

# Audification of Ultrasound for Human Echolocation

by

Theresa Claire Davies

A thesis  
presented to the University of Waterloo  
in fulfillment of the  
thesis requirement for the degree of  
Doctor of Philosophy  
in  
Systems Design Engineering

Waterloo, Ontario, Canada, 2008

©Theresa Claire Davies 2008

## **Author's Declaration**

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

## **Abstract**

Individuals with functional blindness must often utilise assistive aids to enable them to complete tasks of daily living. One of these tasks, locomotion, poses considerable risk. The long white cane is often used to perform haptic exploration, but cannot detect obstacles that are not ground-based. Although devices have been developed to provide information above waist height, these do not provide auditory interfaces that are easy to learn. Development of such devices should adapt to the user, not require adaptation by the user. Can obstacle avoidance be achieved through direct perception? This research presents an auditory interface that has been designed with the user as the primary focus. An analysis of the tasks required has been taken into account resulting in an interface that audifies ultrasound. Audification provides intuitive information to the user to enable perceptive response to environmental obstacles. A device was developed that provides Doppler shift signals that are audible as a result of intentional aliasing. This system provides acoustic flow that is evident upon initiation of travel and has been shown to be effective in perceiving apertures and avoiding environmental obstacles. The orientation of receivers on this device was also examined, resulting in better distance perception and centreline accuracy when oriented outward as compared to forward. The design of this novel user interface for visually impaired individuals has also provided a tool that can be used to evaluate direct perception and acoustic flow in a manner that has never been studied before.

## **Acknowledgements**

First and foremost, I would like to acknowledge the contribution of Shane Pinder. Although not my official supervisor, he dedicated many hours to the development and evaluation of the system, the testing, and the editing and re-editing of this work. He has been a constant source of support both with the thesis and with the children while I was dedicating countless hours to achieving this goal.

Catherine Burns, my official supervisor, stood by me throughout. She took a big risk by taking me on as a student, especially when I moved to New Zealand half way through. She supported the move and visited to check on my progress. She never doubted my dedication and was prepared to trust and support my decisions.

In Auckland, I gained valuable knowledge from the team at the Acoustics Research Centre. George Dodd took me under his wings from the beginning and opened his lab to me. Gian Schmid was a fountain of ideas as well as helping me to dismantle and reconstruct test equipment within the anechoic chamber. I cannot thank these people enough for the never-ending support and guidance.

I am grateful for the considered thoughts of my committee members and for their cooperative spirit and careful reading of my final document. Their comments and suggestions have sparked many ideas regarding both the current work and future projects.

My family has also been a constant source of encouragement and moral support. My mother dedicated many hours as a pilot subject when visiting Auckland, time that could have been spent doing many more interesting things! My parents have always provided a shoulder to cry on when things get rough and were supportive throughout my academic career.

I cannot mention my parents without mentioning the lack of “Mummy time” suffered by my children as a result of me continually having my head in the clouds. I’m sure they’ll turn out okay, but they took a big hit in their early childhood and I hope they’ll be able to forgive me. My son’s last words to me on the way to the airport for my defence, “Mummy, you can only do your best”. I guess he’s learned something over the past few years.

To Nathan and Suzie

## Table of Contents

List of Figures .....	xii
List of Tables .....	xvi
1 Introduction .....	1
2 Background .....	7
2.1 Perception of environmental information.....	7
2.1.1 Direct Perception .....	8
2.1.2 Indirect perception .....	9
2.1.3 Using several senses for locomotive guidance .....	10
2.2 How do visually impaired individuals view their environments? .....	11
2.3 The basics of sound localisation.....	13
2.3.1 Fixed position studies of localisation.....	14
2.3.2 Motional Perception of Localisation.....	19
2.3.3 Distance Perception .....	20
2.3.4 Localisation by individuals who are visually impaired .....	21
2.4 Using sound in the design of auditory interfaces .....	22
2.4.1 Earcons.....	23
2.4.2 Auditory Icons .....	23
2.4.3 Audification .....	23
2.4.4 Sonification.....	24
2.4.5 Behavioural Responses to auditory elements of context .....	24
2.4.6 Multivariate Icons .....	25
2.5 Echolocation .....	26
2.5.1 Echolocating mammals.....	27
2.5.2 Human Echolocation.....	29
2.5.3 Spectral content of echoes .....	30
2.5.4 Factors that affect echolocation .....	33
2.5.5 Echolocation Summary.....	34
2.6 Obstacle Detection.....	34
2.6.1 Primary Mobility Aid.....	35

2.6.2	Devices to increase preview distance – Secondary Mobility Aids .....	39
2.6.3	Device Summary .....	46
2.6.4	Navigational Aids .....	47
2.7	Summary of research to date .....	48
2.8	What are the next steps? .....	51
2.8.1	What must be displayed to a visually impaired traveller to provide sufficient information for obstacle avoidance while attempting to minimise the user’s cognitive load? ..	52
2.8.2	How can auditory information be displayed to allow for direct perception of an environment? .....	53
2.8.3	Can audification in the form of down converted ultrasound provide adequate information for detection and localisation of environmental obstacles? .....	54
2.8.4	Does the use of audification allow for direct perception of environmental obstacles without training? .....	55
2.8.5	How is audification influenced by the manner in which the information is retrieved from the environment? .....	55
2.9	Summary .....	56
3	Designing displays for auditory systems .....	57
3.1	Cognitive systems life-cycle development of auditory display .....	58
3.2	Designing for attention with sound .....	59
3.3	Summary .....	63
4	Development of an auditory display .....	65
4.1	Problem identification .....	65
4.1.1	The vOICE system .....	66
4.1.2	The KASPA .....	67
4.1.3	The Sonic Pathfinder .....	67
4.2	Statement of the problem .....	68
4.3	Needs Analysis .....	69
4.3.1	Work Domain Analysis .....	69
4.3.2	Information Availability .....	73
4.3.3	Recipients of the information .....	74



4.3.4	Level of cognitive control .....	75
4.4	Design Synthesis.....	75
4.4.1	Defining the modalities.....	75
4.4.2	Semantic Mapping .....	76
4.4.3	Attentional Mapping.....	78
4.5	Evaluation.....	79
5	Design of Prototype Device .....	81
5.1	Conceptual analysis .....	81
5.2	Digital Downconversion.....	84
5.2.1	Intentional Aliasing.....	85
5.2.2	Doppler Effect.....	86
5.2.3	Combining intentional aliasing with Doppler to allow for audification .....	88
5.2.4	Does this mask other environmental sounds? .....	89
5.3	Hardware .....	90
5.4	Preliminary system tests .....	92
5.4.1	Evaluating click signals .....	93
5.4.2	Evaluating the echoed signal in an anechoic environment .....	95
5.4.3	Intensity differences on approach of a wall .....	100
5.5	The move to human testing .....	101
6	Comparison of audible echoes to downconverted ultrasound echoes for localisation by moving and stationary human participants .....	105
6.1	Purpose .....	106
6.2	Hypotheses.....	106
6.3	Test Environment .....	106
6.4	Lessons learned from pilot participants.....	107
6.5	Method.....	113
6.6	Results and Analysis.....	115
6.7	Discussion.....	120
6.8	Conclusions .....	128
7	Comparison of downconverted ultrasound echoes to vision for passing through apertures by moving human subjects.....	131

7.1	Purpose .....	131
7.2	Hypotheses.....	131
7.3	Test Environment .....	133
7.4	Method.....	134
7.5	Analysis .....	137
7.5.1	Time to pass through door .....	138
7.5.2	Ability to pass through midline of aperture .....	143
7.5.3	Angle of passage through the aperture.....	145
7.6	Discussion.....	148
7.6.1	Time to pass through aperture.....	151
7.6.2	Ability to pass through centreline .....	153
7.6.3	Angle of passage through the aperture.....	156
7.6.4	General Discussion .....	157
7.7	Conclusions .....	158
8	Comparison of audible signals to downconverted ultrasound signals for localisation .....	159
8.1	Test Environment .....	159
8.2	Localisation signals .....	161
8.3	PART A: Characterising speaker levels to judge direction.....	162
8.3.1	Purpose.....	162
8.3.2	Method .....	163
8.3.3	Analysis.....	163
8.3.4	Discussion .....	163
8.3.5	Conclusion .....	165
8.4	PART B: Human testing of localisation.....	166
8.4.1	Purpose.....	166
8.4.2	Hypotheses.....	166
8.4.3	Method .....	167
8.4.4	Analysis.....	170
8.4.5	Results.....	171
8.4.6	Discussion.....	180
8.5	Conclusions .....	187

9	General Discussion.....	189
9.1	Answering the research questions .....	190
9.1.1	Information requirements.....	190
9.1.2	Auditory displays for direct perception .....	192
9.1.3	Audification to avoid environmental obstacles.....	196
9.1.4	Audification virtually eliminates the need to train participants.....	198
9.1.5	Achieving effective directional audification.....	200
9.1.6	Conclusion: Can audified ultrasound be used for safe and efficient travel?.....	203
9.2	Limitations and future research .....	203
9.3	Summary.....	207
	References.....	209
	Appendix 1.....	224
	Appendix 2.....	227
	Appendix 3.....	230

## List of Figures

Figure 2.1	Distance differences between the right and left ear relative to the sound source.	16
Figure 2.2	A cone of confusion exists such that if the head is stationary, information provided to the ear can be interpreted as coming anywhere along the circumference of the cone. ....	17
Figure 2.3	Ambient sound pressure variations as a function of distance from a wall, for critical frequency bands centred on, in the top panel: 50, 100, and 200 Hz; in the bottom panel: 400, 800, and 1600 Hz. (from Ashmead and Wall, 1999).....	31
Figure 2.4	Feedforward toe clearance as a result of obstacle height detection using a cane. The * represents significant difference at the $p < 0.05$ level and error bars are 1 SEM. (from Patla et al., 2003) .....	38
Figure 2.5	Evaluation of head clearance at five positions of obstacle height, each one ten cm higher than the previous, with height 1 being at the chest of the participant (from Davies and Patla, 2004) .....	43
Figure 2.6	Methods that information can be detected in the environment and how this information is relayed to a user.....	50
Figure 3.1	Cognitive systems life-cycle development of auditory displays (Johannsen, 2004). ....	59
Figure 3.2	Process for the design of auditory displays using EID. Those parts in italics represent the considerations for an auditory model, not typically addressed in EID. SRK=skill-, rule-, knowledge-based behaviour. (from Watson and Sanderson, 2007).....	60
Figure 3.3	Skill, rule and knowledge-based behaviours based on auditory elements of context (Sanderson, 2005). ....	62
Figure 4.1	Work Domain Analysis for secondary mobility system to display spatial information.....	71
Figure 5.1	The similarities among human echolocation, hybrid system echolocation, and phased array radar system (from Pinder & Davies, 2007).....	83
Figure 5.2	The original waveform is in black. The red waveform represents a sampling rate at 2x the sample rate. The frequency is maintained. The green waveform	

	represents a sampling rate that is lower than the Nyquist frequency. The frequency is much lower than the original waveform. ....	85
Figure 5.3	Observed frequency of signal generated by the individual.....	87
Figure 5.4	The top panel represents a 40 kHz signal (black line) which is sampled at 40 kHz (green dots). The result is a line of 0 kHz (green line). The bottom panel shows a 30 kHz signal (black line) sampled at 40 kHz (green dots) which result in the 10 kHz blue signal. ....	88
Figure 5.5	Hardware components in novel system design.....	91
Figure 5.6	PSD of common clicks: clap, click and snap.....	94
Figure 5.7	Testing apparatus used for pilot experiments. ....	96
Figure 5.8	PSD of the reflections on approach of a large metal door (at 0.5 m intervals).....	97
Figure 5.9	A PSD plot of the noise floor with no receivers plugged into the system. ....	98
Figure 5.10	The average power of the signal as the individual approached a wall.....	101
Figure 5.11	Orientation of receivers on representative mannequin. ....	103
Figure 5.12	Examining distance between receivers in the “eye” direction.....	103
Figure 6.1	Apparatus used for testing human ability to determine distance. ....	107
Figure 6.2	Results from pilot subject. Measured distance as compared to actual distance shows trend toward ordinal responses. Constant error scores show judgement as being over for the short distance and under for the longer distance .....	110
Figure 6.3	Demographics of participants in the perception of distance study. ....	113
Figure 6.4	Constant error scores.....	116
Figure 6.5	Interaction effect of Distance and Condition based on raw data. ....	117
Figure 6.6	Main Effect of condition. The * represents significance at the $p=0.05$ level, the **represents significance at the $p=0.10$ level. ....	118
Figure 6.7	Main effect of distance on raw perceived distance scores.....	119
Figure 6.8	Main Effect of distance on constant error scores. The * represents significance at the $p=0.05$ level, the **represents significance at the $p=0.10$ level. ....	120
Figure 6.9	Effect of receiver orientation and rotation of head on the direction of signal. ...	126
Figure 7.1	Experimental apparatus for aperture testing within the anechoic chamber. ....	134
Figure 7.2	Participant demographics for aperture testing. ....	135

Figure 7.3	Mean time for each condition to pass through the aperture with bars representing the standard error of the mean. The * represents significant difference from the other conditions at $p<0.05$ . ....	140
Figure 7.4	Interaction effect of Group $\times$ Condition from the variability of time to pass through the aperture. ....	141
Figure 7.5	Histogram of the distributions in time for the ultrasound conditions. ....	142
Figure 7.6	Interaction effect of group by condition for the two groups based on time through the aperture. The * represents significant difference from the other conditions at $p<0.05$ . ....	142
Figure 7.7	Mean distances and variability from the centre of the aperture based on condition (aperture size is not to scale).....	144
Figure 7.8	Mean distance and variability from centre of aperture with respect to each aperture (aperture size is not to scale).....	144
Figure 7.9	Interaction effect of condition by aperture for the angle of passage through the aperture. ....	146
Figure 7.10	Main effect of condition for the angle of passage through the aperture across all apertures.....	147
Figure 7.11	Interaction effect of group by condition for the variability among participants. ....	147
Figure 7.12	Direction of ultrasound signal upon reflection in each ultrasound condition. The solid lines represent the transmitted ultrasound and the solid, the reflected. ....	152
Figure 8.1	Experimental apparatus for localisation experiment performed within an anechoic chamber.....	160
Figure 8.2	On each speaker mount, a speaker, and LED and an ultrasound transmitter.....	161
Figure 8.3	Sound pressure levels at each of the ten positions.....	164
Figure 8.4	Gain pattern for ultrasonic receiver and transmitter (Air ultrasonic ceramic transducer 400 ST/R 160 from Farnell-in-One datasheet 65326).....	165
Figure 8.5	Demographics of participants .....	167
Figure 8.6	Audiogram of the mean values of hearing level for all participants.....	168
Figure 8.7	This figure represents the three way interaction of Group $\times$ Condition $\times$ Position of the distance from the calibrated centre of the image in the direction of azimuth. Error bars represent standard error of the mean.....	172

Figure 8.8	Interaction effect of group by position.....	173
Figure 8.9	Effect of condition by position of the look direction. The grey circles are numbered in the speaker locations. The coloured symbols represent the mean look direction for all participants in each condition with standard error bars. ....	174
Figure 8.10	This graph shows the main effect of position. Positions 1 and 10 were significantly different from all others ( $p=0.0001$ ) as were positions 5 and 6 ( $p=0.0001$ ).....	175
Figure 8.11	This figure represents the two way interaction effect of position and condition when the distance from the calibrated centre of the image was determined. Error bars are standard error of the mean.....	177
Figure 8.12	Effect of Group by Condition. The mean distance from the centre is represented by the bar accompanied by lines of standard error of the mean. The * represents significantly different distances from group 1 in the forward orientation. ....	178
Figure 8.13	Main effect of position on the distance from the calibrated centre of the image to the LED in the elevation of the look direction.....	179

## List of Tables

Table 2.1	A summary of systems and methods currently used for spatial orientation and localisation. ....	46
Table 4.1	Information requirements for the ultrasound or vision system. ....	74
Table 7.1	Significant differences between groups based on the time to pass through the aperture. ....	141
Table 7.2	Main effect of group with respect to distance from midline.....	143
Table 7.3	Main effect of group with respect to angle of passage through aperture.....	146
Table 8.1	Mean difference in pixels from the calibrated centre to the LED in the look direction in azimuth. ....	174
Table 8.2	Decreased variability in the audition condition relative to the ultrasound conditions when detecting azimuth.....	175
Table 8.3	Mean difference in pixels from the calibrated centre to the LED in the look direction in azimuth. ....	178



# **1 Introduction**

Environmental perception is critical to all aspects of daily living. An individual must understand where they are within the environment to be able to assess options such as reaching, stepping, rotating and locomoting. For most individuals, the first environmental assessment is accomplished with vision. Perception occurs based on light stimulus and action follows. Gibson argues that direct perception resulting from invariants in the light stimulus are cause for action. If individuals have minimal or no vision, hearing can also be used for situating one's self within an environment. Can an individual perceive similar invariants through a sound array to allow for direct perception?

The number of individuals with visual impairments continues to increase as the population ages. The North American definition of individuals with visual impairments includes all individuals who have a visual acuity of 20/40 or worse and at present, Canadian statistics show that approximately 0.95% fit this definition (Maberley et al., 2006). Those who have a visual acuity with the best correction in the better eye as lower than 20/200 are classified as legally blind. Based on a study by Maberley et al. (2006), the age standardized prevalence of legally blind individuals within Canada is 23.6 per 10 000. These individuals who have lost their vision have difficulty maintaining their independence and mobility.

Sensory deterioration makes it difficult for individuals with visual impairments to assess their environments and generally people start to rely on aids to enhance their independence. When an individual is functionally blind, these aids typically include a support person, a guide dog or a long white cane. Before actually taking a step, an individual who has in the past looked around to see if there are obstacles in their path must now make an effort to determine if their safety is at

risk. Environmental assessment and path planning decisions must be made prior to locomotion (Patla, 1997). With vision, this can easily be performed, but after loss of vision, a person must learn a new technique for evaluating their surroundings. Environmental assessment for independent mobility is often achieved by haptic exploration.

Haptic exploration is a task that is easy to learn but difficult to master. The long cane techniques taught by orientation and mobility (O&M) instructors are often modified by the user to enable quicker scanning which in turn allows them to increase their speed of walking. This results in a significant lack of path coverage endangering the individual who is unable to detect environmental changes in the path of travel (Lagrow et al., 1997; Uslan, 1978; Wall & Ashmead, 2002). Other limitations of the long cane include the inability to detect obstacles that are not rooted to the ground such as wall-mounted bookcases and overhead signs. Long canes allow for immediate ground-based obstacle detection, but do not provide sufficient information to accurately perceive the environment.

Several secondary mobility devices have also been developed to enable perception of the environment above waist height. For the most part, auditory “pictures” are presented to the determined, diligent user who has forty or fifty hours to dedicate to training. Mapping of distance or pixel height to pitch and location to sound intensity seemed logical to the sighted inventor (Kay, 2000; Meijer, 1992). A study of individuals with functional blindness trained in the use of secondary mobility devices (Blasch et al., 1989) found that although 86% reported having a device in their home, only 46% had used it in the 30 days prior to the interview. Although no dominant reason for lack of use was given, 21% suggested that design

modifications would improve their usefulness. These devices have often been designed to provide as much information as possible, but require considerable concentration rather than allowing direct perception. Providing information that allows the individual to perceive their environments naturally may increase the usefulness of such devices.

This research seeks to provide insight into one aspect of environmental assessment. Can an obstacle be detected and avoided quickly and efficiently based on auditory stimulus with minimal risk to the user? What information should be provided in an obstacle avoidance auditory display and how is it best provided? Considerable research has focussed on the perceptual information obtained through vision, but how an individual perceives the auditory environment has largely been ignored. This thesis seeks to evaluate auditory perception of environmental obstacles through locomotion. An assumption is made throughout this research that once environmental perception occurs, the brain processes the auditory or visual stimulus for obstacle avoidance in a similar manner. Path planning and other cognitive tasks required for effective locomotion are not examined, rather the ability to perceive and avoid environmental obstacles to mitigate risk along a previously planned route are evaluated. In pursuit of this goal, there are five main questions that have been developed that, when answered, will provide significant insight into how to best display acoustic information to visually impaired travellers.

- 1) What critical information must be displayed to a visually impaired traveller to provide sufficient information for obstacle avoidance while attempting to minimise?

An auditory display is one that transforms data into sound. Being able to hear differences in the environment can enable safer more efficient travel. Design of an auditory display requires an

understanding of the information requirements to achieve these goals. One of the biggest problems with the current technology is the haphazard choice of auditory signals. A more carefully considered approach to design of the display must precede the design.

- 2) How can auditory information be displayed to allow for direct perception of an environment?

Direct perception involves the immediate response to a stimulus without cognitive mapping. In vision, the changing optic array provides sufficient information to move away from an obstacle or toward an aperture. It may be possible that direct perception can also be used to evaluate sound environments to allow visually impaired individuals to move away from an obstacle or toward an aperture. What are the best sound elements to elicit a quick and efficient response?

- 3) Can audification in the form of downconverted ultrasound provide adequate information for detection and localisation of environmental obstacles?

Audification is the direct translation of a waveform into sound (Kramer, 1994). Echolocation is a direct form of audification in that the sounds from the environment are reflected and interpreted by the traveller, but in the noisy environments of today, it is rarely used. Ultrasound can provide the same information based on echoed signals but a method to effectively display this information has not been developed. Audification allows for ultrasound information to be converted into the auditory domain directly, enabling human echolocation from ultrasound.

- 4) Does the use of audification allow for direct perception of environmental obstacles without training?

Audification is thought to be a skill-based behavioural response in that an individual can respond intuitively. According to the SRK taxonomy, individuals should be able to elicit action based on

detection of specific features. If so, an individual with minimal training should be able to effectively use this method to detect and avoid environmental obstacles.

- 5) How is audification influenced by the manner in which the information is retrieved from the environment?

Audification of ultrasound requires that ultrasound information be collected in a manner that allows for effective display. Echolocation is effective and involves the use of two receivers (ears) to collect this information. Should the receptors be placed laterally at the side of head to simulate the ears, or perhaps facing forward in the direction to obtain the same information as the eyes? To provide sufficient information to allow for direct perception, what is the best orientation of ultrasound receivers to enable intuitive response?

Once answered, these questions will provide insight into the auditory information required to allow for direct perception of environmental obstacles during locomotion.

The research presented herein showed that parallels can be drawn between the optic array and the acoustic array suggesting that direct perception of obstacles with audified ultrasound signals is plausible. Perception of distance was better through audification than through audition. Perception of aperture size was similar to that of vision. Localisation using audified ultrasound signals was equivalent to using direct sound signals. Perception of audified sound was linked to the orientation of the receivers suggesting that more accurate perceptual information was present when oriented outward, a design approach that has never been used before.

In Chapter 2 of this thesis there is an evaluation of how perceptual information is retrieved from the environment through sound. This includes direct sound signals, echoed sound signals and signals that have been processed. It also discusses devices that have been developed to provide environmental information to visually impaired individuals. The third chapter proceeds to discuss a framework in the design of auditory systems and processes to provide required information to the user and the fourth applies this framework. Chapter 5 seeks to discuss how to apply the knowledge gained through the development of an auditory interface to the application of a device. The sixth chapter provides insight into how individuals perceive distance through the auditory sense with both auditory and audified reflections. Chapter 7 evaluates how participants use audified signals in a perception-action task of aperture passage. In Chapter 8, a localisation task is performed to determine perceptual accuracy of audification. Finally, a general discussion (Chapter 9) brings all three tasks together to answer the five research questions listed above. This chapter also discusses the limitations and future areas of this research followed by a summary of the results of this research.

## **2 Background**

To be able to effectively address the aforementioned research questions, the first two sections of this chapter enable an understanding of how environmental information is obtained and processed by both sighted individuals and those with visual impairments. Section three of this chapter considers how an individual perceives and interprets direct sound sources from their environment. Section four discusses elements of context from an auditory display perspective to relay environmental information through sound. In the fifth section, reflected sound as a means to obtain environmental information is introduced. By the sixth section, the use of assistive technology to enable environmental observation is examined. Finally, the discussion turns to the directions of the research within this thesis.

### **2.1 Perception of environmental information**

How is spatial information obtained from the environment and how is it processed? Three senses can be used for assessment prior to path planning and include vision, hearing, and touch. The most common method of obtaining environmental information is that of vision. Vision involves the reflection of light off environmental obstacles which is then processed by the brain. Hearing involves the perception of vibrational information through the use of auditory nerve centres. Touch refers to the sensation based on somatosensory information. Once the information has been obtained from the environment, perceiving this information can occur directly or indirectly as suggested by Gibson (1979) and Marr (1976) respectively. Direct perception refers to Gibson's theory that locomotion occurs in response to patterns of light projection. Indirect perception refers to a bottom-up approach using cognitive processing to evaluate the surroundings to create a spatial map. It is also possible that both direct perception

and indirect perception occur either independently or concurrently in the forming of a cognitive map. These two theories are discussed in more detail below.

### **2.1.1 *Direct Perception***

According to Gibson, perception and movement are linked such that pure vision provides enough information to allow a person to process information about the spatial world (Gibson, 1958).

Objects move with reference to other stationary objects. In Gibson's theory of direct perception, the input of light reflected off objects creates an optic array. It is a general theory of locomotor behaviour relative to physical objects and argues that pure visual information alone allows an individual to collect all required information about their spatial environment. The perception of motion is thus a result of discrete perceptions of static positions. Patterns and changes of patterns in the projection of light, known as an optic array, are stimuli for the control of locomotion relative to the objects of the environment.

The optic array lends itself to the concept that object motion is always local within the field of view. When motion occurs, this produces changes in the visual field, causing a stimulus flow in the direction of heading. As one nears the optical centre, expansion occurs such that there is a greater increase in the field of view as one nears the goal. The primary goal is to reach the centre of this field unless an obstacle exists away from which flow must occur. Changes in this pattern are the stimulus for changing movements during locomotion. This approach suggests that no cognitive processing is required and no internal maps need to exist or be built prior to locomoting. Gibson argued that information is received with respect to spatial and temporal methods and can be fine-tuned with processing.



Gibson assumed that all locomotion is controlled by similar principles and this was supported by research conducted by Lee (1976). A visual tau is based on the fact that objects expand more rapidly as they get closer than those farther away. With driving, Lee found that the changing optic array gives information about time to contact while on a collision course and can be used to make judgements about changing path plans (Lee, 1976).

### **2.1.2 Indirect perception**

The other theory of perception involves a burden of computation and is known as indirect perception. This computation can occur with internal models either feedback or feed-forward. One must see an object, and then the brain must process the information before recognition occurs.

David Marr indicates that instead of being able to directly interpret the information from an optic array, the brain draws on learned experiences to create retinal images by the brain in the same way that images are processed in a computer vision scenario. Initially, the brain creates a monochromatic primal sketch which is based on zero-crossings, changes in intensity, and edge detection (Marr, 1976). This is the initial stage in taking the images from the two eyes and drawing a stereoscopic image which also eliminates any false targets. This is replaced with a complete primal sketch which is a result of performing concurrent analyses (crossings, intensity, and edge detection) from multiple orientations. The 2 ½ dimensional sketch is the next processing step in which the surface contours and properties are added (Marr & Hildreth, 1980). Surface information is portrayed using needles to indicate directional characteristics. Finally a 3D model which portrays the volume of the image is created using information about the centre of mass, size and principle axes of symmetry.

Both theories of direct and indirect perception have been developed for and applied to vision. The body of work relating direct and indirect perception with respect to hearing is minimal. After examining collision avoidance in the optic array, Lee conducted studies of bats to determine if a changing acoustic array allowed for similar judgements to be made (Lee et al., 1992; Lee et al., 1995). Lee was able to draw direct comparisons between the visual tau and an acoustic tau suggesting that invariants exist in the acoustic array, similar to the optic array, allowing collision avoidance.

Very few studies have taken this to the human level due to the inability to evaluate differences within the auditory array (Stoffregen & Pittenger, 1995). There has been some evidence of invariants in acoustic arrays that parallel those of the optic array, including acoustic tau and motion parallax, suggesting that perhaps direct perception can also be achieved through sound. As with the bat studies, acoustic tau refers to the time to collision of an observer with an approaching source while the observer maintains a constant closing velocity (Speigle & Loomis, 1993). Motion parallax relates to the ability to perceive the changing azimuth of a sound source when an observer translates forward (Speigle & Loomis, 1993). The research supporting evidence of invariants within the acoustic array has relied on point source sound sources rather than reflections. A more ecological approach to environmental assessment through sound is necessary to further evaluate possible relationships between indirect and direct perception of auditory stimulus.

### ***2.1.3 Using several senses for locomotive guidance***

Gibson argues that optical stimulation is the main control for locomotion because in total darkness locomotion ceases. Gibson's argument that visual perception is inherent to locomotion

can be contradicted by examining environmental evaluation using other senses. Bats and dolphins have the ability to use sonar for identifying objects, and evolution has allowed these animals to develop an auditory perception method (echolocation with sonar) that makes them better able to interpret sounds than humans. In the past, humans without vision developed techniques of echolocation of sound to determine the presence of obstacles. Individuals who are termed totally blind do not have light perception, whereas those with functional blindness are able to detect the presence of light. As the number of noises in the environment increase (traffic, computers, telephones, and higher frequency lines) the audible detection methods of individuals who are functionally blind are slowly becoming obsolete. Other senses can be used for obstacle avoidance by humans and although these can be very precise they are not as accurate as vision (Patla et al., 2004).

As suggested, some individuals have a sensory deficit and must rely on other senses to guide them. The next section discusses how individuals with visual impairments collect information from their environment to enable locomotion.

## **2.2 How do visually impaired individuals view their environments?**

Individuals who are totally blind do not have access to visual stimuli suggested by Gibson (1958), thus one must account for other relationships with nature that provide information. People who do not have the sense of vision have developed methods of evaluating surrounding environments to allow them to locomote effectively. These rely on senses of touch and hearing.

Facial vision (Worchel & Dallenbach, 1947) suggests that the individuals who are totally blind have a sense of their environment which allows them to detect otherwise unnoticeable environmental changes. In the mid-twentieth century, there were two postulates defining facial vision. The first was that pressure changes on the skin, specifically of the ear, could detect small pressure variations in the presence of obstacles, whereas the second option depended on changes in echoed sounds to detect obstacle presence. A series of experiments were conducted with participants who were totally blind and those with full vision to examine the method by which they detected obstacles (Supa et al., 1944a; Worchel & Dallenbach, 1947). These participants walked varied distances and were able to detect an obstacle as far as 24 feet away when they wore shoes on the hardwood surface. When performing the same task with stockinged feet, the ability of all was greatly reduced. On the other hand, while facially masked with cloth, there was only a slight decrease in performance as compared to the shoes condition. It was concluded that in fact auditory echoes were the only means by which a totally blind individual could “see” obstacles and sensitivity to air pressure differences by the skin was not accurate enough to detect obstacle presence.

One reason that individuals with functional or total loss of vision may be more secure in a man made environment is that there are recognisable patterns that provide them with clear paths (Strelow, 1985). Industrial development has created patterns in nature which provide individuals with regularities to enable them to navigate using direct signals. The visually impaired have gained significant independence with the industrial changes in cities (Strelow, 1985). Sounds such as traffic, echoes between buildings, and temperature differentials as one passes the doorway to a building are all indicators of the environmental layout. Cognitive information is

provided by means of spatial layout and redundancy such as parked cars, hedges, and lampposts which form a row to follow. Individuals with vision loss are aware of these continuities, and do not have to concentrate as carefully on sensory information if they can rely on man-made environments.

Both totally and functionally blind individuals learn to follow specific paths and are able to do so with minimal assistance. For the most part, these methods of evaluation require previous knowledge that must be attained to attune oneself with the environment. In environments that are unfamiliar to individuals with functional or total blindness, many have learned to evaluate their surroundings by touch. The problem with using touch is that the person must make contact with an obstacle before avoiding it. If instead, they rely on auditory information to evaluate the environment, information can be processed without the need for physical contact. Sound can be used to localise and provide information about object locations. The following chapter will discuss how sound is localised.

## 2.3 The basics of sound localisation

*If you cause your ship to stop, and place the head of a long tube in the water and place the outer extremity to your ear you will hear ships at a great distance from you.*

— Leonardo da Vinci, 1490.

