orientation are described in section 2.5.4, which show that listeners' head position and orientation shows very little deviation during the task, confirming the efficacy of these methods.

In section 2.4 an experiment was described where the spatial characteristics of stimuli were varied, in an attempt to find the best possible presentation method for spatially localisable static stimuli using the laboratory equipment. It was shown that the angular separation between speakers is important for ensuring that the minimum possible MAA can be achieved. Smaller angular separations are beneficial for measurements of precision. There is, however, little apparent need to reduce these angular separations far below the 7.5° separations between the speakers in the ring array used here.

It was also shown that the spatial width of the stimuli, quantified as the standard deviation of the Gaussian weighting function used, was important. There is a critical value (at or near 0.7 times the angular separation between speakers) above or below which the MAA found using that spatial width will rise. Here the lowest MAAs found were very close to 1° of azimuth, which is the minimum MAA found using point sources. An MAA of 1° is also equivalent to an ITD of 10 μs (Klumpp & Eady, 1956), the limit of the auditory system's sensitivity to timing differences.

This chapter provides a solid grounding for the methods employed in subsequent chapters in this thesis. The findings made here also give guidance for any researchers who seek to use similar methods to present sounds, and, in particular, it highlights the pitfalls that may occur if spatial width and angular separation of sources are not considered.

Chapter 3 – Spatial Precision as a Function of Source Azimuth

3.1 Chapter Overview

Current interest in topics such as immersive VR and ambisonics requires good information about auditory perception in the rear hemifield. Many models used in these areas are based on spherical head models (Algazi et al., 2001; Brungart et al., 1999; Brungart & Rabinowitz, 1999; Kopco & Shinn-Cunningham, 2000) and/or generalised HRTFs (Mendonça et al., 2012; Wenzel et al., 1993; Wightman & Kistler, 2005). Since the head is not symmetric in the front and rear, and generalised HRTFs may diffuse individual specific differences between front and rear regions, these models may be lacking in verisimilitude for sound sources originating behind the listener. Therefore, providing more information about auditory perception in this region of space is very useful for attempts to build more accurate models, particularly given the shortcomings of many localisation studies in this region, discussed in section 1.4.

This chapter will examine spatial precision in both the front and rear hemifields, using a direct assessment of spatial precision (MAA). Unlike indirect measurements of spatial precision from localisation tasks (section 1.4.3), this task can be used without concern of confounds from response mode and does not require complex conceptualisations of auditory space. The patterns of performance by listeners on this task will be used to elucidate how spatial precision changes in the rear hemifield, a previously little studied area.

3.2 Minimum Audible Angle and Source Azimuth

3.2.1 **Introduction.**

As pointed out in Chapter 1, the earliest direct examination of auditory spatial precision comes from Starch (1905), who used a task which measures what would eventually be known as the minimum audible angle (MAA). The invention of this method is more commonly attributed though to Mills (1958), who popularised this task for assessing spatial precision. Both of these studies examined spatial precision on the azimuthal plane, and both found the same pattern of performance from their listener. Precision on the azimuthal plane was best (smallest MAA) for sounds presented directly in front of the listener, while spatial precision in the front hemifield declined (larger MAAs) as sounds were presented further from the midline towards the interaural axis.

At the time of writing, I could only find only two studies have used the MAA task for azimuth in regions behind the listener. Starch (1905) found that spatial precision was worse in the rear hemifield than in front but did not report specific values. Saberi et al. (1991) found that the MAA was the same between 180° and 0° but did not measure any other positions in azimuth in the rear hemifield as their primary concern was spatial precision on oblique planes (i.e., diagonal planes where sources vary in both elevation and azimuth).

In contrast, *indirect* measures of spatial precision are frequently included in studies of auditory localisation accuracy over the complete sphere. However, as discussed in section 1.4.3, these are often confounded with the method of response. Localisation studies which use pointing responses (e.g. Carlile et al., 1997; Majdak et al., 2010; Oldfield & Parker, 1984a) have a confound in the ability to turn the hand or head towards locations beyond 90° (Gates et al., 2016; Swinkels & Swinkels-Meewisse, 2014). Studies which use naming methods (e.g. Wenzel et al., 1993) can have issues of conceptualisation of auditory space, or response coarseness (Nielsen, 1991). Despite this, studies of localisation are consistent in their findings that precision is worse in the rear hemifield than in the front hemifield. Even so, these confounds can provide some obstacle to the inference that can be made.

Given the shortage of direct spatial precision studies for the rear hemifield, and the problems with indirect spatial precision measures from localisation studies, we do not know with confidence whether auditory spatial precision is different in the rear hemifield. It is vital to have good quality, unconfounded measures of spatial precision as predictions about higher-order auditory abilities are dependent on the knowledge we have about the low-level abilities. For example, when considering the contributions of head movement signals to the perception of auditory object motion (discussed in Chapter 5), it is necessary to have a full understanding of the auditory signals before inference can be drawn about the way they combine with other signals such as those coming from the vestibular system. As a second example, if auditory speed judgements are based on successive localisation snapshots (Grantham, 1986), then the precision of the localisation judgements must be known in order to make predictions about the resulting speed judgements.

Using the MAA task described in section 2.2, MAAs are predicted to be smallest (most precise) for sounds located on the midline, and to increase as sounds are located closer to the interaural axis (Mills, 1958). From previous findings (Starch, 1905) we might then expect that spatial precision is worse at positions in the rear hemifield compared to symmetric positions in the front hemifield, hereafter referred to as ‘symmetric position pairs’. Finally, the findings from Saberi et al. (1991) suggest that the spatial precision for sounds presented on the midline (0°, 180°) should not differ.

3.2.2 **Methods.**

3.2.2.1 Participants.

Four listeners were recruited (one female, ages 19-27) who were all familiar with psychophysical experiments, one non-naïve. All self-reported normal hearing.

3.2.2.2 Procedure.

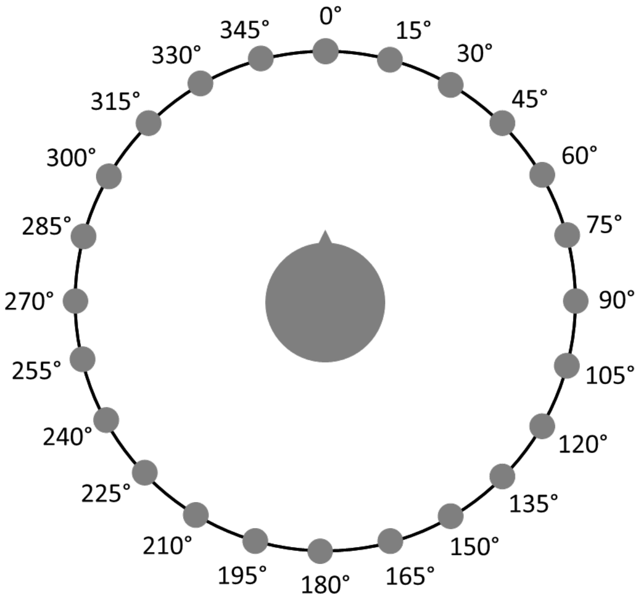

Figure 12. Source location azimuths shown on the speaker array, not to scale.

The MAA task was used as described in section 2.2, with the following changes. The experiment was conducted in three sessions of one hour over several days to prevent listener fatigue. A practice block was provided prior to the start of the experiment to familiarise listeners with the task and stimuli. Different azimuths were then presented in blocks, in a randomised order. Each block presented sounds from one source location only. Twenty-four locations were used, from 0° to 345° in steps of 15° of azimuth (Figure 12).

3.2.3 **Results.**

3.2.3.1 Head position and orientation.

The median absolute deviation of the head position and orientation during each block was calculated and is plotted in Figure 13. It can be seen that the median absolute deviation was low for all axes. Deviation in head position (X, Y, and Z) was less than 1 cm in the majority of blocks. Deviation in head orientation (yaw, pitch, and roll) less than 3° on any axis.

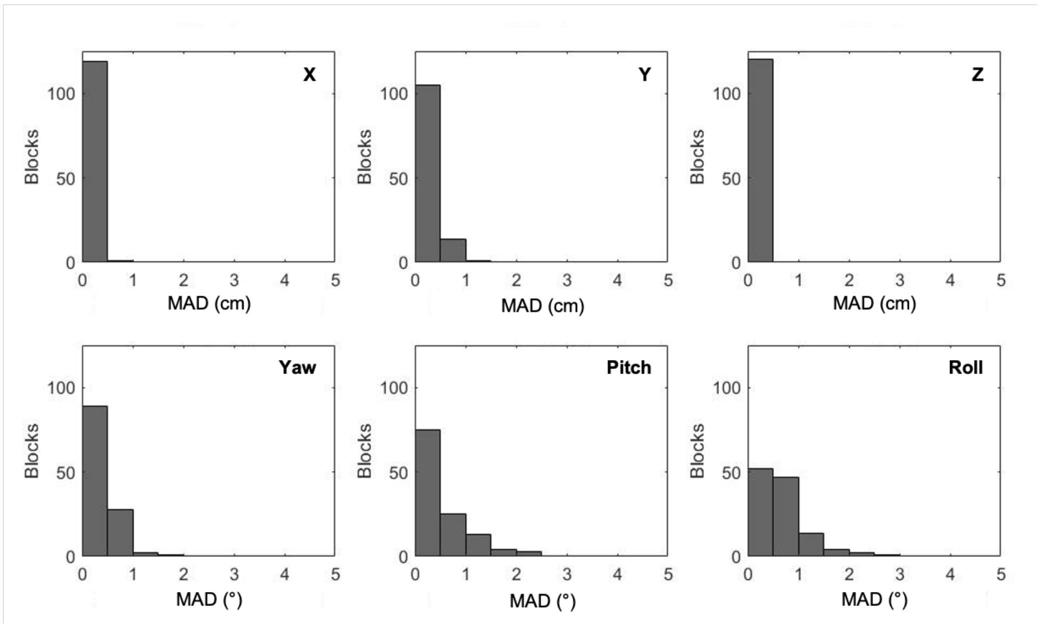

Figure 13. Histograms of median absolute deviation of head movement per block, in positional and rotation axes.

3.2.3.2 Minimum audible angles.

The pattern of MAAs found, and the difference between hemifields, is illustrated in Figure 14. As MAAs were not expected to differ between left/right and front/rear hemifields, these were entered into an ANOVA as factors. This would also permit examination of interactions between hemifield and source azimuth.

A 2x2x5 repeated-measures ANOVA was performed using left/right, front/rear, and oblique source azimuth (15° to 75°) as factors. MAAs did not differ between left and right hemifields, $F(1,7) = 0.222$, $p = 0.652$. MAAs were significantly different between front and rear hemifields, $F(1,7) = 65.472$, $p < 0.001$. MAAs also significantly differed between oblique azimuths, $F(1,7) = 27.858$, $p < 0.001$. There was a significant interaction between front/rear hemifield and oblique azimuth, $F(4,28) = 14.340$, $p < 0.001$. No other interactions were significant.

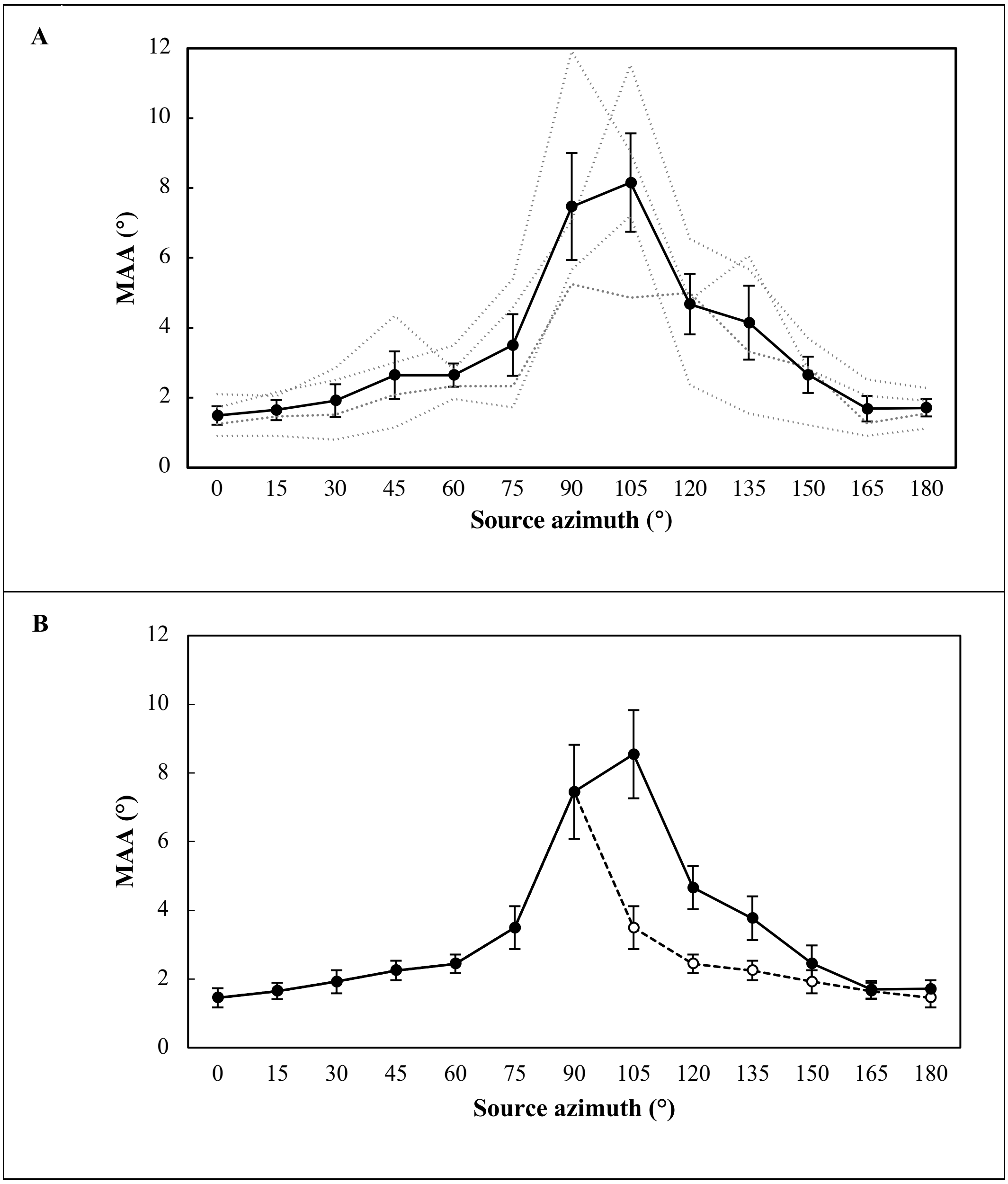

Figure 14. Minimum audible angles for each source azimuth, collapsed over left and right. Solid lines show mean MAAs. A) Dotted lines show individual participants' MAAs. B) To help comparisons, dashed line replots data from positions in the front hemifield that have been mirrored onto the rear hemifield. Error bars are ± standard error.

Simple main effects tests showed that MAAs were significantly greater at all oblique positions in the rear hemifield than at their front symmetric position pairs, $F(1) > 7.6$, $p < 0.05$, for all comparisons except 15° *vs*. 165°, $F(1) = 0.104$, $p = 0.756$. A paired t-test showed that MAAs also did not differ between 0° and 180°, $t(8) = 1.598$, $p = 0.154$.

3.2.4 **Discussion.**

The minimum audible angle was measured for sound sources presented at multiple azimuths relative to the listener, in order to examine how spatial precision changes. In particular, the differences in precision between sounds presented from symmetric position pairs were examined.

It was found that MAAs were smallest (most spatially precise) when sounds were presented on the midline (0°, 180°), and increased as sounds are presented further from the midline towards the interaural axis (90°). Comparison between front and rear hemifields showed that MAAs were larger at most positions in the rear hemifield compared to their symmetric position pairs. This conforms to the predictions made from studies which provide indirect measurements of spatial precision, and to Starch (1905) who provided the only direct measurement of spatial precision across the rear hemifield.

It was predicted that spatial precision would not be different between sounds on the midline in front of the listener (0°) and behind the listener (180°), which was found to be the case here. MAAs were slightly higher for sound sources presented at 180° than at 0°, but this difference did not reach significance. This may be simply due to variability in the data as there was no difference between MAAs for sounds very close to the midline (i.e., 15° *vs.* 165°), and the difference in MAA for sounds on the front and rear midline was very small (<1°).

Indirect measurements of auditory precision, from studies of auditory localisation accuracy, have a consistent finding that spatial precision is worse in the rear hemifield compared to the front hemifield (explored in section 1.3.1). The results from this experiment conform to these general findings, but do not share the same potential confounds that were present in the localisation accuracy studies. Figure 15 replots the data from this experiment against the data from those previous studies. This shows that most indirect measures of precision continue to worsen as sound sources are located further towards the rear, and only improve again very close to the rear midline. Here, the direct measures show that precision does worsen to a peak near the interaural axis, but then improves from just behind the interaural axis towards the midline, rather than continuing to worsen.

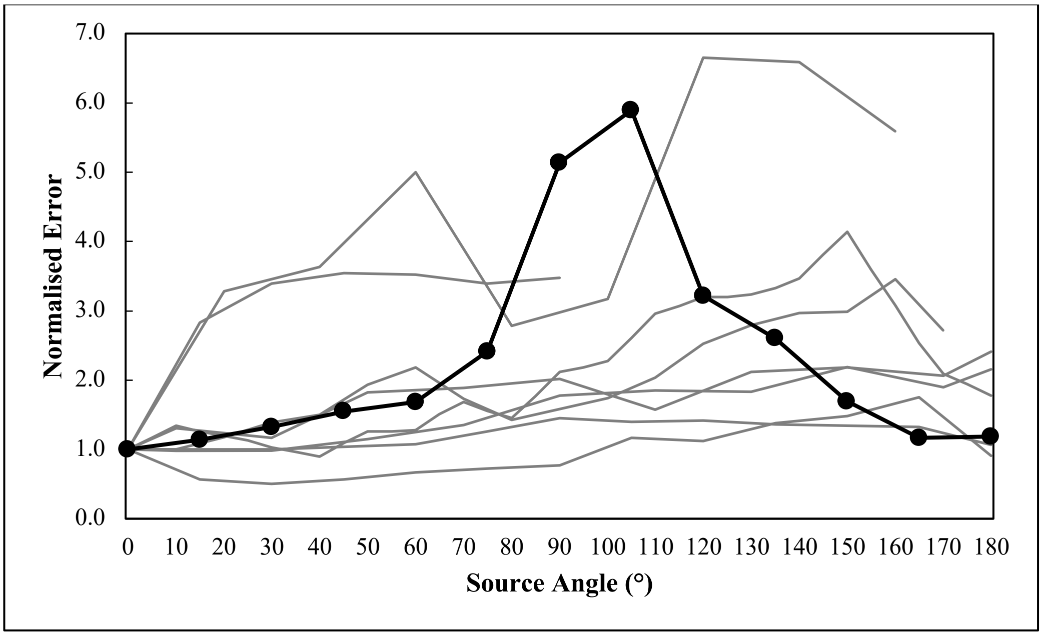

Figure 15. Normalised data from this experiment (black) plotted against the normalised data from indirect measures of precision (grey).

Indirect measures appear to have good estimates of spatial precision for sound sources up to approximately 60°, but then overestimate precision (lower on the y-axis) near the interaural axis and underestimate precision in the rear hemifield near the midline. The confounds discussed earlier in section 1.4.3 could explain the underestimation of precision in the rear hemifield because pointing measures for sources in the rear hemifield necessarily involve extra variability due to sound sources being located outside of the normal range of motion of the head or arm. It is less clear why these studies overestimate precision for sound source positions near the interaural axis, given that additional motor noise will occur for these locations as well.

One possibility why the findings from localisation studies differ from the presented experiments is the correction applied in many of those studies for listeners responding to the incorrect hemifield, that responses showing a ‘front-back confusion’ are flipped to the correct hemifield (e.g. Musicant & Butler, 1984; Oldfield & Parker, 1984a), which would decrease the estimates of error for this region. For example, a response of 92° to a sound source location of 85° would be 'corrected' to a response of 88°. To demonstrate why this is a problem I performed a simulation, the results of which are plotted in Figure 16. One hundred responses (samples) were generated for sound source locations between 45° and 135°, in increments of 0.1°. Responses were taken from a normal distribution with a mean at the sound source location and a standard deviation that varied with azimuth. This standard deviation simulated spatial precision as a normal distribution centred on 90°, which broadly approximates the findings using MAAs from the experiment presented in this section. I then duplicated the responses and applied a 'hemifield correction'. Hemifield corrected responses were mirrored about 90° if the response generated was above 90° when the sound source location was below 90°, and vice versa.

In Figure 16A the spatial precision of the simulated responder is shown as the standard deviation of responses for each sound source location. It can be seen that applying a hemifield correction leads to a large over-estimation of precision (lower standard deviation) at positions close to the interaural axis. Applying this 'correction', therefore, is highly problematic for assessments of spatial precision around 90°. The 'correction' also presents issues for the examination of localisation accuracy as it creates a response bias away from the interaural axis, as can be seen in Figure 16B.

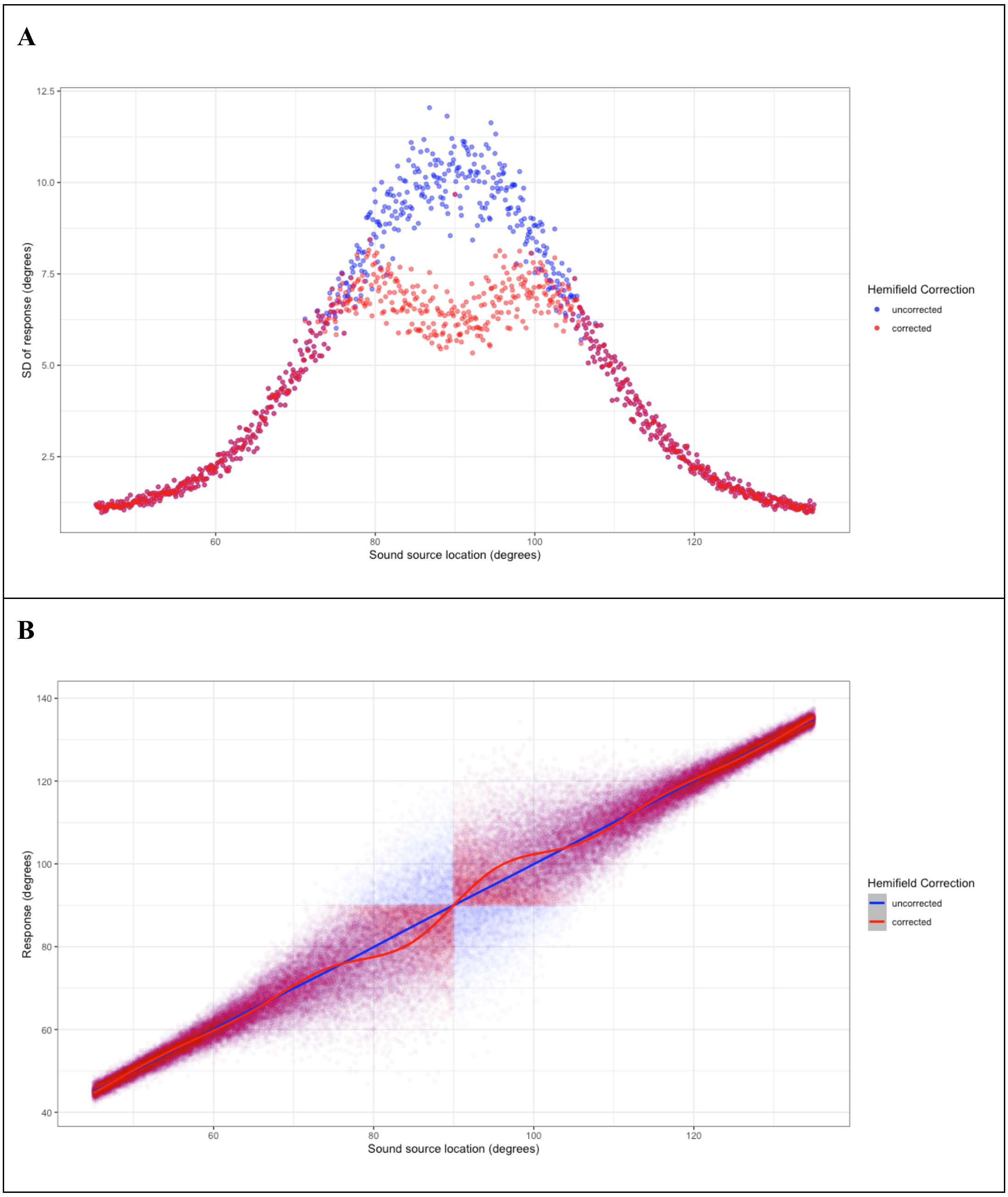

Figure 16. Simulation of the effects of hemifield correction on responses across sound source locations. (A) Response variability plotted as the standard deviation of responses. (B) All responses in this simulation, with points showing individual responses and smoothed lines showing response bias using a fitted general additive model.

Another possibility is that in localisation studies the responses for sound sources located near the interaural axis are benefiting from the limits of the response method. The normal range of motion of the head is ± 80° (Swinkels & Swinkels-Meewisse, 2014). A sound source located at or near 80° azimuth could then be responded to by simply moving the head until the limit is reached, at which point the head stops turning. This is the same position every time, so the variability in responses would be smaller than the actual perceived location of the sound. A similar explanation can be given for pointing with the hand or arm, whose range of motion is larger – but requires the recruitment of more muscles for positions beyond 90°, so provides a similar 'stopping point' (Gates et al., 2016). Hence, indirect measures may be misrepresenting the spatial precision of localisation ability, which is better captured by the indirect measure used here.