Sound is a variation of air pressure resulting from the vibration of an object including human vocal cords or an object moving. This pattern travels in the form of a wave such that the pressure of the sound wave as it travels in the x-direction is as follows:

$$P(x, t) = P_o \sin\left(2\pi f t - \frac{2\pi x}{\lambda}\right) \quad 2.1$$

where  $P_o$  is the pressure amplitude at the sound source,  $f$  is the frequency in hertz,  $t$  is the time in seconds,  $x$  is the distance from the source in meters and  $\lambda$  is the wavelength in metres.

Perception of sound occurs when the varying levels of sound pressure reach the ear drum and cause oscillations which are transduced into electrical signals sent to the brain. The range of frequencies that a young adult can perceive is approximately 20 Hz to 20 000 Hz.

Spatial localisation in a given environment requires both directional and distance parameters. With these, the sound can be used to create a spatial map and thus be used for locomotion and orientation. Localisation studies fall into two broad categories: those that involve the head in a fixed position for localisation and those that allow head movement, or motional studies. These two categories will be further discussed in the next section.

### ***2.3.1 Fixed position studies of localisation***

Fixed position localisation involves maintaining the head in one orientation while detecting and acknowledging the direction of a sound. Localisation requires position information from both the horizontal and the vertical direction. The horizontal localisation (or azimuth determination) is usually achieved with information provided by interaural phase and intensity differences. On the other hand, vertical localisation is achieved through information available from the pinna (Middlebrooks & Green, 1991).

#### **Horizontal Localisation**

Many studies have shown accuracy in the horizontal plane with respect to azimuth. Frontal distance localisation tends to be more accurate than sounds at either ear, especially for obstacles

within the range of a couple of meters (Middlebrooks & Green, 1991). This is typically explained with the aid of the duplex theory.

Lord Rayleigh proposed the idea of the duplex theory in the early twentieth century (Strutt.L.Rayleigh, 1907). This theory is based on localisation of two ears in response to interaural time delay (ITD) and interaural intensity differences (IID). Interaural time delay refers to the difference in time (which also corresponds to a difference in phase) of the sound wave reaching one ear relative to the other. Figure 2.1 shows interaural time delay in that the distance  $d_r$  is shorter than the distance  $d_l$ , thus the time for the sound to reach the right ear will be shorter than the time to reach the left. In addition, the sound waves are louder in intensity at the right relative to the left causing an interaural intensity difference. The intensity of the sound will decrease as the distance from the obstacle increases. If the sound lies on the sagittal plane passing through the centre of the individual's head, the source will be heard at both ears simultaneously with equal interaural intensity and time differences. Lower frequencies (less than 1.5 kHz) are thought to be localised based on ITDs whereas higher frequencies are localised by IIDs. The lower frequencies have a wavelength that is longer than the distance between the two ears thus the same wave can reach one ear at a different time from the other and the difference between the two can be interpreted.

Sound waves can be classified as those that are near-field and those that are far-field. At distances greater than one meter away, sound waves can be termed far-field and modelled as planar. In these situations there are deemed to be no differences in phase between the ears, but intensity is critical. Those sounds which are generated within a meter of the head are termed

near-field and sound waves dissipate spherically and are curved relative to the head. A “shadow” is created which further attenuates the sound as it reaches the opposing ear. When examining localisation in this field, phase differences must be taken into account.

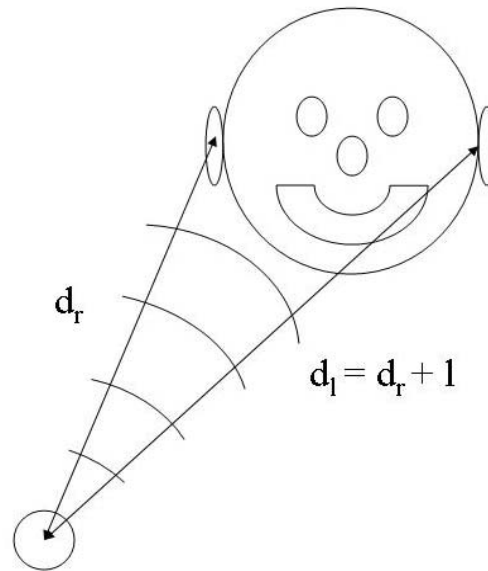


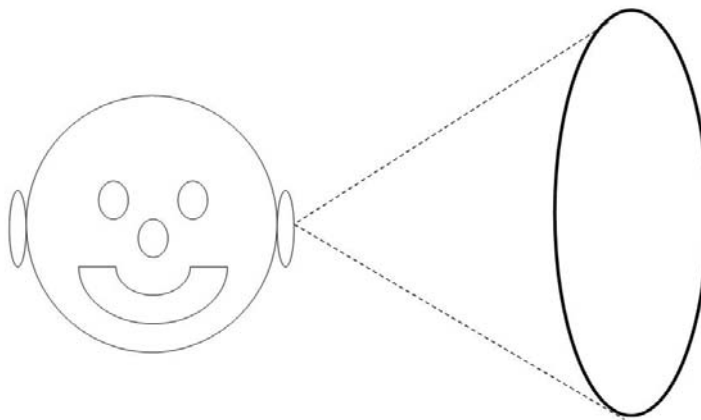
Figure 2.1 Distance differences between the right and left ear relative to the sound source.

Based on the duplex theory, signals in front of the individual will sound the same as signals immediately behind the individual in the vertical plane and the same distance away. The points from which a given perceived sound could originate lie on a circle at a constant distance from each ear. A cone of confusion exists such that a sound can be perceived to be occurring anywhere along this circle and at any range (Figure 2.2). Without additional input of the head movement, this sound is ambiguous (Middlebrooks & Green, 1991). Humans are able to differentiate between front and back when hearing broadband sounds such as clicks or bursts causing many to believe that there is a significant effect of the pinna or outer ear that contributes



to localisation. Head-related transfer functions became promising models to better describe the duplex theory when Batteau introduced the theory that filtering effects occur in the pinna or outer ear (Batteau DW, 1967).

Hearing aid research has shown that azimuth perception can be reversed binaurally, such that the azimuth signal is switched and the sound appears to come from the opposing side. This is achieved by electrically switching the input to hearing aids such that the left head acoustic cues, phase and intensity differences are applied to the right and vice versa (Hofman et al., 2002). Hofman found that after several weeks, localisation performance was similar to that prior to the change. This shows the plasticity of the human brain to be able to adapt to new environments very quickly.



**Figure 2.2** A cone of confusion exists such that if the head is stationary, information provided to the ear can be interpreted as coming anywhere along the circumference of the cone.

## **Vertical Localisation**

Lateral localisation involves pinpointing a sound that is heard in a plane parallel to the ears of the observer with resolution being better in front of the observer than at the periphery. Vertical localisation, on the other hand, is ambiguous especially if the head is maintained in a stationary position rather than turning to disambiguate sounds. The cone of confusion also contributes to difficulties in discerning elevation. The human ear compensates for this ambiguity with the shape of the ear. As suggested for lateral localisation, the shape of the outer ear has filtering effects that allow a user to differentiate between sounds of different vertical position. The effectiveness of the pinna has been shown to enable localisation even in the case of one ear canal being blocked or plugged with a mould, or in the case of an individual deaf in one ear (Hofman & Van, 2003).

Head-related transfer functions (HRTFs) have been proposed to explain the filtering effects of the pinna. Time and intensity differences are a result of differing path lengths, amplitude differences and diffraction due to the torso, head, and external ear and these can be modelled as functions that are dependent on the pinna shape (Brown & Duda, 1998). These transfer functions, one for each ear, give a ratio between the sound pressure level at each of the eardrums and the free-field sound pressure level (that would occur theoretically at the centre of the head without the shadowing effects caused by the head) (Brown & Duda, 1998). HRTFs are specific to each individual (Wightman & Kistler, 1989), but Kulkarni and Colburn have found that effective localisation can be achieved with general HRTFs (Kulkarni & Colburn, 1998). Both Wightman and Kistler (1989) as well as Blauert (1983) found that simulation of the free-field using headphones resulted in increased ambiguity in the front-back confusion which could be

corrected with HRTFs. The one difficulty when applying HRTFs (usually in virtual reality systems) is that as the individual moves the head, the HRTF must be updated to reflect the new ear position measurement relative to the virtual sound arising from the new orientation.

HRTFs can be applied to enable better localisation when the head is in a stationary position, but often people move the head to provide more information about the location of a sound source. The research in localisation while allowing head movement has been limited in comparison to the static hearing research, but can provide a more ecological approach. Motional theories are discussed in the next section.

### ***2.3.2 Motional Perception of Localisation***

All localisation studies presented thus far have presented a stimulus to an observer with the head maintained in a stationary position usually with a device to fix the head in a stationary position. Turning the head toward a source has been shown to be more effective at localising sound sources (Middlbrooks and Green, 1991) than stationary localisation. Naturally, individuals move their head to hear the source of the auditory stimulus (Munhall et al., 2004). Any cone-of-confusion ambiguities can be resolved using head motion. Although considerable research has been done in the area of static localisation due to the increased requirement for surround sound and virtual systems, there has been very limited research in the area of head motion for localisation. Thurlow and Runge performed experiments that evaluated the amounts of rotation, pivot, and tilt (the three possible directions of motion in the Cartesian plane) when localising (Thurlow & Runge, 1967). They found that rotation of the head was most often used to disambiguate sounds followed by pivoting to determine elevation. Blauert (1983) suggests that if a person is given a sound to localise, then the sound is replayed ten seconds later, the

orientation of the head will be in the direction of the sound. Thus, sound localisation based on head movement is thought to be very effective.

Individuals who require hearing aids must be retrained to localise as they do not have the added advantage of the information from the pinna (Neuman, 2005; Sebkova & Bamford, 1981). Since the motion of the head cannot be continuously monitored for people with hearing aids, individualised HRTFs cannot be applied. Retraining of individuals to use head movement for localisation is necessary to achieve localisation. Thus, localisation with hearing aids depends largely on the ability to reorient the head for localisation, rather than the shape of the ear.

Sound is not only used to provide information about the horizontal and vertical location of obstacles, but also the distance away. How individuals use sound to judge distance follows in the next section.

### **2.3.3 Distance Perception**

Visual perception of distance has been measured in many studies (Loomis et al., 1993; Loomis et al., 2001); (Bigel & Ellard, 2000). Loomis *et al.* (1993) evaluated visually guided locomotion of both individuals with functional blindness and those with full vision. This study showed that there was little difference between functionally blind and non-blind individuals when performing navigation tasks including pointing to targets and retracing multi-segmental routes. Visually directed walking toward a target that is previously viewed has also been studied extensively (Beall & Loomis, 1996; Bigel & Ellard, 2000; Philbeck et al., 1997; Philbeck et al., 2001).

Visually directed walking refers to the previous viewing of a target then allowing the participant to walk blindfolded to the target. Philbeck et al. (2001) determined that the accuracy in distance

measurement is dependent on preview using visual sense prior to locomotion. For short distances, it is generally accepted that short visual previews allow sufficient information to estimate distance accurately.

When sound alone is used to judge distance, the results are not as promising as in combination with vision (Zahorik, 2001). In a study by Zahorik, two groups of listeners were asked to judge distance based on sound. The group that was permitted vision in addition to sound was better able to judge these distances. Mershon and King have examined changes in intensity as predictors of auditory information to judge sound cues and found that changes in intensity do not necessarily allow for absolute distance determination (Mershon & King, 1975). Philbeck and Mershon further studied distance determination and found that increased intensity of speech as a signal increased the perceived distance of the source (Philbeck & Mershon, 2002). Generally, distance determination from direct auditory signals can be performed, but is not as effective as with vision.

The discussion of localisation has been broad-based in nature applying to general sound localisation. The next section discusses localisation by those who have visual impairments.

#### **2.3.4 Localisation by individuals who are visually impaired**

It was initially believed that the auditory ability of individuals with functional and total blindness is superior for localisation than individuals with vision (Lessard et al., 1998). Roder *et al.* found that this was the case for totally blind individuals, but only in the peripheral field as opposed to the central auditory space, whereas Zwiers et al. found that the two are comparable (Roder et al., 1999; Zwiers et al., 2001). Lewald argues that absolute reference determination in the vertical

direction by individuals with congenital blindness is not as accurate as those with vision, whereas for relative position localisation the two perform similarly (Lewald, 2002). On the other hand, since horizontal localisation naturally has an absolute position that relates to the median of the head such that the interaural time and intensity differences for a source sound in front are the same, individuals with blindness perform as accurately as their visual counterparts. Since the pinna develops without any reference in the vertical direction, “calibration” for individuals with congenital blindness with respect to elevation does not occur. People with full vision are better at elevation detection as they have the visual horizon to allow them to calibrate their sense of hearing.

Understanding how sound is used to localise both in azimuth and elevation is important to understand in the design of auditory interfaces for those with visual impairments. In addition to understanding localisation, the elements of context to enable effective sound display are further discussed.

## **2.4 Using sound in the design of auditory interfaces**

There are two types of sound that can be used to convey information: speech and non-speech sounds. Speech refers to combinations of sound that work together to convey a specific message. Non-speech sounds are tones that are mapped to data and must be learned. Speech information has been shown to be more difficult to process than non-speech signals (Graham, 1999). Also, speech signals are language specific allowing interpretation by only those who understand it. Further discussion will be limited to those that do not require speech in the interface design.

As the elements of context in auditory interface design differ from that of visual interface design, this section discusses the terms used in the auditory interface domain. As with visual displays that use icons, salience, colour etc. there are forms of reference for auditory interfacing. These include auditory icons, earcons, audification, and sonification which will be discussed below.

#### **2.4.1 Earcons**

Earcons are abstract musical tones that are usually represented in hierarchical form to relay information. These tones must be learned by the user in order to be useful in interpreting the information they represent. Earcons can be used in combinations to display multiple meanings, in the same manner that phrases are a combination of words. Earcons have been shown to improve efficiency of tasks (copying, pasting, editing) when accompanying visual icons in computer displays (Brewster, 1997).

#### **2.4.2 Auditory Icons**

Auditory icons represent a sound “image” of the object or motion to which it is referring. This is a direct comparison to visual icons. A heartbeat sound can be used for monitoring pulse information. The heartbeat is well-recognised, especially by physicians, and provides more interpretable information than that of a number representing pulse (Sanderson, 2005). Testing has shown that learning and response times are faster for auditory icons than for other signals such as speech warning (Graham, 1999).

#### **2.4.3 Audification**

Audification represents direct translation of physical energy into audible sound. A Geiger counter is the typical example as it clicks every time energy is received. Seismic data have been

presented very effectively using audification as the frequency of ground vibrations can be increased to be within the auditory range (Dombois, 2002). Effective presentation of these data with audification may be a result of the fact they share a physical phenomenon (frequency) that is readily displayed in this format (Eldridge, 2006).

#### **2.4.4 Sonification**

Lastly, sonification is the mapping of data streams onto auditory dimensions (Kramer et al., 1997). The use of sonification to detect sleep apnoea has been proposed (Ballora et al., 2000). This system maps heart rate variability to sound. Each interbeat interval is mapped to a specific pitch. As the heart rate increases, the pitch also increases. To notify the medical personnel of larger interbeat intervals (larger than 50 msec), a timbral annotation or “tinkling” sound is heard. This auditory mapping complements the visual display and can be used to better understand how the frequency of heart rate oscillation sounds in a given environment.

Elements of context are methods of applying sound to interfaces, but how does a designer decide which ones would elicit the intended response? The behavioural responses to the aforementioned elements of context are discussed below.

#### **2.4.5 Behavioural Responses to auditory elements of context**

According to Watson and Sanderson (2007), audification and sonification are skill-based behaviours such that they reflect “everyday listening” and are object focussed. These types of sound allow for advanced sensorimotor movements rather than requiring stored rules and planning. Earcons are representative of “analytic everyday listening” and are rule-based behavioural in nature. Finally, there is a realm of “knowledge-based behaviour” based on some



earcons and sonification which represent “musical listening” which is abstract and analytic. These will be further discussed in the third chapter on designing auditory systems.

#### **2.4.6 Multivariate Icons**

Individual elements of context can be used to represent a given environment, but additional information can also be added. Much the same as multivariate icons can be created in visual displays, sounds also can be multivariate in their display. Nomic mappings involve the correlation of a sound with an event (Coward & Stevens, 2004) and can be “loaded” with information dependent on the physics of the situation. Loading the auditory signal can be carried out by mapping iconic categories of acoustic parameters such as changing the tempo, frequency, timbre, amplitude, and spatial cues of the tone. This is similar to creating “multivariate icons” for visual systems as several different variables are displayed by a single sound. Larger objects tend to exhibit low frequency pressure changes are usually perceived as low pitch and slow tempo relative to smaller objects. Tempo can be used in an iconic form to give a sense of urgency. Increased pitch gives the perception of being louder, thus perceptually, pitch and loudness are often related. Timbre is the sound quality of complex sounds. Finally spatial cues can provide information about where to turn, or where to look to attend to audible information.

Auditory localisation can be achieved with directional cues instilled in the design of the system. Coward and Stevens (2004) have found that increased salience is achieved by direction parameters such as IID or ITD which can be used to obtain distinct directional information for localisation. In cases where vision is also present, this can be used to draw attention to a specific focal point.

These available elements of context offer rich displays that provide a significant amount of information especially in combination with visual displays. Careful consideration must be given to the design of the auditory display to ensure that the information is presented in an effective manner and to reduce the possibility of information overload.

The ability to localise based on direct sounds is only one piece of the spatial puzzle. Both individuals with vision and those with loss of vision use auditory echoes to establish location and distances within their environments. The ability to use these reflections for localisation is known as echolocation.

## **2.5 Echolocation**

*[The boat captains] used to get their position by echo whistling. They'd give a short whistle and estimate the distance from the shoreline by the returning echo. If the echo came back from both sides at the same time they'd know that they were in the middle of the channel. They could recognize different shorelines by the different echoes - a rocky cliff, for example, would give a clear distinctive echo, whereas a sandy beach would give a more prolonged echo. They could even pick up an echo from logs.*

*The Vancouver Soundscape, 1974, p. 17.*

Localisation and distance perception from direct signals are useful in the man-made environments of today, but the concept of using reflected echoes to localise and determine distance is a learned technique that individuals with functional blindness have used for centuries. This section discusses echolocation as a means to obtain information from the environment. Animals have sophisticated echolocation techniques and these are briefly discussed. The use of echolocation by individuals who are functionally blind is then examined. Past research on characteristic signals of echolocation is visited in the remaining section.

For detection of sound echoes, there are several requirements: the presence of a transducer, a medium, an obstacle, a receiver, and finally a detector that processes the information. In a typical environment for human echolocation, the transmitter is a foot striking the ground, a vocal click, or a clap that transmits waves omnidirectionally through the air. These waves reflect off walls, the ground, and any other obstacles in the environment before returning to each ear with an independent time and intensity.

Echolocation capitalises on the concept of an optic array as suggested by Gibson (1979) in that environmental differences can be identified based on changes in the sound array when reflected back to the individual. If an individual notices a change in the auditory response of the reflection of sound in the environment, modifications to path planning can be made accordingly. An acoustic tau may be comparable in sound to a visual tau in vision (Lee et al., 1992). Lee et. al. (1992) found that an acoustic time-to-arrival variable affects the braking behaviour of bats when using echolocation to navigate relative to a sound-reflecting surface. An acoustic tau can thus inform a user of the time to arrival of a reflecting surface in a similar way to the visual tau, an optic variable that provides time-to-arrival information about light reflection.

Past research on the abilities of bats to echolocate has provided insight for the development of secondary aids. The use of echolocation by bats and other echolocating mammals is discussed below.

### ***2.5.1 Echolocating mammals***

Echolocation is most often referred to in a sonar sense to discuss the responses of bats. A significant body of knowledge exists identifying the signals that bats use to communicate with

each other (Lee et al., 1995) and the signal characteristics of different species of bats. Bats have also been studied to evaluate the concept of direct perception when responding to a reflecting surface (Lee, 1995). Unlike humans, bats have large pinnae that they can move together or independently to allow them to determine direction and distance.

Another mammal that uses echolocation to catch prey is that of the dolphin. Until recently, limited research has been performed on the sonar techniques of dolphins as the underwater environment poses difficulty to researchers. Dolphins, on the other hand, do not have pinnae that can be moved for directional orientation determination. The inner ear of the dolphin is similar in function to that of the human ear (Au, 2004; DeLong et al., 2007a) though dolphins have a greater ability to interpret higher frequency signals. These mammals send out broadband signals and capitalise on the echoes to catch their prey. Studies have shown that both dolphins and humans can discriminate among objects of different size and shape based on echoed information (DeLong et al., 2007b; DeLong et al., 2007a).

The ultrasound detection system of whales (Au et al., 1985) also relies primarily on Doppler. As an animal swims looking for prey or friends, the whale transmits a signal which echoes back and is used to determine accurate distance. As the animal approaches the prey, it increases the time between clicks to allow for increased resolution of position information.

Though not as developed as that of other mammals, human echolocation allows individuals to spatialise within environments. The following section discusses human echolocation.

### **2.5.2 Human Echolocation**

*Mobility instructors discourage echolocation, especially clicking. While training with my first dog, I forgot myself and clicked to determine if I was near a pole. The instructor told me that my dog would be taken from me if I continued to make "those sounds," that they served no purpose, they made blind people objects of ridicule. And furthermore, I'd confuse the dog. I stopped clicking until I returned home!*

(Feinstein, 2001)

Prior to and including the first half of the twentieth century, individuals with vision deficits were taught how to locomote effectively using echolocation. Echolocation involves the use of reflected sounds to localise obstacles in a given environment. With the advent of the cane and the introduction of guide dogs, this skill is no longer taught and is believed to be “socially unacceptable” drawing too much attention to the individual. The basis for echolocation is that sound made by an observer is reflected off surfaces in the environment. There are several factors that influence the reflections off obstacles including the sound source intensity and duration, the spatial relationship of the obstacle and the presence of background noise (Kish, 1995). The ability to use acoustical perception of obstacles in the environment occurs at a very young age. Congenitally blind children from 5-12 years of age have been found to be able to detect obstacles in their pathway with only the presence of auditory stimuli resulting from their own footsteps and acoustic environment (Ashmead et al., 1989).

Ashmead et al. (1995) found that movement toward sounds increased the ability to localise distances. Rosenblum et al. (2000) also found that moving toward the obstacle while echolocating enabled better distance perception than localising in a stationary position. The study by Rosenblum et al. allowed participants to use their own form of echolocation, taps, clicks or verbal speech. Unfortunately, the training encouraged localisation by moving, thus did not

accurately present a controlled test environment. If training is to occur, it is important to provide similar training in all conditions.

Human echolocation has been shown to be possible, but what characteristic sounds enable effective collection of environmental information? Some research towards answering this question follows.

### **2.5.3 *Spectral content of echoes***

All sounds are made up of sinusoidal waves, the combination of which is often complex and often not periodic. To be able to analyse the contribution of certain components of the sound, Fourier analysis can be used to identify the key harmonics or the spectral content of the sound.

Schenkman and Jansson (1986) evaluated a variety of white canes typically used by individuals with functional and total blindness for detecting obstacles and providing auditory stimulus by tapping. They found that there were no differences in their ability to identify obstacles based on the spectral content of auditory stimuli produced by the cane (Schenkman & Jansson, 1986).

Walraven on the other hand, found that the variation in the spectra was important when different materials were tapped (Walraven, 1982). Schenkman and Jansson (1986) extracted the significant frequency bands of four different materials (concrete, sand, linoleum, and asphalt) and used various canes for detecting the spectral content. They found that canes giving the most acoustical information were those constructed of fewer pieces and longer. Schenkman and Jansson (1986) suggest that studies evaluating the usefulness of sounds for echolocation should include synthetic sounds with systematic variation to identify differences in spectral variation.

Ashmead and Wall were able to show through acoustical modelling that the ambient sound spectrum shifts toward low frequencies near a wall (Figure 2.3) (Ashmead & Wall, 1999). Testing adults and teens, it was shown that at a distance of 0.5 meters from a wall, individuals lacking vision could detect these spectral variations and use the information to follow a path. This study did not attempt to model the responses in the presence of doorways and how an individual might attempt to detect and proceed through a doorway using acoustical information.

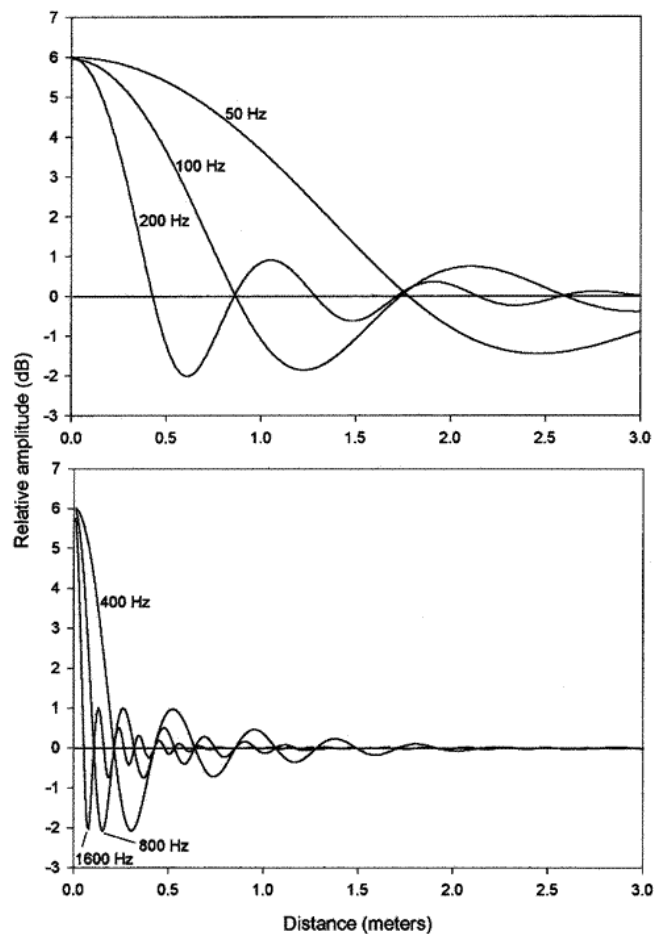


Figure 2.3 Ambient sound pressure variations as a function of distance from a wall, for critical frequency bands centred on, in the top panel: 50, 100, and 200 Hz; in the bottom panel: 400, 800, and 1600 Hz. (from Ashmead and Wall, 1999)

Strelow and Brabyn (1982) studied perceived acoustics in the vicinity of a wall. Functionally blind individuals were asked to follow a wall. Next, the shoreline (wall) was changed to be a series of rows of smaller poles to simulate a row of hedges, or a row of cars. They found that individuals with functional blindness did not perform any better than blindfolded control subjects in the presence of smaller objects, though they did outperform their control counterparts in the presence of a wall (Strelow & Brabyn, 1982). This study showed that individuals who have lost visual perception of objects likely develop a sense for echo skills to achieve locomotor control, but echoes from larger objects may provide better control information than smaller ones. Broad frequency spectrums available through echoes on a wall allow the individual sufficient information for obstacle detection. The lower frequencies available from the reflection of sounds from the poles would be undetectable thus unavailable to aid in locomotion.

There is evidence that different shapes have specific acoustical properties that are distinguishable using echolocation (Kunkler-Peck & Turvey, 2000). Kunkler-Peck and Turvey (2000) showed that suspended thin plates hit with a steel pendulum in circular, triangular and rectangular form are distinguishable by individuals with normal hearing as well as the ability to distinguish among materials including steel, Plexiglas and wood. More recently, DeLong et al. (2007) have shown that humans can discriminate between hollow cylinders and spheres made of steel, aluminum, brass, nylon and glass based on reflected echoes. They suggest that the timbral characteristics, duration of echo and pitch give sufficient information to allow for determination of type of material and shape.



In addition to spectral properties, there are several other factors that affect echolocation and a discussion of these factors can be found below.

#### **2.5.4 *Factors that affect echolocation***

Echolocation can be very effective if used in everyday environments, but can also be difficult to apply to all situations. Kish suggests that there are five key factors that can affect the signal (Kish, 1995). These include the quality of the signal, surface characteristics, ambient noise, quality of hearing and the degree of vigilance. The first three can affect any sound in any environment and are not specific to echolocation. The quality of the signal is highly dependent on the user. The clicks themselves must be loud enough that they can travel a distance and return. Signals that are generated near the ears provide echoes that are most easily interpreted (if a ball is kicked to a wall, it returns to the feet). Sound will reflect to the location at which it is transmitted. Surface characteristics affect the quality of the return signal. If the material is soft, most of the sound will be absorbed rather than echoed. Ambient noise in an environment prevents the echoes from being heard. If an individual is hard of hearing, they typically lose the perception of the higher frequencies. The echoes cannot be fully understood especially as higher frequencies provide the most distinct information. Lastly, being able to echolocate to a person's full potential requires considerable training to hear nuances. Echolocation is often a self learned skill to be able to achieve basic distance understanding, but sophisticated echolocation to enable independent travel can take considerable training.

### **2.5.5 Echolocation Summary**

Generally, if permitted to use echolocation to spatialise, there is evidence that visually impaired travellers can locomote effectively and successfully. Echolocation can provide sufficient information to a traveller who is functionally or totally blind to be able to understand where multiple obstacles are located in a given environment in addition to being able to localise at a distance without the need for contact. Unfortunately echolocation is not promoted as a means to effective localisation as orientation and mobility instructors deem the use of echolocation as socially unacceptable. Echoes can also be highly variable depending on the situation in which they are used. Although many capitalise on this ability, few use it as their primary means of travel due to environmental influences that interfere as well as perceived societal rejection. The primary and secondary devices that are currently used for travel have limitations that may be overcome by providing the same information to the visually impaired individual that would be used for echolocation. These devices and the benefits of each will be described in the following section.

## **2.6 Obstacle Detection**

Although echolocation has been taught in the past, it is becoming obsolete with respect to spatial orientation. One of the reasons is the development of the long white cane and other devices that can be used to detect obstacles in noisier environments. This section explores some of the methods used by visually impaired individuals to evaluate their surroundings to enable them to detect obstacles in the path of travel. First, the use of a primary mobility aid like a cane or a guide dog is discussed. Other devices known as secondary mobility devices, have been developed to detect obstacles above the waist, but these are rarely used due to the difficulties

with interpretation of the information provided to the user (Blasch, 1989). An evaluation of these devices can be found in the section following the primary mobility aids. A summary of all can be found in Table 2.1

### **2.6.1 *Primary Mobility Aid***

The guide dog is one of the most common aids of travel used by visually impaired individuals. Guide dogs can be very effective for some individuals, but many lack the discipline to train and thus effectively utilise these animals to their potential. Also, guide dogs have difficulty identifying objects not rooted to the ground in the same way that a cane does. The discussion of guide dogs will be limited to the information here as guide dogs and their companions often have a relationship beyond what can be presented in scientific terms.

If an individual does not have a guide dog, independent mobility is achieved with haptic exploration, often with the use of a long cane. The long cane is commonly referred to as a primary mobility device as it is a single travel aid that is essential for the purpose of travel. Dynamic touch involves the contribution of muscular effort to grasp the cane and relay sensory information to an individual. The cane can be used to retrieve kinesthetic information from the environment. The kinesthetic tactile sense when using a long white cane is achieved through vibration. Tapping the long cane gives information about surface elasticity and sweeping gives the frictional properties. The sound when tapping can also enable identification of the material properties of the obstacle (Schenkman, 1986; Schenkman & Jansson, 1986). Obstacles are detected by the sudden stop in movement of the long cane. Vibration is sent up the cane shaft and forces the long cane into a more vertical position (Wall & Ashmead, 2002).

Individuals with functional blindness can evaluate their surroundings by contact using cane techniques and must then make changes to their path plan (Ramsey et al., 1999; Schenkman & Jansson, 1986; Wall & Ashmead, 2002). The length of preview for a person using a cane is limited to a couple of meters thus reducing the time available to change the course of locomotion (Schellingerhout et al., 2001). Contact must occur before a change of path can be established. This endangers the individual and makes independent travel difficult. If the traveller detects an obstacle, sufficient time to brake and change direction is essential. Often, effective travellers know their route and have little expectation of additional obstacles, thus have a walking pace fairly similar to those with full vision. This does not allow sufficient time for change of direction when walking in known environments with a cane and as a result, the individual often collides with or trips over any obstacles introduced in the path.