What other explanations of the asymmetry could exist? The binaural cues, interaural differences in sound level and time-of-arrival at the ears should not differ between symmetric position pairs as these cues are determined by the path length from the sound sources to the ears and the occluding effect of the head. Provided that the interaural axis is centred within the speaker array, the path length will not differ between symmetric position pairs. The method of centring of the listener within the ring was established in section 2.2.2 of this thesis, with the low deviation in movement from this position confirmed by the results of the head movement analysis shown in Figure 13. Therefore, it can be concluded that the resultant path lengths did not differ, and so cannot have caused these cues to play a role in the generation of the spatial precision asymmetry found here.

A further possibility is the role of the pinna in spatial localisation. The convolutions in the shape of the pinna reflect and diffract sound before it reaches the ear canal, dependent on the elevation of the sound and whether it originates in the front or rear hemifield (Blauert, 1997). The pinnae are also oriented towards the front hemifield (Shaw, 1974; Starch, 1905), suggesting that there may be some benefits in localisation for sounds located in this hemifield.

The pinna only affects frequencies above approximately 4 kHz (Butler & Humanski, 1992), evidenced by the decline in localisation accuracy displayed when stimuli are low-pass filtered at 4 kHz or listeners' pinnae are occluded (Musicant & Butler, 1984). If the pinnae are causing the spatial precision asymmetry, then the difference between front and rear hemifields should reduce or disappear when stimuli are low pass filtered or the contributions of the pinna are removed or distorted. These predictions are tested in Chapter 4.

It should be noted that the number of participants in this experiment (N=4) was fairly low. All listeners were experienced with psychophysical experiments and were able to concentrate for the long periods of time necessary to complete this experiment. There were also many repetitions of each stimulus condition. Nevertheless, this may have resulted in reduced power. Eta-squared, η^2, for the effect of front-rear hemifield in the ANOVA was 0.127, for source azimuth was 0.324, and for the interaction between front-rear hemifield and source azimuth was 0.099. Partial η^2 for these effects were all greater than 0.74, indicating that they each explained nearly three-quarters of the variance accounting for all other effects. The resulting effect sizes were calculated to be $f = 0.381$, $f = 0.692$ and $f = 0.331$ – classified as medium, large, and medium effect sizes respectively (Cohen, 1988). A post-hoc power analysis using G*Power (Faul et al., 2007) [corr. = 0.75] showed that the power estimates for these effects were 0.324, 0.990, and 0.471, respectively. Hence, while the front-rear difference and its interaction with source azimuth was low to moderate, the power for detecting differences between oblique source azimuths was very high. The same number of participants was used in the second experiment presented in this chapter (Section 3.3), but all further experiments use larger numbers of listeners. Given the size of the effects being studied, this is not expected to cause issues in detecting real effects in these later experiments.

The results of the current experiment show a marked asymmetry in spatial precision between front and rear hemifields. One possibility as to why lies in the acoustics of the lab itself. Although listeners were located in the centre of the speaker array, confirmed using the cord method and head tracking, the laboratory outside of the speaker array was not symmetric relative to the listener, and not entirely anechoic. There was more space to the rear of the listener from the seat to the wall, than in front of the listener to the opposing wall. This possibility is tested next in section 3.3, by rotating the listener 180° from their position in the experiment presented in this section. If the asymmetry reverses, this would be evidence that the lab's acoustics determined the difference between front and rear hemifields.

3.3 Room Acoustics Control

3.3.1 **Introduction.**

In section 3.2 it was found that there was an asymmetry in spatial precision between sounds presented in the front and rear hemifields. Listeners showed poorer precision, as measured using the MAA task, when sounds were presented behind their heads compared to presentation in front of their heads from the symmetric position pair.

Although the array of speakers which were used to present the sound sources was symmetric, the same cannot be assumed for the surroundings in which the array was situated. The laboratory, although sound-deadened, was not anechoic. There was also a larger distance from the listener to the wall behind them than from the listener to the wall in front of them. This asymmetry in the layout of the laboratory may have caused acoustic asymmetries for sound presentation, leading to the perceptual spatial asymmetry found.

In the experiment in this section, the orientation of the listener relative to the laboratory was changed to examine whether this affects the spatial precision asymmetry found. If the asymmetric layout of the laboratory causes of the precision asymmetry, then orienting the listener towards the opposing direction within the speaker array should cause the asymmetry to weaken or reverse. If, however, the laboratory has no effect on the spatial precision asymmetry then there should be no difference in the pattern of MAAs found between listener orientation directions. For efficiency, only the (head-centred) azimuths on the diagonal obliques were compared in the two different listener orientations.

3.3.2 **Methods.**

3.3.2.1 Participants.

Four naïve listeners were recruited (ages 21-25, 2 females). All self-reported normal hearing.

3.3.2.2 Procedure.

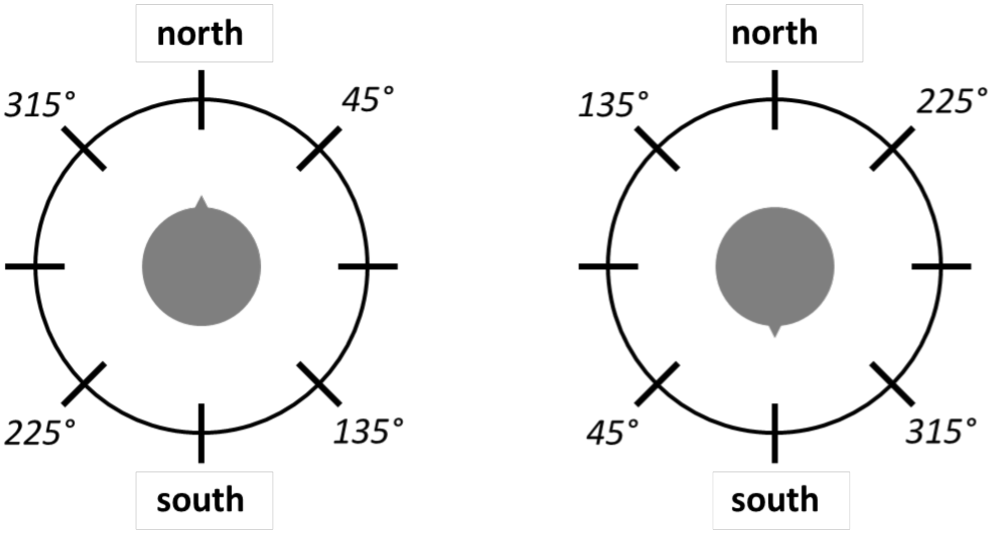

Figure 17. The two listener orientations used in this experiment, not to scale. Numbers in italics indicate the relative source location for the listener. Left and right images are the normal and opposite orientations respectively.

The MAA task was performed as described in section 2.2, with some changes specified below. This is shown schematically in Figure 17. Stimuli were presented in each block from one of four oblique locations (45°, 135°, 225°, 315° with respect to the speaker ring). Listeners were either oriented the same way as in all other experiments (referred to as *north*) or opposite to this (referred to as *south*). For both north and south orientations, listeners heard sounds from all 4 source azimuths being tested, yielding 8 separate conditions that were presented twice in a random order.

3.3.3 **Results.**

Paired t-tests were used to compare left and right positions for both listener orientations. There were no significant differences between any left-right position pairs, so MAAs were collapsed across listener-relative left and right symmetric positions.

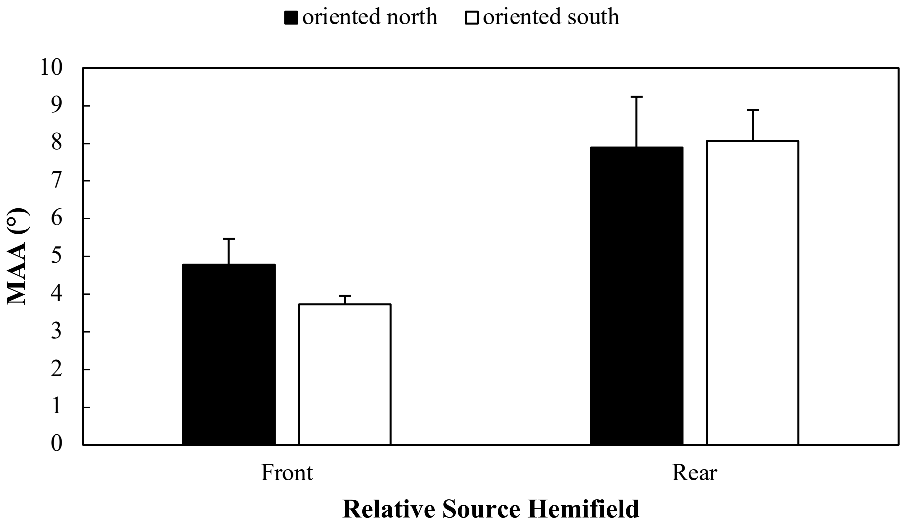

Figure 18. MAA as a function of listener orientation direction and relative source hemifield. Error bars are ± standard error.

The MAAs for the orientation/source azimuth pairs are shown in Figure 18. It can be seen that MAAs are greater for sounds presented to the rear of the listeners compared to sounds presented in front of the listeners, regardless of listener orientation. This was confirmed using a two-way repeated-measures ANOVA. There was no main effect of listener orientation, $F(1,3) = 1.129, p = 0.366$. MAAs were significantly greater in the rear hemifield than the front, $F(1,3) = 22.109, p < 0.05$. There was no significant interaction between orientation and relative source azimuth, $F(1,3) = 0.722, p = 0.458$.

3.3.4 **Discussion.**

The laboratory used in section 3.2 is not symmetric in front of, and behind, the listener. It was suggested that the asymmetry in precision found between front and rear hemifields may have been confounded by the asymmetry in the layout of the laboratory.

This experiment found that this was not the case, as orienting the listener in the opposite direction in the laboratory did not affect the spatial precision asymmetry.

The results found here therefore suggest that the spatial precision asymmetry is dependent on non-environment-specific cues. These may be dependent on stimulus-content or interactions of the sound waves with the listener's head, as discussed above. Before addressing the influences of these spectral changes, the next experiment first investigates whether the asymmetry generalises to sounds that move.

3.4 Minimum Audible Movement Angle and Azimuth

3.4.1 **Introduction.**

The precision of the auditory system for spatial location is best assessed by the smallest detectable difference in location (MAA) (Mills, 1958). Similarly, the precision of the auditory system for spatial motion can be assessed by measuring the smallest detectable arc of movement by a sound source. This is referred to as the MAMA, or minimum audible movement angle (Perrott & Tucker, 1988). A more detailed discussion of the characteristics of the MAMA and how they change with source azimuth and velocity are provided in Chapter 1.

This experiment examined how motion detection, measured using a direction discrimination task, changes as the source azimuth of the sound stimuli was varied across azimuth. MAMAs were expected to increase as sounds were placed further from the midline in the front hemifield towards the interaural axis, consistent with Grantham (1986) and the findings using MAAs presented above. MAMAs should also decrease in the rear hemifield from the interaural axis towards the midline. What is unknown, however, is whether MAMAs in the frontal field are better than those in the rear hemifield.

3.4.2 **Methods.**

3.4.2.1 Participants.

Ten listeners participated in this experiment (ages 20-30, 5 females). All reported normal hearing.

3.4.2.2 Stimuli.

Stimuli consisted of broadband noise generated using the method outlined in section 2.2, and were made to move by updating the location of the centroid on the speaker array at the sampling rate of the head tracker (240 Hz). The amount that the centroid was displaced each refresh was kept constant so that the speed of the stimulus was constant. The duration of the stimulus was fixed at 500 ms. This ensured that only displacement cues were available to the listeners.

The trajectory over which the stimulus moved was centred on the source azimuth selected for a given trial. So, a stimulus would start at a location displaced from the trajectory centre by ½ the total displacement angle, and then move to a location displaced the same angle from the trajectory centre on the opposite side. The start and end points of the trajectory were also varied by ≤ 7.5° (1 speaker separation angle), to prevent these from being used as reliable cues to displacement angle (Carlile & Best, 2002).

Instead of the staircase procedure changing the displacement between sound sources, as in the MAA task, the total displacement angle was changed. The same logarithmic steps were used as in the MAA task. Consequently, the velocity of the stimulus would also change with changes in the staircase value.

Five mean source azimuths were used, from 0° to 180° in 45° increments. Because stimuli moved, these source azimuths therefore represent the mean average location across a stimulus' trajectory. All stimuli were presented on the right side of the listener.

3.4.2.3 Procedure.

Prior to starting the experiment, listeners were centred in the speaker ring using the cord method outlined in section 2.2.2. Listeners were also provided with practice trials to familiarise themselves with the task. Each mean source azimuth under investigation was presented in individual blocks, with blocks being presented in a random order. These random series were presented three times in total.

On each trial the stimulus moved over a trajectory with a total displacement determined by the staircase value for that trial. The movement was either clockwise or anticlockwise, with the direction of travel chosen randomly on each trial. Listeners were required to indicate which direction the stimulus had moved. This decision varied depending on the source azimuth of the stimulus. Stimuli with a source azimuth at 0° or 180° were judged as moving either to the listener's left or right. All other stimuli were judged as moving either towards the listener's front or rear.

3.4.3 **Results.**

One listener's data was removed (N = 9) due to having an average MAMA in multiple conditions above the outlier cut-off for those conditions, using modified Z-scores (Iglewicz & Hoaglin, 1993; Jones, 2016).

3.4.3.1 Head position and orientation.

Head position was tracked prior to the start of every trial, and the median absolute deviation from the centre of the ring was calculated for each block (Figure 19).

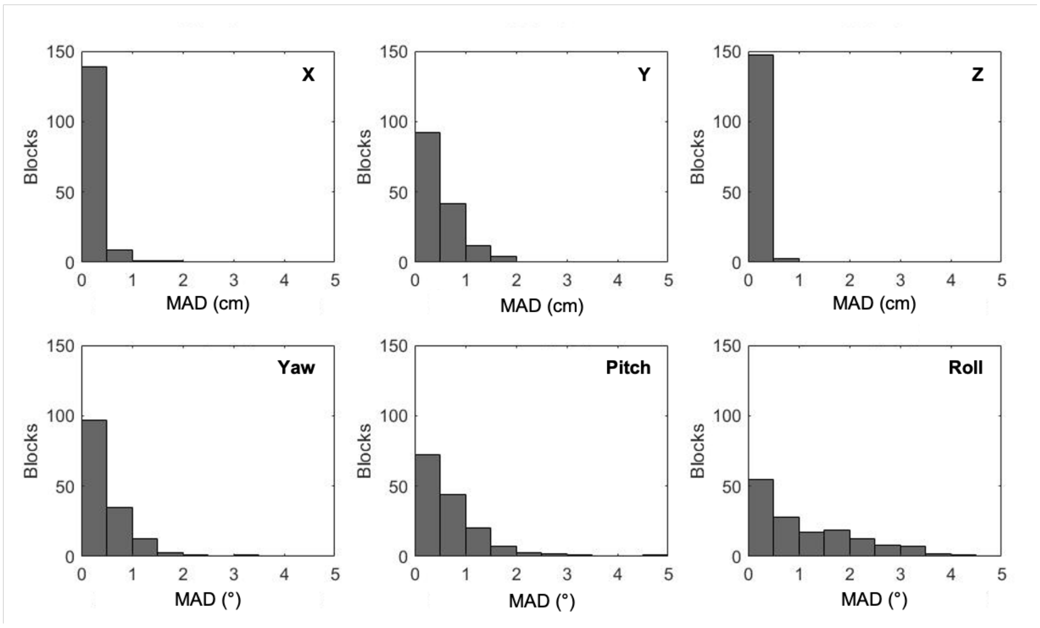

Figure 19. Histograms of median absolute deviation of head movement per block, in positional and rotation axes.

The median absolute deviation from the centre of the ring was less than 1cm in all positional axes (X, Y, Z) for the majority of blocks. Deviation in rotational axes (yaw, pitch, roll) was similarly small, with most blocks showing less than 2° in median absolute deviation.

3.4.3.2 Minimum audible movement angles.

Thresholds for each source azimuth are shown in Figure 20. A repeated-measures ANOVA found that there was a significant main effect of source azimuth on MAMA, $F(4,32) = 47.230$, $p < 0.001$. T-tests found that there was no significant difference in MAMA between midline complementary pairs ($t(8) = 1.092$, $p = 0.307$). MAMAs were significantly greater at the oblique position in the rear than at its complementary pair ($t(8) = 2.566$, $p < 0.05$).

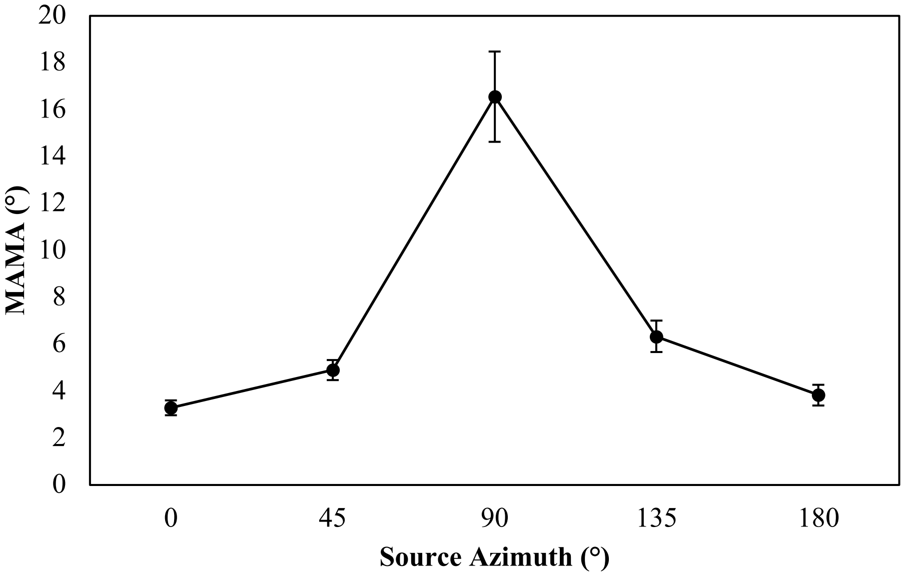

Figure 20. Minimum audible movement angle as a function of source azimuth. Error bars are ± standard error.

3.4.4 **Discussion.**

The asymmetry in spatial precision found in the MAA experiments between front and rear hemifields was replicated here for moving stimuli. The increase in MAMA from midline to interaural axis locations replicates previous MAMA research (Grantham, 1986), while the overall pattern from front to side to rear was predicted based on the MAA data presented earlier in this thesis.

The method of MAMA chosen here was direction discrimination, which is arguably a higher-order perceptual process for the auditory system compared to motion detection per se. For the auditory system to discriminate the direction of motion it must be able to detect motion in a stimulus first. Hence, as the precision asymmetry has been found for the higher-order task, it would seem to be necessarily the case that it should also be found for the low-level task (motion detection). Indeed, in experiments not reported in this thesis this was found to be the case.

It could be that listeners were simply treating the direction discrimination task as if it were an MAA task by comparing the start and end points of the motion trajectory (Perrott & Marlborough, 1989). If this were the case then there would be no need to detect motion to perform this task, and therefore no discrimination of motion direction. Anecdotally, listeners all reported perceiving the stimulus as moving even down to threshold performance, so this is unlikely to be the case. Additionally, Perrott & Marlborough's finding that listeners benefitted from the addition of motion across the whole trajectory suggests that listeners use this information when it is available to improve their performance. So, relying solely on start and endpoints would be a worse strategy than attempting to discriminate direction of motion.

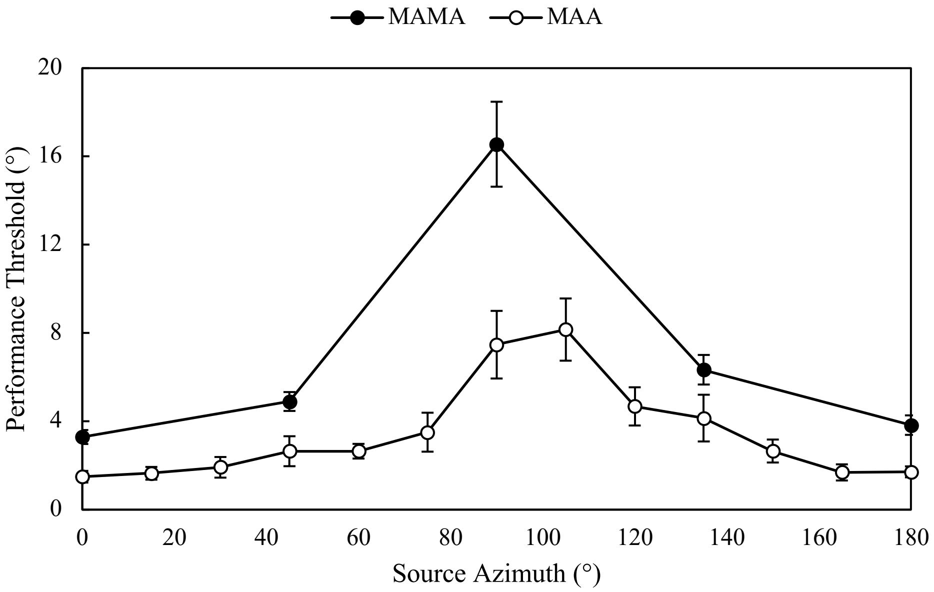

Figure 21. MAMAs from this experiment (filled circles) plotted against MAAs from the experiment in section 3.2 (open circles). Error bars are ± standard error.

Note also that MAAs are typically found to be lower than MAMAs when comparisons are made between the two (Carlile & Leung, 2016). Figure 21 shows that the current data replicates this finding. This is consistent with the predictions from snapshot theory (discussed below): auditory motion perception inherits the variability from the lower level localisation abilities coupled with the variability of the motion analysis process itself. Hence, motion perception is more variable than static localisation, and so motion detection and direction discrimination thresholds (MAMAs). must be greater than localisation thresholds (MAAs).

There are two general theories as to how the auditory system generates the perception of motion of sounds in the spatial field. The first of these is snapshot theory (Grantham, 1986),which proposes that motion is inferred from comparisons of successive 'snapshots' of locations in the spatial field. Changes in the location of a sound between snapshots are inferred to be due to movement of that sound. This can be thought of as analogous to a camera capturing consecutive images of a moving scene. Snapshot theory suggests that these snapshots happen in rapid sequence, with successive snapshots being integrated into a single stimulus. Sounds whose position changes less frequently than this (e.g., due to a low refresh rate on the presentation method) are perceived not as moving, but as 'jumping' between locations or as separate sources appearing and disappearing. Conversely, sounds whose position updates more frequently than this can appear to 'blur' into one larger sound (Griffiths & Green, 1999) as the auditory system cannot generate snapshots fast enough to track the changes in location.

Under snapshot theory higher-order tasks may inherit an increasing amount of variability, because each step in the hierarchy from static localisation to motion perception incurs its own variation. This should have the effect that threshold performance is greater for higher-order tasks than for lower-level ones. This has indeed been found to be the case, with movement detection thresholds being found to be higher than location discrimination thresholds (Chandler & Grantham, 1992; Harris & Sergeant, 1971; Perrott & Tucker, 1988). Similarly, the patterns of spatial precision found with changing stimulus characteristics found with static sources should be reflected in the higher-order tasks as well. Again this has been found to be the case, with movement discrimination thresholds increasing with source azimuth further from the midline in the front hemifield, in the same pattern as MAAs

(Grantham, 1986). These findings are generally taken as strong evidence that snapshot theory is correct in its explanation of auditory motion perception.

The second general theory is direct-motion perception (Perrott & Marlborough, 1989). This theory suggests that the auditory system, much like the visual system, has dedicated low-level detectors for the perception of motion. However, input to the auditory system does not have a direct spatial component, unlike the receptors in vision which preserve spatial relationships. A ‘direct’ perception of motion must somehow be built on low-level cues of timing, level, or spectral content. If such detectors do exist, they are suggested to be located within the temporal lobe (Baumgart et al., 1999). Warren et al. (2002) found several regions which fired selectively to motion in an auditory stimulus but not to a static stimulus, suggesting that this area is specialised to detect motion. However, given that activity in these areas must be preceded by lower-level timing, level, and spectral perception, it may simply be that these areas are the ones which infer that motion is present and process it, and not that there are motion detectors separate of the spatial perception areas of the auditory system. Freeman et al. (2014) found that listeners asked to judge auditory stimulus speed relied on displacement and duration cues in the stimuli, and only used velocity cues when displacement and duration were made uninformative. Additionally, when using velocity to make the speed judgements, their performance was significantly worse than when using lower-level cues. This suggests that if specialised areas for the motion perception abilities (e.g., velocity) do exist, they perform poorly. Alternatively, in Freeman et al. as displacement and duration cues were disrupted but not absent it may have been that the auditory system was forced to rely on poor temporal and spatial information in this condition, and so listeners were not making a direct velocity judgement at all.

Direct-motion theory does not make predictions about whether high-level motion detection and perception abilities should follow the patterns of performance with stimulus characteristics that have been found using static sources. If specialised motion detectors do exist, they need not inherently follow these patterns as they should be independent of spatial localisation abilities. However, as motion perception must be reliant on timing, level, and spectral cues, it is likely that motion perception under a direct-motion model would show similar patterns. This is because the changes in spatial precision are directly related to the informativeness of these cues as they change with, for example, azimuth. The informativeness of these cues would not be expected to be different just because the auditory system is making judgements about motion rather than space. So, it is probable that if the direct-motion theory is correct then the same patterns of performance will be seen in motion tasks as has been found in spatial tasks.