A visually impaired individual is often taught the specific techniques of long cane motion to ensure that the path is evaluated effectively before the person takes a step. There has been recent discussion about the technique used by orientation and mobility instructors as to whether the current technique, the two point touch technique, although biomechanically efficient, is the most suitable technique (Bongers et al., 2002; Ramsey et al., 1999; Wall, 2001; Wall & Ashmead, 2002). This technique involves holding the cane in the dominant hand at the body's midpoint and out from the body so that the arm and cane form a straight line. It is then arced from side to side in rhythm with trailing limb projected footfalls to allow exploration of the surface of the immediately following step. Many people are very meticulous about learning the technique, then adapt it according to their own preferences.

The learning process while using a long cane has not been studied in great detail.

Schellingerhout *et al.* (2001) examined use of a modified cane before and after three weeks of practice and found very little difference in obstacle avoidance, drop-off detection and speed.

This indicates that modifications to the cane do not permit additional information to be detected, thus another method by which obstacle detection can be achieved is important. Wall and

Ashmead (2002) studied the learning effect in blindfolded individuals to evaluate how quickly they become comfortable with the cane. They found that after an hour of practice, individuals could master the basic techniques of cane use specifically hand position and forward velocity.

Learning with a cane comes naturally to both blind and blindfolded users (Patla et al., 2004).

This primary mobility device is a pick up and use device which requires limited learning. An obstacle detection device that could complement cane use and requires limited learning may provide increased coverage allowing risk free travel.

Wall and Ashmead (2002b) have shown that very few people utilise an ideal two touch technique. Instead, the side to side coverage is increased beyond the width of the body to ensure full coverage. As a result, the speed of the individual decreases significantly. Also, the height of the arc is modified such that the user lifts the cane between every tap (Wall & Ashmead, 2002).

This reduces the ground coverage and makes the individual susceptible to small environmental changes like potholes. Drop-offs such as curbs are very difficult to identify using a long cane as depth information is poorly interpreted. Also, the concentration and strength to maintain a good two-point touch technique requires significant energy and many individuals comment that they become tired quickly. As a result, the skills for an ideal technique quickly diminish. The average time that a cane is used in a given day is approximately 1.5 hours (Blasch et al., 1989).

What is used the rest of the time? Although not reported, a form of echolocation is more likely. Even in an environment known to an individual with visual impairments, sound is used to evaluate changes (a window being opened, a door being opened or closed). Capitalising on the sense of sound might reduce the physical stress on the user.

Generally, the long cane does not provide any information about obstacles that are above waist height. The length of the cane affects the ability to detect and avoid obstacles in that the greater the slope relative to the ground, the less preview time for obstacles that are not detected at ground level. Feedforward information using haptics provides sufficient information for obstacle avoidance when stepping over obstacles but the precision is lacking when compared to individuals with full vision (Patla et al., 2004). After an hour and a half of practice, there was a significant difference not only in accuracy (top panel of Figure 2.4), but also in precision (bottom panel).

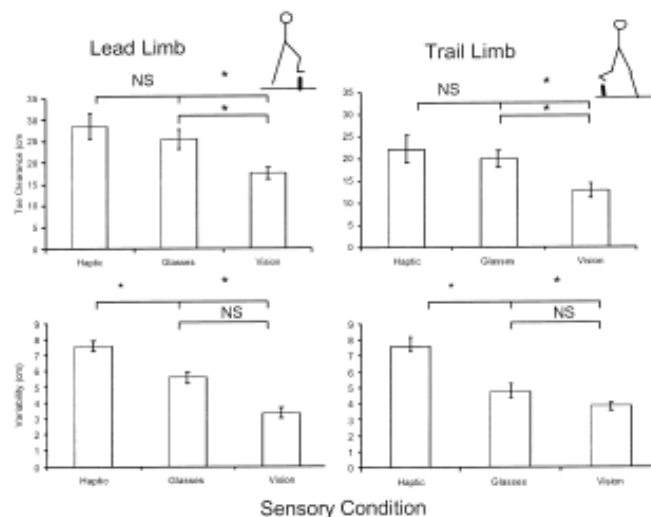


Figure 2.4 Feedforward toe clearance as a result of obstacle height detection using a cane. The \* represents significant difference at the  $p < 0.05$  level and error bars are 1 SEM. (from Patla et al., 2003)

Although haptic information does allow obstacle clearance, it does not address those obstacles that are above waist height and cannot be detected with a cane such as pedestal display cases, wall mounted bookcases and signs that extend beyond the poles on which they sit. The next section details other mobility devices that are used to provide additional information about these extruding obstacles.

### ***2.6.2 Devices to increase preview distance – Secondary Mobility Aids***

Several devices have been developed to help increase the preview distance before physically contacting obstacles, but few have gained acceptance. These devices are used in addition to a long white cane and are termed secondary mobility devices. Devices to increase preview to the user include both sonar and imaging devices (Heyes, 1980; Kay, 2001; Meijer, 1992). Sonar devices include the Trisensor and Sonic Pathfinder and convert echo information into audible information. Imaging devices use mono or stereo-imagery to reproduce an image then convert the image to binaural auditory or tactile information. This section discusses various methods of position data collection and the process by which they are portrayed to the user.

#### **Sonar Devices**

Sonar systems for the visually impaired (Kay, 2000), virtual environments (Waters & Abulula, 2007) and robots (Kuc, 2002) have largely been designed to simulate the responses of bats. Is this the most reasonable approach to take for designing sonar systems for humans? Bats have large pinnae or outer ears that they can move independently to determine direction (Walker et al., 1998). They can send out clicks that are frequency dependent and orient their ears upon approach to most effectively gather the information from the signal. Humans don't have the ability to change the direction of their pinnae, nor is interpretation of frequency sweeps intuitive.

Systems that attempt to simulate bat detection require methods to compensate for this lack of pinnae movement. Dolphins also perform echolocation but they do not have pinnae that change direction. Perhaps a more intuitive manner is required to display information, more similar to that of the dolphins, rather than one that requires processing of the signal to compensate for the limited range of motion of the human ear.

A significant advantage to sonar relative to that of echolocation is that it may not be as sensitive to the factors that affect audible echolocation. For instance, the quality of the signal can be improved by increasing the transmitter power. Since other individuals cannot hear the signal, an increase in the power does not affect other people. Absorption of the ultrasound still exists. In an auditory situation absorption results in a lower sound pressure level, making it too soft to hear. If the transmit power is increased at ultrasound, absorption will occur and decrease the signal strength, but ultrasound receivers are also more sensitive than the ears at picking up these high frequency reflections. As a result, characteristic textures can be detected more easily with ultrasound than within the auditory domain. Ambient noise in the auditory domain does not affect the reflected signals in the ultrasound domain, thus sounds that may not be heard in the auditory domain, will still generate a reflected ultrasound signal that can be relayed to the user.

The basic premise of the current sonar secondary mobility devices is that ultrasound information is transmitted by a wide-angle beam ultrasound transducer and received by another two or three transducers depending on the system. Information from the backscatter of ultrasonic waves is transmitted to the ears. There are two devices that are fairly common when referring to sonar



secondary aids. These are the KASPA developed by Dr. Leslie Kay and the Sonic Pathfinder developed by Dr. Anthony Heyes. Additional focus on these systems is further presented.

The Trisensor, later termed the Sonic Guide and more recently the KASPA, was developed by Kay in 1962 as a “wide-angle binaural” ultrasonic aid (Kay, 2000). A transmitter and two receivers are mounted on the nosepiece of a pair of glasses. Information from the backscatter of ultrasonic waves is transmitted to the ears binaurally using sonification of the signal such that interaural intensity differences represent directional differences, and pitch indicates the distance to an obstacle. This device is a continuous scanning device that provides tones about all obstacles in the environment regardless of motion of the user or look direction. As the sonified signal is not developed to minimise masking, other aspects of the surroundings cannot easily be heard. An individual using this device cannot readily communicate with those around, limiting the device solely to independent travel situations.

Easton (1992) has described the ability of the Trisensor to be more effective at detecting obstacles at distances up to 4.6 m away than sound localisation. He credits this ability to the greater abilities of the sonar to range find. This conclusion was achieved through examining the differences between the Trisensor and sounding objects (objects that emitted sounds through attached speakers). The one problem that Easton sees as a problem in creating a spatial picture using ultrasound is the lack of ability to detect “spatio-temporal invariants across perspective transformations” (surface slope, curvature, irregularities) (Easton, 1992). These invariants may not be detectable using the Trisensor device that Easton used for his testing. The Trisensor only portrays information using interaural intensity differences. This system does not consider the

effects of head related transfer functions and interaural time differences, nor does it portray sufficient information about the spatial environment to be able to effectively identify slopes and curvature, even by an individual with considerable practice. True echolocation enables travellers to interpret the information in reflected signals such that slopes and curvature can be detected by individuals with practice.

The Sonic Pathfinder developed by Heyes in 1984 is currently one of the least expensive available secondary detection systems. This device is a pulse echo digital device which uses a musical scale to represent obstacles in the path of the user (Heyes, 1983). This device uses two transmitter transducers and three receivers to cover the field of view. It prioritises the obstacle immediately in front of the user and does not provide any additional information until the obstacle is beyond the field of view. The Sonic Pathfinder uses earcons in the sense that as a person approaches an obstacle, the tone increases or decreases along an Ionic scale (the most common musical scale). Although this is acceptable for the obstacle nearest the individual, it does not represent the spatial environment as a whole. As the distance from the obstacle in the “frame of reference” decreases, the pitch of the tone decreases.

A recent study comparing directed walking towards a target that has been previously located using the Sonic Pathfinder as compared to vision has shown that the height of an obstacle is not accurately determined (Figure 2.5). The participant increases head clearance while passing underneath the obstacle such that the individual squats as low as possible to get underneath the obstacle regardless of the actual obstacle height (Davies & Patla, 2004). With vision, the participant judges the height of the obstacle and uses feedforward proprioceptive control to pass

just beneath the obstacle. This indicates that information provided to the traveller does not allow accuracy similar to that of individuals with vision. The use of the Sonic Pathfinder for judgement of height is not easily interpreted and does not provide sufficient information for effective obstacle avoidance. The main problem with localisation using this device is the lack of additional information provided to the user. This device only relays distance to the nearest obstacle with a musical tone, but provides no additional information that could be used in path planning such as the distance to the obstacle immediately beyond the present one.

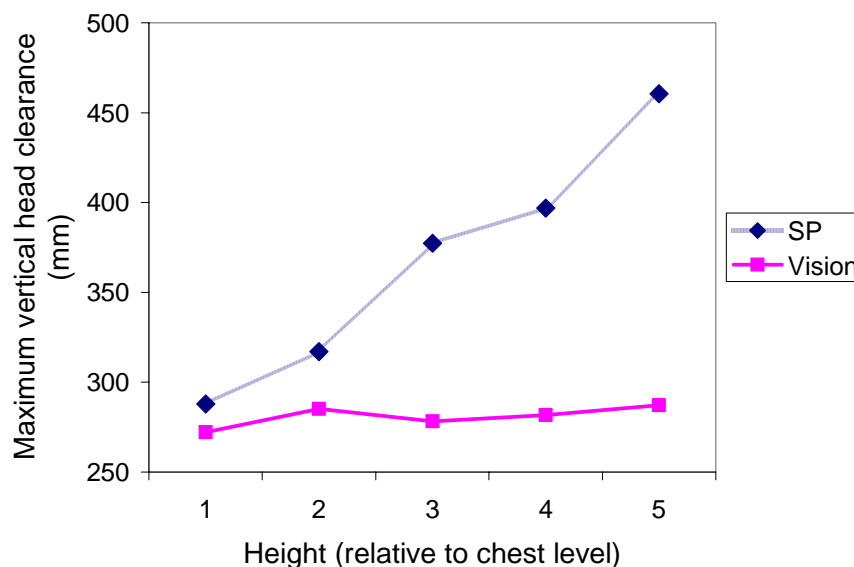


Figure 2.5 Evaluation of head clearance at five positions of obstacle height, each one ten cm higher than the previous, with height 1 being at the chest of the participant (from Davies and Patla, 2004)

Sonar signals are processed before sonifying the information or converting the signal to earcons.

Auditory images can also be sonified to provide information to the user. The next section discusses these auditory image representations.

## **Auditory Images**

There has been a trend toward development of devices to provide individuals with functional blindness with “auditory image representations”. Generally, these devices are developed for the purpose of allowing the blind to “see”. Most of this research has been based on the development of a system by Meijer (1992). This system, which has come to be known as the “vOICe” (OIC is synonymous with Oh, I see) attempts to provide the visually impaired with sufficient information to identify objects through sonification.

Cameras are used to record information in front of the user. The information from the image is then analysed and decoded before relaying the auditory information to the user in a manner determined by the system inventor, apparently without significant thought given to the interface itself. Time multiplexed mapping distributes the images in time (Meijer, 1992). As the image is scanned from left to right, the row of the image is translated to frequency and the amplitude of the sound is dependent on the pixel brightness. Each frame takes one second to scan. The sonification requires the individual to remember the previous “frame” to get the big picture. The flaw to this approach is the expertise required to cognitively analyse the information provided to the user. The vOICe system was used to evaluate the stages of perception to obtain aptitude in localisation which generally took 10 to 15 hours (Auvray et al., 2003). Continuous scanning of the environment from left to right can provide ambiguous information to the user and requires considerable concentration. Meijer himself suggests that the learning process is similar to learning a new language and suggests that a need to understand this process better is “sorely needed” so adaptation can be more efficient (Jones, 2004). Rather than require the user to adapt,

perhaps the inventor should seek to supply the information in a manner that is intuitive to the user.

Rather than capitalising only on the auditory sense, perhaps the sense of touch should also be visited. Some research in this area is now discussed.

### **Vibrotactile Images**

The ability to detect vibratory responses has shown that many variables affect perception. These include frequency, duration, direction, contact geometry, contact area, contact force, state of adaptation, context, temperature, age and pathology (Brisben et al., 1999). Environmental images have been portrayed using stimulation on the tongue (Bach-Y-Rita et al., 1970) but training is required to be able to interpret the image for obstacle avoidance. With respect to pathology, it has been shown that tactile acuity of the fingers of individuals with functional blindness is enhanced, but the actual neural mechanisms behind this response have not been determined (Goldreich & Kanics, 2003). One system developed uses stereo imagery collected by cameras that is converted to vibrotactile information supplied through the hand to enable obstacle detection (Zeilek J.S. et al., 2003). The finger closest to an obstacle produces a vibrotactile sensation which the user must then interpret. In systems that use finger stimulation to elicit a response, careful consideration must be given to the acuity response by each finger. There is a preference toward using the index finger for Braille reading (Goldreich and Kanics, 2003) and may be more sensitive thus vibrotactile forces may be detected in this finger first.

Another difficulty with tactile stimulation is the masking of auditory stimuli. It has been found that the perceived location of obstacles is often overridden by the presence of a vibrotactile

sensation (Caclin et al., 2002). If a user is presented with sound at the right and a central vibration on the fingers, the tactile data biases the perceived location and perceptually the user responds that the sound was heard at the left. This discrepancy appears to be more sensitive to the synchronisation of the input such that if they are presented at different times, they are more distinguishable (Caclin et al., 2002).

### 2.6.3 Device Summary

Method or Device	Input	Output	Strengths	Weaknesses
<b>Echolocation</b>	Clicks, Taps	Auditory echoes	Localise multiple obstacles. Detect obstacles at a distance – increased preview	Reduced effectiveness in noisy environments. Socially unacceptable.
<b>Cane</b>	Touch	Tactile vibration. Sound echoes.	Detect obstacles in given path. Helps braking of individual. In environment with echoes and little noise, can localise.	Preview distance limited. Increased exhaustion from holding cane in position. No detection above waist height.
<b>Sonar Devices</b>	Ultrasound	Ultrasound echoes	Can detect obstacles within 0.3 m. Gives priority information to objects in front.	Requires significant training. Only permits information about one to two obstacles. Relies on musical mapping.
<b>vOICe</b>	Single camera image	Auditory sound based on pitch	Ability to recognise shape using sound.	Requires significant training. Two-dimensional view therefore no depth information.
<b>Visual Stereo Imagery</b>	Stereo Images	Vibrotactile	Allows sense of sound to be unmasked	Requires training to interpret directionality of response

Table 2.1 A summary of systems and methods currently used for spatial orientation and localisation.

The cane has served as a primary mobility for many years. This is one of the most effective and well used travel aids for those individuals with functional blindness, yet cane use does not allow for detection of obstacles above waist height and the average time spent using a cane for travel in a given day is only one and a half hours (Blasch et al., 1989). This suggests that other techniques for orientation and travel are also utilised requiring no external devices, most likely sound and touch. Although secondary devices attempt to relay environmental characteristics to the user, these devices convert the data into tones or vibrations that are not intuitive or natural for the traveller. Capitalising on the localisation information currently understood and learned by all individuals from a young age may simplify the localisation system.

#### **2.6.4 Navigational Aids**

The discussion thus far has been on primary and secondary mobility devices to enable obstacle avoidance for the visually impaired. Navigational aids, on the other hand, refer to those that have been developed for the purpose of orientation and heading. These provide full spatial maps to the user including information about where and when to turn to reach a specific location. The research in this area has often used GPS for navigation (Loomis et al., 2005). There has also been some research on how to relay this information to the traveller, often with the use of speech (Loomis et al., 2002). More recently, Walker and Lindsay (2006) have examined different sonification techniques to provide heading and waypoint information to the user. They examined three different beacons to provide directional information to the user. The most effective beacon appeared to be pink noise (white noise which is low pass filtered to allow only those tones within the audible range to be evident), followed by a single 1000 Hz tone (Walker & Lindsay J., 2006). A study in which ecological interface design was used to develop a sonification interface to provide obstacle avoidance information (Davies et al., 2006) has shown that sonification can be

effective but may not necessarily provide enough information to create a full spatial map. The research into navigational aids is important, but obstacle avoidance research is important in reducing risk. Until the risk to the individual is minimised, research should focus on how to enable individuals to achieve sufficient preview distance to effectively avoid path obstacles.

## **2.7 Summary of research to date**

A summary of this background can be better understood with the aid of a diagram (Figure 2.6).

There are three senses that can be used for obtaining information from the environment, the haptic sense of touch, the sense of vision and the sense of hearing. As a visually impaired traveller, how does one use these senses to detect information from the environment and successfully avoid obstacles?

Somatosensory feedback achieved through cane use allows individuals to use the long cane in detecting obstacles below waist height. This method of obtaining environmental information is simple and convenient, but unless used effectively with a true scan of the environment immediately in front of the individual, it does not allow the individual to perceive all environmental obstructions. It also does not permit detection of obstacles above waist height leaving parks with hanging trees as high risk. Vision systems allow the collection of information from the environment through cameras, but how can this information be relayed effectively? Auditory sonification of images has not been effective and requires many hours of training. Applying vibrations to the hand requires that either the cane cannot be used or the device must be used on the opposite hand, creating a very unnatural gait. Having two hands in front while walking does not permit efficient movement. Sonar systems collect information



through ultrasound but when converted to the auditory domain they either provide information overload or a limited amount of information for obstacle avoidance. Echolocation allows efficient detection of obstacles, but can be ineffective within so many different environments. What is needed is a device that can provide intuitive information to the user in such a manner that it does not affect the individual's ability to use the cane effectively or mask their hearing while also being able to be used in environments where echolocation is ineffective.

As discussed earlier, perception of this environmental information can occur in two ways, directly and indirectly (Gibson, 1979; Marr, 1976; Marr & Hildreth, 1980). Cane use (haptic perception by vibration) allows for either direct or indirect perception. Obstacle avoidance with a cane can result without creating a full spatial map of the environment. The cane hits an obstacle and the individual side steps to avoid it, this is a form of direct perception, a reaction occurs as soon as an obstacle is detected. On the other hand, if the cane hits an obstacle, like a snow bank, and the user has to haptically feel around it, rather than side stepping, a more indirect response is required in that the individual must attempt to create a mental picture of the obstacle size. This task requires indirect perception and takes much longer to accomplish.

Hearing can provide information in a similar way. Direct perception can be achieved by listening to reflections and responding to the acoustic array. This occurs in echolocation such that a person hears changes in the environment and can readily identify where they are coming from without a need to create a mental picture of all environmental features. A spatial map, or indirect perception, could also be created by listening to echoes in the environment and determining how the echoes differ as compared to being in an environment with no reflectors.

This requires the individual to spend more time listening to a specific environment and noting which echoes are more pronounced than others. A spatial map with the location and texture of objects can be determined and the individual can place themselves within the environment. Generally, direct perception can allow for immediate intuitive response, whereas indirect perception requires more time and a mental model to be drawn.

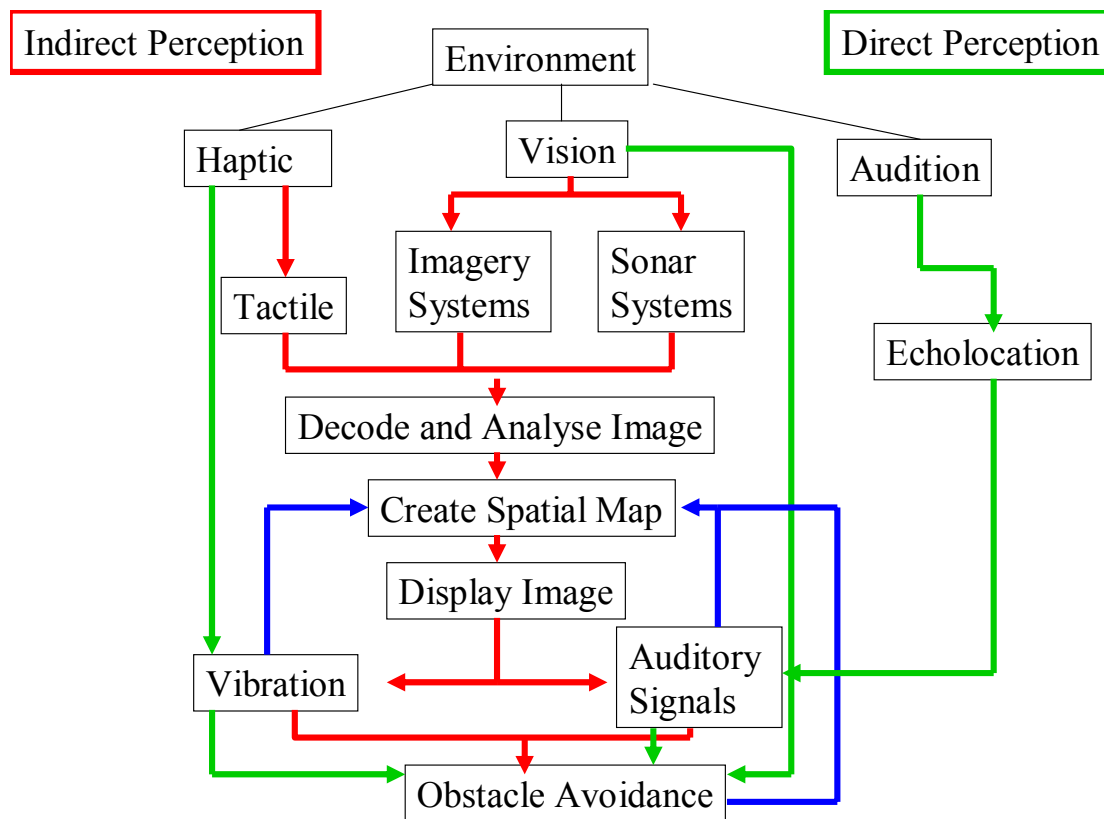


Figure 2.6 Methods that information can be detected in the environment and how this information is relayed to a user.

If these concepts are applied to the mobility systems that are currently used by the individuals with functional blindness, it is apparent that an indirect perception approach is followed.

Information obtained by either cameras or sonar systems is decoded and analysed to create an

image of the spatial environment. This image must then be displayed to the user through either the auditory sense or the sense of touch. The brain of the individual must further process this information to create their own spatial map which can then be used for obstacle avoidance. It is important to put the engineering approach aside and leave the brain to do the processing. A full visual image is not necessary for locomotion, yet engineers attempt to impose a full visual image on individuals who cannot capitalise on the visual sense (Bach-Y-Rita et al., 1970; Meijer, 1992). Individuals who are functionally blind, have developed their own sense of listening and can respond to their own mental models. Why must sighted engineers impose their visual abilities on an individual with functional blindness to accentuate that individual's inabilities? A spatial map does not need to be drawn and interpreted, let the individual draw conclusions based on stimulation directly retrieved and further use that information to indirectly create a cognitive map. Leslie Kay's KASPA (2000) provides a first step, but provides too much information that he feels is necessary. A device that is more intuitive allowing for direct perception without a need to create a full mental image of the environment is necessary to enable safe and efficient travel.

## **2.8 What are the next steps?**

Is it possible for a secondary mobility system for individuals without visual perception (either in total darkness or with functional blindness) to provide information that can be processed directly rather than indirectly? This system would require the use of the brain to process the information and create a spatial map rather than providing a spatial map directly to the user. As discussed, the current auditory interfaces do not provide sufficient information to users about their environment in an easy to interpret manner. This research seeks to examine how to provide

auditory localisation information to a user directly from ultrasound which will then permit the user to interpret the information provided to either avoid obstacles using direct perception or by creating a spatial map for indirect perception approach. How can the design of an auditory system be developed to enable direct perception of environmental obstacles? The answers to five main research questions will be sought throughout this dissertation. These are now described.

### ***2.8.1 What must be displayed to a visually impaired traveller to provide sufficient information for obstacle avoidance while attempting to minimise the user's cognitive load?***

Individuals who use a cane as a primary mobility aid need to use a sense other than tactile to provide additional environmental information. An auditory display is one that transforms data into sound. Auditory displays can offer significant information to a traveller who is visually impaired. Being able to hear differences in the environment can enable safer more efficient travel. Design of an auditory display requires an understanding of the information requirements to achieve these goals. One of the biggest problems with the current technology is the haphazard choice of auditory signals. A more carefully considered approach to design of the display must precede the design. This thesis will seek to determine the best method to display auditory signals to a visually impaired traveller by applying an interface framework (Watson & Sanderson, 2007) in the development of a prototype interface.

### **2.8.2 How can auditory information be displayed to allow for direct perception of an environment?**

Direct perception in a visual sense involves gaining information from ambient light. This results as an individual senses the environment which causes the individual to react to a stimulus. This act does not require cognitive mapping (Strelow, 1985). In vision, the changing optic array provides sufficient information to move away from an obstacle or toward an aperture. An optic array as observed by a moving animal consists of changing perspectives of invariants existent in ambient light (Gibson, 1979). One can draw the parallel to a locomotor's acoustic array that forms through changing perspectives in ambient sound. An argument for direct perception can then be made to allow visually impaired individuals to initiate action away from an obstacle or toward an aperture without the need for a full cognitive map.

Ambient sound bears similar characteristics to that of ambient light in that it reflects off environmental obstacles and if effectively perceived can provide information. Ambient sound on the other hand, is not as powerful as a source of illumination as that of light. The ears are extremely sensitive to certain frequencies (20 Hz to 20 000 Hz), but the wavelength of sound is approximately 0.017 m to 17 m whereas that of light is  $4 \times 10^{-7}$  m to  $7 \times 10^{-7}$  m. The shorter length of the wave provides for more information to be obtained in a shorter time period with vision allowing for more efficient perception than that of sound.

Images from light create two dimensional arrays, yet sound requires a temporal component to elicit a perceived array. As such, sound is transient and cannot be recorded in a similar manner to that of visual light. Sound, is also observed through outward oriented receivers allowing for

omni-directional detection whereas vision permits only the area in front of the individual to be observed. How can an ambient sound array be created to provide information to enable direct perception? What are the best sound elements to elicit a quick and efficient response without the need to create a spatial map? Possible answers to these questions will be sought through the research presented herein.

### ***2.8.3 Can audification in the form of down converted ultrasound provide adequate information for detection and localisation of environmental obstacles?***

Audification is the direct translation of a waveform into sound (Kramer, 1994). Echolocation is a direct form of audification in that the sounds from the environment are reflected and interpreted by the traveller. Those sounds that are most effective in providing this information for echolocation are on the order of 1 – 8 kHz (Kish, 1995) which is a shorter wavelength than most ambient sound. This form of environmental assessment has lost ground in recent years as parents and mobility instructors discourage its use. Ultrasound can provide the same information based on echoed signals but a method to effectively display this sound to a human effectively has not been developed. Ultrasound has a wavelength more similar to that of light and may create a more complete acoustic ambient array to enable direct perception. Audification allows for ultrasound information to be converted into the auditory domain directly, enabling human echolocation from ultrasound. Two tests will be presented to determine the effectiveness of using audified ultrasound in obstacle avoidance tasks.

#### ***2.8.4 Does the use of audification allow for direct perception of environmental obstacles without training?***

Audification is thought to be a skill-based behavioural response in that an individual can respond intuitively. Gibson hypothesised that response to environmental features is an innate behaviour such that perception and proprioception are complementary (Gibson, 1979, p157). If this is the case, auditory direct perception would enable efficient kinaesthetic response to environmental obstacles. An individual with minimal training should be able to effectively use this method to detect and avoid environmental obstacles without the need for a full cognitive map. Like the cane (Patla et al., 2002) , a secondary mobility device should require less than 1.5 hours of training. This research presents three human studies to evaluate the human response to environmental sound, both that of sound reflected off environmental obstacles and that presented by a speaker.

#### ***2.8.5 How is audification influenced by the manner in which the information is retrieved from the environment?***

Audification of ultrasound requires that ultrasound information be collected in a manner that allows for effective display. Since this information is to be provided directly to a traveller who is visually impaired it should simulate the natural listening environment. Echolocation is effective and involves the use of two receivers (ears) to collect this information. To provide sufficient information for direct perception, what is the best orientation of ultrasound receivers to ensure intuitive response? The distance between the eyes and the orientation of the eyes varies among different animals. Humans have eye placement in the forward direction whereas other animals, rabbits and horses for example, have eyes that have evolved more laterally. Unlike bats, humans cannot move their ears to provide additional localisation information. How does the

response of a human differ when presented acoustic arrays perceived through lateral receivers versus frontal receivers? All three human tests of the novel audification display will seek to explore the answer to this question.

## **2.9 Summary**

A new auditory interface will be presented that was designed using audification for obstacle avoidance. This device is intended to be used in addition to a primary device such as a guide dog or a cane to provide information about obstacles that exist above waist height. Since the tactile sensation will already be in use with a primary mobility device, an ability to capitalise on the auditory sense is necessary. Development of a device that allows for audification of reflected ultrasound echoes will be discussed. The evaluation of this device will allow individuals to perceive the distance of obstacles as accurately, if not better than that of the auditory sense with respect to echolocation. A perception-action task will examine participant response to environmental obstacles. Finally, a localisation task will show that localisation is similar to the auditory sense with the new interface. The development of this interface and the testing that surrounds it form the basis for the rest of this thesis.



### **3 Designing displays for auditory systems**

In designing a pick-up-and-use device (Burns & Hajdukiewicz, 2004) that allows for direct perception of obstacles by a user who cannot use the sense of vision (either due to environmental darkness or functional blindness), the needs of the user must be paramount. It must be determined what information should be displayed and by what means. The study by Blasch et al. (1989) provides a good indication of the users' suggestions for improvements to electronic travel aids by functionally blind individuals. These must be incorporated within the needs analysis of the interface design. Since individuals will also use a primary mobility device which will require use of at least one hand, tactile display is not ideal, leaving the sense of audition as the best method to display secondary information. Auditory display of information through simulated sound has become a new area of research over the last two decades (Kramer et al., 1997; Kramer, 1994; Sanderson et al., 2000; Walker & Kramer, 2004; Watson & Sanderson, 1998). As discussed earlier, there are several different elements of context that can then be "loaded" to create auditory signals to enable response. These include auditory icons, earcons, audification and sonification (Neuhoff J.(Ed), 2004). But which elements should be used?

In cases of high mental workload or significant data display, the information relayed through audition could reduce stress. Brewster has performed several studies that have shown that auditory interfacing can decrease mental workload while performing specific computer tasks (Brewster et al., 1994; Brewster et al., 1995; Brewster, 1997; Brewster & Crease, 1999). Tasks such as clicking on buttons (or icons), following menu tasks, and copying and moving files can be performed faster and with less mental workload when auditory information is used to compliment visual information. From these studies it is noted that it is important to reduce the cognitive load of the user.

Elements of context allow sounds to be applied to various systems, but how to best apply them is difficult to determine. A manner to systematically apply auditory information to displays has been largely ignored until recently. Two possible design cycles have been introduced to enable designers to largely follow a formula in the development of an auditory display, with both stemming from the concept of cognitive work analysis (Vicente, 1999). Johannsen (2004) suggests a design cycle based largely on general design cycle principles whereas Watson and Sanderson (Watson & Sanderson, 2007) have described the use of specific stages of cognitive work analysis in defining what sounds to use. These two design frameworks will be discussed below.

### **3.1 Cognitive systems life-cycle development of auditory display**

Johannsen suggests that one apply a cognitive systems life-cycle (Figure 3.1). First, one defines the user and the usability requirements, and follows with an evaluation of the task scenarios and object models. The tasks are those that the user will be performing whereas the object model is imposed by the designer with the intention of creating mental models within the user's mind.

The style guide defines the auditory elements of context to be used in the interface and includes the number of total earcons or auditory icons. The nature of the specifications within this model is developed from evaluation of the user tasks and objects. These specifications can then be used in determining the number of sounds, type of sounds and loading characteristics in the model of the user interface. Iterative modelling through the design and prototype and evaluation stages occurs before a final auditory user interface can be applied. This is a very top level approach to the design of the interface leaving the designer with considerable freedom in the selection of the

various auditory elements of context. Another approach which provides more specific guidelines in the design of the auditory display is that of Watson and Sanderson (2007). This approach is discussed in the following section.

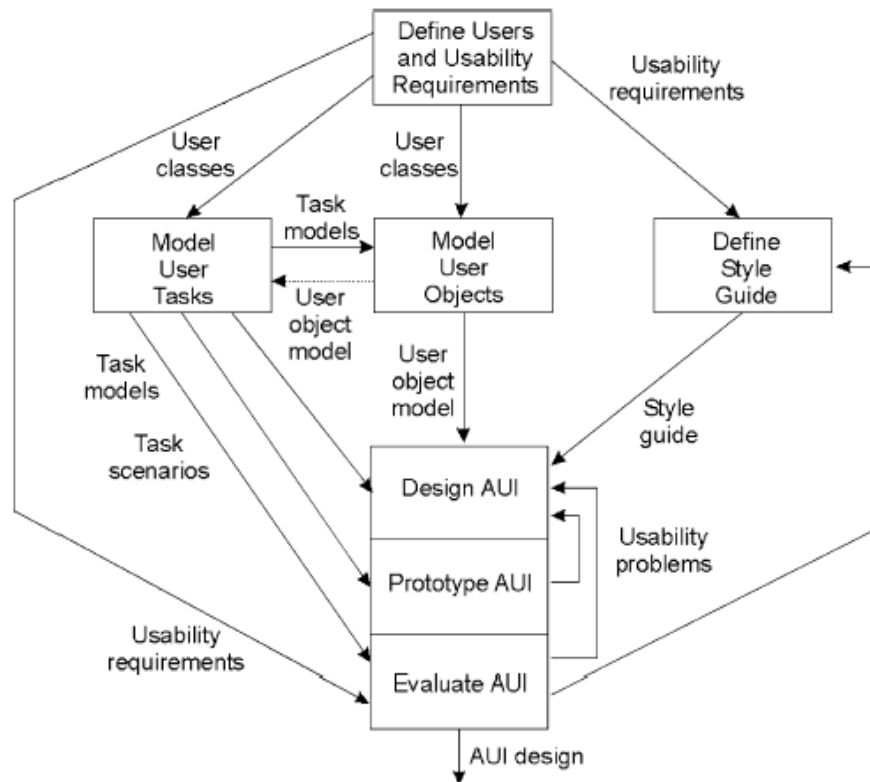


Figure 3.1 Cognitive systems life-cycle development of auditory displays (Johannsen, 2004).

### 3.2 Designing for attention with sound

Watson and Sanderson take a more ecological approach in the design of the application of sounds to auditory systems (2007). This method is based on the ecological interface design of visual systems. Ecological interface design (EID) is a systematic approach to the design of interfaces for complex systems (Burns & Hajdukiewicz, 2004). Figure 3.2 shows the process for design of an EID design accompanied by those aspects that are critical for auditory displays. It

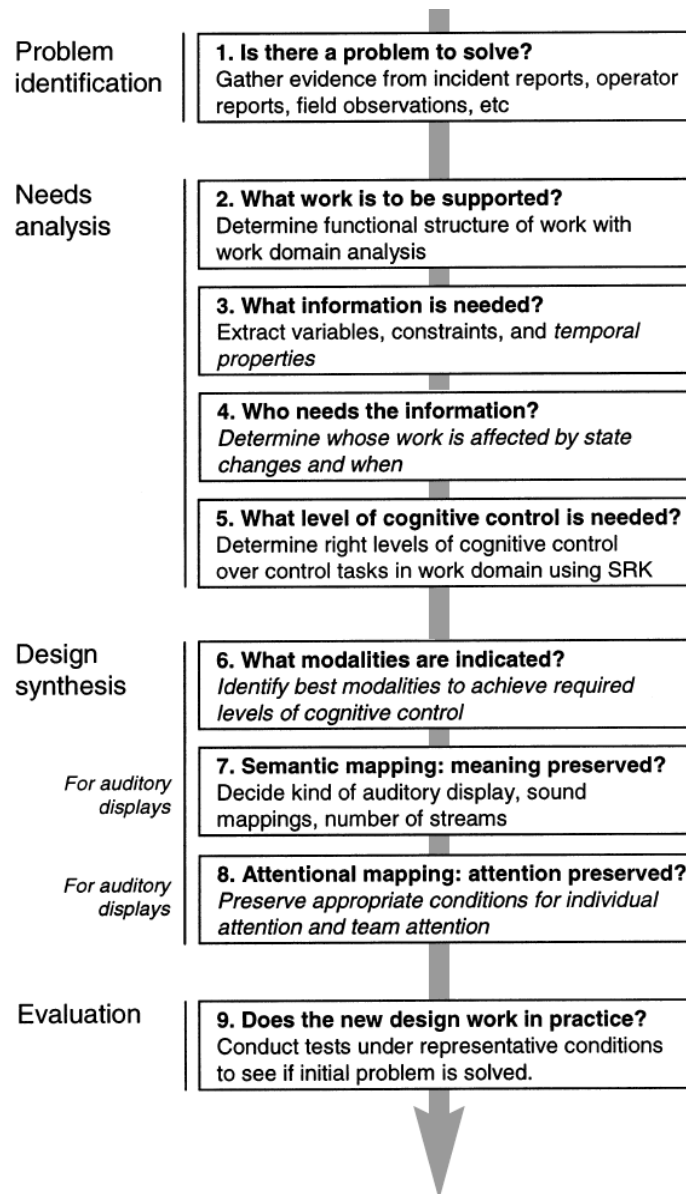


Figure 3.2 Process for the design of auditory displays using EID. Those parts in italics represent the considerations for an auditory model, not typically addressed in EID. SRK=skill-, rule-, knowledge-based behaviour. (from Watson and Sanderson, 2007)

starts with the identification of the problem and is followed by an analysis of the needs requirements of the system and the interface. Within the needs analysis there are four stages. The first is to perform a work domain analysis which enables recognition of the critical relationships through the use of a five-level hierarchy, from the concrete description of the

system components to an abstract evaluation of the overall system purpose (Burns and Hajdukiewicz, 2004). From this a needs requirement is drawn defining the variables, constraints and temporal properties. Next, the concept of whose environment is affected comes into effect. Since sound is omnidirectional and may interfere with other individuals within the environment, the concept of annoyance must be addressed (Sanderson et al., 2000). Where the sound is coming from cannot be limited as vision can by turning away. All information to the individual should provide additional information about the environment.

The next stage of the needs analysis needs to be addressed more clearly as this is the stage in which the specific elements of context to be used in the display are identified. Cognitive control demands are realised and how to reduce the cognitive load is addressed. The best manner to achieve this goal is to attempt to move those tasks which are knowledge-based in nature, down to either a rule-based or skill-based signal (Figure 3.3). Cognitive control can be manipulated with different auditory interfaces (Sanderson, 2005). Knowledge-based behaviour is concerned with analytical problem solving based on symbolic representation, whereas skills- and rule-based behaviour deal with perception and action (Vicente & Rasmussen, 1992). Sonification and earcons are abstract and analytical and generally fall within the knowledge-based behaviour realm whereas auditory icons fall in the realm of those sounds which are rule-based behavioural in nature requiring some analytical deciphering (Sanderson & Watson, 2005). Ideally, a system should be skill-based in nature which requires everyday listening to guide action. This would lead us toward development of an auditory interface that uses audification or rule-based intuitive sonification.

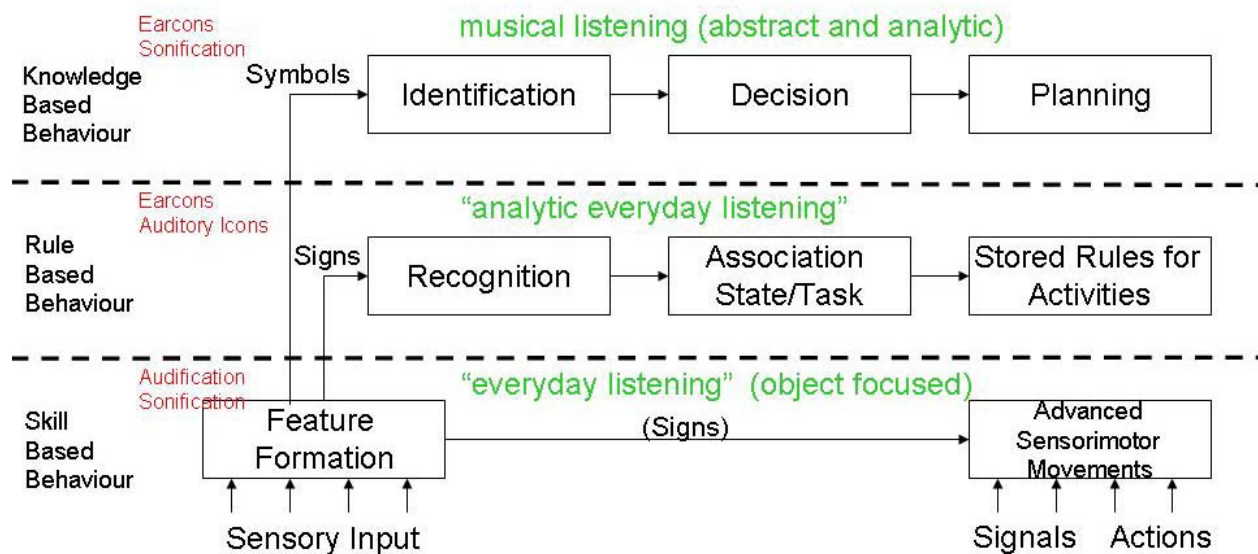


Figure 3.3 Skill, rule and knowledge-based behaviours based on auditory elements of context (Sanderson, 2005).

The next stage of the design cycle is synthesis. The best modalities to relay the most information is discussed at this stage. Would it be best to use another modality in conjunction with audition to reduce the load and allow for more skill- or knowledge-based action? One of the benefits of an auditory display is that it can be used in addition to another device such as the cane or a guide dog. How can sound be best used to complement the other needs of the user? Now, the elements of context that best fill that need can be determined with semantic mapping. This is the stage at which the number of sounds and type of mapping is determined based on the needs analysis of the work domain analysis. It is important to note that specific sound cannot be reproduced in exactly the same way (Sanderson et al., 2000). Even a recorded sound will have a different environment in which it is played every time. This transience requires that sounds are developed to provide quality information to an observer continuously. The final stage in the

design of the auditory interface using this framework is that of application. One applies the auditory interface, trains individuals in its use and evaluates the effectiveness.

### **3.3 Summary**

The two frameworks presented here both draw on the concept that the cognitive load of the user must be minimised. The first is a more general approach but does not give specific guidelines about how to choose the elements of context. The second allows for more of a recipe approach in the design. An auditory interface that requires continuous monitoring rather than intuitive action is one that will likely fail. By following these frameworks in the design process, a designer is more likely to succeed. It is important to note that, although all these cognitive requirements are defined and the interface has addressed these needs, design is an iterative process. The design of the prototype interface must be tested, revised and retested until the users of the system are satisfied that additional improvements are no longer required. The next chapter will draw on the information presented here to develop an auditory interface for visually impaired travellers.





## **4 Development of an auditory display**

The two frameworks previously addressed can now be applied in the design of an auditory display for individuals who cannot perceive visual information. The framework by Watson and Sanderson (2007) incorporates ecological interface design (EID). Those displays that best employ EID are those in which there is a high cognitive load, high requirements for vigilance of the operator and for those data sets that are hard to display (Sanderson et al., 2000). The task at hand is one in which there is no visual information present to the user thus a method that can effectively identify the key characteristics of the interface is essential. EID is very effective at defining what information needs to be displayed to the user and the breadth of the design in the auditory domain (Sanderson et al, 2000). Thus, the design of the present auditory display will be one that follows the steps outlined by Watson and Sanderson (2007). The first stage is to identify the problem. Considering the current systems as presented in the first section of this chapter, a need arises for the design of an auditory interface that provides obstacle avoidance information while minimizing the cognitive load. Once the problem is introduced, the needs analysis is addressed in section 2 and design synthesis of the interface is discussed in section 3.

### **4.1 Problem identification**

As previously indicated, the interface to be developed is one that provides information to allow for avoidance of obstacles that are above waist height. The device must enable effective travel, with a trade-off of speed efficiency and risk mitigation, while the individual is also holding a primary mobility aid such as a cane or a guide dog. This research will seek to explore those systems that relay information through sound rather than those that require haptic information. There are three systems that were already mentioned as being used by individuals with functional blindness, though not extensively. Each one of the interfaces for these devices will now be

discussed in the order of reduced cognitive load based on the elements of context used in the design of each, the vOICe, the KASPA and the Sonic Pathfinder. This will follow with a statement of the problem before going on to discuss the needs analysis.

#### **4.1.1 *The vOICe system***

The vOICe system provides an acoustic rendering of the environment such that all information within the field of view is provided in a manner that Peter Meijer deemed effective (1992). First, an image of the environment is taken. Based on this, sonification is used to map information binaurally to “soundscapes”. A soundscape is a one second representation of the image such that pitch is mapped to the row of the image and the amplitude depends on the brightness of the pixel. If there is a spot on the left hand of the image, you will hear a sound on the left. A straight line across a screen passing from the top left corner to the bottom right corner will sound like a steadily decreasing pitch until the scan reaches the far right. A vertical line would create one quick “bleep” such that all the pitches in that image are superimposed. One image is one “soundscape” that takes a full second to scan. Turning to another location or moving forward will require another second to relay the updated information. This is definitely unrealistic for travel purposes. Although the speed of update can be increased, continuous scanning of the image from left to right can provide ambiguous information to the user and requires considerable concentration (Amedi et al., 2007). Based on the earlier description of behavioural requirements, this system requires significant training. As such, this interface falls within the knowledge-based behaviour realm requiring significant cognitive load. This system is one that has been developed more to create a full visual picture rather than to provide obstacle avoidance and requires about forty hours of training to become efficient (70% effective in a four part multiple choice of identification of simple objects) (Amedi et al., 2007). The ability to use this interface to avoid

obstacles has never been reported. Meijer (1992) specifically notes that the brain can perform tasks much more effectively than a computer system, but chooses to test the perceptive nature of his system by creating a computer image representation based on the output of the signal rather than human testing. Another system that uses sonification is that of the KASPA, the next system to be discussed.

#### **4.1.2 *The KASPA***

The KASPA is another heavily marketed device that provides information through sonification specifically for obstacle avoidance. This device collects sonar information from the environment and uses that information to sonify the sound streams. The distance to an object is represented by pitch, the direction, by a difference in interaural intensity, and this system seeks to provide obstacle texture information by timbre. Thus, if an individual approaches an obstacle to the right, the sound decreases in pitch and is heard on the right side (Kay, 2000; Kay, 2001). The mapping of texture has never been discussed in detail and it has not been tested. After four weeks of training, only 70 % of participants in a study suggested their mobility was better than prior to the device (Kay 2000). A device should be able to be “picked up and used” instead of requiring significant training to understand. An auditory display that uses a systematic approach to develop the display can reduce the need for memory or calculation (Burns and Hajdukiewicz, 2004) and may enable effective use of the device without significant training.

#### **4.1.3 *The Sonic Pathfinder***

Dr. Heyes gave much more consideration to the auditory display for his Sonic Pathfinder which was designed with simplicity in mind (Heyes, 1984). He performed human testing with various earcons to evaluate the best musical scale to use in his system (Heyes, 1980). After testing five

ascending and descending scales, he settled on an Ionian scale which is the most familiar musical scale. He also justified using the descending scale such that the pitch would decrease when nearing an obstacle by the fact that some users might be hard of hearing, especially older travellers and might then be more sensitive to lower frequency sounds. He developed the Sonic Pathfinder to provide distance information based on sonar reflections (Heyes, 1984). If the traveller is approaching an obstacle on the right side, a decreasing pitch is heard every 0.3 m closer to the obstacle from the right ear. This device is restricted to providing information about the nearest obstacle, and provides no additional information about one behind it. If a traveller is passing one obstacle, only to reach another within two feet, the individual does not have sufficient preview time to change the approach path. Earcons are skill-based rather than knowledge-based in nature, so this interface does provide much easier information to understand, but it is limited in its display of multiple objects and the accuracy with which distance can be judged (to within 0.3 m) (Heyes, 1984).

## **4.2 Statement of the problem**

The vOICE system and the KASPA provide environmental information about a fairly large field of view, but lack the easily interpretable interface required from an auditory display. The Sonic Pathfinder provides a more skill-based interface, but is limited in the amount of information it can provide. An auditory interface that can provide information about the whole field of view immediately in front of the traveller in a manner that is easy to interpret and designed with the needs of the user in mind (rather than requiring the user to adapt as Peter Meijer requires (Jones, 2004)), might result in increased acceptance and safer travel by individuals who are visually impaired. Designing such an interface is now discussed starting with the needs analysis.

### **4.3 Needs Analysis**

The needs analysis has four sections to it, the discussion of what work must be supported, the information needed, who needs the information, and at what level of cognitive control. The first section will discuss in detail the work domain analysis, followed by a definition about what information is needed and who needs the information. Finally, the requirements for cognitive control will be discussed.

#### **4.3.1 Work Domain Analysis**

Using work domain analysis in the form of a five-level hierarchy, from the concrete description of the system components to an abstract evaluation of the overall system purpose, the critical relationships can be realized (Burns and Hajdukiewicz, 2004). A full work domain analysis to solve the problem is reported starting with the system boundaries and followed by the five level hierarchy: functional purpose, abstract function, generalised function, physical form and physical components (Figure 4.1).

#### **System Boundaries**

The system boundary in this case includes the individual and a 2m hemisphere above the waist of the individual. A distance of 2m allows sufficient time (1.5 seconds at a typical walking speed of 1.3 m/s) to avoid an obstacle. The main goal of the system is to determine the interaction between the individual and environment while detecting and avoiding obstacles. This system will be used as a secondary mobility device in addition to a cane or guide dog so obstacles below waist height are not modelled in this system. To be able to effectively detect and avoid

obstacles, the individual must have control over velocity of walking, volume of feedback, the frequency of updating and her own orientation in the environment.

### **Functional Purpose**

The question to ask when determining the functional purpose of this system is “what was the device designed to do?”. This device was developed to allow individuals to travel from point A to point B in a safe and efficient manner. In this case, B is a distance that is within two meters of A. The device extends the preview distance of the cane which has a preview distance of less than a meter, while also providing information above waist height. Thus, the criteria that might flow when evaluating the work domain include:

- 1) The individual must leave point A and arrive at point B.
- 2) The individual must arrive efficiently – as quickly as possible, or as quickly as the individual is comfortable with.
- 3) The traveller must take optimal path to achieve efficiency.
- 4) The traveller must arrive safely – without falling or hitting obstacles in the pathway.

Based on these criteria, the functional purposes of the system are to 1) achieve speed efficiency and 2) to avoid obstacles safely. Speed efficiency and risk mitigation are two of the main contributors to effective locomotion (Patla, 2004). These are conflicting purposes in the sense that one must go slower to evaluate obstacles in the path. A balance must be achieved between the two purposes to allow the system to be efficient. These two purposes show the key aspects of the design that are required to enable effective travel from one point to another with or without vision (Figure 4.1).

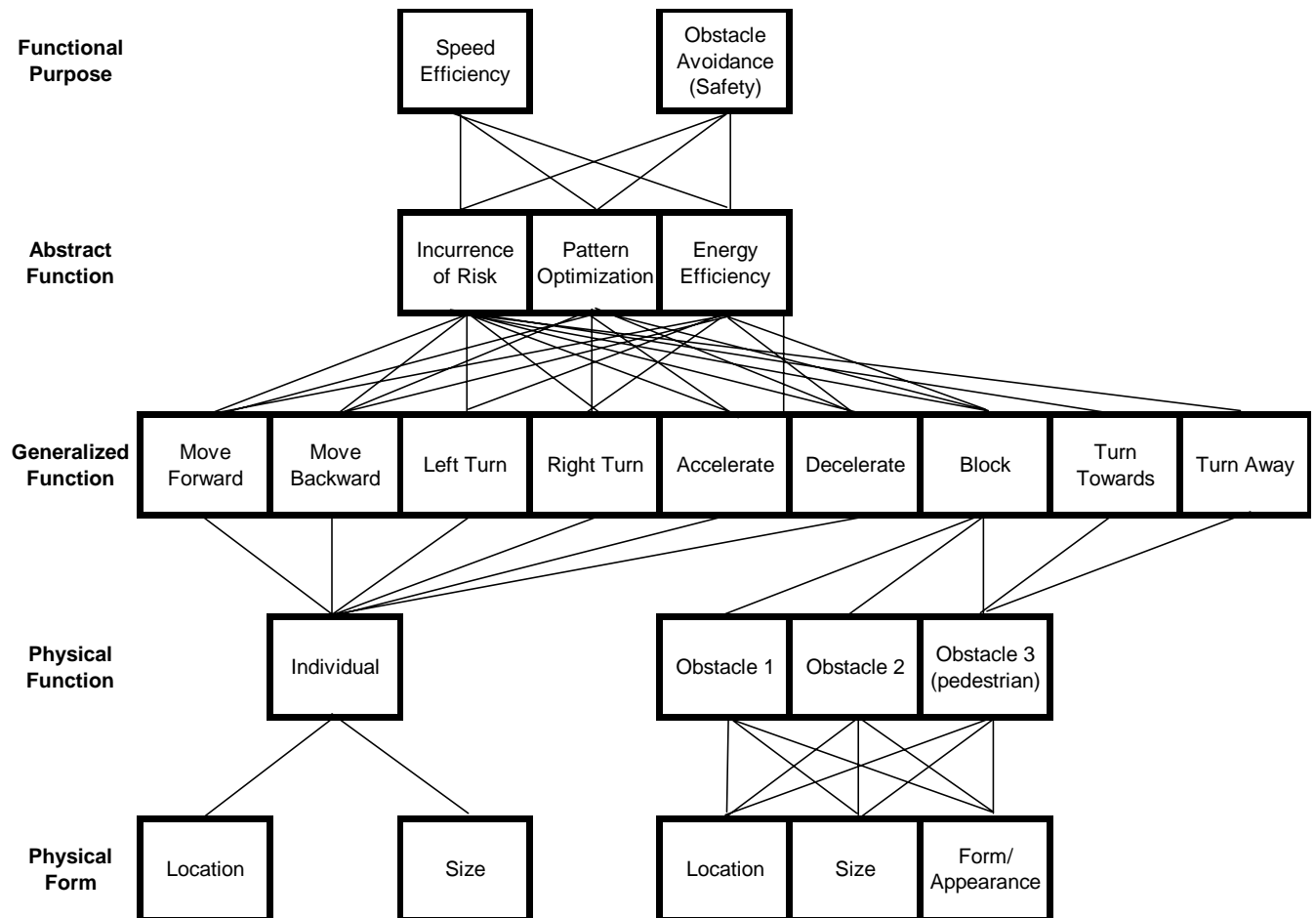


Figure 4.1 Work Domain Analysis for secondary mobility system to display spatial information.

### Abstract Function

At the abstract function level, causal inferences are identified that must occur to achieve the goals set out at the functional purpose level. This particular example uses abstract functions that are both social relationships and physical relationships. The social relation includes the abstract function of acceptable level of risk. The amount of risk is the trade-off between efficiency and safety. The physical relations are those that ensure pattern optimisation and energy efficiency.

The most efficient path must be determined and followed. Efficiency for path optimisation is the

shortest distance available to the user without hitting any obstacles. Energy efficiency also relies on the speed of walking of the individual and the minimal number of stops and starts that the individual has to make. Energy efficiency is achieved by following the shortest path in an effective “smooth” manner without stops and starts to reorient.

### **Generalised Function**

To achieve a trade-off between risk and safety, to optimise the path and to achieve energy efficiency, the mechanisms by which each is achieved are evaluated. For instance, the travelling individual has the ability to move in different directions to avoid an obstacle. These directions include moving forward, backward, left, and right. This permits the individual to effectively avoid the obstacles in her pathway. On the other hand, to achieve these goals efficiently, s/he must also have the ability to accelerate and decelerate. The general function of the obstacles is to block the individual. If there is a pedestrian acting as an obstacle, the pedestrian also has the ability to move. In this model, assume that the pedestrian can turn either toward or away from the traveller.

### **Physical Function**

At the physical function level, the components and capabilities are examined. Thus, the components that will affect the abstract functions include the motions of the individual, as well as stationary obstacles (1-n) and moving obstacles (1 – m). The individual has control over the acceleration, deceleration, and the movements to the right, left, forward and back, thus linking to the general function level. Each obstacle creates a block in the pathway of the moving traveller. The pedestrian also blocks the traveller but also has the ability to move toward or away from the traveller.



## **Physical Form**

At the physical form level, the location and size of the traveller are the critical information. This specifically refers to the relative position of the individual to the other obstacles in the spatial environment. The obstacles themselves will differ in shape, size and location and are the key components at this level that must be understood.

This abstraction hierarchy provides sufficient information to effectively determine a needs analysis for the auditory interface. A description of the information requirements follows.

### ***4.3.2 Information Availability***

From this abstraction hierarchy, a chart summarising the availability of information was developed (Table 4.1). A lot of the information is readily available but the mechanism by which it is presented must be determined during the interface design prototyping. The relative positions must be deduced based on the information presented to the user. Directional cues are the easiest to present to the traveller to allow the individual to determine the relative positions of the obstacles. To ensure that the needs of the traveller are met, the auditory interface design must allow all this information to be presented in a manner that is interpretable by the individual. Temporally, all information is continuously changing. The individual must be able to react instantly and intuitively to changes in the environment. None of the information displayed will be able to be revisited (as the sound is not recorded), so the information must be displayed effectively the first time around. The next section discusses the recipients of the information.

	Not sensed	Not available but can be calculated	Readily available
<b>Functional Purpose</b>	Successful avoidances		Individual's Speed
			Individual's Heading
<b>Abstract Function</b>	Individual's Relative Orientation	Optimal Path Length (shortest path, minimal time)	Rate of approach
<b>Generalized Function</b>	Rate of Obstacle 3 Approach	Individual's Acceleration/Deceleration	Individual's velocity
		Distance deviations from optimal path	Direction of Obstacle 3 Approach
			Obstacle 3 Velocity
			Path length traveled by individual
<b>Physical Function</b>		Relative Location of Obstacle 1 and 2 to Individual	Individual's Direction (left, right, forward)
			Separation Time (individual and <b>each</b> obstacle)
			Separation Distance (individual and <b>each</b> obstacle)
<b>Physical Form</b>	Location Individual relative to obstacles	Shape Obstacle 1 2	Height Obstacle 1 2 3
			Size Obstacle 1 2 3

Table 4.1 Information requirements for the ultrasound or vision system.

### 4.3.3 Recipients of the information

The only recipient that has to be addressed in this scenario is that of the traveller. The device is being designed to enable safe and efficient travel without social interference to other individuals. The system must be designed with the intention to address the needs of that individual. The system must be designed to reduce environmental masking. It is an auditory display, but should not interfere with other environmental sounds that the traveller may also need to respond to (an ambulance approaching at a busy intersection). Although this is a time critical system,

interference must also be reduced. The next section looks at cognitive control and reducing the mental load of the traveller while avoiding obstacles in an environment.

#### **4.3.4 *Level of cognitive control***

As both Johannsen (2004) and Sanderson (2000) suggest, the idea is to move toward auditory systems requiring skill- or rule-based behaviour so that human error can be reduced. The current systems use sonifications that are knowledge-based in nature. A traveller will likely want to talk with friends, navigate or remember the path of travel, while also using a cane and a secondary mobility device. It is best to attempt to reduce the cognitive load of the secondary mobility system by designing it to be a skill or rule-based type of system. Continuous auditory information that does not mask other sound but that could be attended to if an obstacle is encountered may provide an ideal cognitive level of control.

### **4.4 Design Synthesis**

The next three components of the design process are those involved in the design of the display itself. These include defining the modalities, semantic mapping and attentional mapping which will be discussed in the following sections.

#### **4.4.1 *Defining the modalities***

It is important to determine whether the addition of another modality will affect the control of the user. In this case, the user is already using a primary mobility aid, either a dog or a cane. Can an auditory interface be developed that can provide information about obstacles above waist height, yet minimally interfere with other environmental information?

#### **4.4.2 Semantic Mapping**

There is no established manner to achieve effective semantic mapping. If one looks at similar kinds of auditory display and realise that sonification (with the KASPA and the vOICe) is ineffective as it requires too much cognitive load of the user. The earcons from the Sonic Pathfinder do provide additional information, but as a sole source of information, they are quite limited. The first iteration of this interface design was developed using both earcons and auditory icons to provide information about all obstacles in the user's path (Davies et al., 2007). This interface moved the level of behaviour to a skill-based instead of a knowledge-based allowing for an expected reduction in cognitive load. For novice users with limited training, this interface allowed for detection of obstacles 95% of the time and distance estimates were correct 91% of the time. This interface did not allow for effective localisation or size determination. Participants were only able to identify the true direction of the obstacle in 73% of the tasks they were required to perform in test scenarios and could only report size differences in 43% of the cases. Auditory earcons in combination with auditory icons can provide sufficient information if a method of localisation can be achieved. Otherwise, it may be best to move to a more dynamic display that more closely simulates that of echolocation.

Echolocation takes the form of audification as the physical reflected sound can be used directly to guide action. This type of control (as can be seen in Figure 3.3) falls in a category of rule-based behaviour which is object focussed and requires intuition rather than learning. Moving toward an audification system may allow intuitive information to be displayed. Both the systems developed by Kay (2000) and Waters (2007) attempt to simulate the abilities of bats to echolocate. Bats have large pinnae that they move to gather information from their own

reflected signals. Sea mammals like whales or dolphins do not have large pinnae, and their ability to decipher sound (or ultrasound) may more closely resemble that of humans. Whales echolocate by sending out broadband clicks to determine the location of prey and determine Doppler Shift. As they approach the prey, the interval between clicks becomes shorter allowing them to pinpoint the location of their next meal (Johnson et al., 2004; Johnson et al., 2008). Perhaps in designing systems for the visually impaired, a parallel should be drawn to echolocating methods of whales instead of those of bats. A system that gives information about the environment to human listeners based on broadband Doppler signals might provide more information to humans than those that rely on frequency sweeps similar to those of bats.

A system that uses audification from ultrasound echolocation to provide information to the user may provide a more easily interpreted interface. Instead of analysing the environmental information prior to providing it to the user, the ultrasound signal can be applied directly after down conversion to the auditory domain. This would take the form of audification. Since sonar is a higher frequency form of auditory signals, the same physical properties are shared. As with earthquake information being effectively displayed using audification (Dombois, 2002), echoed signals from sonar could similarly be provided and follow a skill-based method of displaying pertinent information. Direct down conversion of ultrasound signals to the auditory domain without additional processing may provide adequate localisation characteristics. The echoed signals would provide information about all obstacles present in the given environment at a distance as far away as can be detected with the ultrasound device. Since auditory response to both direct sound stimulus and echoes is learned from a very young age by individuals with vision and without, it is believed that significant additional training will likely not be required.

True echolocation provides information with direct auditory reflections. The echoes provide information about environmental layout. As an individual moves closer to an obstacle, the intensity of the echoes increases, the larger the obstacle, the more reflected echoes, creating a louder sound. As an individual gets closer to an obstacle the reflections come back much more quickly and the subconscious is able to interpret distance information, thus the rate of approach can be deciphered and velocity can be deciphered. An approaching pedestrian might be heard by direct sound signals or by Doppler shift. Doppler shift at the auditory frequencies is very difficult to identify as the frequency differential is so small, but may possibly be subconsciously processed. The key to the design of a new auditory display is to attempt to provide the same characteristics to an individual with audification as would be evident through echolocation. Semantic mapping leads to the idea that audification would be the most intuitive method of displaying auditory information to the visually impaired. The final stage of the interface design is that of attentional mapping which is now discussed.