The findings from this experiment cannot distinguish between snapshot theory and direct motion perception, as both theories can predict that motion perception may follow the precision pattern across azimuth of spatial perception (Chandler & Grantham, 1992). However, the more parsimonious explanation is given by snapshot theory, based on the correspondence of precision patterns across tasks. Perrott & Marlborough's (1989) finding might also seem to suggest that improvement given by the movement during a trajectory is evidence for direct-motion perception. Snapshot theory, however, proposes that successive localisations can be made at a more rapid interval than the start and end-points of the stimuli in this experiment would permit (Neuhoff, 2004). Consequently, a complete trajectory would provide more snapshots and so provides more information.

As the spatial precision asymmetry has been replicated in simple motion perception, it should generalise to higher-order motion perception contexts, such as judging source movement during a head turn. The MAMA task, being performed with a stationary head, need only track the spatial motion of a stimulus across a static field. However, in normal life the head is rarely still even when visually fixating (Kunin et al., 2007), and so the auditory system needs to be able to integrate information about how the ears move in space, as described in more detail in Chapter 5. As will be argued, this ability must be reliant on the simple motion perception abilities of the auditory system, and so it is expected that the same or similar pattern of precision performance across azimuth will be found.

3.5 Chapter Discussion

Direct measures of spatial precision provide insights that cannot be, or have not been, achieved from localisation accuracy studies. The lack of response bias confounds in the MAA task removes much of the variability seen in indirect measures of spatial precision. This allows for better inference about low-level spatial precision, and so better prediction of higher-level motion perception.

In the experiments in this chapter, auditory spatial precision was tested on the azimuthal plane using the MAA task described in 2.2, and with a MAMA task described in section 3.4.2. The first experiment demonstrated how spatial precision changes as sounds sources are located in different positions in azimuth. The finding of an increase in MAA as sound sources are placed further from the midline towards the interaural axis replicates that found by Mills (1958) for the front hemifield. This pattern of MAAs was also present in the rear hemifield. However, an asymmetry was found between MAAs in the front and rear hemifields as in Starch (1905), such that precision was worse in the rear hemifield than at symmetric positions in the front hemifield.

The second experiment tested the hypothesis that this asymmetry was dependent on the spatial characteristics of the laboratory, by orienting the listener opposite to their orientation in the first experiment. This did not change the asymmetry in the data, confirming that it was not an artefact of the laboratory or speaker array layout.

The third experiment used the MAMA and examined the minimum displacement angle required for the discrimination of motion direction. This experiment found that spatial precision for motion direction discrimination followed the same pattern of spatial precision with source azimuth as has been found using the MAA.

The experiments in this chapter showed that the way precision varies with azimuth is similar for both static stimuli and moving stimuli. This conforms to predictions made by snapshot theory (Grantham, 1986), which proposes that all perceptions of auditory motion are built from successive ‘snapshots’ of the perceived locations of sound sources. Hence, the changes in precision at the lowest level (MAAs) are reflected in higher level, more complex tasks such as motion discrimination (MAMA) and sensory signal integration (head compensation). The competing theory, direct motion detection (Perrott & Marlborough, 1989), suggests that specialised detectors for motion exist within the auditory system, and that these are independent of spatial localisation ability. As these are independent, they need not be susceptible to the same spatial precision changes as the localisation ability of the auditory system. However, as they must be based on the same physical cues (ITD, ILD, spectral content), they may inherit the same precision changes if these are due to sensory signal reliability.

An alternative interpretation is that the spatial maps used by listeners are somehow deficient for locations in the rear hemifield. Auerbach and Sperling (1974) found evidence that auditory and visual stimuli share a common spatial map in the brain, typically believed to be in the superior colliculus (Middlebrooks & Knudsen, 1984; Palmer & King, 1982). Meredith & Stein (1986) found that there were cells in the superior colliculus which showed spatiotopic mapping to auditory stimulation and somatosensory stimulation, as well as cells which responded to combinations of these. This is evidence that the combination of sensory maps occurs in this area. Research into spatial maps has been primarily conducted using invasive methods on animals (Knudsen & Brainard, 1995; Meredith & Stein, 1986; Middlebrooks & Knudsen, 1984), making human replications somewhat harder to perform. Non-invasive imaging methods have, however, been used in humans. Using fMRI, DuBois & Cohen (2000) found retinotopic mapping in the superior colliculus corresponding to a visual map. Giard & Peronnet (1999) and Foxe et al. (2002) found multisensory mapping in humans in the temporal gyrus between auditory and somatosensory stimulation, using electrophysiological and magnetic resonance imaging methods respectively. This suggests that the combination of spatial maps found in animals (Meredith & Stein, 1986) may also be occurring in humans.

However, not all maps can receive inputs from all areas of space. A spatial map for vision must necessarily only include the front hemifield, as this is the region covered by the visual field, whereas a spatial map for hearing can conceivably cover the full sphere of space. Therefore, if there is a common map across sensory modalities then it would contain fewer sources of information for location in the rear hemifield where visual information cannot be accessed. If listeners are reliant on the precision of this map to perform the MAA task, then this may be why they exhibit poorer performance in the rear hemifield. This is, however, outside the scope of investigation of this thesis.

The influence of room acoustics was ruled out in section 3.3. In the next chapter, the focus turns to the location-dependent spectral filtering of the pinna as the reason that spatial precision asymmetry occurs. The pinnae are oriented towards the front hemifield, and are known to diffract and reflect sounds differently based on their hemifield of origin (Musicant & Butler, 1984; Rayleigh, 1907; Shaw, 1974). The pinnae only affect sounds containing frequencies above around 4 kHz (Butler & Humanski, 1992), so removing frequencies above this threshold or removing the ability of the pinnae to contribute spectral cues should disrupt the spatial asymmetry.

Previously the pinnae have only been suggested to play a role in spatial perception for front-back discrimination (Shaw, 1997) and the perception of elevation (Shaw, 1974). It is possible that the results reported in the current chapter suggest that the pinnae may also provide a benefit for spatial localisation in the front hemifield, or that it creates a detriment to localisation in the rear hemifield. If true, this would be a previously unknown application of the pinnae's spectral cues. This suggestion is tested in section 4.4.

Chapter 4 – Spectral Content as a Source for an Asymmetry

4.1 Chapter Overview

The previous chapter demonstrated an asymmetry in the spatial precision of the auditory system between the front and rear hemifields: spatial precision is worse for most positions in the rear hemifield than at symmetric positions in the front hemifield, referred to as symmetric position pairs. It was suggested that changing the precision of the sensory signals may reveal the source of the front/rear asymmetry. The clearest candidate, as proposed in section 3.2, is the pinnae. Sounds originating from symmetric position pairs should not differ in path length, so binaural differences (time, level) should not be different between hemifields. Therefore, it is the spectral content that is most likely to be different, and for this we must consider the pinnae.

This idea was tested in two ways. The experiment in section 4.2 was based on the idea that the pinnae are believed to only affect spectral content of sound above approximately 4 kHz (Butler & Humanski, 1992). If the pinnae are the cause of the asymmetry then using filtered noise that excludes sound in that frequency range should eliminate the differences in precision. As this experiment used free-field listening we cannot definitively state that the pinnae are having no effect on perception. To address this issue, sections 4.3 and 4.4 describe a pair of experiments that use head-related transfer functions to nullify the effect of the pinnae. In the former, measurements of the contribution of the pinnae are made using an acoustic mannikin. From this, a method of bypassing the pinnae using tubes is created. In the latter section this method is tested on human listeners, and the spatial precision asymmetry, or lack thereof, is examined with the contributions of the pinnae bypassed.

4.2 Spatial Precision Asymmetry and Stimulus Frequency Content

4.2.1 **Introduction.**

Previously in this thesis, all examinations of spatial precision have been conducted using broadband noise (200 Hz – 20 kHz). Broadband noise, or white noise, has equal power at all frequencies and so the auditory system should not experience any (abnormal) biases in the cues which it uses to localise sound.

It is known that different auditory cues to spatial location are more or less informative depending on the frequency content of the to-be-localised sound. For binaural cues, interaural time differences are more informative at low frequencies and interaural level differences are more informative at high frequencies (Macpherson & Middlebrooks, 2002). Of course, when listening to broadband sounds both cues are available. But these are not the only cues to spatial location. Monaural cues can be obtained indirectly from the filtering effect of the pinnae, the fleshy outer part of the ear. Pinna cues are most informative for frequencies above 4 kHz (Musicant & Butler, 1984) and show little to no informativeness for frequencies below this threshold (Bronkhorst, 1995).

The stimuli in all experiments presented in this thesis use loudspeaker presentation, so neither primary type of binaural cue (ITD, ILD) can be eliminated entirely, despite this being a direct means of examining the contributions of a given cue to perception. They can be relatively strengthened or weakened by changing the spectral content of the presented stimuli, but they must both be present due to the nature of the movement of acoustic energy through air. Pinna cues, though, can be eliminated for the majority of listeners, by removing high frequencies (Musicant & Butler, 1984; Oldfield & Parker, 1984b). Consequently, by changing the spectral content of the stimuli the relative informativeness of binaural cues versus pinna cues can be changed. This should then have an effect on spatial localisation ability.

In this section the MAA task (section 2.2) was used to examine how the spatial precision asymmetry changes when the frequency content of the sound sources is changed. This asymmetry demonstrates that listeners have worse precision, as indicated by larger MAAs, in the rear hemifield compared to the front hemifield. This asymmetry may be caused by differences in available high-frequency information between the hemifields, most likely due to the orientation of the pinnae towards the front of the head (Shaw, 1974). Consequently, it was predicted that removing high-frequency information from the sound sources would remove any differences between MAAs for symmetric position pairs.

Three noise stimuli were created for this experiment. The first was broadband noise, which should replicate the findings from the previous experiments of poorer precision in the rear hemifield. The remaining stimuli were high-passed and low-passed broadband noise. These were broadband noise stimuli which had any spectral content below or above 4 kHz removed, respectively.

To reiterate the discussion in section 3.2, the use of lowpass filtered noise stimuli should reduce or remove the spatial precision asymmetry that was found for broadband noise. This is based on the front-facing orientation and filtering specificity of the pinnae (Musicant & Butler, 1984; Shaw, 1974; Starch, 1905) The pinnae filter the spectra of sounds depending on their elevation, and also on whether they originate in the front or rear hemifield. The pinnae only have an effect on frequencies above approximately 4 kHz, so lower frequencies in sounds should remain unaffected by the pinnae regardless of the location of their source. If the pinnae provide a different benefit (or detriment) for sounds in the front hemifield compared to sounds in the rear hemifield, then removing the high frequencies should have prevented the pinnae from causing any asymmetry in spatial precision between these hemifields.

4.2.2 **Methods.**

4.2.2.1 Participants.

Ten listeners took part in this experiment (5 females, ages 20-30). All listeners self-reported normal hearing.

4.2.2.2 Stimuli.

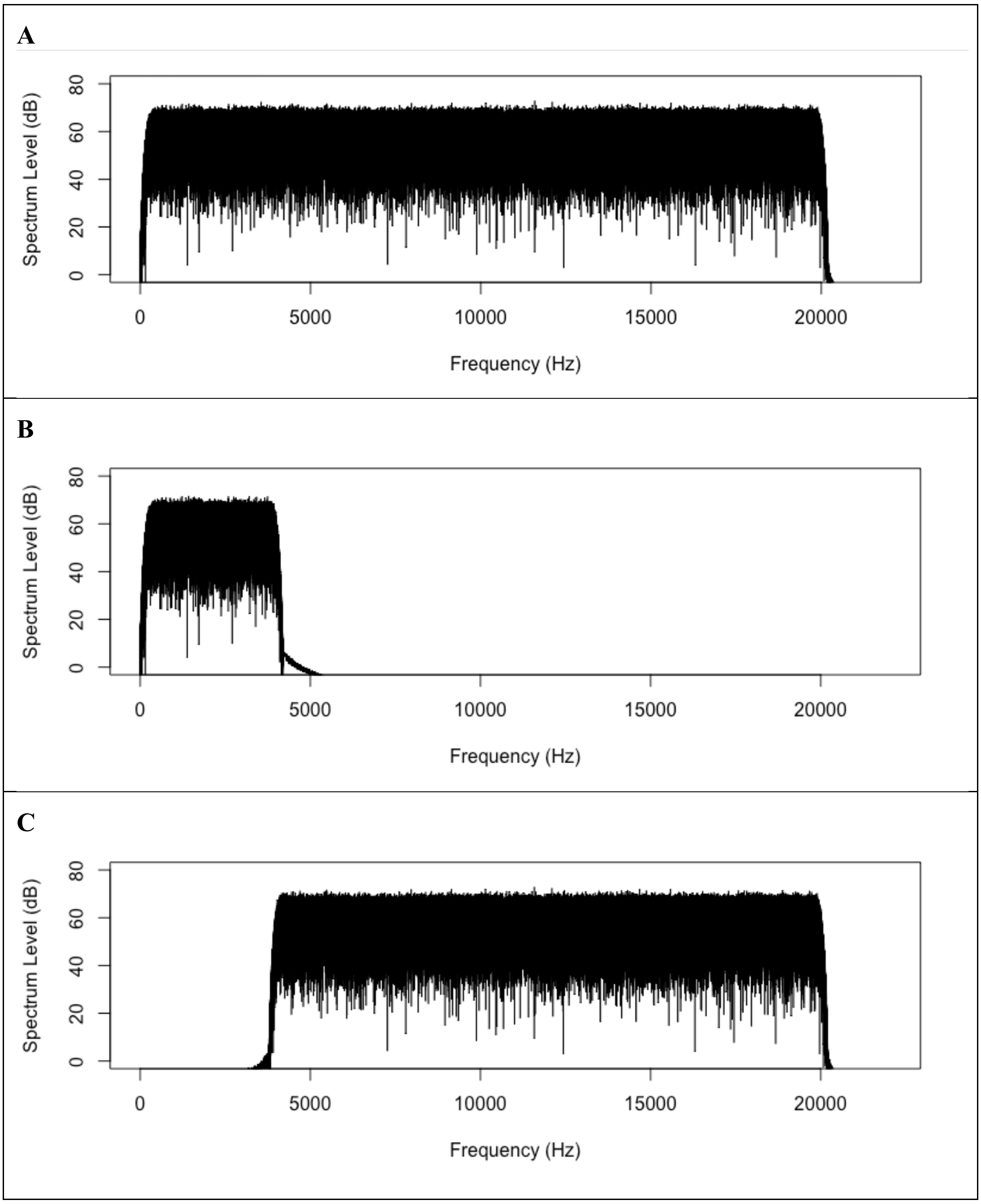

Figure 22. Power spectra of noise stimuli: (A) broadband, (B) low-pass filtered at 4 kHz, (C) high-pass filtered at 4kHz.

Three noise stimulus conditions were used. Broadband stimuli were generated as in section 2.2. Two 512-point FIR filters were then used to create the filtered noise stimuli. These consisted of broadband noise high-pass or low-passed filtered at 4 kHz. The power spectra of all noise stimulus types are shown in Figure 22.

4.2.2.3 Procedure.

The MAA task was used, as specified in section 2.2. Each combination of filter type and sound source location was presented in an individual block. All conditions were presented in a random order, with three repetitions of all conditions. This experiment took place in two sessions of one hour each on separate days to prevent listener fatigue.

4.2.3 **Results.**

4.2.3.1 Head position and orientation.

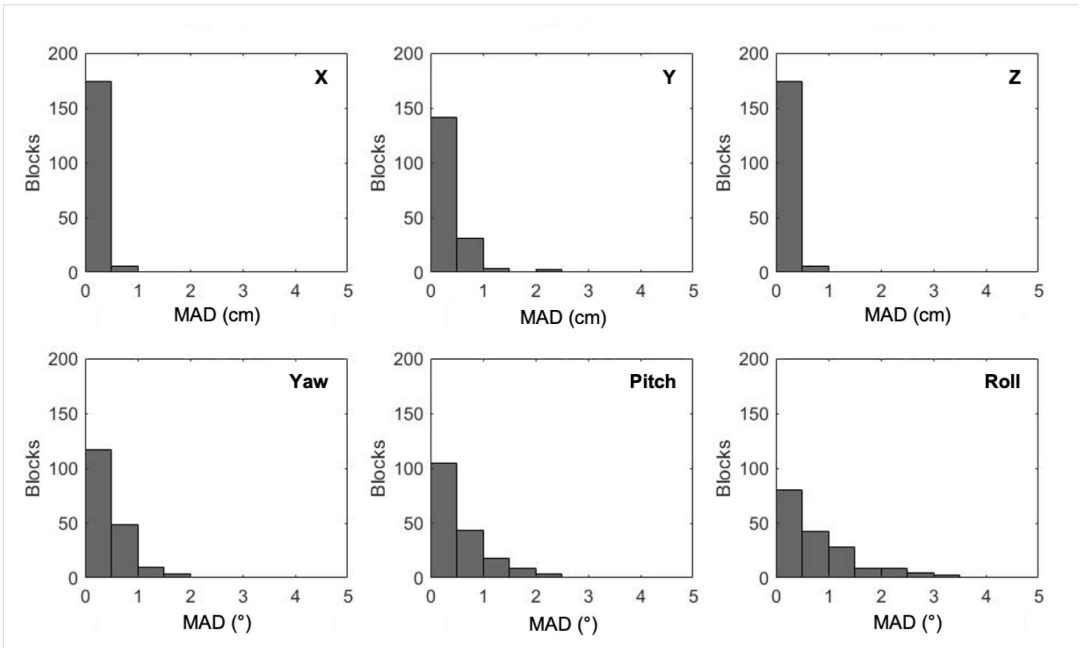

Figure 23. Histograms of median absolute deviation of head movement per block, in positional and rotation axes.

Head position was tracked prior to the start of every trial, and the MAD in head position and orientation was calculated for each block (Figure 23). Deviation was less than

1 cm in the majority of blocks for all positional axes (X, Y, Z), and was less than 2° for almost all blocks in rotational axes (yaw, pitch, roll).

4.2.3.2 Minimum audible angles.

MAAs were averaged across all repetitions and staircases per condition, for each listener. Figure 24 shows that the asymmetry reported in Chapter 3 was replicated with broadband noise (grey bars) and persists for highpass filtered noise (black bars). However, the asymmetry is eliminated when lowpass filtered noise is used (white bars).

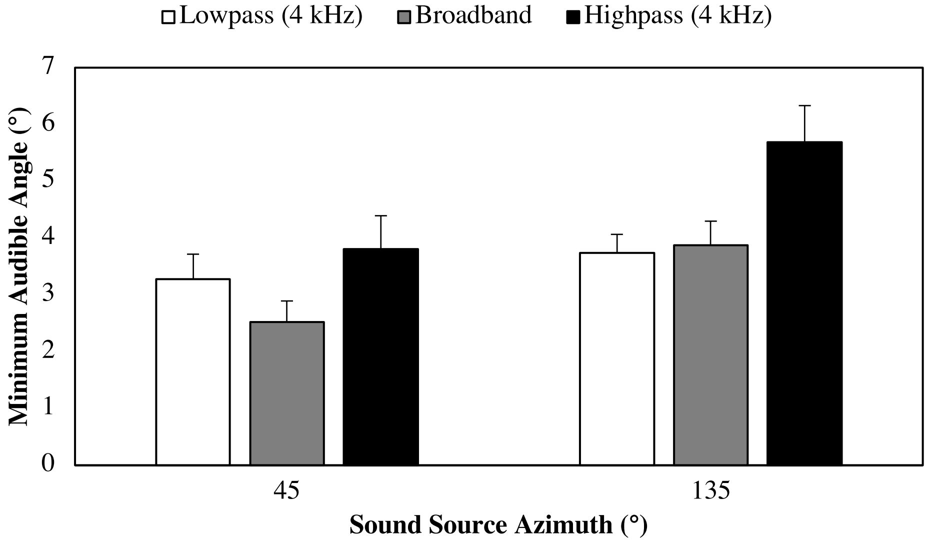

Figure 24. Minimum audible angle as a function of source azimuth and noise filter type. Error bars are standard error.

This was confirmed using a repeated-measures ANOVA. There was a significant interaction between filter type and sound source azimuth, $F(2,18) = 3.886, p < 0.05$. There was a significant main effect of filter type on MAA, $F(2,18) = 27.091, p < 0.001$. There was also a significant main effect of sound source azimuth, $F(1,9) = 7.627, p < 0.05$.

Paired t-tests showed that MAAs were significantly higher when sounds were presented in the rear hemifield than in the front hemifield for highpass filtered noise, $t(9) = 2.572, p < 0.05$, and for broadband noise, $t(9) = 3.124, p < 0.05$. MAAs were not significantly different between hemifields for lowpass filtered noise, $t(9) = 1.235, p = 0.248$.

4.2.4 **Discussion.**

The MAA was measured for sound sources in the front and rear hemifields, using three types of noise. As found in previous chapters, MAAs for broadband noise stimuli showed spatial precision that was worse in the rear hemifield than the front. The same asymmetry was found for highpass stimuli, with overall MAAs in this condition higher than for broadband stimuli. However, the asymmetry disappeared for lowpass filtered sounds. Hence, high frequency cues appear to be the main driver of the asymmetry between front and rear hemifields.

While the pattern of MAAs between hemifields is the same for broadband and highpass filtered noise stimuli, the magnitude of the MAAs measured is not. In both front and rear hemifields, highpass filtered stimuli show larger MAAs than broadband stimuli at the same locations. The sole difference between these stimulus types was the inclusion or exclusion of frequencies below 4 kHz. When listeners had access to low frequency information, their MAAs improved compared to when they had high frequency information alone. It has been found that ITDs dominate over ILDs when the two are in conflict (Wightman & Kistler, 1992, 1997), and ITDs are weighted more heavily for low frequencies than for high frequencies (Macpherson & Middlebrooks, 2002). These highly weighted ITD cues are present in broadband stimuli (and lowpass stimuli) but not in highpass stimuli. Carlile et al. (1999) found that localisation accuracy did not suffer when low-frequency cues were removed, which would seem to argue against the importance of ITDs. However, they note that standard deviations in localisation response (an indirect measure of precision) increase for some rear locations. This was in addition to the increase in error for rear locations usually found in localisation studies (Figure 3). Carlile et al.'s findings are therefore confirmatory evidence that a lack of low-frequency cues is actually problematic for spatial localisation. Recall that an imprecise cue can result in good accuracy if enough samples are taken (dartboard analogy, section 1.4.1), so accuracy can be a poor measure of the precision of cues to spatial location.

In the following section a technique is developed that effectively ‘removes’ the pinna during free-field listening in the speaker ring. The technique is based on the use of tubes inserted into the ears to bypass the pinnae. This is combined with a filter that corrects for the residual effects of the tubes on the sounds delivered by the speakers, based on measurements made using an acoustic mannequin. In the subsequent section (4.4) this technique is then used with human listeners.

4.3 Generation of a Method of Bypassing the Pinnae in Human Listeners

4.3.1 **Introduction.**

In the previous section it was found that lowpass filtering broadband noise removed the spatial precision asymmetry found between hemifields. This was proposed to be due to the removal of high-frequency information from the stimuli used, suggesting that the precision asymmetry may be due to the influence of the pinnae. The pinnae are known to provide direction-selective spectral filtering for elevation (Blauert, 1997; Wenzel et al., 1993), and aid discrimination of front versus back (Musicant & Butler, 1984). There has been no direct evidence that the pinnae provide information on monaural localisation in azimuth within a given hemifield, though some studies have indirectly shown poorer precision in the front hemifield when the pinnae are occluded (Gilse and Roelofs, as cited in Musicant & Butler, 1984).

It has, however, previously been shown that human HRTFs vary substantially with source location on the horizontal plane (Carlile & Pralong, 1994), and that the difference in transfer function gain between front and rear positions is most evident for frequencies above 4 kHz. This is the frequency range in which the pinnae show filtering effects on the auditory spectrum. Hence, it is reasonably to conclude that the pinnae play a role in the front-rear hemifield differences found by Carlile et al. What is unclear is whether removing the effects of the pinnae also removes these qualitative differences, and what effect this might have on localisation performance.

In this section a method is developed to bypass the pinnae and deliver sounds unadulterated to the ear canal. This requires the use of some physical apparatus (section 4.3.2) to acoustically bypass the pinnae, as well as a method of filtering the stimuli so that the more general effects of listening through such apparatus are nullified (section 4.4.2). The auditory spectrum is then measured across azimuth to examine whether qualitative differences between front and rear hemifields exist under these conditions, compared to normal hearing conditions. The method of bypassing the pinnae is then tested on human listeners using the MAA task to examine how this affects spatial precision.

4.3.2 **Tube measurements.**

4.3.2.1 Physical apparatus.

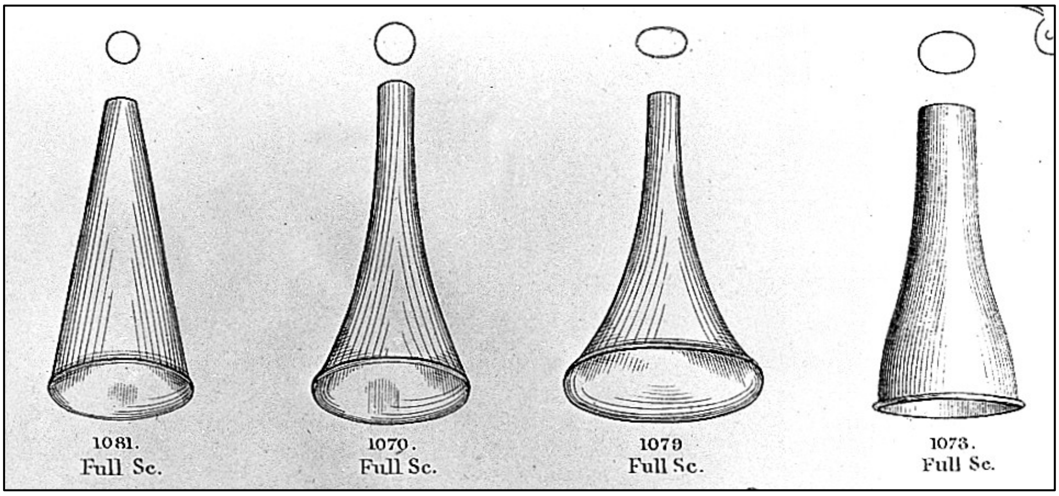

Figure 25. A set of otoscope specula. The specula used in this experiment most closely resembled the second from the left, with a curved taper from the end towards the horn, and a circular opening. Adapted from Dr Siegle's pneumatic otoscope (Weiss, 1889). CC BY 4.0.

In order to bypass the pinnae, modified otoscopic specula were used. These tubes were long enough to extend past the rim of the pinna, without extending too far and creating an artificially large 'head'. This would have had the effect of creating 'supernormal' localisation cues, which listeners find very hard to use (Shinn-Cunningham et al., 1998a, 1998b). They also sat comfortably and securely in the external opening of the ear canal such that they could be worn for moderate periods of time without moving unintentionally.

An otoscope speculum (Figure 25) was chosen as the inner shell of the tube due to its established clinical use within audiology, making it safe to enter the ear canal. It also provides a convenient funnel shape, creating a conical surface to funnel sounds into the ear canal with limited distortion. On the narrow end of the otoscope speculum a silicone earphone bud was attached to enhance comfort. The inner diameter of the earphone bud was chosen to be small enough so that there would be no hard plastic extending beyond the earphone bud that could come in contact with the walls of the ear canal. Similar to otoscope specula, earphone buds are designed to be inserted into the outermost part of the ear canal, which reassured us as to their safety. The silicone material of which these buds were made provided a close and secure fit to the otoscope tips. They were purchased in a variety of shapes and sizes, which provides a range for listeners to choose from that best suit their own ear shape. When assembled and inserted, the complete tube apparatus extended less than 1cm past the rim of the pinna.

4.3.2.2 Acoustic dummy measurements.

The effect of the tubes on the sound received at the eardrum was assessed electroacoustically. A KEMAR acoustic dummy (GRAS Sound & Vibration), with full head and torso, was used to record head-related impulse responses (HRIRs) from each speaker in the circular array. The HRIR measurements used a stimulus which sweeps from the lowest to the highest frequency, and so can record the excitation at an ear drum analogous point for all frequencies.

The acoustic dummy was positioned in the centre of the speaker array, this time using the laser method (not the cord method) detailed in section 2.2.2. Three sets of recordings were then made. The first was a ‘normal listening’ condition, in which a pair of silicone outer ears were attached to the dummy. The second was a ‘pinnae removed’ condition, where the outer ears were detached from the dummy, leaving the opening to the ear canal flush with the air. To check that the use of specula would not distort the proximal stimulus beyond utility for localisation, in a third condition the outer ears were re-attached and tubes were placed into the meatus of the ear canal ('tube bypass'), to replicate the tubes being worn by a human listener. In this condition the tubes were moved by the experimenter until they securely pointed directly out from the head as perpendicularly as possible. The importance of the exact angle is assessed in section 4.3.2.4.

4.3.2.3 Results.

Figure 26 shows the excitation patterns for each condition (normal listening, pinna removed, tubes inserted) across all azimuths. In Figure 26(A) the excitation pattern shows an asymmetry in the excitation between 15°-90° (front hemifield) and 90°-165 (rear hemifield), primarily between 4 kHz and 10 kHz. Figure 26(B) shows no such asymmetry, implying that this is due to the presence of the pinnae in the *normal listening* condition. Figure 26(C) also shows no front/rear asymmetry, suggesting that the use of the tubes has successfully bypassed the pinnae. There is still a difference however between the *pinnae removed* and *tubes bypass* conditions. There is an attenuation of sound level between 1.5 kHz and 5 kHz, and an increase between 650 Hz and 1500 Hz. This means that using tubes to bypass the pinnae is not the same as removing the pinnae completely. This is unsurprising as listening through tubes introduces a new acoustic environment and may increase or reduce sound level for the coupling of the ear canal to the surrounding air for different regions of the frequency spectrum.

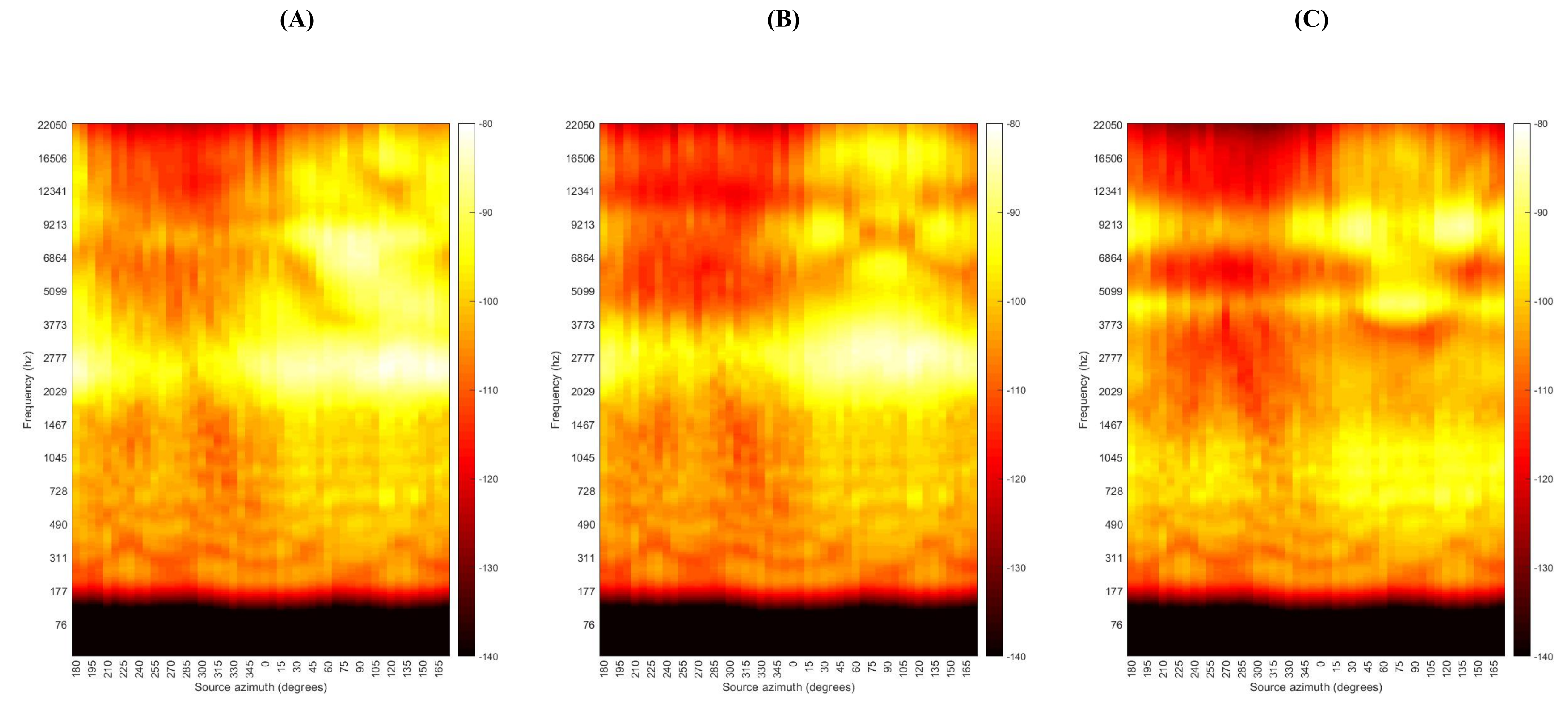

Figure 26. Excitation patterns recorded using an acoustic mannikin, colour scale shows dB. (A) excitation with the pinnae attached to the mannikin [*normal hearing*]. (B) excitation with the pinnae removed from the mannikin [*pinnae removed*]. (C) excitation with pinnae attached and tubes inserted into the ear canal [*tube bypass*]. Excitation is shown for the right ear only.

4.3.2.4 Tube position angle.

The orientation of the tubes relative to the head may be very important in this experiment. If the tubes point towards an azimuthal location which does not lie on the interaural axis, then there may be an artificial benefit in the collection of sound waves from this hemifield. This would then affect the HRIRs that are gathered and may confound the results.

In order to test the acoustic effect of the orientation of the tubes, a series of recordings was made with the tubes oriented at different angles on the horizontal plane. Examples of how a tube could be oriented in the ear canal are shown in Figure 27.

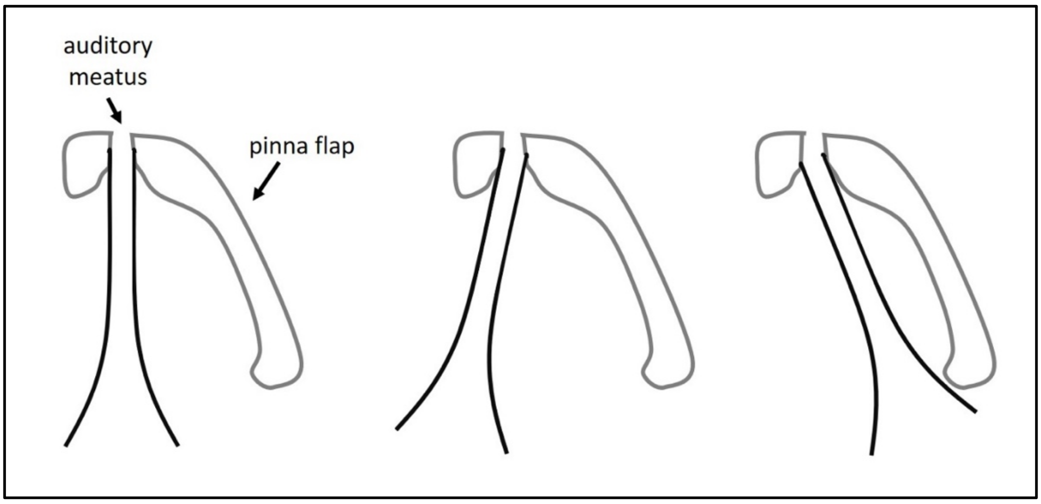

Figure 27. Three possible positions for the tube (black) to be oriented in the ear canal, with a horizontal cross-section of the pinna (grey). The left image shows a tube oriented directly outwards from the head at a perpendicular angle. The central image shows a tube offset towards the front hemifield. The right image shows a tube offset as far as possible towards the rear hemifield, nearly touching the pinna.

Measurements were made with the tubes pointed directly out at a perpendicular angle from the dummy’s ears, as far towards the rear as possible so that the outer material of the tubes was against the pinnae, and at an intermediate angle between these two positions. It was not possible to orient the tubes further towards the front field without disrupting the coupling of the tubes to the ear canal, so this orientation was not measured. Power spectra were compared between these orientations visually (Figure 28). Inspection did not reveal prominent differences in the spectra between orientations. From this it was inferred that there was a reasonable range of angles around a perfect perpendicular tube orientation, within which there would be no substantial difference in the filtering response received at the ear canal.

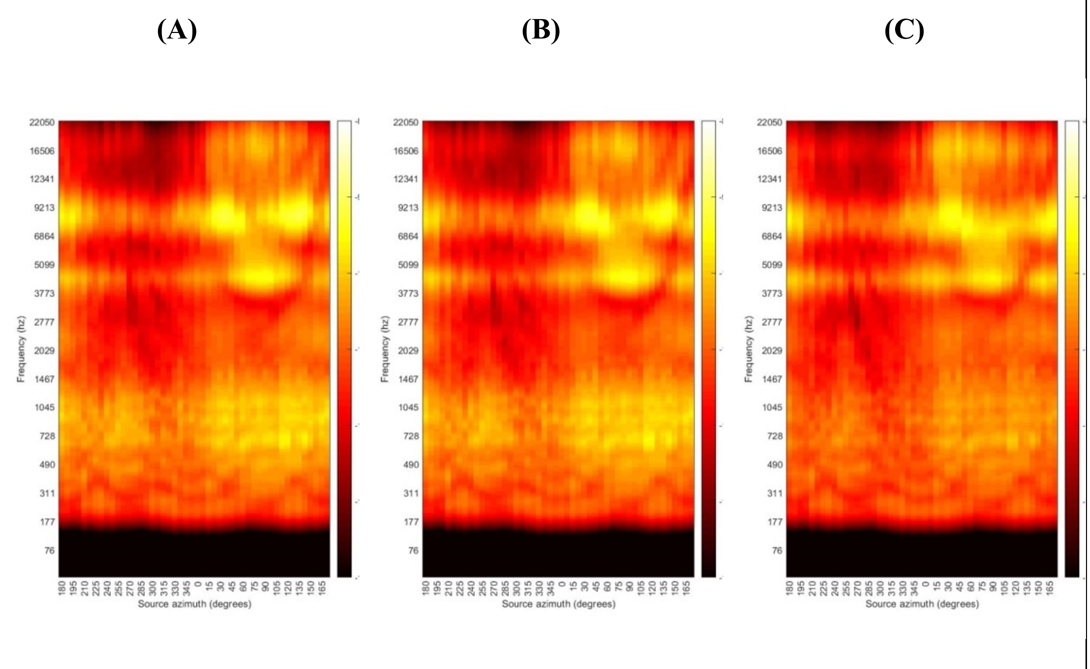

Figure 28. Excitation patterns for three different orientations of the tube apparatus in the ear canal of the acoustic mannequin: (A) directly out from the head, (B) slightly towards the rear, (C) further to the rear, nearly touching the pinna flap.

4.3.3 **Discussion.**

In this section it was shown that the pinnae are the probable cause of the front/rear spatial precision asymmetry found in previous experiments. By removing the outer ears of an acoustic mannikin, a comparison could be made between power spectra across azimuths with and without pinnae. Under normal hearing conditions, qualitative differences were observed between front and rear hemifields for frequencies above approximately 4kHz, replicating the findings from Carlile and Pralong (1994). When the pinnae were removed these differences disappeared, suggesting that the directional filtering of the pinnae was the cause of the asymmetries seen. Bypassing the pinnae using a tube apparatus also removed the differences between front and rear hemifields. This represents a promising method of removing the directional effects of the pinnae in human listeners without surgical intervention – which is prohibitive to perform in human listeners for a multitude of reasons. Use of the tube apparatus to bypass the pinnae did generate spectral changes in the proximal stimuli, however, but this effect is controlled for in the experiment in the following section.

4.4 Bypassing the Pinnae to Remove Spectral Filtering Cues

4.4.1 **Introduction.**

Based on the data presented in the previous sections, it was hypothesised that spectral filtering from the pinnae is the cause of the spatial precision asymmetry established in this thesis. Hence, this experiment will examine spatial precision in the front and rear hemifields using the MAA task, with normal (unaltered) pinnae and with bypassed pinnae.

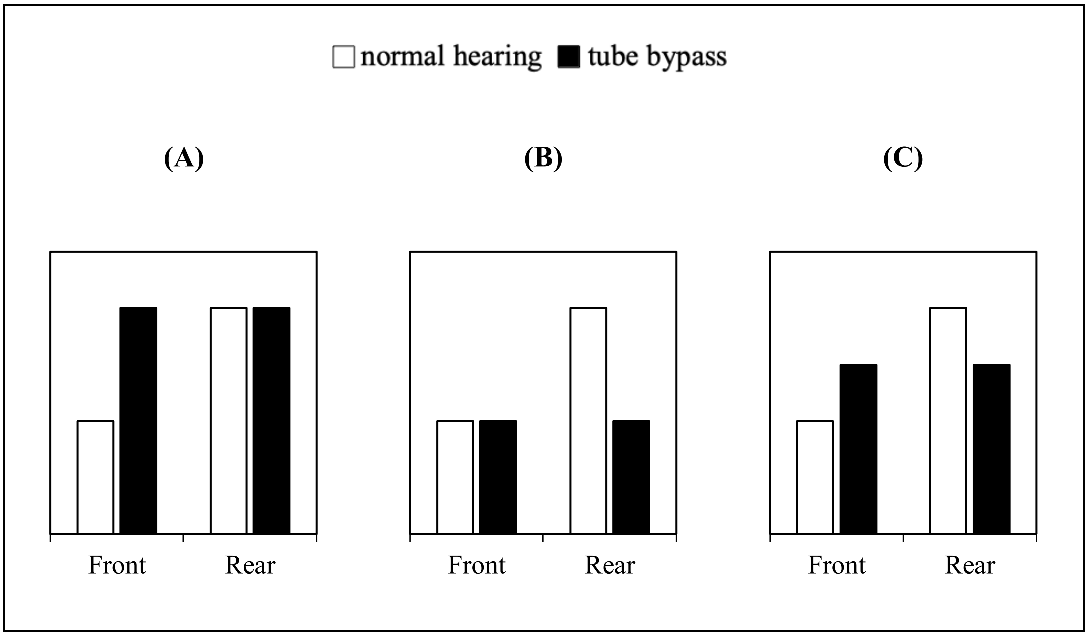

Figure 29. Hypothetical patterns of performance by listeners on an MAA task. (A) The pinnae provide a selective benefit in the front hemifield. (B) The pinnae provide a selective detriment in the rear hemifield. (C) The pinnae provide both a selective benefit in the front hemifield and a selective detriment in the rear hemifield.

If the use of a method to bypass the pinnae does remove the spatial precision asymmetry between hemifields, there are three possible patterns of performance that may be found (Figure 29). If the pinnae provide a selective benefit for sounds originating in the front hemifield and have no effect on sounds originating in the rear hemifield, then MAAs in the *tube bypass* condition should be worse than the *normal hearing* condition in the front hemifield but no different in the rear hemifield (A). If the pinnae provide a selective benefit in the rear and have no function in front, then MAAs with the *tube bypass* should be as good in front as the *normal hearing* condition and better than *normal hearing* in the rear hemifield (B). Finally, if the pinnae provide both a benefit for sounds in front and a detriment for sounds in the rear, then MAAs with the *tube bypass* in both hemifields should be at a level somewhere between *normal hearing* performance in the front and rear hemifields (C).

If the pinnae do provide a hemifield-selective benefit to auditory localisation, this would be a novel finding. It is well-established that the pinnae play a role in the perception of elevation (Blauert, 1997) and for the discrimination between front and rear hemifields (Shaw, 1974). However, there has been no formal assessment of its contributions to localisation in azimuth when binaural cues are available. Though many authors have alluded to a hemifield-selective role in localisation (e.g. Rayleigh, 1876, 1907; Shaw, 1974; Starch, 1905), there has been little or no evidence forthcoming.

4.4.2 **Methods.**

4.4.2.1 Participants.

Ten listeners took part in this experiment (3 females, ages 22 – 27). All self-reported normal hearing.

4.4.2.2 Stimuli.

To counteract the diffuse-field acoustic effects of listening to sounds through the tube apparatus, a filter was generated to be applied to stimuli in the *tube bypass* condition. Use of the tube apparatus creates broad spectral differences in the power spectrum, beyond just removing the directional filtering of the pinnae. For example, in Figure 26(A) there is a band of excitation between 2 kHz and 3.5 kHz that is consistent across azimuth (with the exception of ipsilateral presentation of the sound at 270°). In Figure 26(C) this band has disappeared, and a new (weaker) band of excitation can be seen between 0.7 kHz and 1.5 kHz. This is due to the resonant frequency of the ear canal and the ear canal plus tube apparatus, respectively. By inserting the tube apparatus, the ear canal is artificially lengthened and so the resonant frequency will be lowered. To counter this and similar broad spectral changes introduced by use of the tube apparatus, a simple method was used to generate a filter.

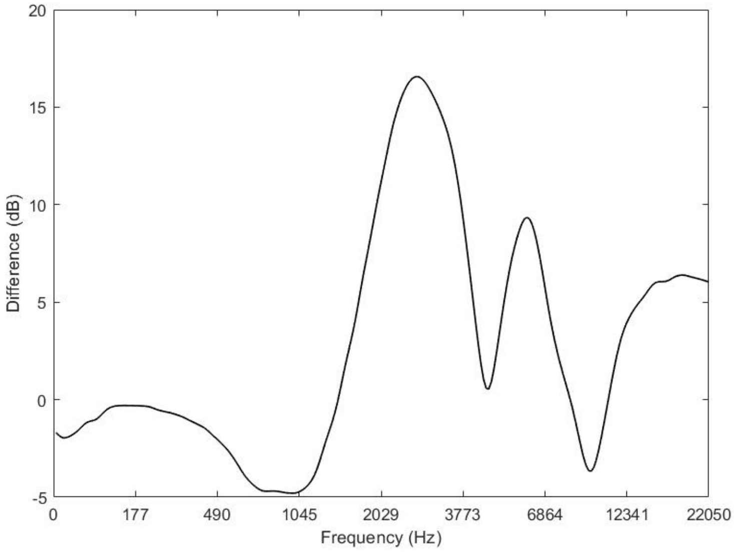

Figure 30. Difference in average sound level between the normal hearing and tube bypass conditions plotted as a function of frequency.

The average excitation value was calculated for each frequency across azimuths in the *normal hearing* and *tube bypass* conditions. The difference between these values was calculated (Figure 30), and the resulting function was applied to all stimuli presented when listeners used the tube bypass method.

Two types of noise stimulus were used in this experiment, paired with a pinna condition (*tube bypass* or *normal hearing*). Broadband noise was generated in the standard method as described in section 2.2. The filter described above was applied to broadband noise stimuli in the *tube bypass* condition to create the noise stimuli for this condition.

4.4.2.3 Procedure

Prior to starting the experiment, participants were allowed to choose from a range of earbuds to attach to the otoscope specula. Once they had made their choice, the earbud was securely attached to the tip of the specula to create the personalised tube apparatus. Listeners were instructed to insert the tube into their ears to a comfortable position. This was then visually inspected by the experimenter and, if necessary, corrections were made by the listener to orient the tubes within a reasonable margin of error from perpendicular (see section 4.3.2.4). Tubes were removed or inserted, and realigned, as necessary for each block.

Listeners were centred within the speaker array using the cord method as specified in section 2.2.2. Listeners' head position and orientation was tracked in six degrees of freedom throughout the experiment.

The MAA task was used in this experiment, with two source locations (45° and 135°) and two pinna conditions (*normal hearing* and *tube bypass*). All four conditions were presented in a random order. Pilot work found that in the *tube bypass* condition some listeners experienced front-back confusions. If listeners heard the sounds from the incorrect hemifield due to disruption of pinna cues, then a front vs back judgement would be reversed and so responses would be recorded incorrectly. Use of a left-right judgement was therefore implemented to avoid this problem.

4.4.3 **Results.**

4.4.3.1 Head position and orientation.

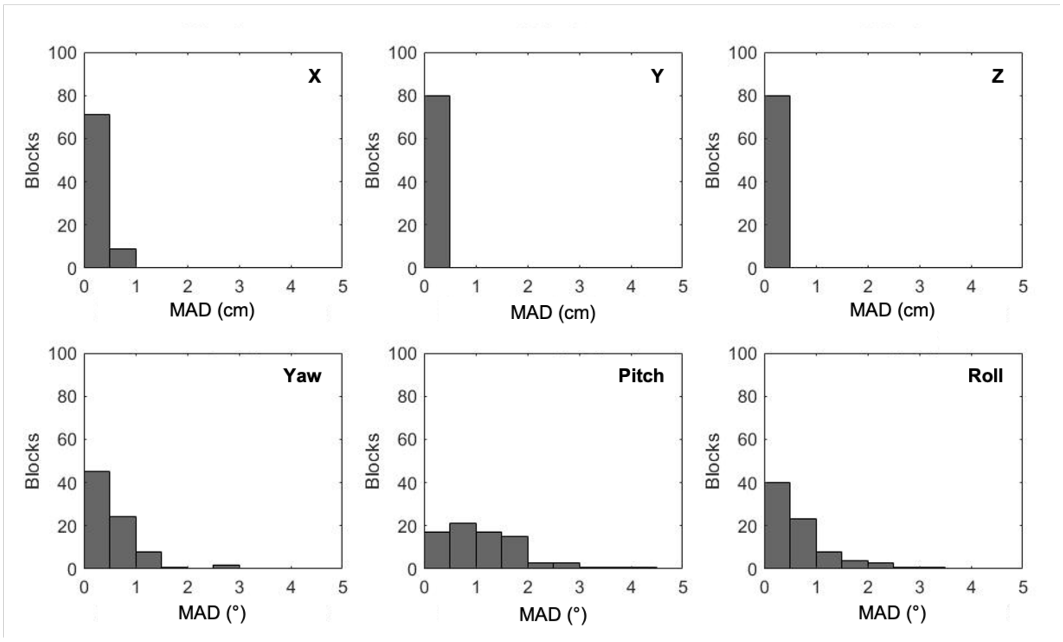

Figure 31. Histograms of median absolute deviation of head movement per block, in positional and rotation axes.

Median absolute deviation for positional axes (X, Y, Z) was less than 1 cm in all blocks (Figure 31). Deviation in yaw and roll was less than 1 degree in the majority of blocks. The median absolute deviation in head position was higher for pitch, but still below 2 degrees in most blocks. It is unclear why this is higher than in previous experiments, but as this did not qualitatively differ between conditions, and is on the vertical plane, it is unlikely to have been a confound with the MAAs in this task.

4.4.3.2 Minimum audible angles.

Anecdotal data collected during practice confirmed the existence of front-back confusions when pinna cues were removed. During two practice blocks (one per pinna condition), listeners were asked to identify the hemifield of origin of sound sources. They were not provided with feedback. Nine (90%) of the listeners misidentified the hemifield of the stimuli in the *tube bypass* condition. For five (50%) of the listeners, sounds were consistently perceived to be in the rear hemifield during this condition, regardless of true source hemifield. For the other two (20%) listeners, this misidentification was intermittent and changed on each stimulus presentation. No listeners reported a misidentification of hemifield during the *normal hearing* condition. These data confirmed the need to use left-right as opposed to front-back responses in the main experiment.