#### ***4.4.3 Attentional Mapping***

Attentional mapping seeks to identify those characteristics that will influence the actions of the user (Sanderson et al., 2000). In this scenario, sound is being provided to the individual rather than transmitting the sound to a broader environment. Thus, the only concern is the mapping of the individual. Consideration must be given to auditory signals from the system as compared to other environmental sounds. Masking must not occur preventing an individual from responding to other tasks at hand.

## **4.5 Evaluation**

The design cycle enables the identification of the key characteristics that must be displayed to the user of the device. Based on this information a new interface can be designed that ideally provides all the information required for a new user of the system to “pick up and use” the device. The development of the interface has now been suggested, but a user system is not available to which this interface can be effectively applied. The next section of this thesis discusses the development of a device that can enable audification. Once a system that can enable audification is produced, further work to evaluate the interface design is discussed. The development of this interface and testing of this prototype form the basis for the experimental work discussed in this thesis.





## **5 Design of Prototype Device**

An interface that purports to be intuitive and easy to use has been proposed. This interface was designed with the aid of a framework identifying the key features that can enable safe and efficient response to environmental obstacles. Unfortunately, the systems that are commercially available do not allow for the application of this interface. Instead, a new system was developed that allows for audification. This chapter discusses the development of a novel system that was designed in collaboration with Defiant Engineering and resembled a phased-array radar system. This device relies on Doppler to provide information about obstacles while the individual is moving (audification). A click signal can also be imposed on the signal to provide information while the individual is stationary. First the conceptual design is discussed. In the next section, the theory behind the conceptual design is detailed. Following the theory section, a description of the actual system design is presented. Finally, preliminary testing of the usefulness of the device from a practical perspective is examined.

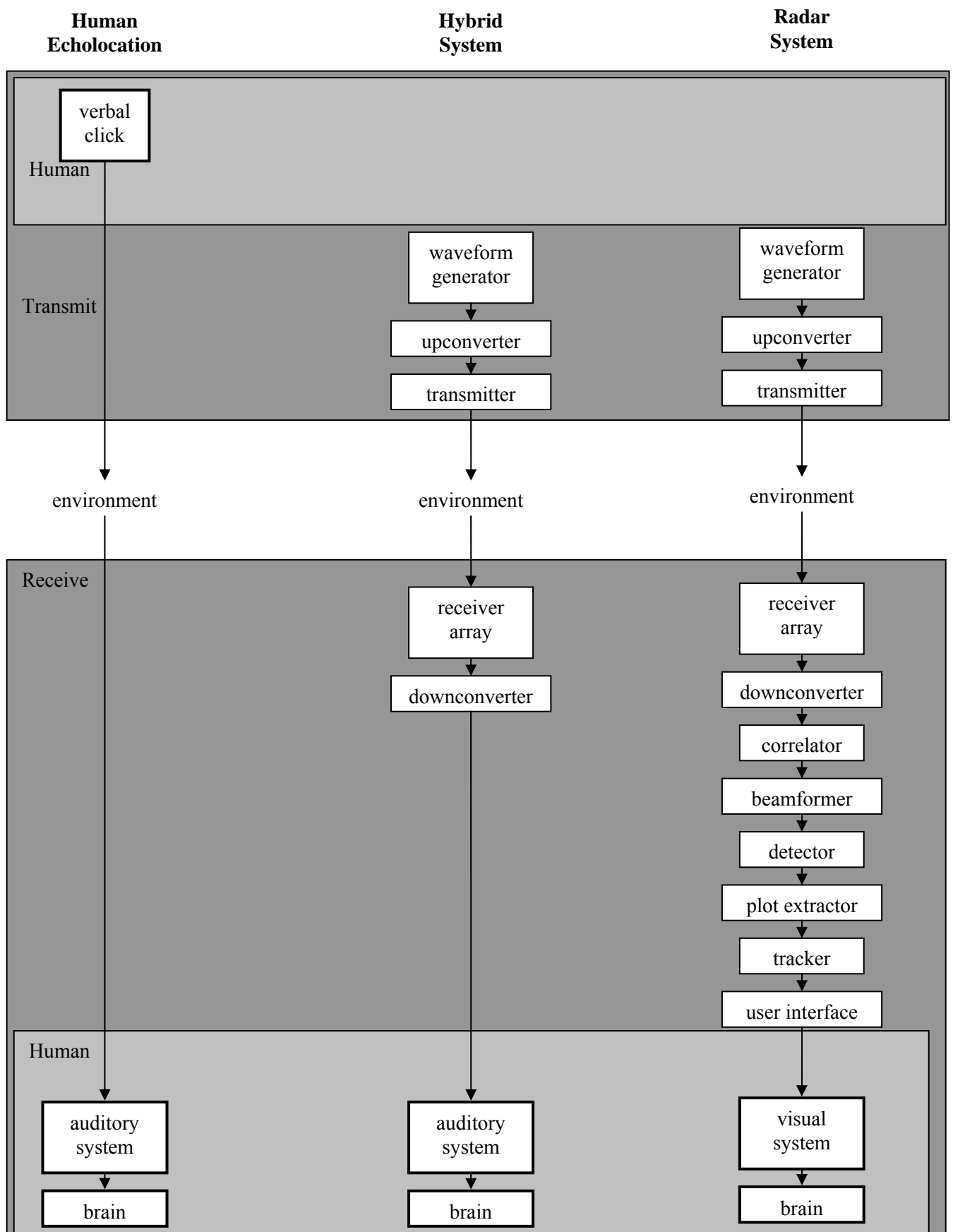
### **5.1 Conceptual analysis**

Since existing systems require a full spatial map to be displayed to the user, rather than allowing the environmental reflection to be heard, it was necessary to develop a system that would allow for direct downconversion of the ultrasound signals. This system would require a transmitter (as with dolphins or humans when they are echolocating) and two receivers (ears). The signal would be sent at ultrasound to minimise the annoyance of surrounding individuals but still allow visually impaired individuals to develop a spatial map based on reflected sound or echolocation.

The original concept for the device was based on the model and circuitry of bat detection devices. Detection devices only receive information from bats and do not require the pinna

scanning mechanisms for simulating bat echolocation. There are various hobbyists who have developed ultrasound receiver devices that can be used to detect bats in the environment. These individuals seek to find bats and listen to them, but do not require a time-sensitive system. An obstacle avoidance system would require real-time processing.

Phased-array radar systems provide real time data about reflections using downconversion of radar signals. This provides a good basis for the design of a human system to allow for audification. Perhaps audification can be achieved by the direct downconversion of ultrasound signals. A hybrid system is proposed taking into account human skills and an audified interface. This system would enable detection of environmental obstacles in the same manner as phased-array radar, but allow the human brain to perform many of the complex tasks (Pinder & Davies, 2007). In the Hybrid System shown in the centre column of Figure 5.1, the verbal click of human echolocation is replaced with the waveform generator of the radar system, though the waveform generated could be identical to conventional echolocation (a click, hiss, or clap). Transmission of the waveform at an ultrasonic centre frequency provides several advantages, apart from the elimination of the social disturbance. For instance, Doppler is proportional to transmit frequency, and therefore more pronounced, allowing perception of the rate of closure with an object by a novice user. Furthermore, the degree to which the surroundings can be illuminated is limited only by the power and dynamic range of the hardware, providing the opportunity for greater preview than is available from human echolocation.



Copyright 2006, Defiant Engineering

Figure 5.1 The similarities among human echolocation, hybrid system echolocation, and phased array radar system (from Pinder & Davies, 2007).

As described in the work domain analysis, the main purpose for the system is to achieve efficient, risk-free locomotion. In this case, a system with an ability to transmit sound while both stationary and moving was developed. Studies by Ashmead *et al.* (1995) and Rosenblum *et al.* (2000) found that movement toward an obstacle increased the ability to localise distances as compared to stationary echolocation. Continuous transmission of ultrasound signals in this system may allow for a greater Doppler while an individual is in motion. As within the auditory range, movement toward an obstacle with the novel system complements a click with Doppler signals which should also enable more efficient localisation.

When the traveller stops moving, the clicking sound becomes more prominent. Naturally, individuals move their head to hear the source of the auditory stimulus (Munhall *et al.*, 2004). Turning the head toward the source has shown to be most effective at localising sound sources (Middlebrooks & Green, 1991). Dynamics of the head are intuitively incorporated in the Hybrid System in exactly the same way as long as the necessary instrumentation moves with the observer's head. Localisation is achieved directly through head motion. This closely resembles an echolocation click that visually impaired individuals use and can allow for fine tuning the information received through ultrasound. The theory behind intentional aliasing and Doppler shift will now be discussed ending with a description of how these are applied in the Hybrid System.

## **5.2 Digital Downconversion**

The concept of “intentional aliasing” is commonly used in radar systems to convert a signal at a high intermediate frequency to a useable digital signal. Aliasing occurs when the sample rate is

lower than can be fully representative of the transmitted signal. The Doppler Effect provides changes in the frequency of the signal and allows for perception of environmental obstacles. Although this is not often observed in the auditory domain, once a Doppler-shifted ultrasound signal is intentionally aliased (or downconverted), the user can hear environmental changes. The next two subsections will discuss intentional aliasing and how it can allow for auditory perception of the Doppler shift.

### 5.2.1 *Intentional Aliasing*

The Nyquist Theorem states that the highest frequency that can be accurately represented by sampling a signal is less than one half of the sampling rate. This allows for the presence of sufficient information to draw out all components of a frequency spectrum. Aliasing occurs when the sample rate falls below the Nyquist frequency. Aliasing can be described by examining a simple sinusoid (Figure 5.2). In this case, at the Nyquist frequency two sample points would be gathered for each period (the red line). The red line has the same frequency as

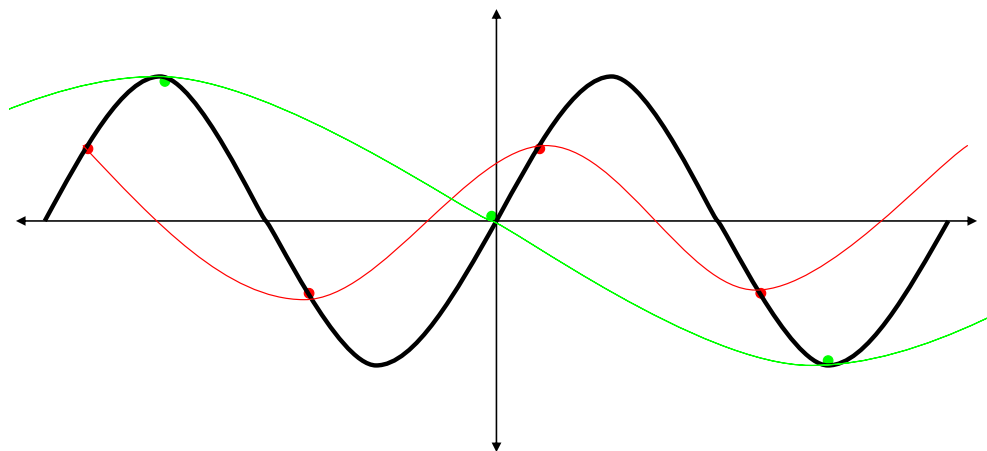


Figure 5.2 The original waveform is in black. The red waveform represents a sampling rate at 2x the sample rate. The frequency is maintained. The green waveform represents a sampling rate that is lower than the Nyquist frequency. The frequency is much lower than the original waveform.

that of the original black line, though it is of lower amplitude. If sampled at a frequency lower than Nyquist, there will be fewer points per period and it will appear as though the points came from a sinusoid that was of a lower frequency (the green line). By sampling the ultrasound signal at a frequency lower than Nyquist, the ultrasound signal can be represented by a signal within the auditory range.

Before applying this theory to the ultrasound system, the Doppler Effect must be discussed, after which all will be drawn together to describe how the signal is audified.

### **5.2.2 Doppler Effect**

This system uses Doppler to provide information about the general environment during locomotion. The Doppler Effect describes a change in the frequency of a signal as a result of the relative motion between the transmitter (or a reflector) and the observer. Figure 5.3 shows a transmitter-person generating a clicking sound with the tongue while moving forward, relative to a stationary observer. The waves in front of the person have a higher observed frequency than those behind.

The transmitter-person, of course, is also a receiver. If an obstacle is located in front of the person, such that the waves are reflected, a Doppler shift can be observed by the person, provided the person has sufficient sensitivity to the frequency shift. The obstacle is stationary with the transceiver walking toward it. Thus, the received frequency,

$$f = f_o \left( 1 + \frac{v_o}{v} \right)^2 \quad 5.1$$

where  $f_o$  is the transmitted frequency,  $v$  is the speed of wave in the medium,  $v_o$  is the rate of change of distance between the transceiver and the obstacle.

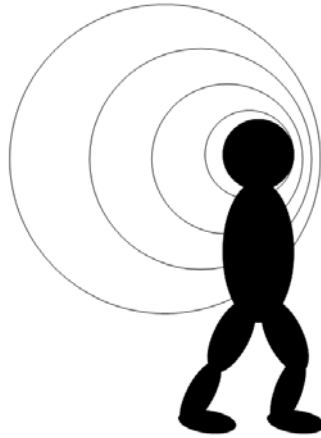


Figure 5.3 Observed frequency of signal generated by the individual.

This individual is performing echolocation in the auditory range. Assume that sound travels at 344 m/s in air. If the person sends out a signal at 100 Hz while travelling directly toward a stationary obstacle at an average walking speed of 1.3 m/s the perceived frequency of the return signal will be 100.76 Hz. This difference in frequency from the initial signal will be virtually undetectable to the untrained observer. It is almost impossible to detect Doppler shift by normal human movement. Kish suggests that higher frequency clicks are more easily distinguished (Kish, 1995). If the individual clicks at 1000 Hz, the perceived frequency of the return signal will be 1007.6 Hz, thus a difference of 7.6 Hz, more perceptible, but not for the average individual.

A combination of the use of intentional aliasing in addition to the Doppler Effect is now warranted. The next section describes the combination of the two to enable audification of ultrasound signals.

### 5.2.3 Combining intentional aliasing with Doppler to allow for audification

For the purpose of this thought experiment, the transmitted signal is at a frequency of 40 kHz.

Assume that an individual with this system is stationary. The sine wave in the top panel of Figure 5.4 represents one tenth of a millisecond of a 40 kHz signal. This signal is transmitted out to the environment and reflected off obstacles. When the return signal is sampled at 40 kHz (the green dots) a signal of 0 kHz will result (the green line) as there is no motion. Nothing can be heard. For the echo that is unmodified as a result of non-movement, the difference between the two signals results in a value of zero hertz.

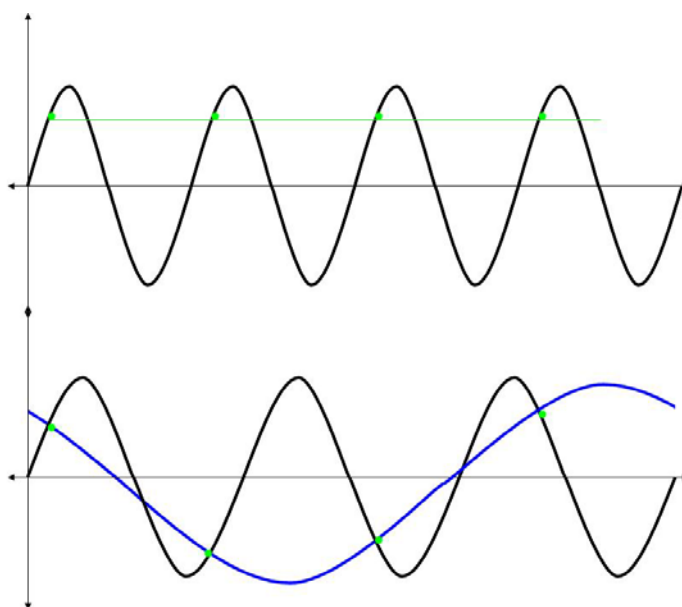


Figure 5.4 The top panel represents a 40 kHz signal (black line) which is sampled at 40 kHz (green dots). The result is a line of 0 kHz (green line). The bottom panel shows a 30 kHz signal (black line) sampled at 40 kHz (green dots) which result in the 10 kHz blue signal.

Now there is movement in the environment, such that a 30 kHz signal is received (black sinusoid in lower panel of Figure 5.4). Sampling of this signal at 40 kHz (green dots) results in the blue sinusoid which is a 10 kHz signal, which is an alias of the original. A 10 kHz signal can be



heard in the auditory range. This is an extreme example for diagrammatic purposes. Now let's visit a more realistic scenario.

The individual performing echolocation in Section 5.2.2 was walking toward an obstacle at 1.3 m/s. If instead of generating a signal at 100 Hz or 1000 Hz, the individual can now generate a signal at 40 000 Hz, what is the result? The signal is transmitted to the environment at 40 kHz, the individual moves toward it at 1.3 m/s, thus using equation 5.1 again, the signal that is “heard” is 40 303 Hz. This signal cannot be perceived by a human. If intentional aliasing is now applied to the return signal such that it is sampled at 40 000 Hz, there are two components to the signal one that is 40 303 Hz and another that is 303 Hz. As the first component is beyond the perceptual limit of the human ear, it is naturally filtered away. The other component, the 303 Hz signal can be heard by even a hard of hearing individual. The information from the 40 kHz transmit signal has been reflected and intentionally aliased such that audification has occurred. This is a very simple digital down conversion with the sample frequency the same as the local oscillator frequency.

As the user walks through a room with multiple obstacles, each one creates a different Doppler shift depending on its location relative to the traveller. One can hear the reflections off the walls based on the distance away from them and the direction of travel. The result itself is a soft “noise” with frequencies dependent on the location of the obstacles and the motion of the user.

#### **5.2.4 Does this mask other environmental sounds?**

This process works effectively to audify the ultrasound signals, but what about using this proposed device to walk alongside a friend when talking? The frequency of voices is between 85

Hz and 255 Hz. The voice signal is collected by the receiver and sampled at 40 000 Hz.

Returning to the Nyquist Theorem, when a signal of 255 Hz is sampled at 40 000 Hz, there are more than enough samples to represent the 255 Hz signal. The human ear can hear frequencies as high as 20 000 Hz. The Nyquist sampling frequency to allow for all frequency information in those signals to be determined is 20 000 Hz. Thus, all signals that would otherwise be heard in the auditory range, such as talking or other environmental sounds, are still present. Intentionally aliasing the high-frequency ultrasound signals does not alias those at lower frequencies, so masking of other sounds does not occur.

This section has shown that intentionally aliasing those Doppler-shifted signals created by movement allows for audification of ultrasound signals without masking other environmental sounds. Now the specific application of this theory to the system design must be visited.

### **5.3 Hardware**

The system developed is essentially a direct conversion receiver which takes the information obtained from the echoed ultrasound signal and performs intentional aliasing to within the auditory range (Figure 5.5). It consists of a processor which houses a waveform generator and a downconverter. There are two receivers which can be moved to test different directions of signal arrival as well as the transmitter. Output from the system is received by the user from the processor with headphones. Further detail of the system components can be obtained from Defiant Engineering.

### *Waveform Generator*

The waveform generator is an oscillator which generates a signal with a square wave at a constant frequency of 40 kHz. A square wave was used in this circuit to allow for effective intentional aliasing and to ensure the phase differences were maintained.

### *Downconverter*

The purpose of the direct conversion receiver is to demodulate the signals. It takes the information from the ultrasound echo, performs intentional aliasing and provides the signals to the headphones. The ear naturally filters any portion of the signal above 20 kHz allowing only

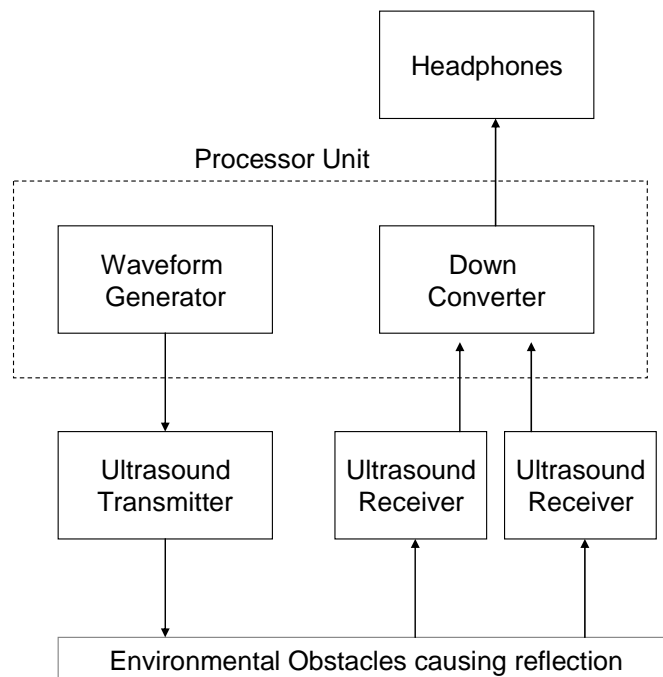


Figure 5.5 Hardware components in novel system design.

the auditory signal to be heard by the user. As the sampling frequency is exactly the same as the transmit (or carrier) signal, it is removed entirely when there is no Doppler shift. More

generally, the downconverted signal is a mixture of Doppler and the transmit signal, which the ear naturally filters, allowing the user to hear only the Doppler.

#### *Transmitter*

The 40 kHz ultrasound transmitter takes the signal from the processor and transmits it into the environment. The receivers then collect the echoed information for processing.

#### *Receivers (A and B)*

The receivers are designed to receive and amplify the ultrasound signal. Each has an ultrasound receiver with a frequency centred at 40 kHz. The signal from the receiver enters a preamplifier followed by an amplifier before entering the synthesizer for intentional aliasing of the signal.

The theory of the system and the system design suggests that audification of ultrasound can occur within the system. Now, some basic preliminary testing must be performed before performing human tests.

### **5.4 Preliminary system tests**

The theory behind the development of this device has been suggested, but the system must be tested to ensure it operates as expected. As mentioned, the aliasing should not mask other environmental sounds, yet the system should provide sufficient information within the auditory domain to enable localisation and distance determination. First, this theory is testing by collecting direct auditory click signals from the receivers. Next, various frequency components

within an ultrasound reflection are examined. Finally, an evaluation of the intensity of the signal on approach of a wall is reported.

#### **5.4.1 *Evaluating click signals***

Although several studies have used the ability of participants to self-generate clicks while echolocating to detect obstacles, a detailed description of the method used is not presented (Rice et al., 1965; Rosenblum et al., 2000). In these studies, the clicks were highly variable ranging from tongue clicks and hisses to words like “hello”. Other studies have shown that for localisation, sounds must be short with a broad band of frequencies (Neuhoff J.(Ed), 2004). Although sounds that elicit localisation responses have been studied, specific sounds which provide accurate information for echolocation have not received much attention. There are several techniques that are known to be useful to individuals who use echolocation, though the specifics of the clicks themselves have not been studied. Kish (1995) reviews artificial and organic clicks and indicates that the benefit of artificial clicks is the ability to be repeatable both in length and spectral content. Organic (oral or self-generated through claps or footsteps) clicks tend not to be repeatable and are often user-specific. Kish also reports that a self-generated oral click with a frequency range of 900 Hz to 8 000 Hz is typical with the duration being 6.6 to 20 ms (1995).

To better understand a small sample of the sounds that are typically used in echolocating as well as to gain an understanding of the spectral signature of direct signals in an anechoic environment, sample sounds of claps, clicks, and snaps were sampled at 96 kHz. The direct signals in an anechoic environment were collected with the receivers in the aforementioned system.

The duration of these clicks was between 1.7 ms (click) and 8.2 ms (clap). Although the duration of the click is less than indicated by Kish, oral clicks are not completely repeatable. The range of oral clicks within the five sample data collections was 1.7 ms to 4.7 ms. Figure 5.6 shows sample power spectral density (PSD) plots of these signals (the frequencies to a maximum of the human range of hearing are shown, from 0 Hz to 20 000 Hz).

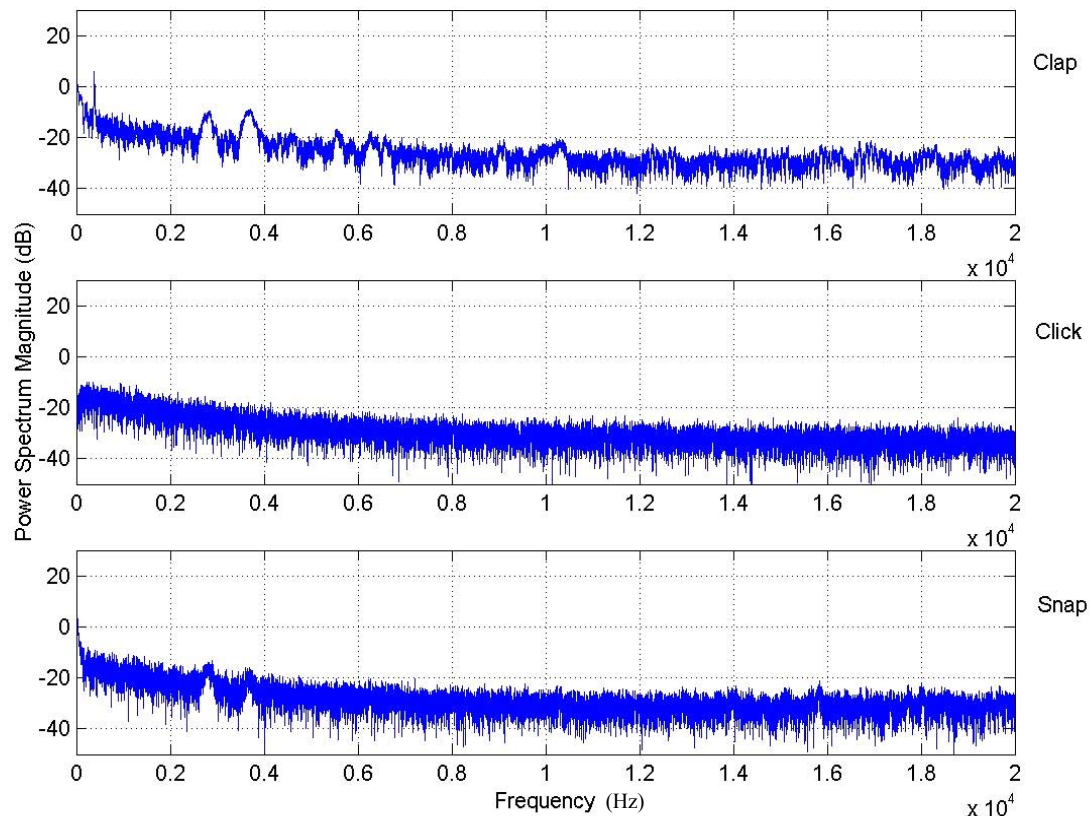


Figure 5.6 PSD of common clicks: clap, click and snap.

The clap has a strong power magnitude spike at 380 Hz, and distributed components from 2600Hz to 3000 Hz and from 3500 to 3900 Hz with the rest of the signal fairly randomly distributed. The snap has a broad peak of higher power from 2600 Hz to 3000 Hz and another

from 3600 Hz to 3800 Hz. Otherwise the signal is well distributed across all frequencies. The PSD plot of the click sound displays a fairly constant distribution of frequency components in the lower end of the spectrum gradually decreasing to 10 000 Hz. Based on this preliminary observation, one can see that all three signals provide broadband audible signals to the user at the lower end of the frequency spectrum. The click has the most random distribution of these broadband signals. Many individuals who use echolocation use a mouth click to “illuminate” their surroundings with echoes, though even this is highly variable in the tones and pitch they self-generate. Since this signal also comes from a source that is a constant distance from the ears, as well as being easily generated even while holding a cane, it is likely one of the best methods of localisation using echolocation. It is also consistent with studies suggesting that broadband signals are more easily localised. The next task is to examine reflected echoes and the power spectrum associated with those echoes.

#### ***5.4.2 Evaluating the echoed signal in an anechoic environment***

This pilot testing involved orienting the receivers on a bicycle helmet pointing outwards with the transmitter strapped to the top (Figure 5.7). This orientation will provide the least signal strength (thus most conservative results) as the receivers are perpendicular to the transmitted signal.

The large metal (bank vault) door to the anechoic chamber was closed and the reflected signals from a 40 kHz square wave click were recorded at 96 kHz at seven different distances from the door. The increased recording rate allowed for an evaluation of the effect of the intentional aliasing. For an ideal 40 kHz square wave, sampled at 96 kHz, one would expect to see peaks that represent the alias frequencies of the higher harmonics at 24 kHz, which would be 9 dB lower than the 40 kHz peak, at 8 kHz, which would be 12 dB down from the 40 kHz peak, and at

16 kHz and 32 kHz, which would each be 35 dB down. A power spectral density plot was generated as seen in Figure 5.8. The axes on the figure are power spectral density (dB/Hz) and frequency ( $\times 10^4$  Hz). The top figure represents the signal closest to the door and the seventh is farthest from the door.

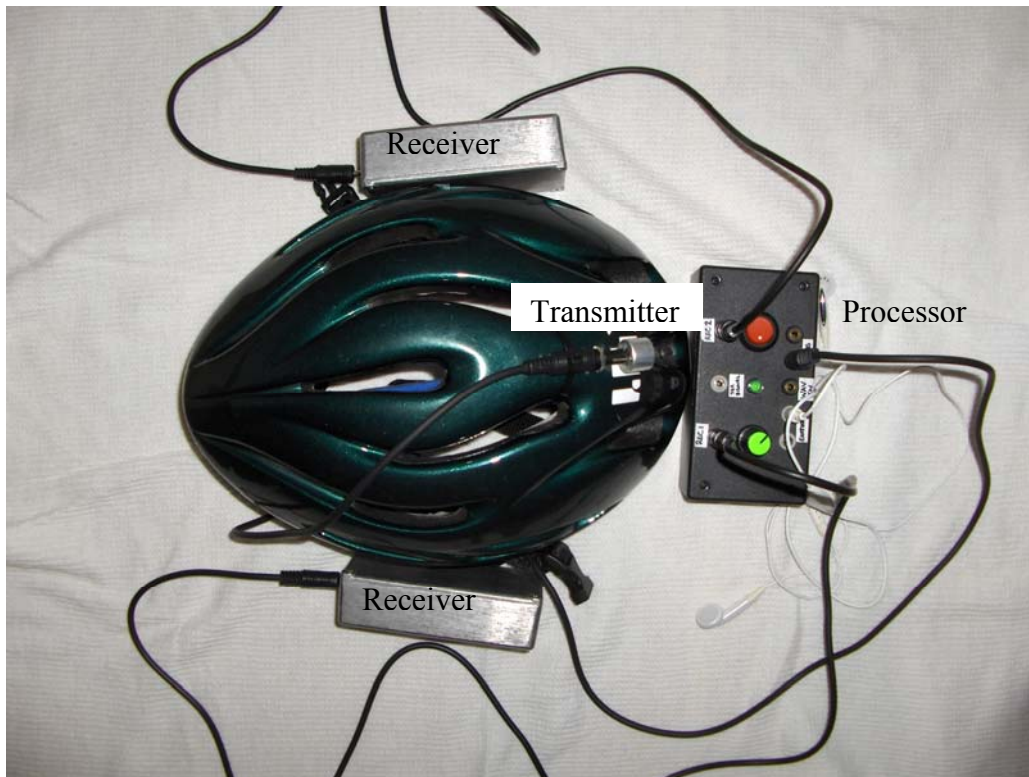


Figure 5.7 Testing apparatus used for pilot experiments.