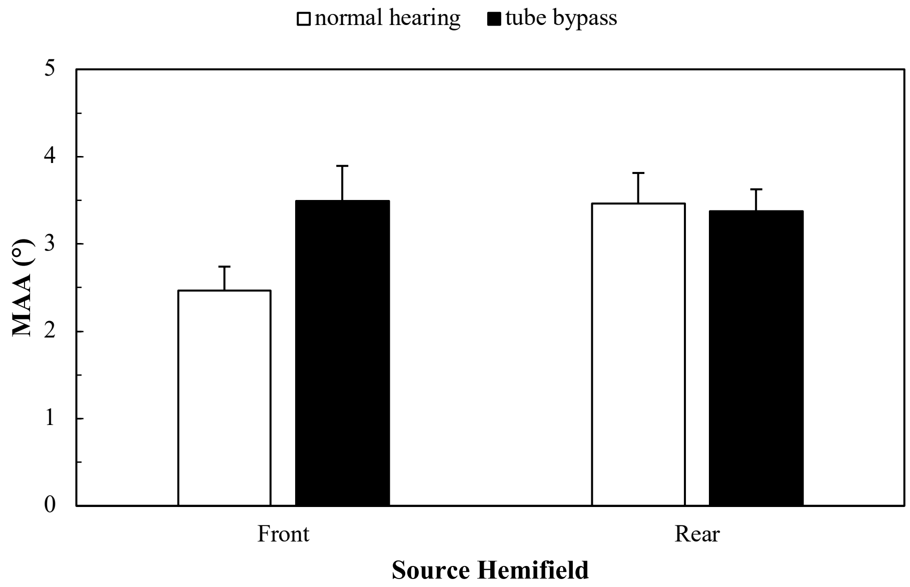

Figure 32. MAA as a function of source hemifield and pinna condition. Error bars are standard error.

Figure 32 shows the mean MAA across listeners for front and rear hemifields and the two types of filter. They closely resemble the prediction made in Figure 29(A), namely that the pinnae confer an advantage to the processing of sounds in the front hemifield. This was tested using a repeated-measures ANOVA which found a significant interaction between source hemifield and pinna condition, $F(1,9) = 6.56, p < 0.05$. There was no significant effect of pinna condition on MAA, $F(1,9) = 2.10, p = 0.181$. There was also no significant effect of source hemifield on MAA, $F(1.9) = 2.381, p = 0.157$.

Paired t-tests showed that MAAs were significantly greater in the rear hemifield than in the front hemifield in the *normal hearing* condition, $t(9) = 2.415, p < 0.05$, but were not significantly different between hemifields in the *tube bypass* condition, $t(9) = 0.397$, $p = 0.700$.

Comparing listening conditions, in the front hemifield MAAs were lower in the *normal hearing* condition than the *tube bypass* condition, $t(9) = 2.270$, $p < 0.05$, but were not significantly different between listening conditions in the rear hemifield, $t(9) = 0.283$, $p = 0.784$.

4.4.4 **Discussion.**

The pinnae were predicted to be responsible for the spatial precision asymmetry between front and rear hemifields, based on previous experiments presented in this thesis. This was found to be the case, as the precision asymmetry disappeared when listeners performed the task with tubes inserted into their ears that bypassed their pinnae. MAAs in the front hemifield were better in the *normal hearing* condition than the *tube bypass* condition. MAAs in the rear hemifield did not differ between conditions.

Three patterns of performance were proposed prior to this experiment which suggested three ways in which bypassing the pinnae would interact with the spatial precision asymmetry (Figure 29). The results found here closely resemble that of the first pattern (Figure 29(A)). The interpretation for this pattern is that the pinna provides a selective benefit for localising sounds whose origin is in the front hemifield. Under this interpretation the pinnae provide neither a benefit nor a detriment to spatial localisation in the rear hemifield. This can be seen from the data, which shows that when the pinnae are bypassed performance in the front hemifield worsens to a par with that in the rear hemifield, which does not differ depending on whether the pinnae have been bypassed or not. Therefore, the pinnae must be providing a benefit in the front hemifield that is removed in the *tube bypass* condition. Since there is no benefit or detriment in the rear hemifield to remove, performance does not change with condition for sound sources at this location.

The preliminary testing in the practice blocks of this experiment also revealed an interesting finding. Nine of the ten listeners reported misidentification of the sound source hemifield during the *tube bypass* condition. Five of the nine hemifield-misidentifying listeners reported that sounds whose true origin was in the front hemifield appeared to be in the rear hemifield. The remaining four reported an intermittent switching of hemifields. No listeners reported stimuli in the rear hemifield consistently appearing as in the front hemifield.

The perceived switching of source hemifield may be explainable when the role of the pinna is considered. Along with the perception of elevation, one of the roles of the pinna is in front-back discrimination (Shaw, 1974). When the pinnae are occluded or their cues otherwise disrupted, listeners are liable to make confusions between stimuli in the front or rear hemifields (Oldfield & Parker, 1984b). However, studies which have disrupted the function of the pinnae have not accounted for the changes in spectral content that occur because of this disruption. Oldfield & Parker (1984) used moulds in the convolutions of the outer ear, which would have changed the reflections due to the absorbance of the material. Shaw (1997) also used tubes but did not change the sound stimuli used. As noted in section 4.3.2, inserting tubes also has an effect of changing the proximal stimuli at the ear, in particular the level. In this experiment we accounted for those changes by creating a filter that was applied to stimuli in the *tube bypass* condition.

If, as appears to be the case, the pinnae provide a benefit in the front hemifield then this is likely to be in the form of an additional cue. Bypassing the pinnae and correcting for spectral changes from the tube apparatus, creates a proximal stimulus that is spectrally ‘normal’ but does not contain this cue. This is functionally the same as the proximal stimulus created by a sound in the rear hemifield. Hence, in this scenario, it would be appropriate for the auditory system to make the assumption that this stimulus must originate from the rear hemifield. This would then explain why many listeners experienced a constant misidentification of sounds as originating in the rear hemifield.

4.5 Chapter Discussion

In the first experiment in this chapter, the spectral content of stimuli was changed to examine the effect of highpass and lowpass filtering broadband noise on spatial precision, as measured using the MAA. It was shown that the front/rear asymmetry found in previous chapters remained for highpass filtered noise stimuli but using lowpass filtered stimuli removed the asymmetry. This was taken as evidence that the asymmetry is caused by some aspect of high frequency information.

The second experiment used an acoustic mannikin to model the effects of removing the pinnae, which was hypothesised to be a source of high frequency, asymmetry-causing, cues. It was shown that removing the pinnae entirely removed broad asymmetries measured using head-related impulse responses. However, this would not be practical to examine in vivo, so a method of bypassing the pinnae was developed instead. This involved using tubes to bypass the outer ear and funnel sounds directly into the ear canals. A filter was also used to negate the spectral effects of listening to sounds through tubes.

In the third experiment the tube bypass method was tested using human listeners and compared to normal hearing conditions. It was found that bypassing the pinnae, and so removing pinnae cues, removed the front/rear asymmetry. Additionally, it was shown that the pinnae provide a selective benefit to sound localisation in the front hemifield and do not affect sound localisation in the rear hemifield.

The results found in the first experiment in this chapter can now be explained in the light of the findings from the third experiment. The lowpass filtered stimuli showed no asymmetry because the pinnae only interact with frequencies above approximately 4 kHz (Butler & Humanski, 1992), ergo there would be no benefit provided in the front hemifield. The highpass filtered and broadband noise stimuli contain frequencies above this threshold and so the pinnae do provide a benefit in the front hemifield. Broadband stimuli are then generally better than high-pass stimuli solely because in this condition listeners have access to all frequencies and so there are more available cues to location.

The most important finding from this chapter is that in addition to discrimination of elevation and front-back discrimination, the pinnae also appear to provide a benefit for sound localisation in the horizontal plane in the front hemifield. This is a novel finding and provides new insight for models of human hearing.

The following chapter examines how auditory precision changes when listeners make judgements about moving stimuli while their own heads are moving. This is a more complex task that requires the auditory system to integrate information from multiple sensory modalities to make an accurate percept of auditory motion.

Chapter 5 – Head Compensation for Auditory Motion

5.1 Chapter Overview

In the previous chapters it was established that there is an asymmetry in the spatial precision of the hearing system between front and rear hemifields. In Chapter 4 the location-dependent filtering of the pinnae was indicated as a cause of the spatial precision asymmetry. The asymmetry was assessed using experiments in which the listeners kept their heads stationary while making judgements about the stimuli. In normal life this is rarely our experience. Sound sources are often moving, and we ourselves move within space. Even when we attempt to remain stationary, there are small movements that we make in order to preserve visual fixation or balance (Kunin et al., 2007). It is important therefore to know how the low-level findings of previous chapters, with static sources and a stationary head, apply to more 'realistic' tasks involving movement of the listeners. In this chapter, the ability of the auditory system to compensate for the effect of head rotation on the auditory spatial field is examined, using sound sources placed at different azimuths.

5.2 Introduction

When we move our heads, our auditory field changes accordingly. If, for example, the head rotates 30° to the right about the vertical plane then an external static sound source will be displaced 30° to the left in the auditory field. In order to perceive that source as stationary with respect to the world, the auditory system must be able to compensate for the effect of the movement of the head. The same principle applies to moving sound sources, as their motion in the auditory field may be accelerated, decelerated, their trajectory altered, or some combination of these effects when the head moves. Hence, the auditory system needs to be able to not just compensate for the movement of static sources in the auditory field, but also to compensate for velocity and trajectory changes of moving sources too.

The ability of the auditory system to compensate for head movement, specifically head rotation, has been examined by Freeman et al. (2017). They found that the auditory system was able to perform some compensation for head rotation, but that this compensation was incomplete: a static sound source was perceived as ‘moving’ relative to the world during a rotation of the listener’s head. The direction of perceived movement was opposite to the performed head rotation. This, Freeman et al. inferred, was due to a mismatch in the signals from the auditory system about the speed of the sound source in the auditory field, and signals from the proprioceptive and vestibular systems about the speed of movement of the head (Figure 33). They termed these ‘cochlear’ and ‘extra-cochlear’ signals, analogous to ‘retinal’ and ‘extra-retinal’ signals thought to be used by the visual system to compensate for eye movements (Freeman et al., 2010; Von Holst, 1954). They noted that the incomplete compensation for head movement by the auditory system is similar to that found for smooth pursuit eye movements in vision (Filehne, 1922; Freeman & Banks, 1998; Furman & Gur, 2012).

Freeman et al. reasoned that for the static sound sources to be perceived as moving in the same direction as head rotation, the cochlear speed estimate must be greater than the extra-cochlear speed estimate. This is illustrated in Figure 33. Their results showed that sound sources needed to be moved in the same direction as listener’s head rotations, at approximately 15% of their head speed, in order for the sources to be perceived as static. This was the case regardless of the speed that listeners moved their heads; listeners were trained to rotate their heads at 20°/s, 60°/s or 100°/s. The speed of the sound source required to cancel out its perceived motion was found to be a constant percentage of the head speed.

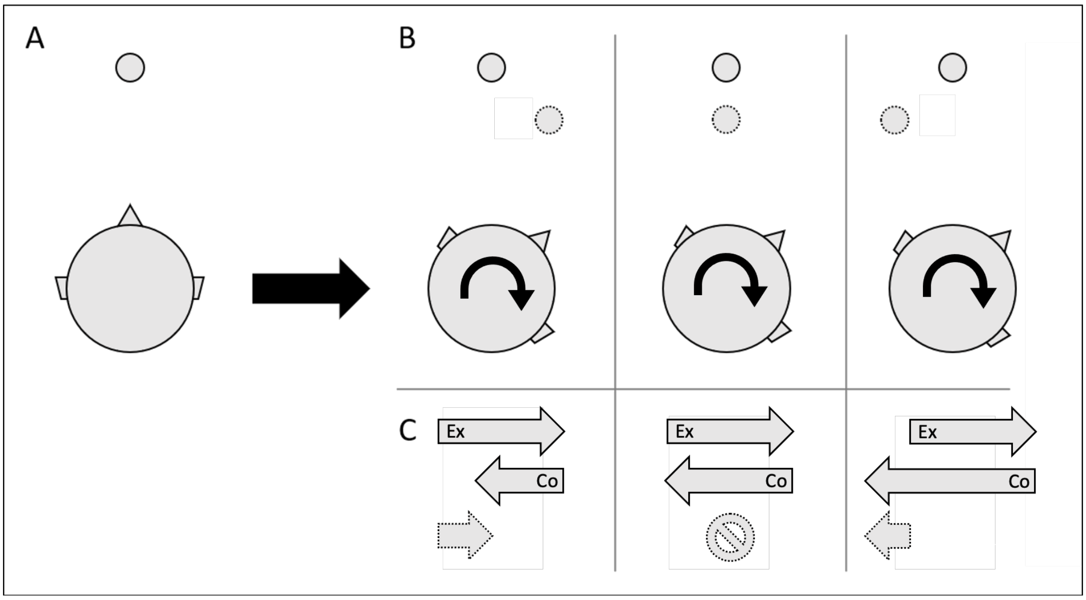

Figure 33. (A) Listener with static source, prior to head rotation. (B) Three potential outcomes of compensation, with actual source azimuth (solid outline) and perceived source azimuth (dotted outline): left shows incomplete compensation with perceived source motion in the same direction as head rotation, middle shows complete compensation with source perceived as world-stationary, right shows incomplete compensation with source motion perceived in opposite direction to head rotation. (C) Relative motion estimates (dotted arrow) from cochlear [Co] and extra-cochlear [Ex] signals (solid arrows).

Two models which have been used to account for the incomplete compensation of the auditory system for head movement are the Bayesian model and the topological account. The Bayesian model proposes that the estimate of speed produced by a sensory signal is dependent on two factors, the precision of that signal and the prior belief about sound source movements. For movement the prior belief is usually stated to be that objects usually do not move, or move very slowly (Senna et al., 2015). These are referred to as the no-motion or slow-motion prior, respectively.

Under the Bayesian model, when a sensory signal (likelihood) is less precise the perceptual estimate (posterior) that is produced will be biased more towards the prior belief (Figure 34B). Conversely if the sensory signal is high in precision then the posterior will be biased less (Figure 33A). Spatial precision has been shown to be better at the midline than at the interaural axis, and so estimates of speed should also be higher for sources near the midline than for sources near the interaural axis. Consequently, compensation should worsen as sources are moved away from the midline.

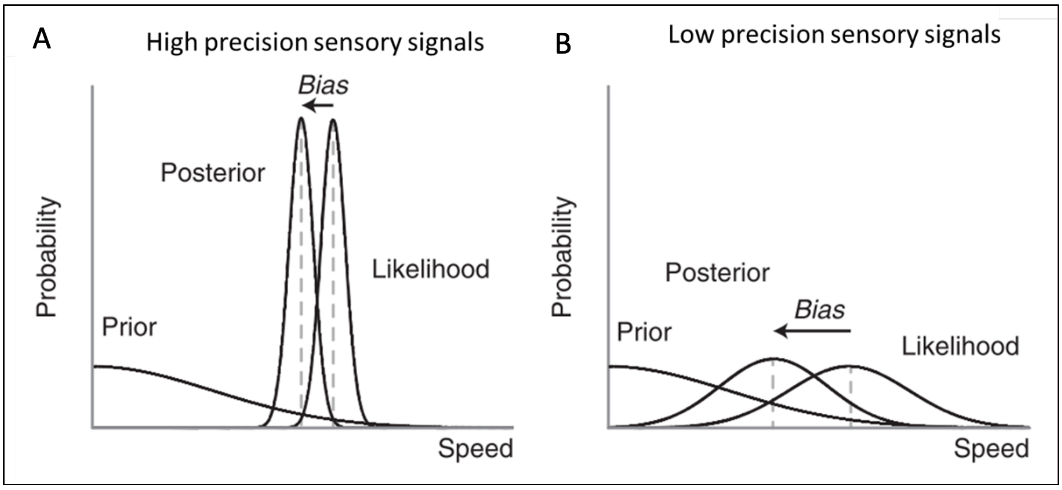

Figure 34. Comparison of high precision and low precision sensory signals, under the Bayesian model. Adapted from Senna et al. (2015).

The topological account (Brimijoin, 2018) proposes that auditory compensation is incomplete because of a perceptual expansion and compression of auditory space. According to this account, the dilation of space occurs because the spatial resolution changes as a function of azimuth. Spatial resolution is greatest where MAA and MAMAs are smallest, which occurs at the midline. Perceptual space at the midline is therefore expanded, while at the interaural axis it is compressed because MAAs and MAMAs are so much larger (see Figure 35). This is, in effect, a type of Fechnerian scaling (Fechner, 1851, 1987), although Brimijoin does not frame his model in these terms.

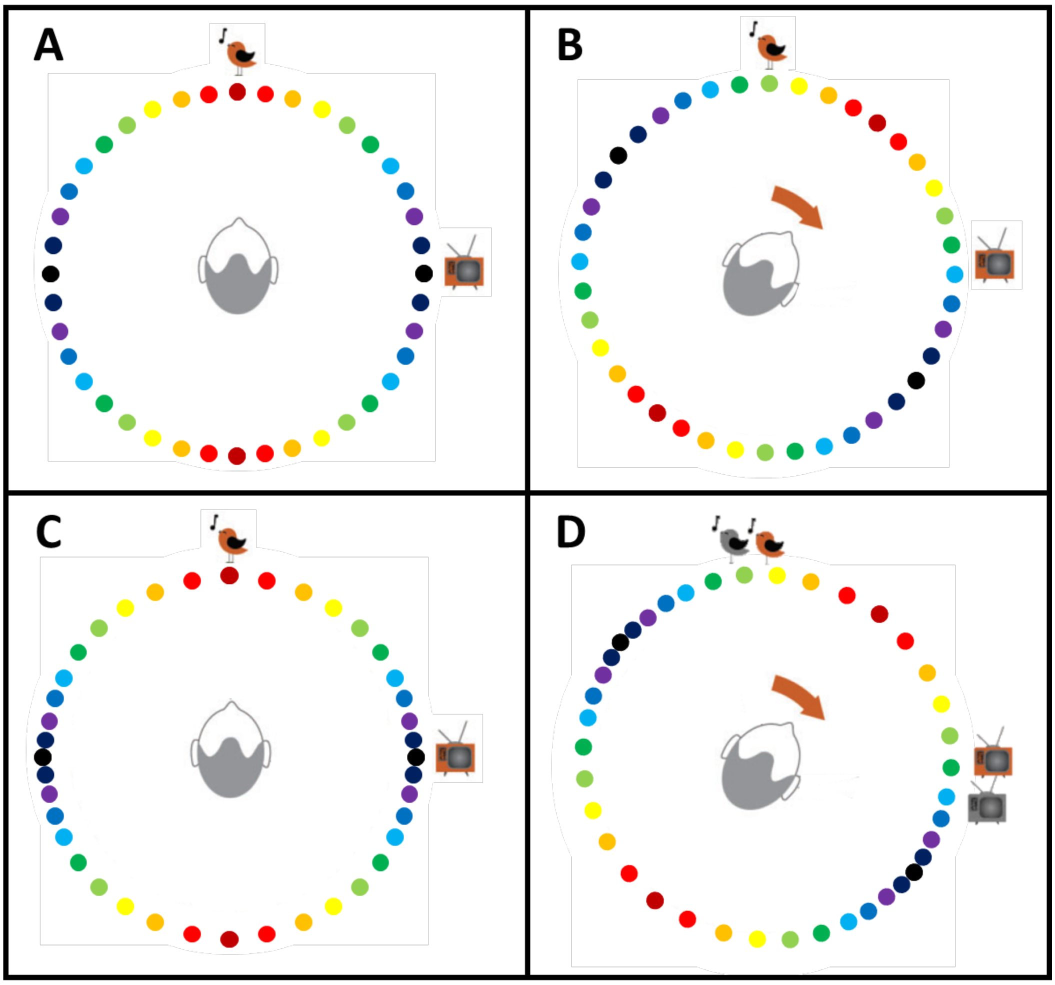

Figure 35. Illustration of the expansion and compression of perceptual space in the topological account. Left panels show a listener prior to head rotation. Right panels show the listener after a head rotation to the right. (A) dots represent the real locations of 40 sources, also pictured are two world stationary stimuli. (B) dots represent the real locations of the sources if they move perfectly with the listener, and the two world stationary stimuli. (C) dots represent the perceived locations of the sources, and the two world stationary stimuli. (D) dots represent the perceived location of the sources if they move perfectly with the listener, the perceived locations of the world stationary stimuli (grey) and their real locations (colour). Adapted from Brimijoin (2018).

When a source moves relative to the listener, whether through its own motion or the listener moving, it moves through different regions of expansion and compression in perceptual space. This leads to perceived acceleration and deceleration, which causes the perception that world stationary objects move during head rotation. This is analogous to an object appearing to accelerate and decelerate as it is moved linearly under a magnifying lens. An object that moves away from the midline (relative to the listener) should appear to move in opposition to the listener as it moves from a region of expansion to one of compression.

Conversely an object that moves away from the interaural axis should appear to move in the same direction as the listener as it moves into a region of space that is expanded (Figure 35(D)).

Though the mechanisms through which compensation for head movement changes with azimuth are different to the Bayesian model, the topological account is also defined by spatial precision, which form ‘perceptual units’ (Fechner, 1851, 1987). Both models, therefore, provide explanations linking compensation to spatial precision. So, by measuring the ability of listeners to compensate for their head movements, it should be possible to estimate the spatial precision of the cochlear signals.

In this experiment, the head compensation task used by Freeman et al. will be performed over a wider range of source azimuths, including positions in the rear hemifield, because azimuth has already been demonstrated to change auditory spatial precision by large amounts in earlier chapters. This will permit an examination of how auditory spatial precision changes in a more complex task, and whether the patterns of performance from lower-level tasks are replicated. It will also allow for the examination of how head compensation ability is affected by spatial precision changes with azimuth.

This task uses a Method of Constant Stimuli (MoCS), unlike the other experiments described in previous chapters. The resulting psychometric function produces two measures, the point of subjective equality (PSE) and the slope. The PSE represents the point at which complete compensation occurs, i.e., the cochlear and extra-cochlear signals are equal and so the source appears to be stationary. If the PSE for a source azimuth occurs when the sound source is stationary, then the cochlear and extra-cochlear speed estimates are equivalent – that is, they both produce the same estimate of speed for auditory field and head motion respectively. If the PSE for an azimuth occurs when the sound source is not stationary then the cochlear and extra-cochlear estimates are not equivalent; they produce different speed estimates. In Freeman et al. (2017) the sound source needed to be moved slightly with the head for the percept of a stationary source to occur. The PSE was a small positive value, indicating that cochlear speed estimates were greater than extra-cochlear speed estimates.

The slope of the psychometric function is a measure of the precision of the listeners' ability to compensate for their head rotation. This is the key measure for this chapter as the slope is determined by the precision of both the cochlear and extra-cochlear signals. Assuming that the precision of the extra-cochlear signals from head rotation is equivalent across all conditions, changes in the slope of the psychometric function are attributable to changes in the precision of the cochlear signals only. From the findings in previous chapters, auditory (cochlear) precision worsens towards the interaural axis. This should be reflected in the slope becoming shallower (worse overall precision) as the source azimuth is moved towards the interaural axis. Additionally, the slope should be shallower in the rear hemifield than in the front hemifield, showing the front-rear hemifield asymmetry found previously.

5.3 Methods

The data in this experiment were collected by an undergraduate student in the School of Psychology, Mark Smith, under the supervision of the author.

5.3.1 **Participants.**

Fifteen listeners participated in this experiment (7 females, ages 20 – 26). All self-reported normal hearing.

5.3.2 **Stimuli.**

Eight source azimuths were used in this experiment, from 0° to 315° in increments of 45°. All source azimuths were presented in individual blocks, presented in a random order. There were two repetitions of these random series.

Stimuli were yoked to the ongoing head movement by multiplying the source motion by a given percentage of the listeners' head rotation speed. These are defined as 'gains', though they are signed as explained below. The gain values therefore controlled the correspondence between the movement of the listeners' head and the movement of the sound source. A positive gain meant that the source moved in the same direction as the listener's head rotation, a negative gain meant that it moved in the opposite direction to the listener's head rotation, and a gain of zero meant that the sound did not move at all.

Sound sources were made to move by sampling the orientation of the listener's head in yaw (the horizontal plane) using a head tracker (Polhemus Liberty) with a sampling rate of 240 Hz. The orientation of the listener's head was used to move the centroid of the stimulus by an angle, $\Delta\theta$, calculated using the difference in head orientation between the current update, θ_n, and the previous update, θ_{n-1}, multiplied by the gain, β, for that trial, as shown in (8). The resultant change in azimuth per sample interval is the speed of the stimulus.

$$\Delta\theta = \beta(\theta_n - \theta_{n-1}) \tag{8}$$

The stimulus was only presented when listeners' head movement speed was above a velocity threshold. By preventing the stimulus from being present when the listener's head is stationary, the listener is prevented from using localisations of a static stimulus in order to make a judgement about the movement of the stimulus. For example, if a listener had rotated their head 30° anti-clockwise and then judged a static stimulus to be originating from 20° to the listener's right (clockwise) then the stimulus must have moved by a small amount in the same direction as the listener's head rotation. The minimum-speed requirement meant that this strategy was unavailable to the listeners.

A side effect of using a minimum-speed threshold combined with real-time differentiation is that the stimulus can appear 'crackly' when listeners are moving their heads near the threshold. This is due to noise from the head tracker itself. This noise may act to decrease or increase the calculated speed for a given interval, creating sharp differences in speed that result in a ‘crackly’ sound. This problem was solved using a ‘leaky integrator’. This is a low-pass filter that fully weights current input and gradually ‘leaks’ past input over time. It is simply implemented by adding the current head-speed value at each update cycle to a running total that has been multiplied by a decay factor (0.5). A similar model is used in neuroscience to model the spiking rate of neuronal populations (Eliasmith & Anderson, 2004). This filtering had the effect of ‘smoothing’ the sound near to the head speed threshold, while rapidly fading out the stimulus if the head movement falls below threshold in sufficient time for the perception of a static stimulus to occur.

For each block, the origin for each sound was placed close to the source azimuth. This origin was varied by ± 7.5° (one speaker separation angle) randomly on each trial to reduce the predictability of the sound location. Presentation only began when the listener’s head was oriented towards the 0° location on the speaker array. Hence stimuli were first played at the origin azimuth for that trial, and then moved in accordance with the listener’s head rotation for the rest of the trial.

As listeners were free to move their heads (albeit subject to correction by the experimenter, see below), there was no constraint on the maximum trajectory over which the stimulus could travel. Pilot work found however, that the majority of listeners did not move their heads more than 60° from the origin in either direction. This is only slightly less than the maximum range of rotation that the head can perform in the horizontal plane, 80° (Swinkels & Swinkels-Meewisse, 2014), which suggests that listeners will move their heads within a large extent of the maximum range when unconstrained or un-instructed.

The gain values controlling the source motion were determined by a method of constant stimuli (MoCS) as opposed to a staircase. This method had the advantage that the range of gain values can be constrained in advance. This helped prevent the stimuli from moving over too large a trajectory and thereby invalidating the source azimuth measurement if stimuli overlapped too far into adjacent regions of auditory space. Based on Freeman et al. and pilot work, the centre of the psychometric function (in other words the PSE) was near to 0, where a gain of 0 represents perfect compensation and deviation from this value indicates incomplete compensation. Linear steps were used about 0 to allow the gain values to take on positive and negative values. In pilot work it was found that the granularity needed in the set of gain values used (step sizes) varied depending on the source azimuth. To a large extent this anticipates the prediction made earlier that head movement compensation is also governed by the spatial precision of the auditory field that varies with azimuth. In order to keep the number of gain values at seven while fully capturing the psychometric function at each azimuth, the total range of gain values presented was made dependent on the location of the source azimuth. For stimuli presented on the midline (0°, 180°) gain values were between -0.45 and 0.45 in steps of 0.15. Stimuli on the interaural axis (90°, 270°) were presented using gains between -0.9 and 0.9 in steps of 0.3. All remaining stimuli were presented using gains between -0.625 and 0.625 in steps of 0.225. All gain values were presented ten times in a random order for each source azimuth.

With the largest gain, ± 0.9, and the maximum normal head rotation angle, 60°, the largest displacement angle that a stimulus could cover from its source azimuth origin would be 54°. As the centre of the azimuthal regions being tested were 45° apart, only the largest gain values and the largest head rotations would cause a salient overlap. But at this gain value listeners' performance should be asymptotic (i.e., at 100% accuracy), so this was not anticipated to cause problems for the interpretation of the data.

5.3.3 **Procedure.**

Due to the necessity of listeners moving their heads freely during this task, they could not remain perfectly centred throughout the experiment. Prior to starting, however, their heads were centred within the ring using the cord method described in section 2.2.2. This ensured that they remained close to the centre.

During the experiment, the position of the listeners' head was monitored using an onscreen visual available only to the experimenter. This provided a graphic that depicted the speaker array and the head tracker from a top-down perspective. Concentric rings were overlaid that depicted circles of 2 cm, 5 cm, 10 cm, and 15 cm radii from the centre of the ring. The location and yaw angle of the head tracker could be seen on this graphic, which allowed the experimenter to see if the listener moved outside of a reasonable area around the centre. If this occurred, the listener was verbally corrected to move their head back to the centre area before continuing with the task.

Listeners' task on each trial was to determine whether the sound moved 'with' or 'against' their head movement. This response was made on a handheld controller, using two shoulder buttons (NB. this is the same controller described in 2.2, but different buttons were used). The duration of the trial was not constrained, and listeners could make their response at any time once the stimulus had begun. Once the listener had made a response, the stimulus ceased. The next trial then began once the listener's head was oriented towards the speaker directly in front. This was determined using the head tracker, with a requirement of 10 consecutive samples of azimuth angle recorded within ± 7.5° of directly forwards.

5.4 Results

A psychometric function was fit to the data for each block, an example of which is given in Figure 36. The fitting procedure used maximum likelihood estimation to fit a cumulative Gaussian function to the probability of the listener selecting 'with the head' across the gain values investigated.

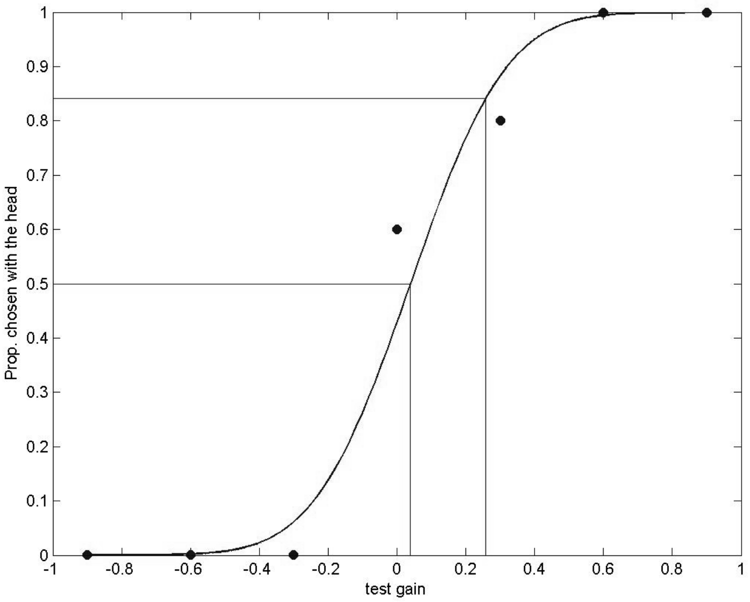

Figure 36. An example psychometric function. The left vertical line indicates the PSE, the right vertical line indicates 84% performance, equivalent to 1 SD from the mean. Gain values were presented ten times each. Proportion chosen could therefore vary in steps of 0.1.

From the psychometric functions two values were calculated for each listener per block, the PSE and the slope. The PSE was taken from the 50% point on the psychometric function. The slope was calculated as the difference between the 50% and 84% probabilities, which is equivalent to the SD of the cumulative Gaussian function.

One listener did not complete all blocks, so their data were removed from the dataset (N = 14). Two listeners were removed due to having at least one PSE or slope outside the cut offs, defined using the modified Z score method (N = 12) (Iglewicz & Hoaglin, 1993; Jones, 2016).

5.4.1 **Compensation.**

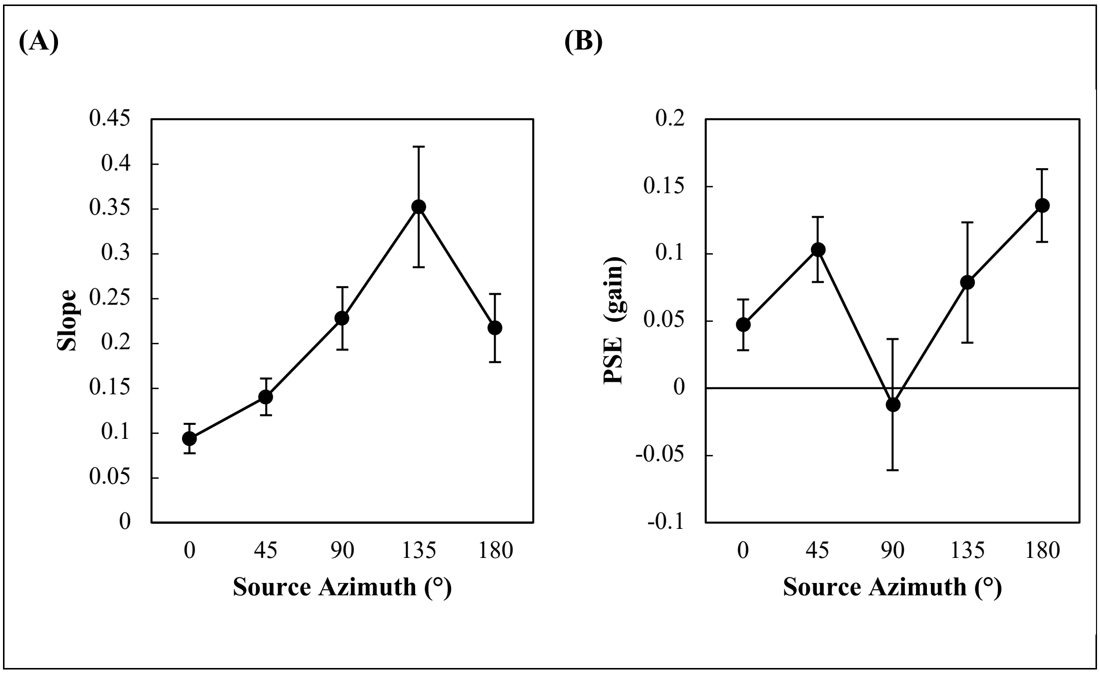

Figure 37. Average slopes (A) and PSEs (B) as a function of source azimuth (collapsed across left/right). Error bars are ± standard error.

Data were not found to be significantly different between left (315°, 270°, 225°) and right (45°, 90°, 315°) positions for psychometric function slope, $F(1,11) = 0.683$, $p = 0.426$, or for PSE, $F(1,11) = 4.564$, $p = 0.056$. Consequently, data were averaged across symmetric left/right positions. Data are presented here using azimuth values on the right side (0° - 180°) for consistency with other experiments presented in this thesis. Average slopes and PSEs for each source azimuth (collapsed across left/right positions) are shown in Figure 37.

One-way repeated-measures ANOVAs were used to examine the effect of source azimuth on the slopes and PSEs. There was a significant effect of source azimuth on slope, $F(4,44) = 11.484$, $p < 0.001$. Paired t-tests found that the slope was greater at 180° than at 0°, $t(11) = 3.950$, $p < 0.01$, and at 135° than 45°, $t(11) = 3.843$, $p < 0.01$. This indicates greater variability in the sensory signals for source azimuths in the rear hemifield than in the front hemifield.

There was also a significant effect of source azimuth on PSE, *F*(4,44) = 4.300, $p <$ 0.01. Paired t-tests found that PSEs were greater at 180° than 0°, *t*(11) = 2.765, $p < 0.05$, but were not different between 45° and 135, *t*(11) = 0.620, $p = 0.548$. Using a one-sample t-test, the PSE at 90° was not significantly different from zero, *t*(11) = 0.251, $p = 0.806$.

5.4.2 **Head movement analysis.**

As discussed above, the slope of the psychometric function represents variability from all sensory signals involved in this task. If the variability from the extra-cochlear signals is established to be equivalent for all source azimuths, then any changes in the slope must be due to changes in the precision of the cochlear signals. To test this, it is necessary to examine the way in which listeners moved their heads in this task, as differences in head movements between source azimuths may indicate changes in extra-cochlear signal precision.

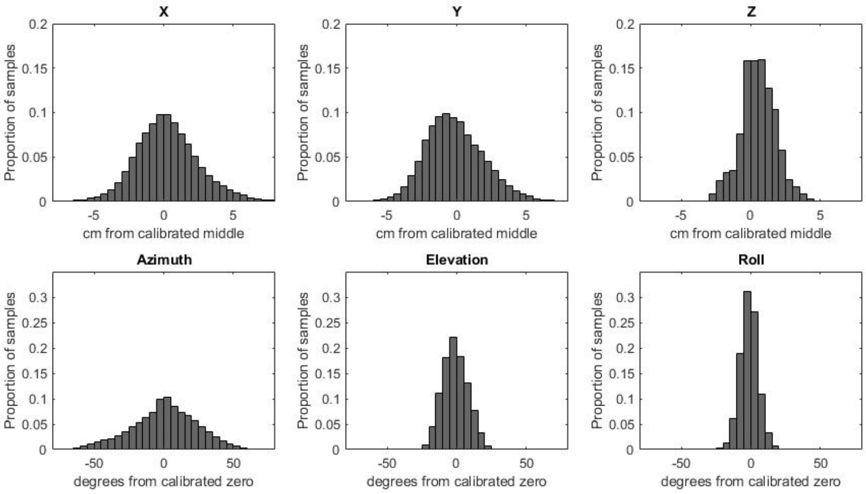

Figure 38. Histograms of head position and orientation for all samples, for the six recorded axes.

Firstly, the distribution of head positions and orientations across all listeners and trials was examined. Figure 38 plots histograms of all head tracker samples collected, for head positions in X, Y, and Z, and head orientations in azimuth, elevation, and roll. The distributions are approximately normal in all cases, with the average head position being at the calibrated centre of the ring, oriented towards 0° azimuth. The most relevant figure is the bottom left histogram, which shows the azimuth of listeners' head orientations.

The maximum extent of listeners' azimuthal head movements was approximately ± 50° from directly forward. This is well within the normal range of cervical rotation for this age group, 80° (Swinkels & Swinkels-Meewisse, 2014). It can also be seen from the histograms for elevation and roll that there was some rotation around these axes during movement, which is consistent with the findings of Kunin et al. (2007) who showed that rotation around the yaw (azimuth) axis also involves both pitch (elevation) and roll rotations.

The second analysis looked at the speed with which listeners moved their heads. Speed was calculated as the difference between sequential samples from the head tracker in the azimuth axis. The distribution of speeds found was positively skewed (Figure 39(A)). Hence, a non-parametric measure of central tendency was required to properly assess any average differences between conditions. Median head speeds were consequently calculated for each trial and means of these were taken for each listener and source azimuth.

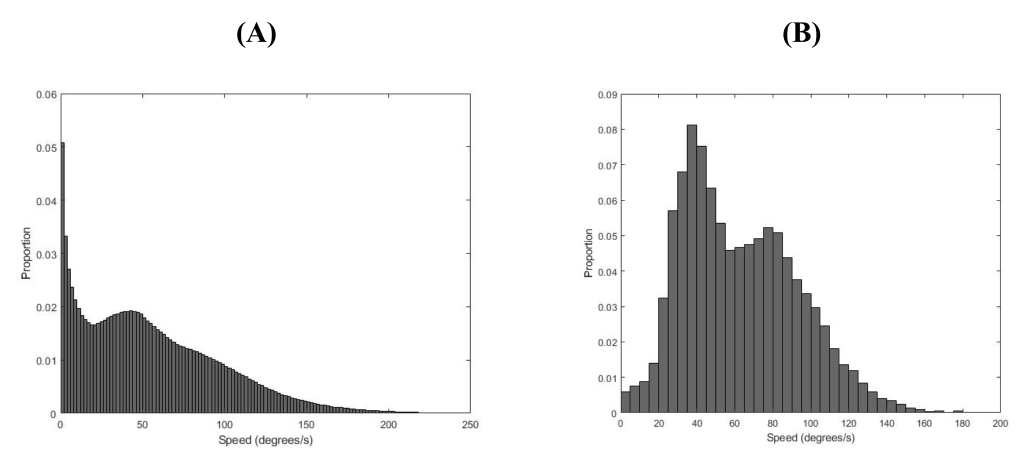

Figure 39. Distributions of listener head movement speeds plotted (A) for all sample intervals and (B) as median speed on each trial.

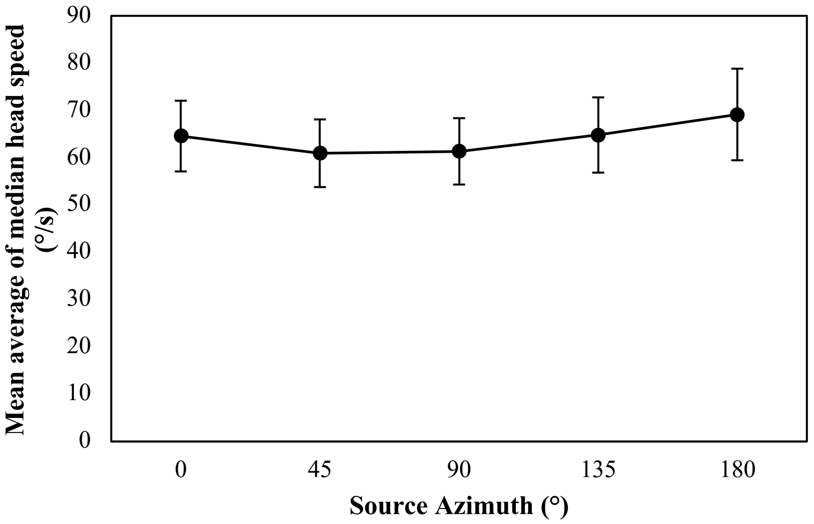

Figure 40. Mean average of median head speeds as a function of source azimuth. Error bars are ± standard error.

The mean of median head speeds is shown for each source azimuth (collapsed over left/right positions) in Figure 40. A repeated-measures ANOVA found that there was a significant difference in head speed between source azimuths, $F(4,44) = 3.424$, $p < 0.05$. However, given the small magnitude of difference in head speeds between conditions and the large overlap in the error bars of head movements this was not regarded to be a meaningful difference.

The third analysis was to examine the shape of the trajectories of head movement in azimuth of listeners. Examination of trajectories for each stimulus condition showed that behaviour could be grouped into three types, examples of which are given in Figure 41:

A. Movement in left and right fields, with no starting direction preference
B. Movement in left and right fields, with a starting direction preference
C. Movement in one field only

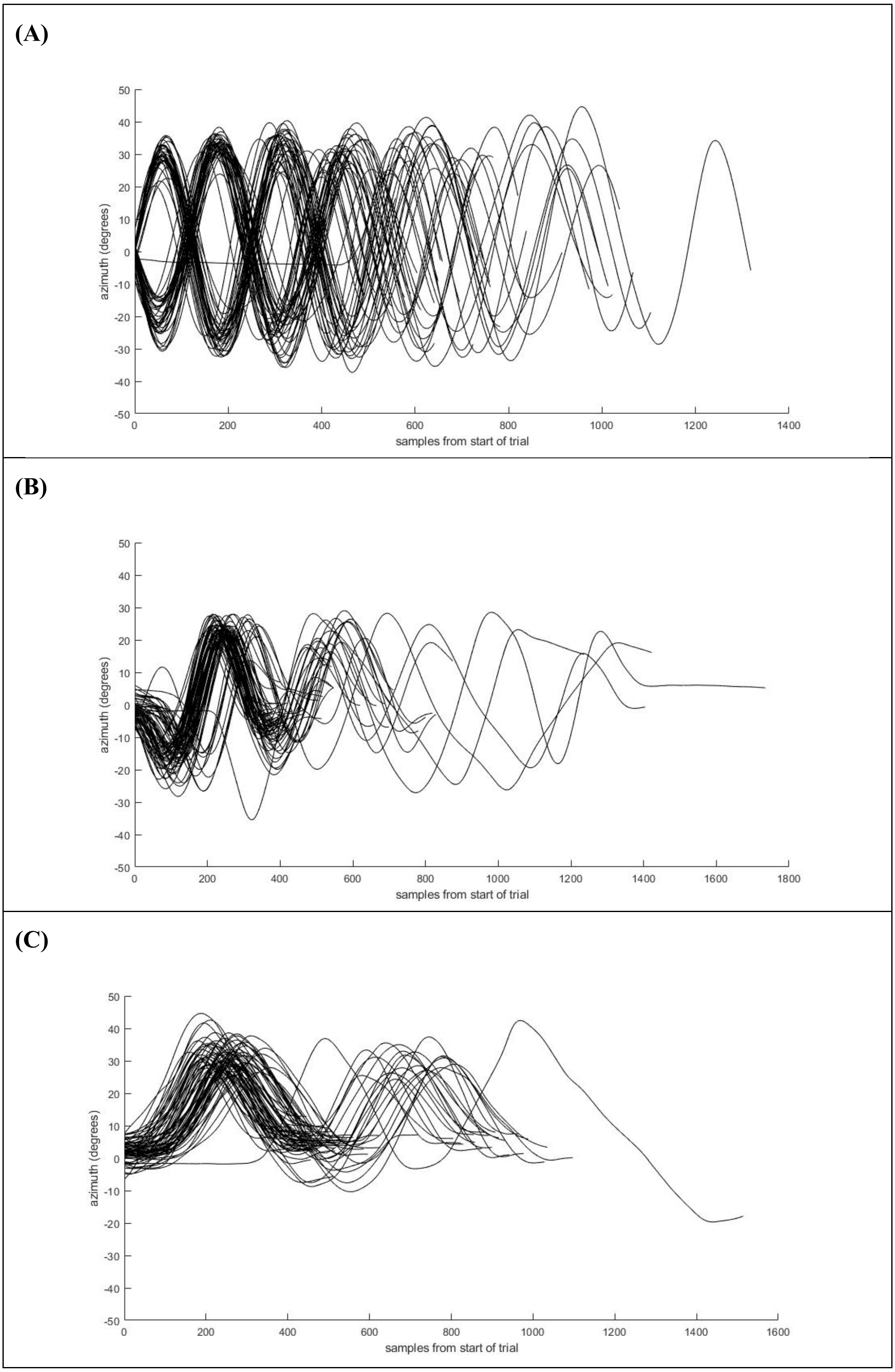

Figure 41. Examples of trajectory types. Each figure depicts all trials in a single block, each line is a separate trial. (A) is a listener showing movement in left and right fields, with no starting direction preference, (B) is a listener showing movement in left and right fields, with a starting direction preference, and (C) is a listener showing movement in one field only.

When these trajectories were broken down by listener, two listeners showed predominantly type C, and the remaining ten showed predominantly types A and B. Of these ten, two strictly adhered to type B, one to type A, and the rest mixed types A and B. No differences in task performance were seen between listeners based on their head movement trajectory type. There was also no correspondence seen between sound source azimuth and choice of trajectory type.

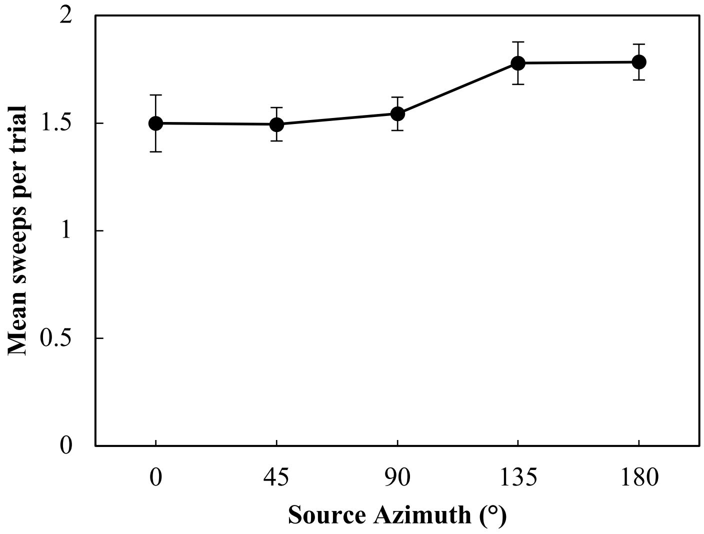

Figure 42. Mean number of sweeps made on a trial, per source azimuth. Error bars are ± standard error.

The number of sweeps made in these trajectories was examined for each source azimuth (Figure 42). Sweeps were calculated using the number of peaks (positive) and troughs (negative) in the yaw (azimuth) axis on a trial. These were defined as a peak in absolute excursion from 0°, calculated using the `findpeaks` functionality of MATLAB (MathWorks). To avoid incorrectly categorising small movements as peaks, a minimum peak height of 10° and a minimum distance between peaks of 100ms (24 samples) were used. The number of peaks was calculated for each trial, and this number was halved to approximate the number of full sweeps (left *and* right excursion). Number of sweeps was then averaged over all trials for each listener in each source azimuth condition. A repeated-measures ANOVA showed that there was a significant difference in number of sweeps made between source azimuths, $F(4,44) = 10.403, p < 0.001$. Paired t-tests found that listeners made significantly more sweeps for source azimuths in the rear than in the front, for 0° *vs.* 180° ($t(11) = 3.878, p < 0.001$) and 45° *vs.* 135° ($t(11) = 4.979, p < 0.01$).

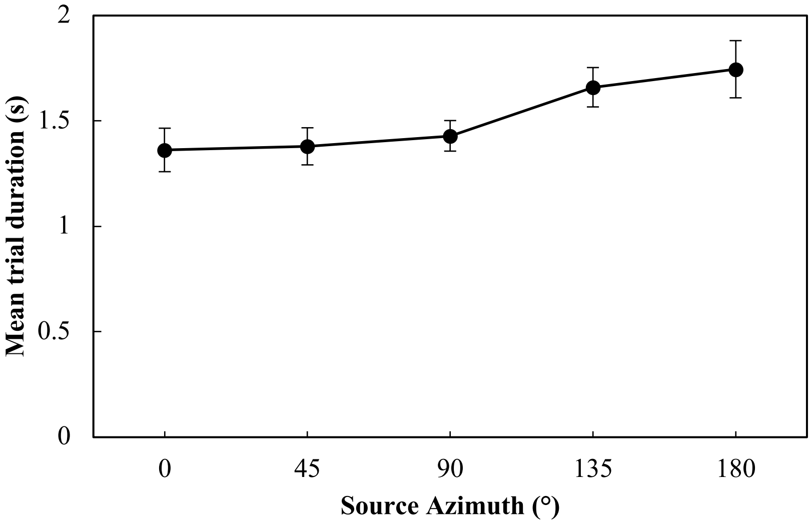

Figure 43. Mean trial duration per source azimuth. Error bars are ± standard error.