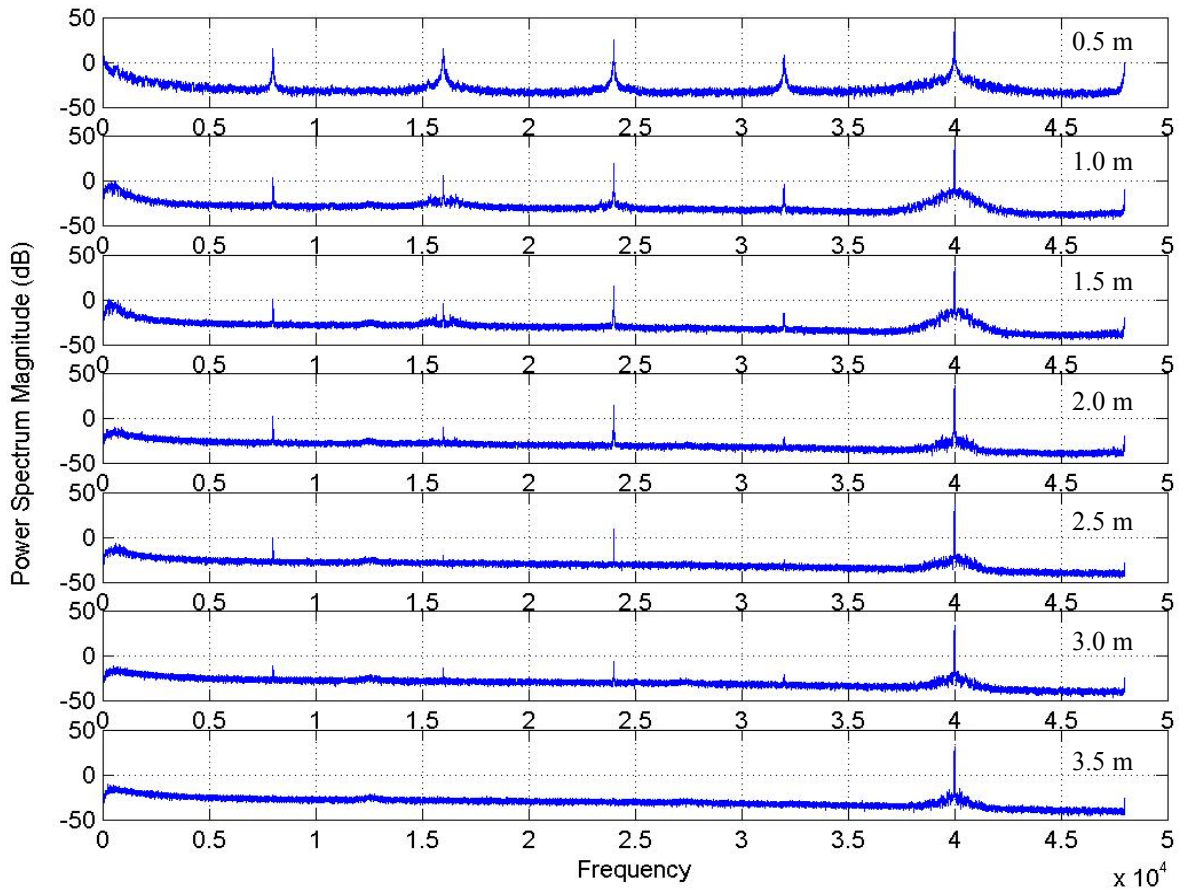


Figure 5.8 PSD of the reflections on approach of a large metal door (at 0.5 m intervals).

These plots show that the system can be used to resolve distances based on the power of the aliased signals. Important to note in this particular example is that the aliased signals appear to be stronger as the distance to the door increases. One would expect that the relative signal strengths would be maintained. There are three possibilities for this observation. These will be discussed below.

- 1) Bleed from the timer oscillator within the processor. To test this hypothesis, the recorder was plugged into the processor and the “noise without receivers” was collected. In this case (Figure 5.8), although there were bumps evident at both the fundamental frequency and that of the aliased frequencies, all lay below the noise floor that is evident in the receiver (Figure 5.9). The noise floor of the receivers was approximately -35 dB/Hz whereas the range of these peaks in the “no receiver” condition went from -100dB/Hz to approximately -40 dB/Hz. These peaks are not additive with the other peaks received during echoing task (logarithmic scales are not additive unless discussing a multiplicative function). An increase in power as a result of these would not be discernable relative to the transmitter/receiver system, so this is likely **not** the cause of the additional harmonics.

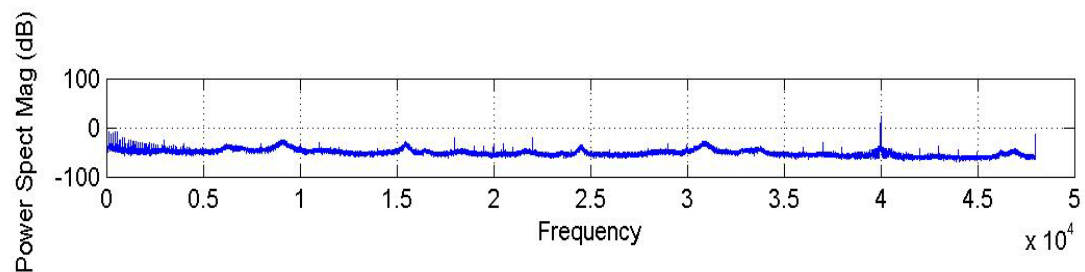


Figure 5.9 A PSD plot of the noise floor with no receivers plugged into the system.

- 2) Aliases as a result of reflected echoes of the ultrasound system. In the ideal system, the ratio of the harmonics to the fundamental frequency would be maintained at the various distances from the door. The fundamental frequency is that of 40 kHz which is at approximately 70 dB above the noise floor in the receiver system (though the noise floor gradually drops off at higher frequencies, thus less power at higher frequencies). The sample rate was 96 kHz, thus the first harmonic (aliased frequency) is that of 24 kHz which should be about 9 dB less than the fundamental based on an ideal system. The

following harmonic should be visible at 8 kHz and should be 12 dB less than the fundamental. As all these aliases are present in the PSD plots closest to the door, multiple components of the fundamental frequency are being transmitted and received. This results from either the reflective signal or bleed from the direct signal contributing to these higher powered aliases.

- 3) Bleed from the direct ultrasonic signal. Since the signal farthest from the door does not include any of these aliased signals, it is apparent that the direct signal from the transmitter is not causing additional peaks at the aliased frequencies. This is likely due to the geometry of the system. In this case, the transmitter is acting in a direction perpendicular to that of the receivers. This is a factor that must be considered further as it shows that the directionality of the ultrasound transmitter and receivers are very small band. This may hinder further experimental work.

There is very little evidence of the direct signal contributing to the harmonics. Nor is there any evidence that the harmonics are a result of the oscillator circuit. The aliased signals are most likely due to the nature of the echoes off the door.

Since the echoes are contributing to the different aliasing effects, different materials will reflect the various harmonics to a different degree. Although not evident with the chirp, the Doppler shift should allow for the detection of different materials. This was not anticipated in the development of the device. There will be different amounts of reflectivity of the ultrasound signal depending on the type of surface. Those that are highly reflective will result in an

increased number of aliased signals (harmonic frequencies) to be present in the tone, resulting in different tones depending on the material properties of the system. Although the ability to evaluate texture is present in this system, it will not be studied within the context of this thesis.

The actual signal components while approaching a door have been examined with clicks, but the Doppler Effect of the system has not been examined. The last task will involve evaluating the intensity of the signal on approach to a wall by a human participant.

#### ***5.4.3 Intensity differences on approach of a wall***

The same system as Figure 5.7 was used to evaluate the intensity of the signal during a 3-metre approach to a wall that was 1.2 m wide by 2.4 m high within the anechoic chamber. The information was recorded while a person travelled toward the wall. The data shown are intensity levels with one reading calculated per second (Figure 5.10). Each point represents the average intensity level over 0.5 seconds.

One would expect a gradual increase in intensity as the individual approached the door. As the individual approached the wall, there was an increase in the sound pressure level of 1.21 dB over the distance of 3 m. Although not linear, this particular graph was gathered from information of a human participant gradually walking toward a “wall” from 3 m such that it took 10 seconds. A constant velocity device may produce more linear results, but this ultrasound system will be used for human approach and a more accurate determination of intensity relative to distance is not required. The increase in intensity over this distance would be discernable by the human ear especially as the individual can hear the gradual increase on approach.

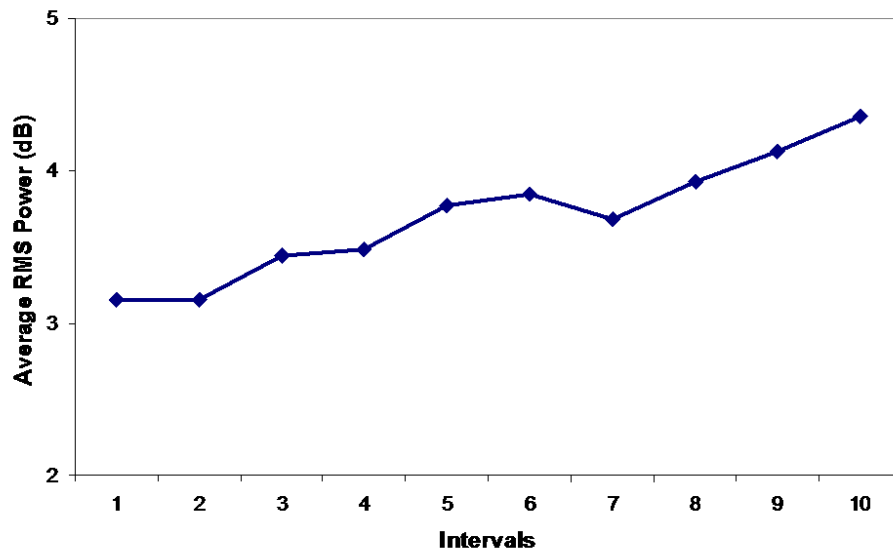


Figure 5.10 The average power of the signal as the individual approached a wall.

These three elementary tests of the system showed that the intentional aliasing of the system did not mask other sounds within the environment, the frequency of the signal provided information about the aliased signal both in the auditory and in the ultrasound domain, and the auditory intensity of the signal increased on approach to a wall. Now, the human testing of this device can be performed to evaluate its true effectiveness of providing audified signals reflected in the ultrasound domain.

## 5.5 The move to human testing

This chapter started with a discussion about the conceptual design of a system that would allow for audification of ultrasound. The theory behind intentional aliasing and the Doppler Effect as related to this system was then presented. The development of the system incorporated this

theory to enable audification of the ultrasound signal. Finally some elementary testing of the device is presented. These tests have shown that the system is operating as expected and now the system can be applied to human tests.

Three tests were undertaken to evaluate the human ability to use this system effectively. All three evaluated two possible orientations of the receivers. The first represented the direction of the ears and the receivers were faced outward. The other represented that of the eyes with the receivers faced forward. These two orientations were chosen to provide similar signals as those that would be received by the ears or the eyes to enable brain processing. This would minimise the amount of training required to use the system effectively.

Another question drawn from the determination of receiver orientation is as follows, how does the distance between receivers affect the intentionally aliased sound? Two distances between the receivers can be readily supported (Figure 5.11). These include placing the receivers:

- a) 6.3 cm apart (the average distance between eyes), or;
- b) 17.5 cm apart (the average distance between the ears).

In the most conservative case, the direct signal will reflect off a flat surface and be directly sent to the receiver. This can be shown with Figure 5.12 which is a top view of a transmitter, a reflector, and two receivers. To get an estimate of the ability to decipher sounds differently between cases A and B, assume a distance of 1 m between the source and the reflector. The distance from the reflector to the location of the receiver when the receiver is 3.15 cm from the transmitter (6.3 cm eye spacing) is 100.05 cm. In the case where they are 16.5 cm apart, the

reflected signal travels 100.38 cm. The difference between the lengths of the reflected signals is 0.33 cm. If the speed of sound is assumed to be 343 m/s, then the time required to travel 0.33 cm is  $9.62 \times 10^{-6}$  s. The rise time for a digital signal would be more than this especially after being transmitted from a speaker. Once downconverted, there would be no discernable difference between the two distances. This calculation shows that there is no need to test the receivers at different distances in addition to examining different orientations. All human testing examined the difference between the outward and forward orientations, but did not examine different distances between the two.

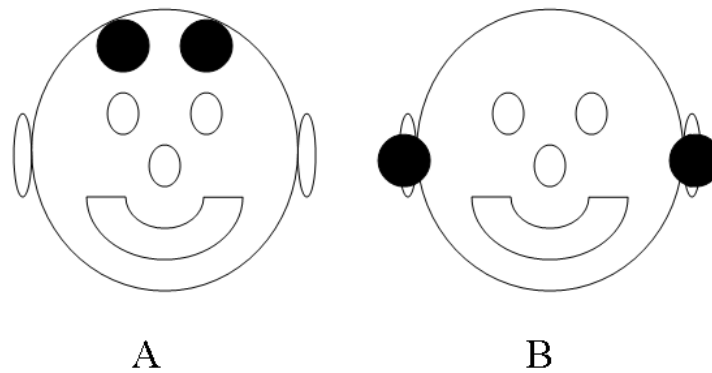


Figure 5.11 Orientation of receivers on representative mannequin.

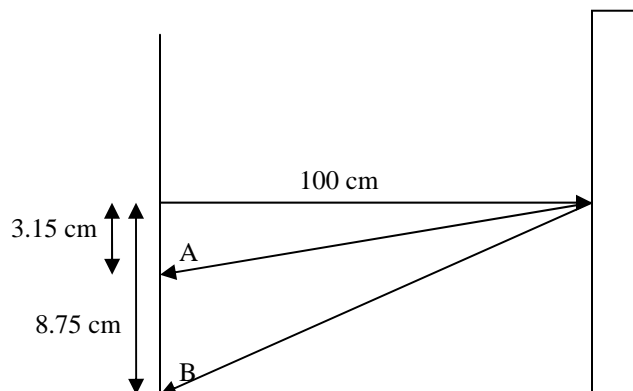


Figure 5.12 Examining distance between receivers in the “eye” direction.

The three human tests were chosen as those that could provide insight into the ability of the user to effectively use audification for obstacle avoidance. The first test compared the ability of human participants to judge distance using the ultrasound system with the receivers in the forward facing direction and the outward direction and compared these to auditory echolocation. Next, a perception-action test was performed comparing the ability of participants to pass through apertures of different sizes with the ultrasound system as compared to vision. Finally, a localisation task showed the similarities among the three conditions for determining the azimuth and elevation of auditory sound sources. These three tests will be presented in the following three chapters. The first of the three is the test that examined distance determination.



## **6 Comparison of audible echoes to downconverted ultrasound echoes for localisation by moving and stationary human participants**

Now that the design of a device that provides audified signals has been discussed in detail, the human testing of the prototype device and the ability to use audification to determine distance will be examined.

Several studies have observed the ability to judge distance using echolocation dating as far back as the mid-twentieth century (Supa et al., 1944b; Worchel & Dallenbach, 1947) and these have shown that individuals with functional blindness use sound to perceive their environments and can detect walls based on acoustic reflections. More recently, Speigle and Loomis (1993) and Ashmead et al. (1995) have looked at the ability of individuals to judge distance based on direct sounds rather than reflected ones. These studies have shown that moving toward the source is marginally better at allowing perception of distance than listening to sound sources from a stationary position. Only one recent study has examined distance perception through echolocation. Rosenblum et al. (2000) found that sighted listeners were able to make ordinal distance measurements, but absolute accuracy was not achieved.

This study seeks to use a similar approach to that of Rosenblum et al. (2000) such that individuals will use echolocation cues to provide distance information. It will also examine an effect that has never been studied before. Doppler shift in the auditory domain is virtually undetectable to the untrained observer, but the novel ultrasound system can provide insight into how an individual can use Doppler shift to determine distance of surrounding obstacles.

## **6.1 Purpose**

This study identified similarities and differences in individuals with no training in echolocation of determination of distance using audible echoes from the ultrasound system relative to audified ultrasound echoes by.

## **6.2 Hypotheses**

H1: The accuracy of distance measurements will be similar in both the auditory and audified ultrasound conditions.

P1: Since the echoed signal will be considerably different from the audible echo, localisation will not come naturally to most people. The ability to hear Doppler will increase the ability to judge distance, but make it comparable to the auditory condition with which individuals are more familiar. There is a possibility that a learning effect may occur and this should also be analysed. By examining the localisation accuracy of the first set of trials to that of the other sets, the effect of learning can be determined.

H2: The ability to judge distance will be more accurate with the receivers oriented forward rather than outward.

P2: Since the forward direction provides the ability to collect the most information from the transmitter, the forward direction will enable better distance judgement.

## **6.3 Test Environment**

To reduce the effect of external environmental sounds and uncontrolled reflections, an anechoic chamber was used. The supporting floor of the chamber was an acoustically transparent mesh below which the structural floor was covered with absorbing foam wedges, thus the only

reflections in the environment resulted from the wall. The wall consisted of a sheet of medium density particle board that was 1.2 m in width by 2.4 m in height. This wall was attached to a metal moving trolley to enable ease of movement to the different locations (Figure 6.1).

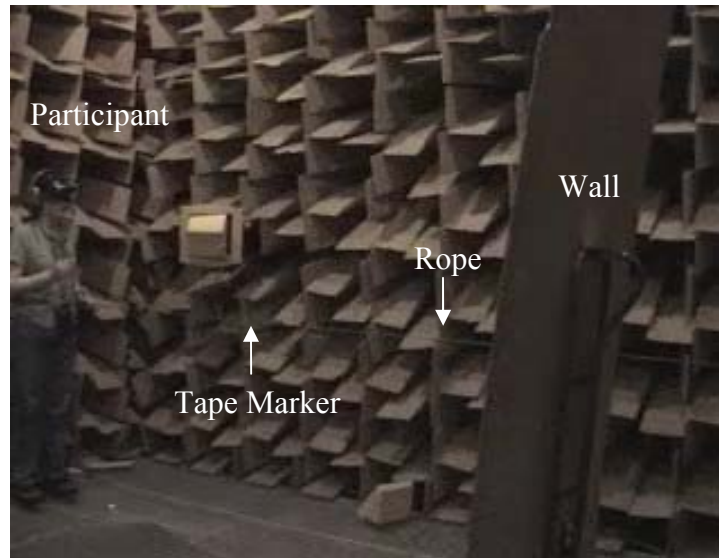


Figure 6.1 Apparatus used for testing human ability to determine distance.

#### 6.4 Lessons learned from pilot participants

Although the Doppler system would provide information about movement without additional input, a comparison to human echolocation was being drawn. A signal that would provide information for echolocation needed to be selected. Speigle and Loomis (1993) used a 20 Hz pulse train signal for 3 seconds for stationary trials or while the individual was moving toward the sound source for a time that corresponded to the distance travelled over 2 m or 4 m.

Ashmead et al. (1995) used broadband white noise for 1500 ms, also while the participant was heading toward the source. Rosenblum et al. allowed the individuals to use any signal they chose. A controlled signal that would allow for effective human echolocation needed to be

chosen. Kish (1995) suggests that to obtain the most information from echoed signals the click must be of high frequency (0.9 to 8 kHz) between 6.6 and 20 ms in length. On the basis of previous studies as well as the suggestions by Kish, four square wave transmit signals were selected. Square wave signals have the advantage of multiple frequency components. These included:

- 1) 1000 Hz for 10 ms
- 2) 100 Hz for 10 ms
- 3) 500 Hz for 20 ms
- 4) 50 Hz for 20 ms

Before the study, four individuals were surveyed to determine the type of click that they would prefer to hear if attempting human echolocation. They all chose a 1000 Hz square wave click for 10 ms. This click was within the range suggested by Kish in both duration and frequency, suggesting that sighted individuals also realise that higher frequency short clicks may provide more information than longer, low frequency clicks.

Two pilot participants were tested to evaluate the study plan. These pilots provided input about possible changes to the proposed test procedure. Three changes were made to the original protocol including changing the number of trials, the method of determining distance, and the method of training.

The original protocol was to perform twenty-five trials in each of six conditions. These conditions included stationary and moving trials of the audified ultrasound with the receivers pointing outward and the receivers pointing forward as well as auditory conditions using the

receivers in the two orientations (to allow for direct comparison between ultrasound and auditory received in the same orientations). Three different locations of the wall were tested. The participant followed a guide string alongside the wall to ensure that a straight path was maintained. For the moving trials, the participants were permitted 10 seconds to walk 1.5 meters toward the wall. Although this is a shorter distance than the Rosenblum et al. (2000) study, this is the first to be conducted in an anechoic environment which is limited by the size of the chamber (5 m x 5 m). It is also similar to the 2 m distance walked in the Speigle and Loomis (1993) study. The wall was moved to a distance 3.5 m away, the individual then walked to the perceived location of the wall. For the stationary trials, the participant used echolocation for 10 seconds, or used the system for ten seconds to judge the distance of the wall. The wall was then moved out of the way and the participant walked to the perceived location of the wall. The experimenter used an infrared rangefinder to determine the distance walked from a set location. This information was recorded and the individual returned along the guide rope to the initial location.

The number of trials for both stationary and moving conditions took twelve hours (over two days) to complete with the first pilot subject. As a result, the subject remarked several times throughout the procedure “this trial was off, I was thinking about something else”. On the first day, the participant performed auditory distance measurements with the receivers in the outward orientation. She commented that the task was very difficult and she had no idea if she was performing as expected. After four hours of this testing, she was allowed to perform the auditory echolocation task without the receivers. She indicated that this method was much easier to perform. After the first day, it was decided that the auditory condition would be performed

without the receivers. Also, due to the time required to perform the test, it was determined that only the moving conditions were needed. Since prior studies (Ashmead et al., 1995; Rosenblum et al., 2000) have shown that movement increases the ability to judge distances marginally, the experiment was continued with only the moving condition. Also, Doppler is only evident when moving, contributing the argument to only use moving conditions. The results of this pilot subject showed that the individual did perform similarly to the study by Rosenblum in that a trend towards ordinal judgement was seen and that overestimates occurred for shorter distances and underestimates for larger ones, but no conclusions can be drawn from one individual (Figure 6.2).

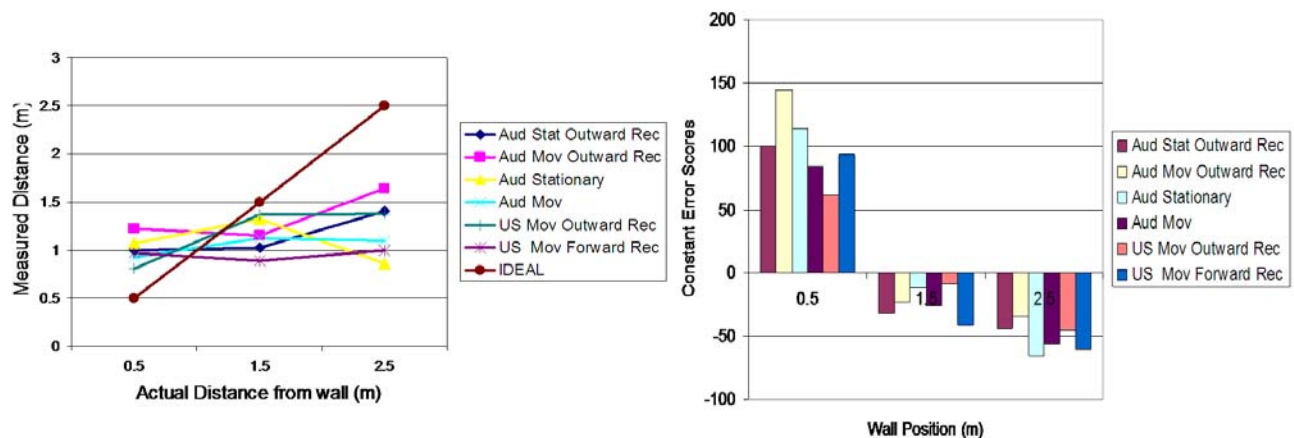


Figure 6.2 Results from pilot subject. Measured distance as compared to actual distance shows trend toward ordinal responses. Constant error scores show judgement as being over for the short distance and under for the longer distance

Distance determination was to be evaluated by locomotion of the participant to the perceived location of the wall after the wall had been moved out of the way as per Rosenblum et al. (2000). This proved to be difficult because the time required to move the wall allowed the individual to forget the initial response (based on comments such as “I can’t remember where it was” and “can

I try that one again?”). Ashmead et al. (1995) suggest that they had more accurate distance responses based on a movement task that required continuous movement rather than a task that required a stop before walking to the perceived sound source location. In the Ashmead et al (1995) study, individuals walked while the stimulus was on, the stimulus was then turned off and the speaker swung out of the way as the individual approached. For this pilot work, the “wall” could not easily be swung out of the way. Since the individual had difficulty remembering the location while waiting for the wall to be moved out of the way, another test method using a rating scale was developed.

The rating scale was developed to allow the individuals to provide a personal judgement on the relative distance of the wall. This is a method of direct psychophysical scaling based on perception of a stimulus rather than the actual number that the stimulus represents. As with all three of the earlier studies (Ashmead et al., 1995; Rosenblum et al., 2000; Speigle & Loomis, 1993), the expectation was that ordinal properties would be maintained regardless of the number the individual chose especially after several trials had been performed. Participants were not told that there were only three distances, allowing for some error in judgement to occur. A seven-point rating scale was developed to correspond with the average number of steps a participant would need (0.33 m blindfolded per step) between the first distance (0.5 m) and the final distance (2.5 m). The expected responses based on the target distances were 1 for the 0.5-m distance, 4 for the 1.5-m distance and 7 for the farthest, 2.5-m, distance. Although these were the expected responses, each individual developed a personal scale rather than being provided an absolute scale. They were told however, that 1 represented the distance closest and 7 represented the distance farthest away. This is a form of direct scaling based on linear intervals (Allard, 2001).

After giving the judgement response, participants were permitted to continue walking to the location of the wall. This form of training allowed for self-calibration instead of information provided by the experimenter or no information. This is similar to the method used by Hughes (Hughes, 2001) in a study of the KASPA to judge apertures such that self-feedback lends itself to a more ecological approach to learning how the system functions than with experimenter feedback or no feedback. It also allowed the participants to change their scale on future trials if they felt it was necessary based on additional feedback.

Rosenblum et al. (2000) also trained individuals with a set of training trials which allowed the presentation of feedback but provided no additional feedback of their success once testing began. Pilot testing of this experiment showed that individuals naïve to the project became frustrated without any feedback about the actual location. The second pilot participant commented that it left “no reward” to him, suggesting that he wanted to know how successful he was. With feedback, the individuals had more of an incentive to continuously improve their calibration (especially on those occasions when an individual ran into the wall as they had completely misjudged the distance).

The input from the pilot participants required several significant changes to the protocol. The training procedure was reduced from 10 trials per condition to five minutes, the distance estimation became a scaled judgement rather than a distance measured, and the individuals were continuously allowed to interpret feedback for future judgements. These changes allowed the experimental sessions to be reduced from twelve hours to three hours as well as simplifying the



protocol to allow for training while participating. The next section discusses the exact method used for the test protocol.

## 6.5 Method

Nine individuals with no known deficiencies of hearing were tested (Figure 6.3). Participants read and signed a consent form approved by the University of Waterloo Human Research Ethics Committee and the reviewed by the University of Auckland Human Participant Ethics Committee (Appendix 1). The total testing time for each participant took approximately three hours.

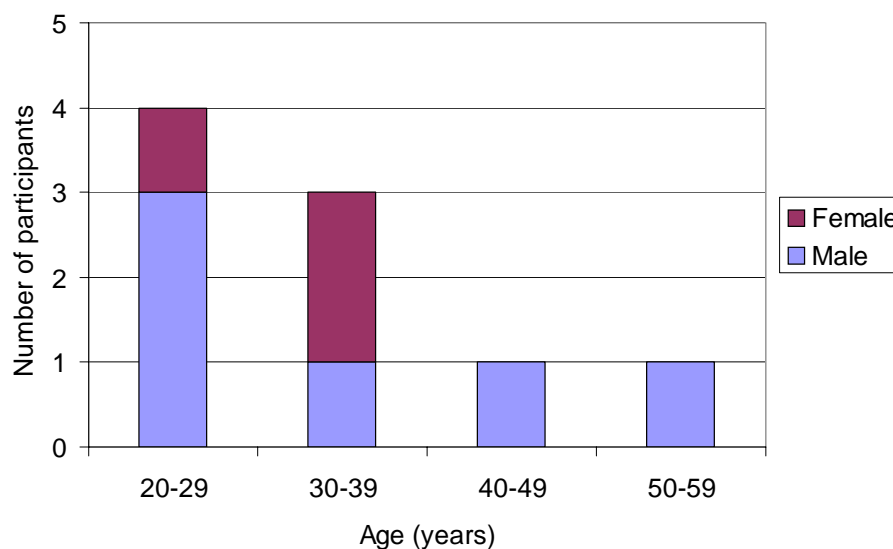


Figure 6.3 Demographics of participants in the perception of distance study.

Hearing ability was evaluated using a generic hearing test of each ear with headphones using 6 tones stepped from 250 Hz to 8000 Hz at three different volume levels (10 dB, 30 dB, and 50 dB). Two participants were not able to hear the higher frequency signals (8000 Hz) at 10 dB, but since the perceived Doppler shift was not expected to reach signals that high (the individual

would have to travel at a speed of 70 m/s), the participants were still permitted to complete the experiment. Each subject was blindfolded and the device was fitted in one of three conditions. The three conditions included: auditory echolocation, device echolocation with receivers oriented forward (same direction as eyes), and device echolocation with receivers oriented outward (same direction as ears). The auditory echolocation involved a speaker generated square wave click of 1000 Hz for 10 ms. The speaker was mounted on the participant's head and the participant used their own ears to judge distance. For the ultrasound conditions, the individuals used the ear defender mounted device which provided information in Doppler. All conditions were counterbalanced among participants to account for any order effects.

After being fitted with the device, the subject was taken into the anechoic chamber and led to a rope which crossed the chamber (Figure 6.1). This rope acted as a guide along which the subject walked for all experimental trials. The subject was permitted 5 minutes with the wall in place (in a different location than the test locations) to explore the nuances of the device after which testing began. The individual faced the wall, turned on the device, walked to a point 1.5 m from the start (as indicated by a piece of tape on the rope), and provided a judgement. They were permitted up to ten seconds to walk the 1.5 m and then provide their judgement. If the judgement was not made within ten seconds the experimenter said "I need your judgement now" and took the response. If one was not made, the trial was repeated later in the session. This happened on two trials of only one participant. Once the judgement was made, they were allowed to walk to the actual location of the wall and based on that information permitted to recalibrate for future trials.

The subject performed ten trials at each of three distances, 0.5 m, 1.5 m, and 2.5 m, in a random order with the only condition being that no three consecutive trials would be the same distance. The subject was then guided out of the chamber and permitted a rest prior to refitting with the device in the next orientation. This process was repeated for each condition.

## **6.6 Results and Analysis**

For each subject the perceived distance score was subtracted from the target distance score for each trial and for each individual. If the individual had been performing an action-oriented guidance task, the magnitude from the perceived distance would have been subtracted from the magnitude of the target distance (as with the pilot subject) so this calculation compared directly to the action-oriented guidance task.

Relative differences between the actual position of the obstacle and the estimated distance of the wall in three conditions were calculated. The trial effects were also evaluated. In addition, accuracy as determined by constant error was also evaluated as with Ashmead (1995) and Rosenblum (2000) such that  $(\text{distance estimate}/\text{target distance}-1)*100$  gave a percent constant error. These scores indicate a tendency to undervalue or overvalue the actual distance. Positive results indicate a tendency to overshoot, whereas negative indicate a tendency to underestimate the target distance (Figure 6.4).

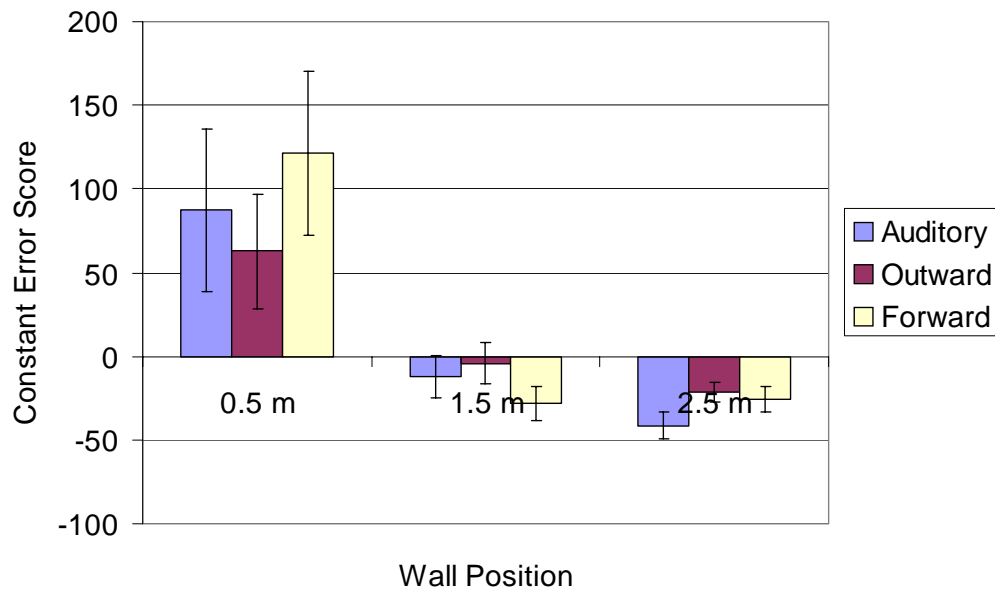


Figure 6.4 Constant error scores.