The average trial duration was also calculated for each source azimuth, collapsed across left and right positions (Figure 43). A repeated-measures ANOVA found that the trial duration differed between source azimuths, $F(4,44) = 10.403, p < 0.001$. Paired t-tests found that trial durations were significantly longer for source azimuths in the rear than in the front, for 0° *vs.* 180° ($t(11) = 6.408, p < 0.001$) and 45° *vs.* 135° ($t(11) = 4.197, p < 0.01$).

The results of this final head movement analysis therefore question that tacit assumption that the head movements were the same at all azimuths tested. There is a clear difference in the number of sweeps per trial and mean trial duration between source azimuths, which means that we cannot conclude that the variability in the extra-cochlear sensory signals was equivalent in all conditions. For this particular experimental design, changes in the slope of the psychometric functions appear not to be driven solely due by changes in the spatial precision of the cochlear signals.

5.5 Discussion

A head compensation task was used to examine how auditory spatial precision changes when additional sensory signals from extra-cochlear sources are included in the perception of auditory motion. In Freeman et al. (2017) it was found that compensation was incomplete for sources at 0°, 45°, and 90°, with gain value of approximately 0.15 required for the perception that a source was stationary relative to the world (PSE). The PSEs found in this experiment do not conform to the same pattern and show differences between source azimuths, because the PSEs are somewhat smaller and fall to zero at 90°. The slope of the psychometric functions also differs between source azimuths, but this does conform to Freeman et al.'s findings in the front hemifield.

Listeners showed complete compensation at 90°, indicated by the PSE not being different from zero. At all other source azimuths, the PSE was a non-zero value, indicating incomplete compensation. They were also positive, which indicates that a world stationary source would be perceived as moving in opposition to the movement of the listener's head. Compensation was slightly better at 0° than at other non-zero positions, and was smaller than that found by Freeman et al., with the PSE at around 0.05. Given that this experiment was conducted very similarly to that of Freeman et al., it is unclear why these differences have occurred in the pattern of compensation.

The difference in PSE between 0° and 90°, and 90° and 180° follows the predictions made by both the Bayesian model and the topological account, that compensation should be best at the interaural axis and worse further from the midline. However, the PSE at 45° does not conform to this prediction as compensation was worse at 45° than at 0°. It is also unclear why this might have occurred, and why this is only the case in the front hemifield.

Alternatively, 0° may be the aberrant condition as this shows an improvement over 45° that is not seen for 135° versus 180°. This also raises questions about the use of this paradigm to assess compensation for head movement, as pointed out above.

The slope of the psychometric functions indicates the overall precision of the listeners for each source azimuth. The variability that this slope represents is produced by the variability of the cochlear and extra-cochlear signals. Listeners' head movements did not substantially differ in their speed, total range, or typical trajectory between source azimuths. They did, however, show a difference in the average number of sweeps made on each trial and the average trial length, with both of these being greater for rear hemifield positions compared to positions in the front hemifield. As the contribution of the extra-cochlear signals (proprioceptive, somatosensory, vestibular) was not equivalent across conditions, it cannot be asserted that the precision differences seen between conditions are solely the result of changes in the precision of the cochlear signals.

Completing more sweeps should provide the listeners with more sensory information, and so their precision ought to increase for rear hemifield source azimuths, yet this was not the case. The increased number of sweeps and longer trial length for rear hemifield source azimuths may instead represent an attempt by the listener to offset the poorer cochlear precision, in that listeners are compelled to complete more sweeps to garner enough information to feel confident in making a response. Nevertheless, even with the increase in information gained for these trials, listener's overall precision (slope) for rear hemifield source azimuths did not reach the level found for source azimuths in the front hemifield. Increasing the amount of information gained did not improve overall precision, so the poorer precision in the rear hemifield must be driven by the precision of the cochlear cues.

In Chapter 4 it was suggested that the precision asymmetry is due to the location-dependent spectral filtering of the pinnae. This was based on direct measures of spatial precision. Using the head compensation task as an indirect measure of auditory spatial precision, it can be seen that spatial precision is broadly similar to that established in previous experiments in this thesis. Precision is better at the midline than at the interaural axis. It is also worse in the rear hemifield than at the symmetric position pair in the front hemifield. However, the worst precision is found here at a position just to the rear of the interaural axis rather than at the interaural axis itself. This may be explicable when the dynamics of the task are considered. When a listener makes a head movement, the sound source roves over an area of their auditory field. Hence, at different positions during its sweep the precision of the signal will be better or worse depending on its subjective azimuth. A source at 135° will likely rove over more areas of worse precision (interaural axis, rear hemifield) than a source at 45° (interaural axis only). It will then experience worse precision overall than the source at the front hemifield symmetric position pair. The same argument can be applied to 180° and 0°. Sources at the interaural axis will rove partially over the front hemifield and so may be given a benefit that is denied to a source whose origin is located to the rear of this. As only the overall precision for each source azimuth is available, this cannot be analysed but it may explain why there is a 'shift' in the worst precision compared to previous experiments.

It should also be noted that in this experiment, an indirect measure of spatial precision was performed that does not share the confounds of auditory localisation studies discussed in section 1.4.3. Although it has been well established previously in this thesis why direct measures of precision are preferable over indirect measures, this does illustrate that it is possible to create an indirect measure of spatial precision without introducing any such confounds.

Chapter 6 – Thesis Discussion

6.1 Summary of Findings

Our auditory system is unique among the senses in its ability to provide location information about distant stimuli that originate from any direction relative to our bodies. Using our hearing we are able to localise sources of sound in areas that other senses cannot, e.g., the rear hemifield. While there has been much research that has examined the accuracy of our localisation ability in both front and rear hemifields, there has been comparatively little research into the precision of our localisation ability in either hemifield but especially in the rear hemifield. Given that this area of relative space is uniquely accessible by our auditory system, it is important that we have a full understanding of our ability to localise sounds in this area.

The aim of this thesis was to examine spatial precision on the horizontal plane at a range of positions across both front and rear hemifields. As argued in Chapter 1, studies of spatial localisation accuracy can provide indirect measures of spatial precision, but these are often confounded with response method and so do not provide good metrics of precision. By using a direct measure of spatial precision, the MAA, these confounds can be avoided.

In a series of experiments directly examining auditory spatial precision, precision was found to be best at the midline and worse towards the interaural axis in the front hemifield, conforming to findings in previous literature (Mills, 1958). Importantly, precision was also found to be worse in the rear hemifield than in the front hemifield at complementary positions. This asymmetry in precision was also present for moving stimuli, shown using measurements of MAMA, and when listeners made judgements about stimuli while they were also moving themselves, an indirect measure of spatial precision.

The precision asymmetry between front and rear hemifield was hypothesised to be caused by differences in high-frequency cues related to the filtering properties of the pinna. When the contributions of the pinnae were removed by high-pass filtering, or by acoustically bypassing the pinnae with extension tubes, the asymmetry disappeared, supporting this hypothesis. It was also shown that the pinnae provide a selective benefit for spatial localisation precision in the front hemifield.

6.2 Implications for Methodology

This thesis highlights many of the challenges that a researcher may face when attempting to measure the spatial localisation ability of the auditory system. Section 2.2 outlines the method used throughout this thesis to administer the MAA task, and the testing that was performed to refine the parameters of this method.

The research described in this thesis used multiple concurrent sources to generate noise stimuli. This method has the advantage of being able to create 'phantom' stimuli which appear to be located between the real sources of sound. This is analogous to creating a visual stimulus whose centre appears to be located 'between' pixels on a screen. In this way, sounds can be generated at any position along a speaker array. Using multiple sources concurrently generates stimuli which are 'wider' than the effective point sources created by single-source methods. This creates an additional variable of spatial width that must be controlled. In section 2.4 the effect of changing this variable on measured spatial precision was demonstrated. It was shown that the narrower the spatial width (specified as the standard deviation of the Gaussian filter applied to make sounds localisable), the more precise listeners' spatial localisations became. However, this was true only down to a point which varied with the density of sources used. This 'critical' value was found to be 0.7 times the separation angle between sound sources. Using sounds with a spatial width lower than the critical value artificially inflated spatial precision estimates due to inconsistent power between sound source locations. Increasing the density of sources improved the estimates of spatial precision. The maximum density available was shown to produce MAAs equivalent to the smallest MAAs found using single-source methods, and to the minimum ITD detectable by the human auditory system (Brungart & Rabinowitz, 1999).

The multiple concurrent source method used within this thesis, and by other researchers (Perrott & Saberi, 1990; Saberi et al., 1991), is therefore capable of producing equivalent performance to single-source methods. Using concurrent sources also requires fewer physical sources to reach a required spatial resolution than single source arrays due to not needing a physical source at every intended location and does not require the complex mechanics of moveable source methods. Nor is it susceptible to problems of mechanism or wind noise that may occur using physically moving sources.

It was shown in Chapter 1 that for many tasks measuring spatial localisation accuracy there is a confound of response measure. For many studies this response is a pointing of the nose or hand, which requires additional movements for regions outside of the normal range of motion, as well as adding motor noise to location estimates. For other studies there was a problem of conceptualisation or quantisation of space, such that the minimum measurements that could be made were far coarser than the theoretical minimum for similar tasks. The direct precision measurements used in this thesis (MAA, MAMA) avoid these confounds. Additionally, the indirect method of measuring precision (head compensation) also avoided these confounds. These methods therefore provide avenues for researchers interested in measuring auditory precision without response measures becoming confounded with their variables of interest.

In section 4.4 experiments were performed that involved bypassing the listeners' pinnae using tubes. Many methods of occluding, disrupting, or bypassing the contributions of the pinnae have been shown to cause listeners to misperceive the hemifield of origin of presented stimuli (Musicant & Butler, 1984; Oldfield & Parker, 1984b; Shaw, 1997). Often this has led to researchers to 'flip' their participants' data into the correct hemifield before continuing analysis. The relative localisation method used in this experiment (MAA) did not require that this be performed due to the leftward/rightward or forward/backward judgements made by listeners. These judgements were the same for listeners regardless of the hemifield from which they perceived the sounds to originate, as these judgements were made ‘relative to the listener’ (e.g., more to the listener’s own left or right). This avoided the need for time-intensive work to ‘correct’ listeners’ data, and thereby also avoided the potential for mistakes to be introduced by the experimenter post-hoc.

Finally, there is a novel methodology introduced in this thesis. This is the filter that is applied to the noise stimuli when presenting sounds to listeners who are wearing tubes in their ears, which counteracted the spectral changes introduced by listening through the apparatus. Previous studies using moulds (Musicant & Butler, 1984; Oldfield & Parker, 1984b) or simple tubes (Shaw, 1997) have had the issue that the use of moulds inherently changes the spectral qualities of sound. In this thesis this has been avoided by using a filter in combination with the tubes. The tubes themselves were simple to construct and implement, being simply made of an otoscope speculum combined with a silicone earphone bud. This tube method is a simple and effective way of bypassing the pinnae which minimises the potential danger for participants. Combining this with the filter creates a very promising method for further research.

6.3 Implications for Spatial Localisation

It has long been acknowledged that spatial localisation is poorer in the rear hemifield, but very few direct measures of the precision of this ability have been performed. The original findings of Starch (1905) have been largely overlooked, and more recent work has only examined in detail the front hemifield. Given the importance of hearing in providing omnidirectional spatial information about the environment, this is a gap in our collective knowledge. Moreover, our understanding of higher-level abilities such as the perception of motion and of the motion of sources relative to our own movements is reliant on our understanding of low-level functionality.

It has also been known for a long time that the pinnae provide information about the elevation of sound sources and allow for front-back discrimination of sources. In section 4.4

evidence was presented that under normal binaural hearing conditions, the pinnae also provide a selective benefit for the localisation in azimuth of sounds in the front hemifield in the horizontal plane. It has often been suggested that one of the purposes of front-back discrimination from the pinnae is to alert the listener to sound sources (e.g. Musicant & Butler, 1984), and so to allow them to orient the more spatially acute visual system towards the source of sound, which may pose a threat or represent an opportunity. Now, we can infer that once the listener has oriented towards the source the pinnae continue to provide information that refines the precise localisation of that source.

6.4 Applications of this Work

The findings from the experiments in this thesis may be beneficial for applied research. For instance, hearing impaired persons who rely on hearing aids or cochlear implants (CIs) will not be able to access the benefits of the pinnae in the front hemifield as both of these devices bypass the pinnae and deliver sounds directly to the ear canal and cochlea respectively (Jones et al., 2016). In Chapter 4 it was shown that bypassing the pinnae led to consistent misidentifications of the hemifield of a source, confirming findings from previous literature (Butler et al., 1990; Carlile et al., 1999). Hearing impaired persons might therefore benefit from simulation of pinna cues when a sound is located in the front hemifield. This could be achieved by applying a generalised HRTF to the presented stimuli. Such a system would require automatic localisation of the sound source of interest from the sound field in order to select the appropriate HRTF, which is complex, but current work using gaze detection and beam forming (e.g. Kidd et al., 2013; Valin et al., 2007) poses a potential solution to this problem. A simpler alternative is the placement of the system microphones in the ear canal or meatus (Jones et al., 2016), which then enables the user's own pinnae to provide spectral information. Jones et al. found that the use of ITE (in the ear) microphone placement resulted in some improvement in azimuthal localisation for bilateral cochlear implant users over more standard BTE (behind the ear) or SHD (shoulder) microphone placements. This supports the idea that re-introduction, or here preservation, of spectral cues from the pinnae can be beneficial for users of hearing aids and cochlear implants.

In a rather different area, there is the field of workplace safety. Recently some organisations and areas have begun recommending that plant vehicles use broadband sounds for reversing signals (Scott, 2014; Vaillancourt et al., 2014) instead of the familiar beeping tone. Research has found that these broadband signals are better localised by workers than the previous tones (Catchpole et al., 2004; Withington, 2004). This finding is supported by the current literature and by the work in this thesis. However, the proposal fails to take into account that plant workers typically wear full ear protection or ear plugs, both of which can disrupt or even totally remove pinna cues.

Most ear protection (e.g. over-ear, ear plugs) has the primary objective of reducing sound level to safe levels for human hearing (Dobie, 1995). Using ear protection can have the consequence of reducing availability of cues to location generally, which is detrimental to localisation ability (Sabin et al., 2005), but can also greatly reduce the informativeness of spectral cues. This is because the sensitivity of the auditory system, i.e. the detection threshold, varies with frequency (Glasberg & Moore, 1990), such that detection thresholds are greater (worse) for higher frequencies. High frequencies (and so spectral cues) are, therefore, most likely to be affected by overall sound level attenuation introduced by the use of ear protection.

Using ear protection also diminishes front-back discrimination ability (Giguère et al., 2011), and increases the likelihood of front-back errors (Abel et al., 2007, 2009). If the use of ear protection removes or worsens the ability of the auditory system to use spectral cues for front-back discrimination, this may also affect localisation ability in the front hemifield. Removal of spectral cues would remove the additional information provided by the pinnae for the front hemifield, worsening localisation to the level found for sounds originating in the rear hemifield – as in section 4.4.

Consequently, if the sound level reduction while using ear protection is too great users may have poorer localisation in front-back discrimination, elevation discrimination, and front hemifield localisation. While the spectral contributions of the pinnae are removed or worsened, binaural cues may still remain. Access to these cues this could give rise to an over-confidence as users can still localise sounds in azimuth, leading them to over-estimate their abilities for other dimensions. This is something that should be considered when providing advice on localising sounds to users of ear protection. While implementation of broadband signals for alerting workers is certainly better than the previously used tones, a behavioural modification is also necessary for workers to be fully alerted. Workers should be encouraged to use head rotations to localise alerting or reversing signals, to prevent misidentification of the hemifield of origin of those signals, which could represent important hazards.

6.5 Limitations

The explanation for the front-rear asymmetry provided in this thesis is that the cues to horizontal location are better in the front hemifield than in the rear, due to an additional set of cues from the pinnae. This bottom-up explanation is more parsimonious than a spatial-maps explanation. However, there is still evidence that implicates spatial maps, specifically the integration and/or calibration of these maps with visual information.

The ventriloquist effect (Pick et al., 1969) occurs when a less precise sensory signal is perceived to be co-located with a more precise signal cross-modally (Alais & Burr, 2004). Although this is an illusion it demonstrates the ability of the sensory system to integrate information to improve, or attempt to improve, localisation ability. This effect has only been shown to occur in the front hemifield as vision cannot provide information about locations in the rear hemifield. However, vision does appear to still have an effect on localisation in the rear hemifield. Lewald and Ehrenstein (2001) found that visual gaze affected the localisations of sounds on the rear midline, such that sounds appeared to be slightly to the left of the midline when fixation was made to the right and vice versa. In their study, auditory localisation was suggested to be affected by vision via a shift in the integrated spatial map used for localisation judgements (Auerbach & Sperling, 1974).

Additionally, vision is believed to be used to calibrate some auditory localisation judgements. Although visually impaired listeners often perform as well as, or even better than, normally sighted listeners on simple localisation tasks (Ashmead et al., 1998; Lessard et al., 1998), on more complex tasks they can show an underperformance. Zwiers et al. (2001) found that early-blind and normally sighted listeners showed the same deterioration in performance in azimuthal judgements with decreasing signal to noise ratio. But, for elevation judgements, early blind listeners showed greater deterioration with decreasing signal to noise ratio than the normally sighted group. The researchers suggested that this was due to poorer calibration of elevation cues from the pinnae, compared to the more easily calibrated binaural cues. Similar findings were made by Gori et al. (2014) who showed that on complex localisation tasks, congenitally blind listeners performed worse than normally sighted listeners.

Systematic distortion of the visual field has also been shown to affect auditory localisation. Owls raised with prismatic lenses show a systematic distortion of their auditory localisation in the direction of their prisms (Knudsen & Knudsen, 1989). Similar, temporary effects have been shown in humans using lens which compress visual space (Zwiers et al., 2003). This demonstrates that integrated spatial maps can be re-calibrated by visual information even after they have been learned.

While in this thesis the conclusion is that the precision asymmetry is caused by sensory signals – and indeed this remains the conclusion – it is clear that spatial maps and sensory signals affect each other in a reciprocal process of constant updating. The sensory signals for sounds in the rear hemifield may be collectively poorer due to a lack of pinna cues that are present in the front hemifield. But, at a spatial map level there is also a lack of visual information in the rear hemifield that can be used to calibrate the sensory signals. So rear regions of an integrated spatial map are adversely affected, compared to front regions, by a lack of information from two sensory modalities.

Therefore, while the front/rear asymmetry found for auditory localisation in the rear hemifield is undoubtedly due in part to effects of the pinnae, there is likely to also be a component of visual information worsening this asymmetry. Notably, in 4.4 this asymmetry did disappear when tubes were used to bypass the pinnae, which can also be seen from a lack of asymmetry in the excitation patterns (section 4.3). So, if there is a benefit of visual information for spatial maps it may not be sufficient to cause an asymmetry in auditory localisation.

6.6 Future Research

There are immediate extensions of this work that would simply involve performing the MAMA and head compensation experiments again, but with a comparison between a *normal hearing* condition and a *tube bypass* condition as in section 4.4. This would provide reassurance of the claims made here about the relationship between pinna cues and spatial precision and demonstrate whether (as in previous experiments) changes to spatial precision at the lowest level are reflected in higher-order tasks.

Further exploration of the head compensation findings from section 5.4would also be interesting, given the lack of relationship between compensation gain and precision found in this experiment. This cannot be predicted by either of the theories discussed in this section, so certainly warrants more attention.

There are also some interesting areas that could be explored with visually impaired listeners. As discussed above (section 6.5), early-blind or congenitally blind individuals may lack the visual calibration required for assessments of elevation, which is informed by the pinnae (Zwiers et al., 2001). So, it would be interesting to determine if visually impaired listeners experienced any front-rear asymmetry in spatial precision. If they showed no asymmetry between front and rear hemifields this would suggest that this selective benefit of the pinnae in the front hemifield is learned in conjunction with, or calibrated by, visual information. If they did show an asymmetry then this would suggest that there is some innate biological advantage to having pinnae, independent of any cross-modal learning that may occur.

References

Abel, S. M., Boyne, S., & Roesler-Mulroney, H. (2009). Sound localization with an army helmet worn in combination with an in-ear advanced communications system. *Noise & Health*, *11*(45), 199–205. https://doi.org/10.4103/1463-1741.56213

Abel, S. M., Tsang, S., & Boyne, S. (2007). Sound localization with communications headsets: Comparison of passive and active systems. *Noise and Health*, *9*(37), 101. https://doi.org/10.4103/1463-1741.37426

Alais, D., & Burr, D. (2004). The Ventriloquist Effect Results from Near-Optimal Bimodal Integration. *Current Biology*, *14*(3), 257–262. https://doi.org/10.1016/j.cub.2004.01.029

Algazi, V. R., Avendano, C., & Duda, R. O. (2001). Estimation of a spherical-head model from anthropometry. *Journal of the Audio Engineering Society*, *49*(6), 472–479.

Ashmead, D. H., Wall, R. S., Ebinger, K. A., Eaton, S. B., Snook-Hill, M.-M., & Yang, X. (1998). Spatial Hearing in Children with Visual Disabilities. *Perception*, *27*(1), 105–122. https://doi.org/10.1068/p270105

Auerbach, C., & Sperling, P. (1974). A common auditory-visual space: Evidence for its reality. *Perception & Psychophysics*, *16*(1), 129–135. https://doi.org/10.3758/BF03203265

Baumgart, F., Gaschler-Markefski, B., Woldorff, M. G., Heinze, H.-J., & Scheich, H. (1999). A movement-sensitive area in auditory cortex. *Nature*, *400*(6746), 724–726. https://doi.org/10.1038/23390

Best, V. (2004). *Spatial hearing with simultaneous sound sources: A psychophysical investigation* [Doctoral Thesis, The University of Sydney]. https://ses.library.usyd.edu.au/handle/2123/576

Blauert, J. (1997). *Spatial Hearing: The Psychophysics of Human Sound Localization*. MIT Press.

Brimijoin, W. O. (2018). Angle-Dependent Distortions in the Perceptual Topology of Acoustic Space. *Trends in Hearing*, *22*, 233121651877556. https://doi.org/10.1177/2331216518775568

Bronkhorst, A. W. (1995). Localization of real and virtual sound sources. *The Journal of the Acoustical Society of America*, *98*(5), 2542–2553. https://doi.org/10.1121/1.413219

Brungart, D. S., Durlach, N. I., & Rabinowitz, W. M. (1999). Auditory localization of nearby sources. II. Localization of a broadband source. *The Journal of the Acoustical Society of America*, *106*(4), 1956–1968. https://doi.org/10.1121/1.427943

Brungart, D. S., & Rabinowitz, W. M. (1999). Auditory localization of nearby sources. Head-related transfer functions. *The Journal of the Acoustical Society of America*, *106*(3), 1465–1479. https://doi.org/10.1121/1.427180

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. https://doi.org/10.3758/BF03212242

Butler, R. A., Humanski, R. A., & Musicant, A. D. (1990). Binaural and Monaural Localization of Sound in Two-Dimensional Space. *Perception*, *19*(2), 241–256. https://doi.org/10.1068/p190241

Carlile, S., & Best, V. (2002). Discrimination of sound source velocity in human listeners. *The Journal of the Acoustical Society of America*, *111*(2), 1026–1035. https://doi.org/10.1121/1.1436067

Carlile, S., Delaney, S., & Corderoy, A. (1999). The localisation of spectrally restricted sounds by human listeners. *Hearing Research*, *128*(1), 175–189. https://doi.org/10.1016/S0378-5955(98)00205-6

Carlile, S., Leong, P., & Hyams, S. (1997). The nature and distribution of errors in sound localization by human listeners. *Hearing Research*, *114*, 179–196. https://doi.org/10.1016/S0378-5955(97)00161-5

Carlile, S., & Leung, J. (2016). The Perception of Auditory Motion. *Trends in Hearing*, *20*, 2331216516644254. https://doi.org/10.1177/2331216516644254

Carlile, S., & Pralong, D. (1994). The location-dependent nature of perceptually salient features of the human head-related transfer functions. *The Journal of the Acoustical Society of America*, *95*(6), 3445–3459. https://doi.org/10.1121/1.409965

Catchpole, K. R., McKeown, J. D., & Withington, D. J. (2004). Localizable auditory warning pulses. *Ergonomics*, *47*(7), 748–771. https://doi.org/10.1080/00140130310001629739

Chandler, D. W., & Grantham, D. W. (1992). Minimum audible movement angle in the horizontal plane as a function of stimulus frequency and bandwidth, source azimuth, and velocity. *The Journal of the Acoustical Society of America*, *91*(3), 1624–1636. https://doi.org/10.1121/1.402443

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Academic Press.

Divenyi, P. L., & Oliver, S. K. (1989). Resolution of steady-state sounds in simulated auditory space. *The Journal of the Acoustical Society of America*, *85*(5), 2042–2052. https://doi.org/10.1121/1.397856

Dobie, R. A. (1995). Prevention of Noise-Induced Hearing Loss. *Archives of Otolaryngology–Head & Neck Surgery*, *121*(4), 385–391. https://doi.org/10.1001/archotol.1995.01890040011002

DuBois, R. M., & Cohen, M. S. (2000). Spatiotopic Organization in Human Superior Colliculus Observed with fMRI. *NeuroImage*, *12*(1), 63–70. https://doi.org/10.1006/nimg.2000.0590

Ege, R., Opstal, A. J. V., & Wanrooij, M. M. V. (2018). Accuracy-Precision Trade-off in Human Sound Localisation. *Scientific Reports*, *8*(1), 16399. https://doi.org/10.1038/s41598-018-34512-6

Eliasmith, C., & Anderson, C. H. (2004). *Neural Engineering: Computation, Representation, and Dynamics in Neurobiological Systems*. MIT Press.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433. https://doi.org/10.1038/415429a

Evans, M. J. (1998). Obtaining accurate responses in directional listening tests. *Audio Engineering Society*, 1–12.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.

Fechner, G. T. (1851). *Zend-Avesta, oder, Über die Dinge des Himmels und des Jenseits*. Lepold Voss. https://catalog.hathitrust.org/Record/008409270

Fechner, G. T. (1987). Outline of a new principle of mathematical psychology (1851). *Psychological Research*, *49*(4), 203–207. https://doi.org/10.1007/BF00309027

Feddersen, W. E., Sandel, T. T., Teas, D. C., & Jeffress, L. A. (1957). Localization of High-Frequency Tones. *The Journal of the Acoustical Society of America*, *29*(9), 988–991. https://doi.org/10.1121/1.1909356

Feinstein, S. H. (1973a). Acuity of the human sound localization response underwater. *The Journal of the Acoustical Society of America*, *53*(2), 393–399. https://doi.org/10.1121/1.1913335

Feinstein, S. H. (1973b). Minimum audible angle underwater: A replication under different acoustic and environmental conditions. *The Journal of the Acoustical Society of America*, *54*(4), 879–881. https://doi.org/10.1121/1.1914341

Filehne, W. (1922). Uber das optische Wahrnehmen von Bewegungen. *Zeitschrift fur Sinnesphysiologie*, *53*, 134–145.

Foxe, J. J., Wylie, G. R., Martinez, A., Schroeder, C. E., Javitt, D. C., Guilfoyle, D., Ritter, W., & Murray, M. M. (2002). Auditory-Somatosensory Multisensory Processing in Auditory Association Cortex: An fMRI Study. *Journal of Neurophysiology*, *88*(1), 540–543. https://doi.org/10.1152/jn.2002.88.1.540

Freeman, T. C. A., & Banks, M. S. (1998). Perceived head-centric speed is affected by both extra-retinal and retinal errors. *Vision Research*, *38*(7), 941–945. https://doi.org/10.1016/S0042-6989(97)00395-7

Freeman, T. C. A., Champion, R. A., & Warren, P. A. (2010). A Bayesian Model of Perceived Head-Centered Velocity during Smooth Pursuit Eye Movement. *Current Biology*, *20*(8), 757–762. https://doi.org/10.1016/j.cub.2010.02.059

Freeman, T. C. A., Culling, J. F., Akeroyd, M. A., & Brimijoin, W. O. (2017). Auditory compensation for head rotation is incomplete. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(2), 371–380. https://doi.org/10.1037/xhp0000321

Freeman, T. C. A., Leung, J., Wufong, E., Orchard-Mills, E., Carlile, S., & Alais, D. (2014). Discrimination Contours for Moving Sounds Reveal Duration and Distance Cues Dominate Auditory Speed Perception. *PLoS ONE*, *9*(7), e102864. https://doi.org/10.1371/journal.pone.0102864

Furman, M., & Gur, M. (2012). And yet it moves: Perceptual illusions and neural mechanisms of pursuit compensation during smooth pursuit eye movements.

Neuroscience & Biobehavioral Reviews, *36*(1), 143–151. https://doi.org/10.1016/j.neubiorev.2011.05.005

Garcia, S. E., Jones, P. R., Rubin, G. S., & Nardini, M. (2017). Auditory Localisation Biases Increase with Sensory Uncertainty. *Scientific Reports*, *7*(1). https://doi.org/10.1038/srep40567

Gates, D. H., Walters, L. S., Cowley, J., Wilken, J. M., & Resnik, L. (2016). Range of Motion Requirements for Upper-Limb Activities of Daily Living. *The American Journal of Occupational Therapy*, *70*(1), 7001350010p1-7001350010p10. https://doi.org/10.5014/ajot.2016.015487

Giard, M. H., & Peronnet, F. (1999). Auditory-Visual Integration during Multimodal Object Recognition in Humans: A Behavioral and Electrophysiological Study. *Journal of Cognitive Neuroscience*, *11*(5), 473–490. https://doi.org/10.1162/089892999563544

Giguère, C., Brammer, A. J., Vaillancourt, V., & Yu, G. (2011). *Advanced hearing protection and communication: Progress and challenges*. 225–233.

Gilkey, R. H., Good, M. D., Ericson, M. A., Brinkman, J., & Stewart, J. M. (1995). A pointing technique for rapidly collecting localization responses in auditory research. *Behavior Research Methods, Instruments, & Computers*, *27*(1), 1–11. https://doi.org/10.3758/BF03203614

Glasberg, B. R., & Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research*, *47*(1), 103–138. https://doi.org/10.1016/0378-5955(90)90170-T

Good, M. D., & Gilkey, R. H. (1996). Sound localization in noise: The effect of signal-to-noise ratio. *The Journal of the Acoustical Society of America*, *99*(2), 1108–1117. https://doi.org/10.1121/1.415233

Gori, M., Sandini, G., Martinoli, C., & Burr, D. C. (2014). Impairment of auditory spatial localization in congenitally blind human subjects. *Brain*, *137*(1), 288–293. https://doi.org/10.1093/brain/awt311

Grantham, D. W. (1986). Detection and discrimination of simulated motion of auditory targets in the horizontal plane. *The Journal of the Acoustical Society of America*, *79*(6), 1939–1949. https://doi.org/10.1121/1.393201

Grantham, D. W. (1997). Auditory motion perception: Snapshots revisited. In *Binaural and Spatial Hearing in Real and Virtual Environments* (pp. 295–313). Psychology Press.

Griffiths, T. D., & Green, G. G. R. (1999). Cortical Activation during Perception of a Rotating Wide-Field Acoustic Stimulus. *NeuroImage*, *10*(1), 84–90. https://doi.org/10.1006/nimg.1999.0464

Guski, R. (1992). Acoustic Tau: An Easy Analogue to Visual Tau?: Ecological Psychology: Vol 4, No 3. *Ecological Psychology*, *4*(3), 189–197. https://doi.org/10.1207/s15326969eco0403_4

Haber, L., Haber, R. N., Penningroth, S., Novak, K., & Radgowski, H. (1993). Comparison of Nine Methods of Indicating the Direction to Objects: Data from Blind Adults. *Perception*, *22*(1), 35–47. https://doi.org/10.1068/p220035

Harris, J. D., & Sergeant, R. L. (1971). Monaural/binaural minimum audible angles for a moving sound source. *Journal of Speech and Hearing Research*, *14*(3), 618–629. https://doi.org/10.1044/jshr.1403.618

Hartley, R. V. L., & Fry, T. C. (1921). The Binaural Location of Pure Tones. *Physical Review*, *18*(6), 431–442. https://doi.org/10.1103/PhysRev.18.431

Heffner, H. E., & Heffner, R. S. (2005). The sound-localization ability of cats. *Journal of Neurophysiology*, *94*(5), 3653–3655. https://doi.org/10.1152/jn.00720.2005

Hürlimann, F., Kiper, D. C., & Carandini, M. (2002). Testing the Bayesian model of perceived speed. *Vision Research*, *42*(19), 2253–2257. https://doi.org/10.1016/S0042-6989(02)00119-0

Iglewicz, B., & Hoaglin, D. C. (1993). *How to detect and handle outliers* (Vol. 16). ASQC Quality Press.

Jones, H. G., Kan, A., & Litovsky, R. Y. (2016). The effect of microphone placement on interaural level differences and sound localization across the horizontal plane in bilateral cochlear implant users. *Ear and Hearing*, *37*(5), e341–e345. https://doi.org/10.1097/AUD.0000000000000297

Jones, P. R. (2016). A note on detecting statistical outliers in psychophysical data. *BioRxiv*. https://doi.org/10.1101/074591

Kidd, G., Favrot, S., Desloge, J. G., Streeter, T. M., & Mason, C. R. (2013). Design and preliminary testing of a visually guided hearing aid. *The Journal of the Acoustical Society of America*, *133*(3), EL202–EL207. https://doi.org/10.1121/1.4791710

Klumpp, R. G., & Eady, H. R. (1956). Some Measurements of Interaural Time Difference Thresholds. *The Journal of the Acoustical Society of America*, *28*(5), 859–860. https://doi.org/10.1121/1.1908493

Knudsen, E. I., & Brainard, M. S. (1995). Creating a Unified Representation of Visual and Auditory Space in the Brain. *Annual Review of Neuroscience*, *18*(1), 19–43. https://doi.org/10.1146/annurev.ne.18.030195.000315

Knudsen, E. I., & Knudsen, P. F. (1989). Vision calibrates sound localization in developing barn owls. *The Journal of Neuroscience*, *9*(9), 3306–3313. https://doi.org/10.1523/JNEUROSCI.09-09-03306.1989

Kopco, N., & Shinn-Cunningham, B. G. (2000). Spatial unmasking of nearby pure-tone sources in a simulated anechoic environment. *The Journal of the Acoustical Society of America*, *107*(5 PART 2), 2849.

Kunin, M., Osaki, Y., Cohen, B., & Raphan, T. (2007). Rotation Axes of the Head During Positioning, Head Shaking, and Locomotion. *Journal of Neurophysiology*, *98*(5), 3095–3108. https://doi.org/10.1152/jn.00764.2007

Langendijk, E. H. A., & Bronkhorst, A. W. (2002). Contribution of spectral cues to human sound localization. *The Journal of the Acoustical Society of America*, *112*(4), 1583–1596. https://doi.org/10.1121/1.1501901

Lessard, N., Paré, M., Lepore, F., & Lassonde, M. (1998). Early-blind human subjects localize sound sources better than sighted subjects. *Nature*, *395*(6699), 278–280. https://doi.org/10.1038/26228

Levitt, H. (1971). Transformed Up-Down Methods in Psychoacoustics. *The Journal of the Acoustical Society of America*, *49*(2B), 467–477. https://doi.org/10.1121/1.1912375

Lewald, J., Dörrscheidt, G. J., & Ehrenstein, W. H. (2000). Sound localization with eccentric head position. *Behavioural Brain Research*, *108*(2), 105–125. https://doi.org/10.1016/S0166-4328(99)00141-2

Lewald, J., & Ehrenstein, W. H. (2001). Effect of gaze direction on sound localization in rear space. *Neuroscience Research*, *39*(2), 253–257. https://doi.org/10.1016/S0168-0102(00)00210-8

Macaulay, E. J., Hartmann, W. M., & Rakerd, B. (2010). The acoustical bright spot and mislocalization of tones by human listeners. *The Journal of the Acoustical Society of America*, *127*(3), 1440–1449. https://doi.org/10.1121/1.3294654

Macpherson, E. A., & Middlebrooks, J. C. (2002). Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited. *The Journal of the Acoustical Society of America*, *111*(5), 2219. https://doi.org/10.1121/1.1471898

Majdak, P., Goupell, M. J., & Laback, B. (2010). 3-D localization of virtual sound sources: Effects of visual environment, pointing method, and training. *Attention, Perception, & Psychophysics*, *72*(2), 454–469. https://doi.org/10.3758/APP.72.2.454

Makous, J. C., & Middlebrooks, J. C. (1990). Two-dimensional sound localization by human listeners. *The Journal of the Acoustical Society of America*, *87*(5), 2188–2200. https://doi.org/10.1121/1.399186

Mason, R. (2000). Verbal and Non-Verbal Elicitation Techniques in the Subjective Assessment of Spatial Sound Reproduction. *Journal of the Audio Engineering Society*, 25.

Mendonça, C., Campos, G., Dias, P., Vieira, J., Ferreira, J. P., & Santos, J. A. (2012). On the Improvement of Localization Accuracy with Non-Individualized HRTF-Based Sounds. *Journal of the Audio Engineering Society*, *60*(10), 821–830.

Meredith, M. A., & Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, *56*(3), 640–662. https://doi.org/10.1152/jn.1986.56.3.640

Middlebrooks, C., & Knudsen, E. I. (1984). A neural code for auditory space in the cat's superior colliculus. *The Journal of Neuroscience*, *4*(10), 2621–2634.

Mills, A. W. (1958). On the minimum audible angle. *The Journal of the Acoustical Society of America*, *30*(4), 237–246. https://doi.org/10.1121/1.1909553

Moore, J. M., Tollin, D. J., & Yin, T. C. T. (2008). Can measures of sound localization acuity be related to the precision of absolute location estimates? *Hearing Research*, *238*(1–2), 94–109. https://doi.org/10.1016/j.heares.2007.11.006

Musicant, A. D., & Butler, R. A. (1984). The influence of pinnae-based spectral cues on sound localization. *The Journal of the Acoustical Society of America*, *75*(4), 1195–1200. https://doi.org/10.1121/1.390770

Neuhoff, J. G. (2004). Auditory Motion and Localization. In *Ecological Psychoacoustics* (pp. 87–111). Elsevier. https://doi.org/10.1016/B978-012515851-0/50005-9

Nielsen, S. H. (1991, February 1). *Depth Perception-Finding a Design Goal for Sound Reproduction Systems*. Audio Engineering Society Convention 90. http://www.aes.org/e-lib/browse.cfm?elib=5624

Oldfield, S. R., & Parker, S. P. A. (1984a). Acuity of Sound Localisation: A Topography of Auditory Space. I. Normal Hearing Conditions. *Perception*, *13*(5), 581–600. https://doi.org/10.1068/p130581

Oldfield, S. R., & Parker, S. P. A. (1984b). Acuity of Sound Localisation: A Topography of Auditory Space. II. Pinna Cues Absent. *Perception*, *13*(5), 601–617. https://doi.org/10.1068/p130601

Oldfield, S. R., & Parker, S. P. A. (1986). Acuity of Sound Localisation: A Topography of Auditory Space. III. Monaural Hearing Conditions. *Perception*, *15*(1), 67–81. https://doi.org/10.1068/p150067

Olsen, W. O. (1998). Average Speech Levels and Spectra in Various Speaking/Listening Conditions: A Summary of the Pearson, Bennett, & Fidell (1977) Report. *American Journal of Audiology: Perspective*, *7*(2), 21–25.

Palmer, A. R., & King, A. J. (1982). The representation of auditory space in the mammalian superior colliculus. *Nature*, *299*(5880), 248–249. https://doi.org/10.1038/299248a0

Perrett, S., & Noble, W. (1997). The contribution of head motion cues to localization of low-pass noise. *Perception & Psychophysics*, *59*(7), 1018–1026. https://doi.org/10.3758/BF03205517

Perrott, D. R. (1984). Concurrent minimum audible angle: A re-examination of the concept of auditory spatial acuity. *The Journal of the Acoustical Society of America*, *75*(4), 1201–1206. https://doi.org/10.1121/1.390771

Perrott, D. R., & Marlborough, K. (1989). Minimum audible movement angle: Marking the end points of the path traveled by a moving sound source. *The Journal of the Acoustical Society of America*, *85*(4), 1773–1775. https://doi.org/10.1121/1.397968

Perrott, D. R., & Musicant, A. D. (1977). Minimum auditory movement angle: Binaural localization of moving sound sources. *The Journal of the Acoustical Society of America*, *62*(6), 1463–1466. https://doi.org/10.1121/1.381675

Perrott, D. R., & Saberi, K. (1990). Minimum audible angle thresholds for sources varying in both elevation and azimuth. *The Journal of the Acoustical Society of America*, *87*(4), 1728–1731. https://doi.org/10.1121/1.399421

Perrott, D. R., & Tucker, J. (1988). Minimum audible movement angle as a function of signal frequency and the velocity of the source. *The Journal of the Acoustical Society of America*, *83*(4), 1522–1527. https://doi.org/10.1121/1.395908

Pick, H. L., Warren, D. H., & Hay, J. C. (1969). Sensory conflict in judgments of spatial direction. *Perception & Psychophysics*, *6*(4), 203–205. https://doi.org/10.3758/BF03207017

Rayleigh, Lord. (1876). On Our Perception of the Direction of a Source of Sound. *Proceedings of the Musical Association*, *2*, 75–84. JSTOR.

Rayleigh, Lord. (1907). XII. On our perception of sound direction. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *13*(74), 214–232. https://doi.org/10.1080/14786440709463595

Saberi, K., Dostal, L., Sadralobai, T., & Perrott, D. R. (1991). Minimum audible angle for horizontal, vertical, and oblique orientations: Lateral and dorsal planes. *Acta Acustica United with Acustica*, *75*(1), 57–61.

Sabin, A. T., Macpherson, E. A., & Middlebrooks, J. C. (2005). Human sound localization at near-threshold levels. *Hearing Research*, *199*(1), 124–134. https://doi.org/10.1016/j.heares.2004.08.001

Scott, T. (2014, January 24). *Why Do Reversing Trucks Not Beep Any More?* [YouTube]. https://www.youtube.com/watch?v=fa28lIGuxq8

Senna, I., Parise, C. V., & Ernst, M. O. (2015). Hearing in slow-motion: Humans underestimate the speed of moving sounds. *Scientific Reports*, *5*(1), 1–5. https://doi.org/10.1038/srep14054

Shaw, B. K., McGowan, R. S., & Turvey, M. T. (1991). An Acoustic Variable Specifying Time-to-Contact. *Ecological Psychology*, *3*(3), 253–261. https://doi.org/10.1207/s15326969eco0303_4

Shaw, E. A. G. (1974). The external ear. In W. D. Keidel & W. D. Neff (Eds.), *Handbook of sensory physiology* (Vol. 1, pp. 455–490). Springer-Verlag.

Shaw, S. J. (1997). Directional Perception in the Human Auditory System. *The Journal of Undergraduate Sciences*, *3*, 135–140.

Shinn-Cunningham, B. G. (2000). Distance cues for virtual auditory space. *Proceedings of the IEEE-Pacific Rim Conference*, *1*, 227–230.

Shinn-Cunningham, B. G., Durlach, N. I., & Held, R. M. (1998a). Adapting to supernormal auditory localization cues. I. Bias and resolution. *The Journal of the Acoustical Society of America*, *103*(6), 3656–3666. https://doi.org/10.1121/1.423088

Shinn-Cunningham, B. G., Durlach, N. I., & Held, R. M. (1998b). Adapting to supernormal auditory localization cues. II. Constraints on adaptation of mean response. *The Journal of the Acoustical Society of America*, *103*(6), 3667–3676. https://doi.org/10.1121/1.423107

Starch, D. (1905). Perimetry of the localization of sound. *The Psychological Review: Monograph Supplements*, *6*(5), 1–45. http://dx.doi.org/10.1037/h0093058

Stevens, S. S., & Newman, E. B. (1934). The Localization of Pure Tones. *Proceedings of the National Academy of Sciences of the United States of America*, *20*(11), 593–596.

Stevens, S. S., & Newman, E. B. (1936). The Localization of Actual Sources of Sound. *The American Journal of Psychology*, *48*(2), 297–306. https://doi.org/10.2307/1415748

Swinkels, R. A. H., & Swinkels-Meewisse, I. E. J. C. (2014). Normal Values for Cervical Range of Motion. *Spine*, *39*(5), 362–367. https://doi.org/10.1097/BRS.0000000000000158

Vaillancourt, V., Néllise, H., Laroche, C., Giguère, C., Boutin, J., & Laferrière, P. (2014). *Safety of Workers Behind Heavy Vehicles—Assessment of Three Types of Reverse Alarm* (Studies and Research Projects R-833; Mechanical and Physical Risk Prevention). Institut de recherche Robert-Sauvé en santé et en securité du travail. https://www.academia.edu/34638937/Safety_of_Workers_Behind_Heavy_Vehicles_-_Assessment_of_Three_Types_of_Reverse_Alarm

Valin, J.-M., Michaud, F., & Rouat, J. (2007). Robust localization and tracking of simultaneous moving sound sources using beamforming and particle filtering. *Robotics and Autonomous Systems*, *55*(3), 216–228. https://doi.org/10.1016/j.robot.2006.08.004

Von Holst, E. (1954). Relations between the central nervous system and the peripheral organs. *British Journal of Animal Behaviour*, *2*, 89–94. https://doi.org/10.1016/S0950-5601(54)80044-X

Wallach, H. (1939). On Sound Localization. *The Journal of the Acoustical Society of America*, *10*, 270–274. https://doi.org/10.1121/1.1915985

Warren, J. D., Zielinski, B. A., Green, G. G. R., Rauschecker, J. P., & Griffiths, T. D. (2002). Perception of Sound-Source Motion by the Human Brain. *Neuron*, *34*(1), 139–148. https://doi.org/10.1016/S0896-6273(02)00637-2

Warren, R. M. (1982). *Auditory perception: A new synthesis*. Pergamon Press.

Weiss, J. (1889). *Dr Siegle pneumatic otoscope. Illustrated catalogue... Surgical instruments* (L0002446) [Surgical illustration]. Wellcome Library, London.

Wenzel, E. M., Arruda, M., Kistler, D. J., & Wightman, F. L. (1993). Localization using nonindividualized head-related transfer functions. *The Journal of the Acoustical Society of America*, *94*(1), 111–123. https://doi.org/10.1121/1.407089

Wightman, F. L., & Kistler, D. (1997). Monaural sound localization revisited. *The Journal of the Acoustical Society of America*, *101*(2), 1050–1063.

Wightman, F. L., & Kistler, D. (2005). Measurement and Validation of Human HRTFs for Use in Hearing Research. *Acta Acustica United with Acustica*, *91*(3), 429–439.

Wightman, F. L., & Kistler, D. J. (1992). The dominant role of low-frequency interaural time differences in sound localization. *The Journal of the Acoustical Society of America*, *91*(3), 1648–1661. https://doi.org/10.1121/1.402445

Wightman, F. L., & Kistler, D. J. (1999). Resolution of front–back ambiguity in spatial hearing by listener and source movement. *The Journal of the Acoustical Society of America*, *105*(5), 2841–2853. https://doi.org/10.1121/1.426899

Withington, D. J. (2004). Reversing goes broadband. *Quarry Management*, *31*(5), 27–33.

Zakarauskas, P., & Cynader, M. S. (1991). Aural intensity for a moving source. *Hearing Research*, *52*(1), 233–244. https://doi.org/10.1016/0378-5955(91)90203-L

Zwiers, M. P., Van Opstal, A. J., & Cruysberg, J. R. M. (2001). A Spatial Hearing Deficit in Early-Blind Humans. *Journal of Neuroscience*, *21*(9), RC142–RC142. https://doi.org/10.1523/JNEUROSCI.21-09-j0002.2001

Zwiers, M. P., Van Opstal, A. J., & Paige, G. D. (2003). Plasticity in human sound localization induced by compressed spatial vision. *Nature Neuroscience*, *6*(2), 175–181. https://doi.org/10.1038/nn999

Appendix

Table 3

Auditory localisation experiments not fitting the exclusion criteria.

Authors	Year	Exp	N	Stimulus	Response Method	Precision presented as	Reason for exclusion
Oldfield & Parker	1984b	--	8	White noise	Gun held in hand	Mean absolute error	Pinna occluded - not normal hearing
Oldfield & Parker	1986	--	8	White noise	Gun held in hand	Mean absolute error	Monaural only - not normal hearing
Haber et al.	1993	--	20	1 kHz, pulses at 2 Hz	Pointing with nose, chest, finger, cane, or stick; waist-mounted dial, table-mounted dial; drawing on board; clock-face verbal	Mean unsigned error (degrees)	Sounds played during response
Wenzel et al.	1993	1	16	0.2 - 14.0 kHz, train of eight 250 ms pulses	Verbal, degrees of azimuth and elevation	Mean unsigned error (degrees)	Precision reported regionally only
Wenzel et al.	1993	2	16	0.2 - 14.0 kHz, train of eight 250 ms pulses	Verbal, degrees of azimuth and elevation	Mean unsigned error (degrees)	Virtual sound presentation, precision reported regionally only
Bronkhorst	1995	1	8	Harmonic, 250 Hz fundamental	Head pointing	Absolute error (degrees)	No data at 0° azimuth
Good & Gilkey	1996	--	3	0.53-11.0 kHz, 25 μsec pulses at 100 Hz	Stylus touch to GELP sphere	Circles with variable radius, plotted around 'truth' line	Data were unrecoverable from graph
Perrett & Noble	1997	1	12	White noise, 0.5 sec or 3 sec	Laser pointer held in hand	Horizontal error	Data were reported anecdotally only

Langendijk & Bronkhorst	2000	1	8	0.2 - 16.0 kHz, 200 ms, spectral cues removed in some conditions	Acoustic pointer, controlled by joystick	Unclear	Virtual sound presentation only
Lewald et al.	2000	1A/1B	9	1.0-3.0 kHz, 10 sec duration	Head pointing (A), Laser mounted to head (B)	Deviation from truth (degrees)	Head moves while sound is playing, data presented with zero error at 0°
Lewald et al.	2000	2A/2B	9	1.0-3.0 kHz, 2 sec duration	Head pointing (A), Laser mounted to head (B)	Deviation from truth (degrees)	No data at 0° azimuth
Lewald et al.	2000	3A/3B	5	1.0-3.0 kHz, 2 sec duration	2AFC, Left/right of straight ahead using laser	Deviation from truth (degrees)	Virtual sound sources used, simulated HRTFs
Lewald et al.	2000	4	5	1.0-3.0 kHz, 8 sec duration	Swivelling pointer and head point simultaneously	Deviation from truth (degrees)	Head moves while sound is playing, data presented with zero error at 0°
Lewald et al.	2000	5	5	1.0-3.0 kHz, 8 sec duration	Swivelling pointer	Deviation from truth (degrees)	Data presented with zero error at 0°
Majdak et al.	2010	1	10	White noise, 500 ms duration	Head pointing, hand pointing (visualised as crosshair in VR)	Standard deviation of lateral error (degrees)	Virtual sounds used, grouped presentation angles over 20° sections (e.g., 0° to 20°, 20° to 40°)