An ANOVA ( $3 \times 3 \times 10$ ) was performed on the relative estimated distance with Condition (auditory, ultrasound forward receivers, ultrasound outward receivers)  $\times$  Location (1, 4, and 7)  $\times$  Trial number (1 through 10). For any significant differences such that  $p < 0.05$ , a Tukey post-hoc analysis was performed to report any specific differences. In addition, an ANOVA based on the standard error scores was also performed with the Condition  $\times$  Location  $\times$  Trial number. Tukey post-hoc analyses were conducted on those significant results as well.

An analysis of variance showed that there was an interaction effect of condition and distance  $F_{(4, 720)} = 14.61$ ,  $p = 0.0001$  (Figure 6.5) of the estimated distance. There were significant differences among each distance suggesting that for each condition, ordinal results for distance were obtained. At the distances of 0.5 m and 1.5 m, there was no significant difference in perceived distance response among the three conditions. At the farthest distance (2.5 m), there

was no significant difference in estimation using either of the ultrasound conditions, but that of the auditory condition was significantly lower ( $p=0.0001$ ).

This was also qualitatively observed during the testing as many commented that they could hear the change in pitch allowing them to determine the closest location in the auditory condition.

Some individuals also commented that the outward facing receivers in the ultrasound condition provided more accurate information so they could better perceive the differences in distance.

The linear nature of the yellow curve supports this finding. Six individuals also adopted a rocking strategy with both the auditory signal and the ultrasound signal to permit them additional information while “moving” over the ten second observation period.

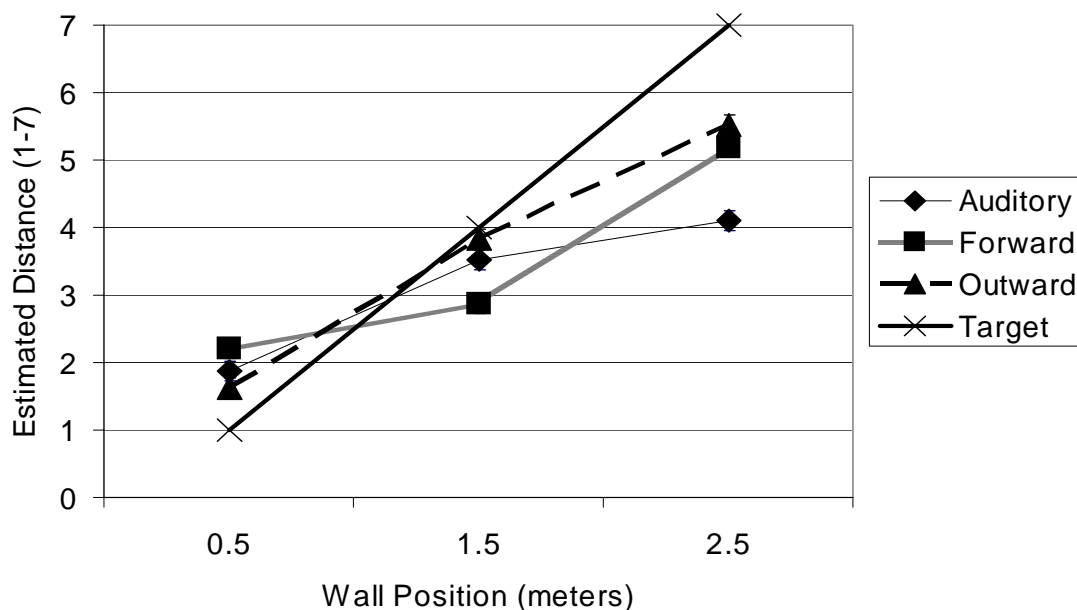


Figure 6.5 Interaction effect of Distance and Condition based on raw data.

There was also a significant main effect of the condition  $F_{(2,720)}=8.12$ ,  $p=0.0003$  (Figure 6.6). The effects of condition showed a better distance estimate with the receivers oriented in the outward direction as compared to the auditory condition. The average response over all trials would be expected to be four. It also showed marginally better responses with the forward direction of receivers as compared to the auditory sense.

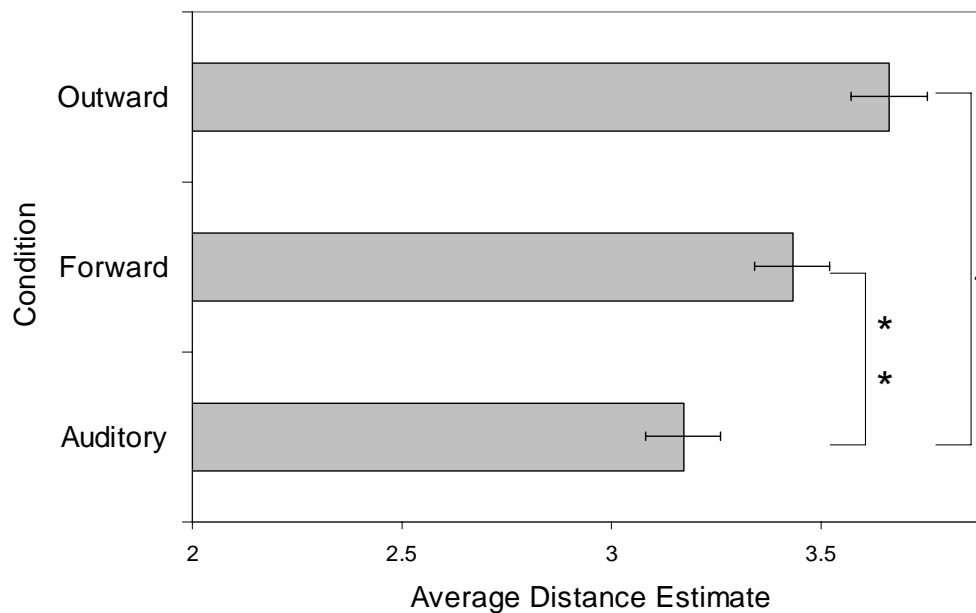


Figure 6.6 Main Effect of condition. The \* represents significance at the  $p=0.05$  level, the \*\* represents significance at the  $p=0.10$  level.

There was a significant main effect of the estimated distance  $F_{(2, 720)}=314.27$ ,  $p = 0.0001$  displaying an ordinal response (Figure 6.7). The judgement of each distance was significantly different from the other two ( $p=0.0001$ ).

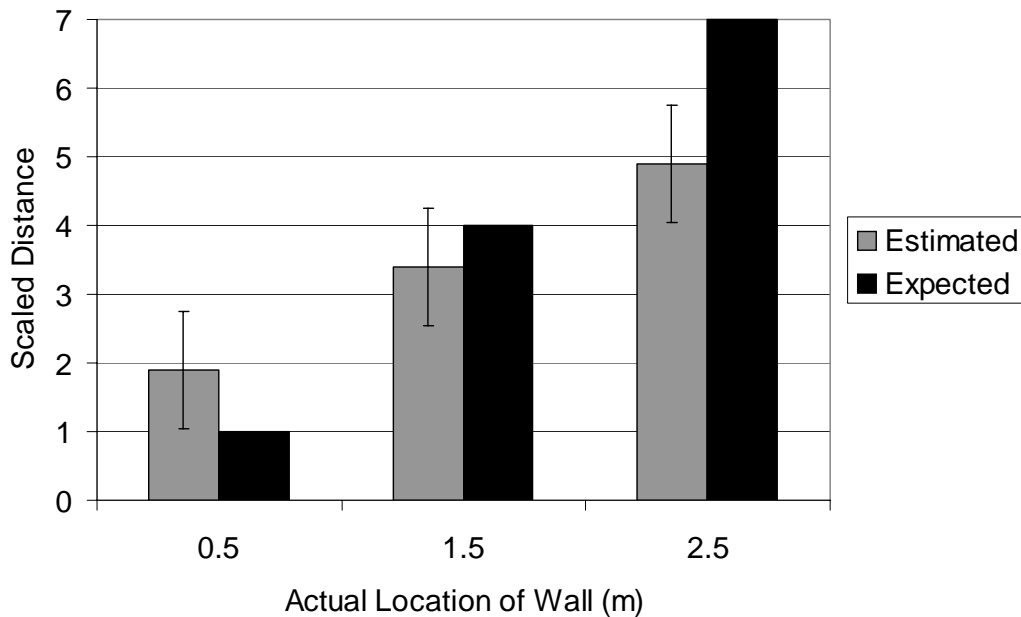


Figure 6.7 Main effect of distance on raw perceived distance scores.

The constant error results showed the interaction effect of condition and location  $F_{(4,712)} = 7.63$ ,  $p = 0.0001$  (refer back to Figure 6.4). There were significant differences at each of the locations, but not among the conditions. At the 0.5-m distance overestimates were observed, and at the 1.5-m distance the results were almost zero, representing a fairly accurate estimate. Finally, at the 2.5 m distance, participants tended to underestimate the location of the wall. There was also a main effect of wall location  $F_{(2,712)} = 193.06$ ,  $p = 0.0001$  (Figure 6.8).

Neither the raw data ( $F_{(9,712)} = 0.99$ ,  $p = 0.4444$ ) nor the constant error data ( $F_{(9,712)} = 1.67$ ,  $p = 0.0929$ ) showed any evidence of trial order effects. As individuals were given five minutes to practice prior to starting the test trials, it is possible that this amount of time was sufficient to learn enough about the device to be able to range-find effectively.

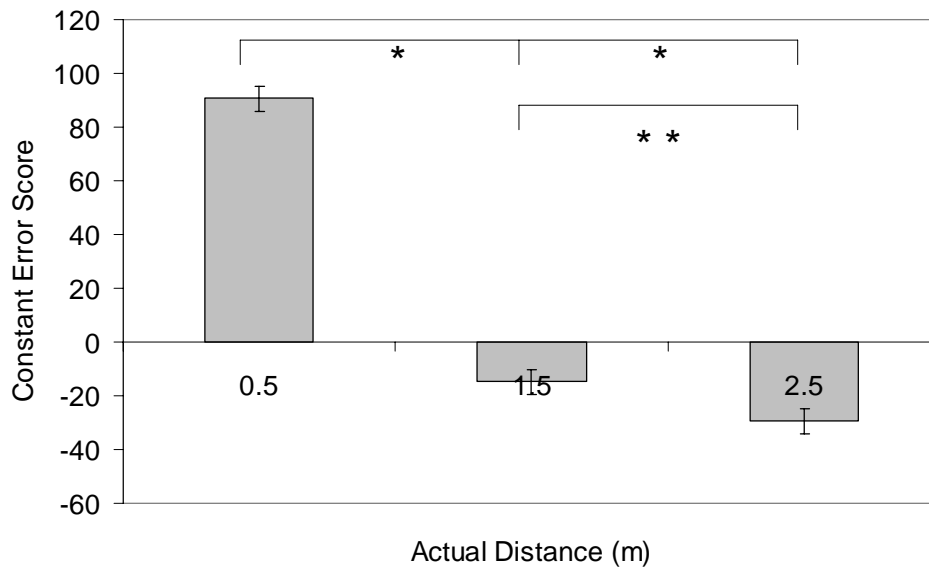


Figure 6.8 Main Effect of distance on constant error scores. The \* represents significance at the  $p=0.05$  level, the \*\* represents significance at the  $p=0.10$  level.

Qualitatively, participants enjoyed the reward of “checking” their responses by walking to the position of the wall. Many participants used comments like “oops, no, that was a five” if they were incorrect in their observation and made expressions like “yes!” and “just right” if they were correct. This reward system led to the participants actually enjoying the testing rather than treating it as a difficult task.

## 6.7 Discussion

Gibson suggests that starting to move creates an optic flow outward, such that the ambient optic array flows away from the point to which one is travelling. The point to which one is travelling is invariant in the midst of centrifugal flow (Gibson, 1979). From the perspective of Doppler originated sound, an individual only senses the obstacle in the environment when movement is initiated. If movement is initiated toward the obstacle (in this case the wall), the obstacle is



magnified in the field of view such that the acoustic stimulus increases in intensity. The increase in intensity can be related to magnification in vision, such that an obstacle is perceived as being closer when the intensity seems to reach an asymptote.

Rather than requiring participants to walk to the location of the wall during testing to estimate the distance as with Rosenblum et al. (2000), the individual was allowed to give a scaled response of distance from one to seven. This was a personal scale, but due to the personal nature of the task which was highly dependent on factors such as the speed of approach, the actual position of the receivers and the ability to perceive sound, this scale was reliable for testing differences between conditions while performing the same task. The individual was permitted a chance to “check” the result after announcing a number allowing for calibration of future trials (if the individual felt that it was necessary). This is similar to a feedback control system, where the ability for the system to calibrate is provided and the system can apply that feedback. Feedback results from purposive action when an individual is locomoting to a specific location (Gibson, 1979) and thus is not a perceptive form of kinesthesia.

The lack of change on a trial-by-trial basis during the test supports the possibility of direct perception when making conscious decisions based on acoustic flow. The first few judgements of distance would be entirely perceptive as the feedback information was not sufficient to make conclusive judgements. In a system that uses adaptive learning to increase the precision, many trials are used to provide step increases in accuracy and precision. If, after feedback, there was a change in the conscious response indicative of increased accuracy or increased precision, one

could conclude that the feedback provided information that was not already present in the acoustic array. In this case, there was statistical evidence that the participants did not perform more accurately or with reduced constant-error scores. This is consistent with direct perception of the acoustic array enabling similar judgements throughout the experiment, though by no means conclusive.

Previous studies have shown that people with no training in echolocation can determine distances using echoed signals (Ashmead et al., 1995; Rosenblum et al., 2000; Supa et al., 1944b). Rosenblum et al. (2000) evaluated distances as far away as 3.66 m and found constant-error scores to be similar to the ones observed in this test. Since the constant-error scores (which are percent values) for perceived distance in the auditory condition in this experiment were consistent with those in other studies, the method of using scaled judgements was an adequate measure for distance determination.

It did appear important to allow individuals feedback to have an idea of their success. They showed an appreciation of their success by the comments that were made. Visually impaired individuals using this system to provide spatial information would use haptic feedback in the learning of this device and be able to use it for future reference. The issue of how much and what type of feedback to provide individuals is a topic of debate. If the experimenter provides feedback, is the individual calibrating their scale to make the experimenter happy or are they actually learning to use the system effectively? By allowing personal feedback on a personal scale, individuals were learning how the system responds, but as suggested earlier, they may

have become more comfortable with the system, but evidence showed that they perceived the wall at similar distances before and after feedback.

Although ordinal response was evident in the auditory condition, the results from the ultrasound system showed a more accurate judgement of perceived location. The interaction effect of condition and distance indicated that at the distance farthest away (2.5 m), the judgement of perceived distance was significantly lower in the auditory condition than that of the ultrasound conditions. The observed non-linear effect of distance judgement based on sound in the auditory condition has been shown in the past (Ashmead et al., 1995). Sound level falls off by 6 dB for every doubling of distance and this tends to invoke a perception of shortened distance (Ashmead et al., 1995; Speigle & Loomis, 1993). Generally, though not always the case (see Ashmead et al., 1995) the judgement of distance is based on signal intensity. The hypothesis predicted that the three conditions would be comparable. The results show an increased ability to perceive greater distances with the ultrasound system. A more linear perceptive response was evident with the ultrasound system. Since the frequency of the transmit signal is higher than perceptible hearing, the intensity of the outbound signal can be greater than that of an auditory signal (many decibels higher). The result is an increased intensity of reflections. The results suggest that the system allows for perception of intensity that closely resembles audition at short distances, but increased intensity is perceived relative to the auditory range at greater distances. Since the ultrasound responses were more accurate those of the auditory condition, the conclusion can be drawn that at these short distances the ultrasound system provides a better indication of distance perception than that of the auditory condition.

The main effect of condition of the perceived distances indicated a significantly more accurate response on average with the receivers oriented in the outward direction than in either the forward orientation or the auditory condition. Participants also indicated a preference for the orientation of the receivers in the outward direction. They suggested that it was easier to determine a personal scale based on sounds from the outward facing receivers. This is in contrast to what was expected. All other sonar systems orient the receivers in the forward direction to gather as much of the reflected signal as possible. The ultrasound condition with the receivers pointing forward was successful at judging distances but was found to be less accurate than the receivers pointing outward. Individuals commented that it was more difficult to get a good estimate of distances using the receivers in this orientation and also indicated that it was highly dependent on movement of the head. As Gibson suggests “the world is revealed as the head turns” (p111, 1986), but the sensitivity to the perceptive environment appears to be dependent on the reception of acoustic flow available during this process. Future studies might include an evaluation of head movement while assessing the distance to allow for a better understanding of environmental scanning with and without the ultrasound system.

The observation that the receivers oriented in the outward direction may allow better depth perception is in conflict with past theory. Psychophysicists have argued that depth perception is achieved through the disparity of two eyes. The overlapping images provide stereoscopic images which relay depth information. Although stereopsis is known to allow for depth perception, the process used by the brain to achieve this goal is not well understood. Gibson argues that depth information is present in an optical array such that invariants exist to allow this perception. In this study, with the receivers oriented outward, there was very little overlap (if any) between

them that would allow for stereoscopic depth perception, yet this method was more accurate at distance judgement. An acoustic array may provide time-to-collision invariants similar to those in the optic array that would allow for depth to be perceived directly. Future studies evaluating the ability of an individual to perceive collision through auditory information would enable testing of this hypothesis.

Theoretically, range finding is more easily performed with a signal that enters the receiver directly rather than in a direction perpendicular. When the receiver is pointed forward in the same direction as the transmitter, both are at their maximum in the gain pattern. Maximum information is available at a direction immediately in front of the observer's look-direction. This makes the system highly dependent on the direction to which the individual is facing. If the individual looks slightly off to one side of the obstacle the signal is not as strong as if directly facing the obstacle (Figure 6.9). Each participant began the trial facing the wall and was guided toward the wall by a rope. As they could draw on these cues for directionality, they were not expected to turn. Although the body rarely shifted from facing the wall, the participants were permitted to move their head relative to their body. If the observer chose to scan the wall from left to right, it is possible that misinformation could have been observed with respect to distance. In the outward direction, the gain pattern is perpendicular to that of the transmitter. A change in the look direction does not significantly affect the reflected echoes. This suggests that for distance determination, the orientation of the receivers in the outward direction is more accurate than those in the forward direction.

Although time to respond, and time to approach were not measured (as they were not expected to be an influence during the design stage of the study), qualitatively participants performed much more slowly in the auditory condition than in the two ultrasound conditions. As they were only permitted 1.5 m (due to the limitation of the size of the chamber) of moving before indicating a judgement, this stationary echolocation was permitted. Previous studies (Rosenblum et al., 2000) have only allowed five seconds, but participants were also in control over their own echolocating signal. This study provided participants with the signal that they were required to use and this signal was repeating every second. Thus individuals were only permitted ten clicks before a judgement was required. Several participants commented a faster click would have been preferred. Previous studies did not comment on the number of times a participant used a

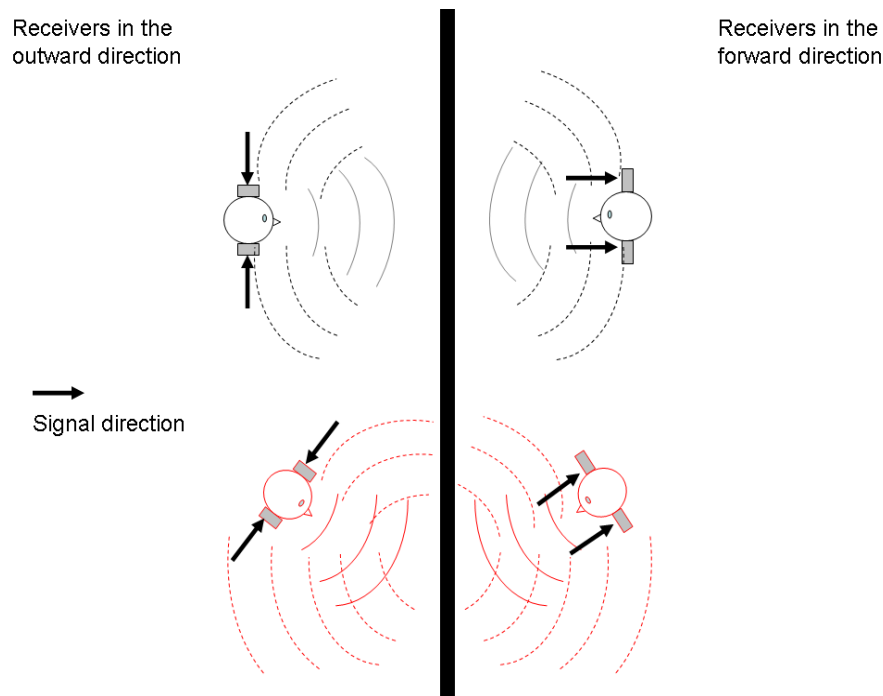


Figure 6.9 Effect of receiver orientation and rotation of head on the direction of signal.

signal while approaching the wall, thus no comments can be made on whether additional information was provided by allowing a longer time to judge the distance. The Doppler signal in the ultrasound condition changed in pitch depending on the speed of the participant and in intensity depending on the relative distance from the wall. In this sense, the moving condition was much more valuable in predicting distance than a stationary condition as more information was available. This additional information provided by the Doppler system is likely the reason that the response seemed quicker in these conditions than in the auditory condition.

Rosenblum et al (2000), Ashmead et al. (1995), and Speigle and Loomis (1993) found that in a moving condition, participants were better able to judge distance than in a stationary condition. In the present test, participants seemed to rock back and forth to obtain additional information in both the ultrasound conditions and the auditory condition even after the 1.5-m approach. This would indicate that moving provided more information than the stationary localisation. This is similar to previous studies by Ashmead and Rosenblum, but rocking has not been suggested by earlier studies as a means to provide motion to help in determining distance. Earlier studies do suggest that Doppler shift and ripple noise pitch may have contributed to enhanced judgements. Rosenblum *et al.* acknowledge that these may not be perceptible per se, but the acoustic components may be used in creating an auditory tau that allows participants to judge the rate of gap closure directly. Since Doppler is more pronounced using this ultrasound system the concept of an auditory tau may be more easily realised, but was not specifically studied in this experiment.

This task which allowed individuals to approach the wall and receive haptic feedback is more ecological or “service of action” (Stoffregen & Pittenger, 1995) oriented than the method used by Rosenblum et al. (2000) for their testing. Action related tasks are those that have specific purposes. Rosenblum moved the wall in that experiment to ensure that additional echoic information would not be provided while the individual was performing the task. By requiring the participants in this study to provide a distance in numbers before actually going to the location of the wall, the participants were provided with further information required for distance judgement. It allowed for recalibration of the auditory sense while allowing individuals a chance to explore the sound further. It also permitted the participants a chance to explore additional perceptual information received during approach such as expansion that may allow them better judgement for future trials. Arguably, this is a better form of judgement of the specific action related task of relative distance judgement.

## **6.8 Conclusions**

The results from this experiment suggested that in all three conditions, ordinal results of distance were observed, such that participants could differentiate between obstacles close to them and obstacles farther away. Also, participants were able to identify distances more effectively using the ultrasound system than the auditory condition and appeared to rely on movement, which resulted in acoustic flow, to make their determination. The orientation of the receivers in the outward direction provided for significantly better results overall indicating that receivers oriented in the same direction as the “ears” may provide better judgement of distance of environmental obstacles. No evidence of a training effect was observed for either the auditory



condition or the conditions with the ultrasound receiver suggesting that in fact direct perception of environmental obstacles through the acoustic array was achieved.

The results of this first study showed that the ultrasound system could be used to make relative distance judgements, but this is a perception report type task. A more dynamic approach to evaluating closure rate with a gap may provide more insight into how an individual perceives differences in the environment with this system. The next experiment requires the movement of an individual between two obstacles and provides for an understanding in a more ecological test environment. This test is presented in the following chapter.



## **7 Comparison of downconverted ultrasound echoes to vision for passing through apertures by moving human subjects**

This experiment was designed to allow for evaluation of perceptual accuracy during a perception-action event (Stoffregen and Pittenger, 1995). Generally, individuals are not consciously aware of their responses to echoes and Stoffregen and Pittenger indicate that behaviour in an ecological perception-action event may be more indicative of perceptual response than perception and conscious report. The previous study examined the behaviour of individuals in their ability to consciously report the distance to a wall. This conscious response requires that the individual perceive then calculate, thus endeavours to examine indirect perception, rather than direct perception of auditory arrays. On the other hand, passage through an aperture only requires direct perception to be able to traverse through.

### **7.1 Purpose**

This study examined localisation performance of audified ultrasound echoes in two orientations by moving individuals as they passed through an aperture. These two conditions were compared to a vision scenario.

### **7.2 Hypotheses**

H1: The time to pass through the aperture will be much shorter in the vision condition than the audified ultrasound conditions which will be similar to each other.

P1: The participants have had considerable practice using vision as the main mechanism for guidance control. Although this would lend itself to testing participants who are functionally blind, the information presented through Doppler shift is new to even those who are visually

impaired. Auditory Doppler shift is not pronounced enough to rely primarily on Doppler shift for travel, thus even individuals with visual impairments would likely have a reduced speed when they first pick up the device. Another factor in the ability to perceive the aperture is that the ambient light in the optic array will be reduced by using a small headlamp but reflections from the floor will contribute additional information to the visual array. The prototype transmitter is limited in the area that it can illuminate due to its size and cannot provide as large an acoustic array as available from the optic array. This additional information will allow for increased speed in the optic array. In the previous study, there were no differences in the ability to judge the distance between the two ultrasound conditions, thus the time to perceive and react to the closing gaps will be similar between the two orientations.

H2: Vision will allow for a more accurate passage through the centre of the aperture, but the ultrasound receivers oriented in the outward direction will be better than the forward direction.

P2: The use of vision will allow the individual to see through the door and aim toward the centre. The orientation of the receivers in the outward direction will permit comparable centreline results since the sides of the aperture will easily be perceived. The orientation in the forward direction will be ambiguous making it difficult to determine the exact position of each of the sides of the aperture.

H3: Smaller gap sizes will require more time to negotiate, but the ability to pass through the centreline will be improved compared to larger gap sizes.

P3: Smaller gaps require the ability to detect the aperture more precisely than the walls in the previous study. Smaller gaps will require increased precision and more care will be taken on the approach. This will result in a more accurate passage through the centreline.

H4: Shoulder rotation through the aperture for the vision condition will be greater for the narrower apertures than the wider ones. The angle of rotation in the audified ultrasound condition with the receivers oriented in the outward direction will be more similar to vision than the receiver orientation in the forward direction.

P4: With vision, individuals tend to rotate the shoulders to pass through apertures that are 1.3 times the shoulder width when normalised (Warren, Jr. & Whang, 1987). The size of the gap will be accurately determined with the receivers oriented in the outward direction, thus the angle of passage will be closer to that of vision than the forward orientation.

### **7.3 Test Environment**

To reduce the effect of multiple reflections, an anechoic chamber was used. For this experiment, a floor structure was installed in the chamber. This structure consisted of medium density particle board mounted on a scaffolding under which structural floor was covered with absorbing foam wedges, thus there were possible reflections from the scaffolding floor as well as the aperture structure. Since the ultrasound transmitter was mounted on the participant's head, the reflections from the floor in the ultrasound condition would only be detected if the individual looked down. The aperture was created by two sheets of medium density particle board that was 0.6 m in width by 1.2 m in height. These sheets hung from a cord-drawn curtain rail mounted on a scaffolding structure (Figure 7.1). These sheets could swing freely if participants ran into them during the test procedure. Two lengths of webbing were crossed at the back to ensure stability of the scaffolding structure. These were mounted high enough that they would not interfere with the reflections from the transmitter.



Figure 7.1 Experimental apparatus for aperture testing within the anechoic chamber.

#### 7.4 Method

Twelve individuals with no known deficiencies of hearing participated in this study. The demographics are shown in Figure 7.2. There were five participants (2 female) that had participated in the earlier study. Each participant read and signed an informed consent approved by the University of Waterloo Human Ethics Committee and reviewed by the University of Auckland Human Ethics Committee (Appendix 2). Hearing ability was evaluated using a generic hearing test of each ear with headphones using 6 tones stepped from 250 Hz to 8000 Hz at three different volume levels (10 dB, 30 dB, and 50 dB). All tones were played on an audio player and the participant would indicate “ok” if the sound was heard. No difficulties with hearing were observed. A measurement of shoulder width was taken to determine whether shoulder rotation would be expected through the apertures.

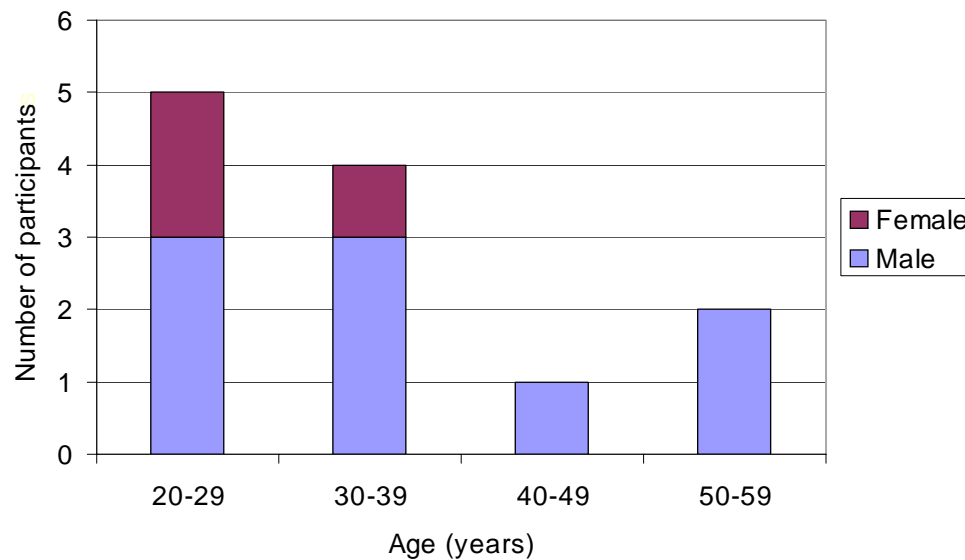


Figure 7.2 Participant demographics for aperture testing.

A similar approach to doorway passability was followed as with Davies and Patla (2004) such that the participants used sound to detect a door opened to one of five apertures and then passed through it. The earlier experiment involved blindfolding the individuals and several commented that the blindfold was uncomfortable. As this testing required movement and assessment of apertures, the testing was conducted in complete darkness rather than expecting the individuals to wear a blindfold. Each participant wore a retro-reflective safety vest which could be viewed using infrared light. A video camera operating in infrared light was used to collect motion of the subject and was mounted immediately above the aperture.

Three conditions were examined; audified ultrasound echoes with the receivers oriented in the eye direction, with the receivers oriented in the ear direction, and finally vision. The previous experiment which required individuals to echolocate a wall showed that the ultrasound system was better than the auditory sense at detecting walls. With auditory echolocation, both doors

would have to be detected independently by the user. Previous studies (Supa et al., 1944b) as well as the first study, showed that blindfolded individuals with limited training have difficulty detecting obstacles with audible echolocation and often run into them. It was determined that aperture passability with audible echolocation alone was a task too difficult for novices to perform. Since the ultrasound system appeared to be working effectively based on the previous test, the audified ultrasound signal information was compared to that of vision. Although performance in the ultrasound condition was not expected to be as accurate as that of vision, vision was used as a “gold standard” for testing. For the vision condition, the participants used a light located at the top of the head, like a spelunker. This allowed a similar area in front of the participants to be illuminated as that of the ultrasound transmitter.

Five doorway apertures from 55 cm to 95 cm were used to evaluate the ability to negotiate different gap sizes (Davies and Patla, 2004). Each participant negotiated the gap five times for each aperture and in each condition. Thus there were 25 trials for each condition. All conditions were counterbalanced among individuals to eliminate the crossover effects between conditions.

In each of the two ultrasound conditions, participants were given five minutes using the device in the environment with the lights on and the aperture opened to a distance of 75 cm. They were permitted to visually guide themselves through the aperture during practice and use whatever techniques they chose. After five minutes of self-directed training, the room was darkened and the trials began.



A trial started with the participant facing the wall opposite to that of the aperture at a distance of 3 m (the distance walked was limited by the size of the 5 m x 5 m anechoic chamber). This allowed the experimenter to select the correct door aperture without the individual sensing the size of the aperture in advance of the start of the trial. As the experimenter indicated “go” to the participant, the participant turned around 180° independently choosing to turn clockwise or counter-clockwise, walked toward the aperture and passed through the middle of it. After passing through, a light was turned on and the individual turned and returned to the initial location. This technique allowed some learning to occur in the direction opposite to that of the trial as they could observe information about the aperture size and their passage through. No information was given to the user about the indicators (time, centreline accuracy and angle of rotation) used to evaluate ability to pass through the aperture to avoid any experimenter influence of the results.

## **7.5 Analysis**

Video analysis was performed using Matlab. The brightest image from each trial in the immediate vicinity of the aperture was the image chosen to perform further processing. Since the camera with the infrared sensor was mounted immediately above the aperture at the centreline, the brightest image would be the image of the person just prior to passing through the aperture.

The time to pass through the door from being told to “go”, relative differences between the centre of the door and the midline of the individual, and shoulder rotation through the aperture were calculated in all conditions. A repeated-measures ANOVA using the two groups (novice

and experienced), three conditions (auditory, outward orientation of receivers, forward orientation of receivers), trial order (1, 2, 3, 4, 5) and the relative differences at the five apertures (55 cm, 65 cm, 75 cm, 85 cm, and 95 cm) was performed ( $2 \times 3 \times 5 \times 5$ ). In addition, the precision was calculated as the standard deviation among trials by each subject and statistics performed on these measures (Patla et al., 2004) such that there was a Group  $\times$  Condition  $\times$  Aperture comparison ( $2 \times 3 \times 5$ ). This statistical evaluation allows for differences in precision to be determined in addition to those in accuracy (Patla et al., 2004). Prior to statistical analysis, standardized residuals and Cook's statistics (to determine outliers) of the raw data were conducted to evaluate the presence of non-constant variance or non-normality of the data. Evidence of non-normality required transformation of the data. The approach for data transformation was determined by the transformation with the weakest effect to those transformations that resulted in stronger effects, for example, square root, then  $\log_{10}$ , and finally the inverse transformation. After applying each transformation, Gaussian distribution was evaluated (repeating the standardized residuals and Cook's statistics). If normality was not achieved, the next transformation was applied. Once the effective transformation was determined, the repeated-measures ANOVA was conducted. A Tukey post-hoc test was performed on any differences identified by the ANOVA as having a p-value of less than 0.05.

### ***7.5.1 Time to pass through door***

Since there was non-constant variance and non-normality in the time data, conversion of the data was required to achieve constant variance and normality. The reciprocal-time data were used for the statistics calculations. The time to pass through the door was affected by the condition, the order of passage and the group. There was a main effect between the vision condition and the ultrasound conditions, but no differences existed between the two ultrasound conditions ( $F_{(2,823)}$

= 530.68,  $p = 0.0001$ , Figure 7.3). The main effect of order showed a learning effect such that the first trial took significantly longer than the other trials for each aperture ( $F_{(4,823)} = 4.13$ ,  $p = 0.003$ ). There were no other trial effects among the remaining trials. There was an effect of group such that those individuals who had not participated in the earlier experiment were significantly faster at passing through the aperture than those who had used the device previously ( $F_{(2,823)} = 14.19$ ,  $p = 0.0002$ , Table 7.1). The variability showed a significant interaction effect of Group  $\times$  Condition ( $F_{(2,150)} = 4.42$ ,  $p = 0.0136$ , Figure 7.4) such that vision was significantly less ( $p=0.0001$ ) for both groups than the other two conditions. In general, the novice participants had decreased precision with respect to time with the outward orientation of the receivers, whereas the opposite was true of the experienced group. This may be indicative of a crossover effect from the previous study of distance judgement. The experienced group had indicated a preference for this orientation of receivers and this may have enabled them to be more precise in their timing during the perception action task of aperture passage. The novice participants had no preconceived preference for a specific orientation and may have found the information from the forward facing receivers more decipherable. There was also a main effect of condition as variability of vision was also significantly less than the variability of the other two conditions ( $F_{(2,150)} = 530.68$ ,  $p = 0.0001$ ).

Returning to the original data and evaluating Figure 7.5, it may be suggested that the reciprocal transformation was perhaps too strong for the time data and there was possibly evidence of a bimodal distribution of the time through the passage (Figure 7.5) for the ultrasound conditions. Further analysis was conducted by separating the times into two groups, those below 20 seconds and those above. An independent  $2 \times 3 \times 2 \times 10$  ANOVA was conducted (Group  $\times$  Distance  $\times$

Ultrasound condition  $\times$  Trial order) examining the differences between the two groups. A significant difference was evident in the interaction of group by condition ( $F_{(1,294)} = 9.40$ ,  $p = 0.002$ ) and group by trial order ( $F_{(4,294)} = 4.16$ ,  $p = 0.003$ ) and the main effect of group ( $F_{(1,294)} = 171.34$ ,  $p = 0.0001$ ). Tukey post-hoc tests were performed. These showed that for the group that took a shorter time to pass through, there were no significant differences between conditions or trial order, but for the group that took longer, there were significant differences between the conditions (Figure 7.6) and the trial order. The trial order effect showed that the first trial and the last trial took significantly longer than the second and the fourth.

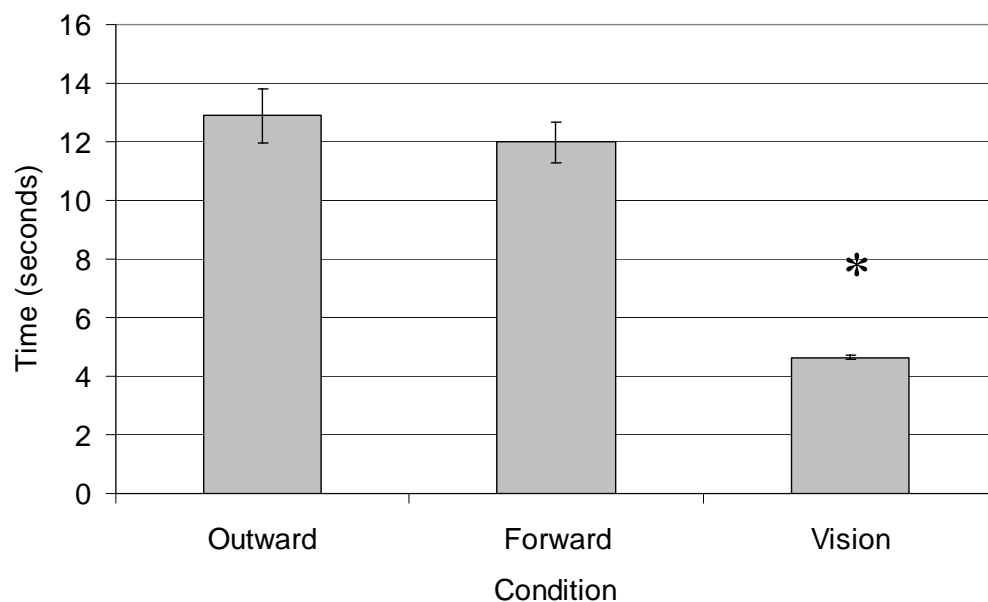


Figure 7.3 Mean time for each condition to pass through the aperture with bars representing the standard error of the mean. The \* represents significant difference from the other conditions at  $p < 0.05$ .

GROUP	TIME THROUGH APERTURE	STANDARD ERROR
Previous practice	10.96 seconds	<i>1.09 seconds</i>
No practice	9.08 seconds	<i>2.01 seconds</i>

Table 7.1 Significant differences between groups based on the time to pass through the aperture.

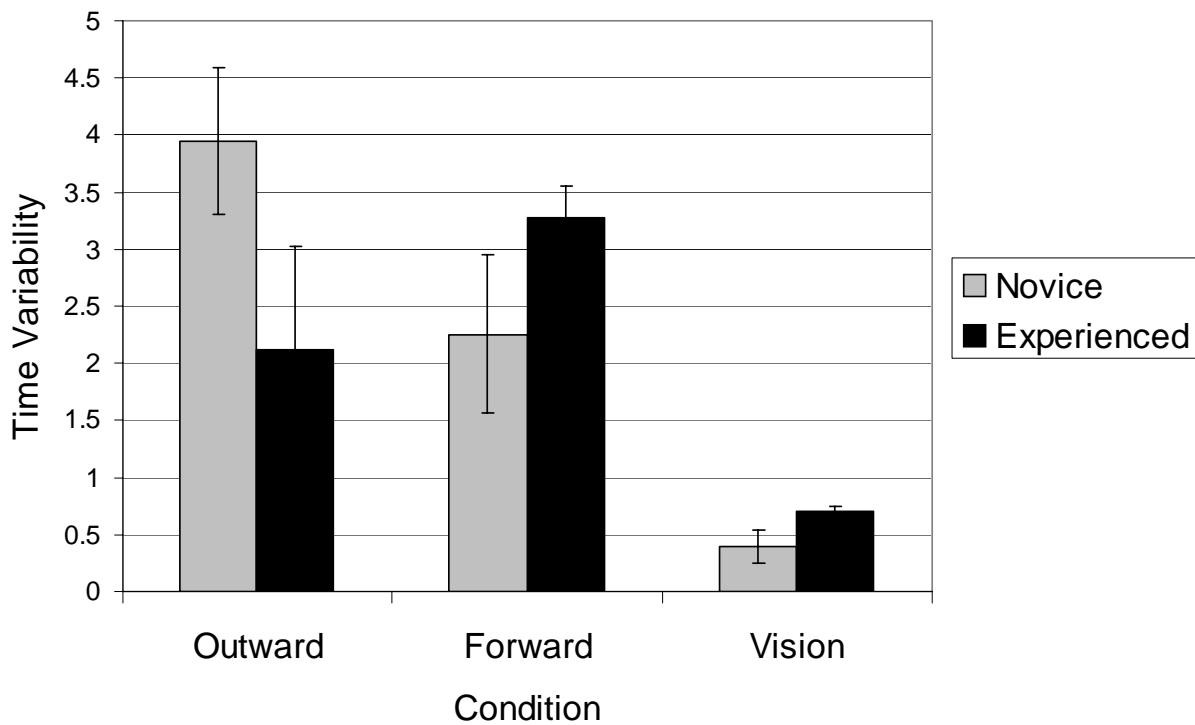


Figure 7.4 Interaction effect of Group  $\times$  Condition from the variability of time to pass through the aperture.

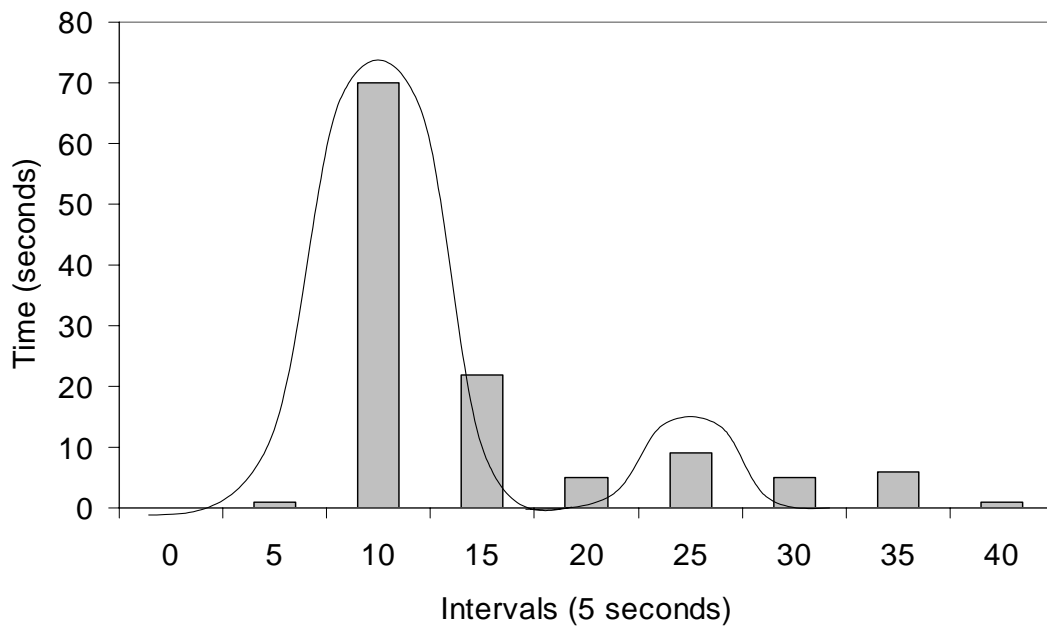


Figure 7.5 Histogram of the distributions in time for the ultrasound conditions.

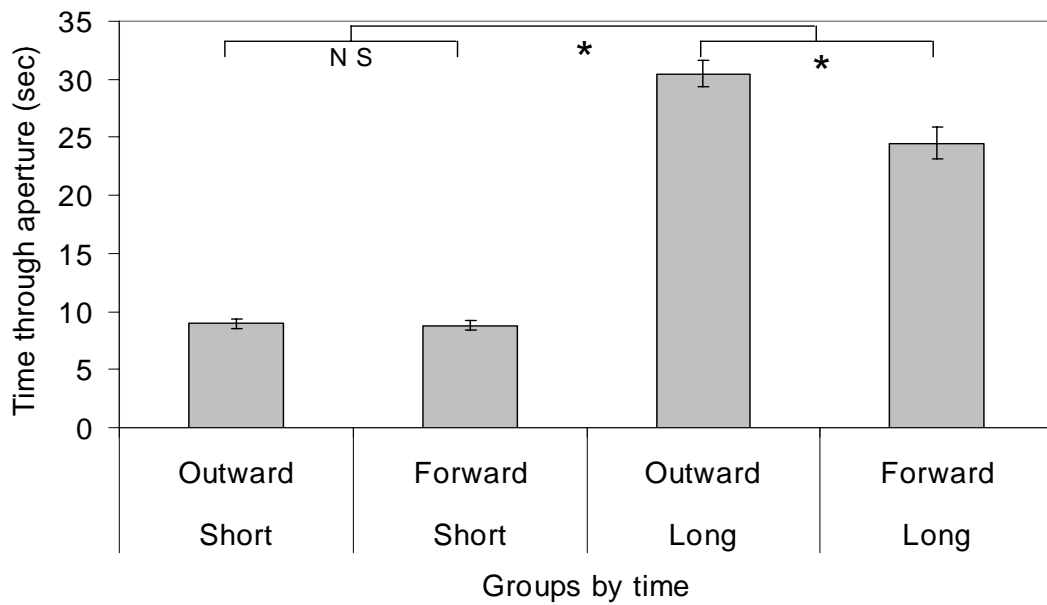


Figure 7.6 Interaction effect of group by condition for the two groups based on time through the aperture. The \* represents significant difference from the other conditions at  $p < 0.05$ .

### 7.5.2 Ability to pass through midline of aperture

The only significant difference in the distance from the midline of the aperture was that of the group effect. The novice group was significantly farther away from the midline of the aperture than the experienced group ( $F_{(1,814)} = 23.95$ ,  $p = 0.0001$ , Table 7.2). There also appeared to be passage through the left side rather than the right, though it must be noted that the mean value of this difference from the centre was only 3.6 cm for the novice participants. The statistics on the variability of passage through the midline showed main differences in the variability for the conditions and the aperture sizes. The variability was significantly greater for the ultrasound conditions than that of vision ( $F_{(2,150)} = 23.95$ ,  $p = 0.0001$ , Figure 7.7). Also, the variability was dependent on the aperture size, the greater the aperture the greater the variability among passage through the centre ( $F_{(4,150)} = 2.7$ ,  $p = 0.03$ , Figure 7.8).

GROUP	DISTANCE FROM MIDLINE	STANDARD ERROR
Experienced	5.5 pixels (0.69 cm)	4.0 pixels (0.5 cm)
Novice	28.6 pixels (3.6 cm)	4.7 pixels (0.6 cm)

Table 7.2 Main effect of group with respect to distance from midline.

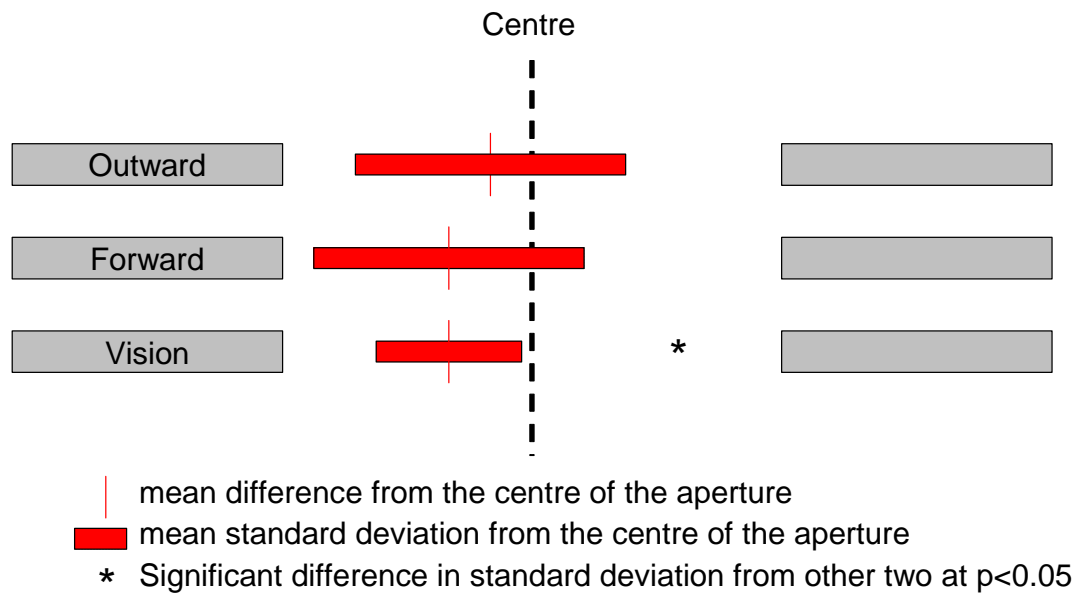


Figure 7.7 Mean distances and variability from the centre of the aperture based on condition (aperture size is not to scale)

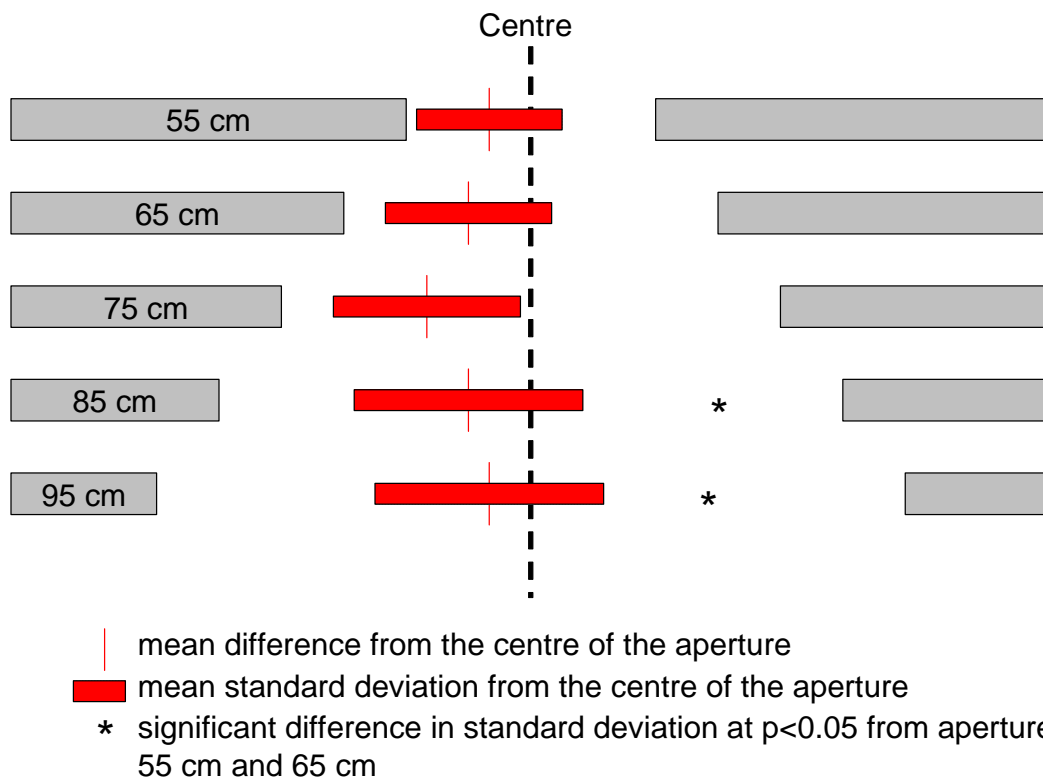


Figure 7.8 Mean distance and variability from centre of aperture with respect to each aperture (aperture size is not to scale).



### **7.5.3 Angle of passage through the aperture**

The amount of turning to pass through the aperture as defined by the angle of the shoulders with the line passing between the two boards was determined. Warren and Whang (1987) found that in a vision condition there was a critical ratio,  $\pi$ , between the aperture width and the shoulder width of the participant of 1.3, below which the individual would rotate the shoulders to pass through the doors. Based on this ratio, the only aperture at which rotation would occur is the 55 cm aperture (mean  $\pi = 1.20$ , minimum  $\pi = 1.10$ , maximum  $\pi = 1.34$ ).

The statistics on this data set showed non-normality and the data were transformed using the square root to be able to apply statistical tests. The dominant effect was the interaction effect between condition and the size of the aperture ( $F_{(8,825)} = 4.71$ ,  $p = 0.0001$ , Figure 7.9) such that the angle through the smallest aperture, 55 cm, was significantly greater than the angle through all other apertures. There were also main effects of both condition ( $F_{(2,825)} = 4.55$ ,  $p = 0.01$ , Figure 7.10) such that the condition of the receivers facing in the eye direction was less than the other two conditions and aperture ( $F_{(4,825)} = 4.01$ ,  $p = 0.003$ ) when collapsed across all apertures. There was also a main effect of group ( $F_{(1,814)} = 7.44$ ,  $p = 0.007$ ) such that the novice group had a larger angle than that of the experienced group. With respect to the variability in the angle as the participant passed through the aperture, there was an interaction effect of Group  $\times$  Condition ( $F_{(1,150)} = 4.92$ ,  $p = 0.009$ , Figure 7.11). This interaction is difficult to describe as it suggests that the two groups rotate with similar precision among the different apertures in the ultrasound condition, but there is decreased precision by the experienced participants in the vision condition relative to the novices. Perhaps these participants were more comfortable with the experimental environment (as they had been inside the anechoic chamber for the previous test) and did not

focus on the possibility of contact with the sides of the aperture as closely as the novice participants.

GROUP	ANGLE THROUGH APERTURE	STANDARD ERROR
Experienced	6.74°	1.41°
Novice	8.11°	1.62°

Table 7.3 Main effect of group with respect to angle of passage through aperture.

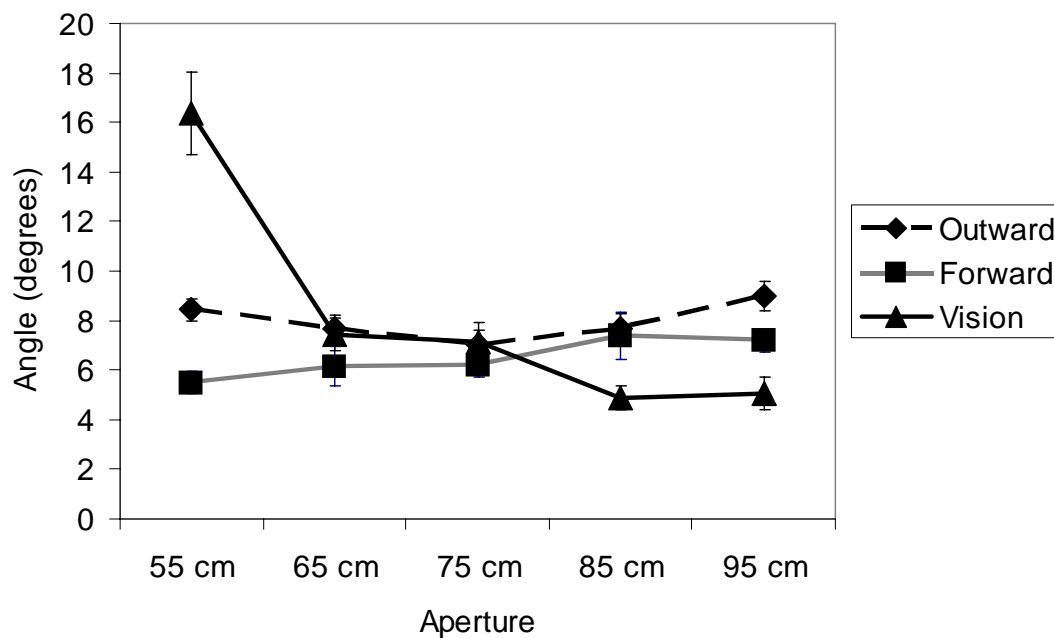


Figure 7.9 Interaction effect of condition by aperture for the angle of passage through the aperture.

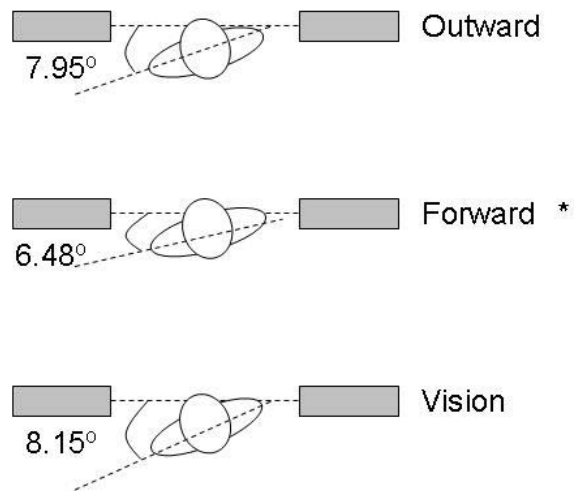


Figure 7.10 Main effect of condition for the angle of passage through the aperture across all apertures.

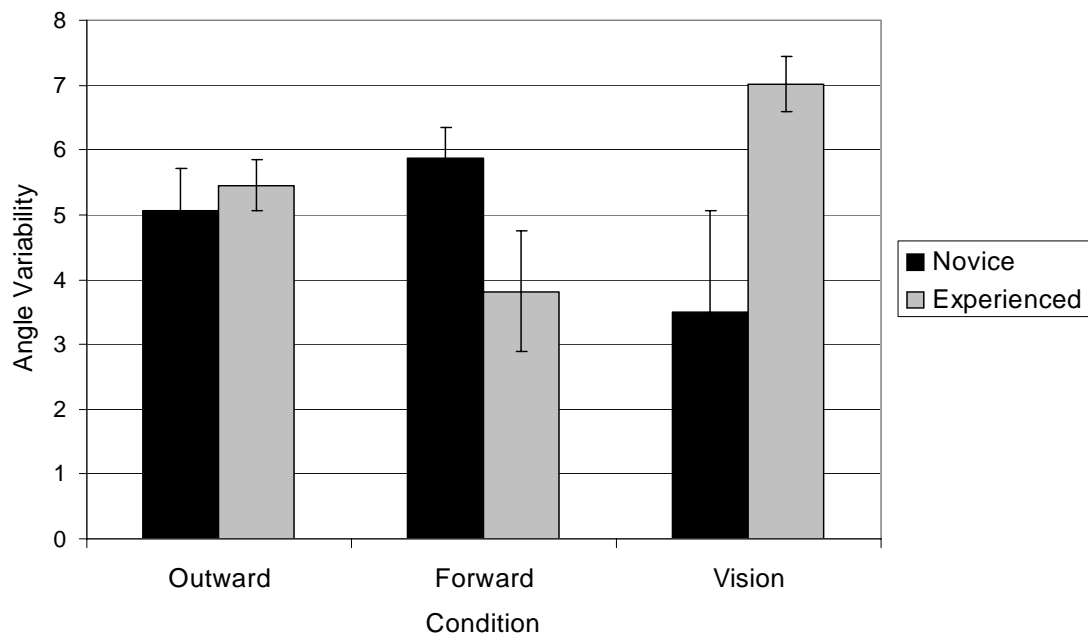


Figure 7.11 Interaction effect of group by condition for the variability among participants.

## 7.6 Discussion

In this anechoic environment, the only element contributing to the noise in the environment was that of Doppler shift. As this signal can only be heard during locomotion, acoustic flow is created as it is in vision. Gibson suggests that the flow of an ambient array in visual light specifies locomotion whereas lack of flow specifies stasis (1979). The same phenomenon is observed with this system. With the Doppler system, acoustic flow is perceived when moving and this concept is grounded in the physical nature of the system. Doppler by definition results from the perceived change in frequency as a result of locomotion.

In an ideal scenario, the individual would hear the sides of the aperture immediately after turning around at the start of the trial. Outflow of the acoustic array would occur on approach as it does in the optic array. Gibson suggests that a move toward the aperture would result in a gain of structure within the opening such that one could see past the structure (p.229, 1979). In this ideal scenario, the gain of structure results in the presence of no additional acoustic array as there are no reflectors on the opposite side. One could, on the other hand argue that the straps beyond the aperture could then be sensed and indicate to the individual to stop. At least one tall participant mentioned the ability to detect these.

In this case, the two receiver orientations affect the acoustic array in different ways. If the individual was centred on the aperture with the receivers oriented in the forward direction, the signals on approach would gradually be reduced in each ear. If the individual approached one side of the aperture, the signal from that side would be slightly louder in that ear. This draws a parallel with the rules for steering in an optical sense. Gibson suggests “to steer, keep the centre

of outflow outside the patches of the array that specify obstacles” (1979, p.233). The individual would then move in the opposite direction to minimise the volume of the signal in both ears. Since the receivers were relatively close together and collecting similar information in this orientation, there would be very little difference between the sounds from each receiver. Once the individual was a short distance in front of the aperture the sound would disappear as there would be no additional reflectors on the previously occluded (or opposite) side.

With the receivers oriented outward, the sound would not be as strong at the start of the trial when the individual turned around. If the individual approached one side of the aperture, the reflection from that side would be considerably louder than the other. Once again, the individual would reorient herself until the sounds were equal in each ear. As the individual passed through the aperture, the individual would still be able to hear reflections from the side of the aperture as though passing through a corridor. This is an ability that humans do not typically sense in that it is the concurrent outflow and inflow at the same time. The placement of receivers outward from the head allow this perception to occur as it does with animals. Once safely through the aperture in either condition, no sound from the acoustic array would be present.

Based on the bimodal distribution of time, there appeared to be two methods of approach to the aperture (see Figure 7.12), individuals may have used either direct or indirect perception in perceiving the aperture. There were individuals who appeared comfortable with the concept that “no noise is good noise” with respect to the acoustic flow passing outward. The system only responds to reflections in the environment, thus if a participant was on the centre path through the doors, they continued through with minimal signals from the acoustic array. These

participants reported moving toward one side or the other based on the acoustic information from the binaural signal signalling the presence of an obstacle to one side. This method is efficient and allows for the individuals to approach without stopping which is indicative of a perception-action coupling through direct perception in that they steered to one side or the other to minimise the signal from the acoustic array to focus on a point beyond two obstacles. There were no differences in the time through the aperture between the two conditions and they performed significantly more quickly than those who used an indirect approach.

Two participants appeared to “move toward the sound”. These participants identified each edge of the aperture before attempting to pass through the centre. This follows along the concept of indirect perception in that they wanted to process the environment before making a decision. These individuals did not appear as confident in the system to provide judgement to them. As these participants wanted to hear the edge location of each side of the aperture before passing through, they took longer. Instead of relying on the acoustic array for steering, they needed to fully process the environment. Instead of looking for a gain in the information (silence) beyond the aperture, they approached the obstacle on each side of the aperture perceiving the obstacle (magnifying or increasing the intensity of the sound within the acoustic array) and processed the location of each before passing between them. These individuals would likely require more training than those who approached directly as they were inconsistent in their approach times. Trials one and five were significantly longer than trials two and four at each aperture for these participants. This variability suggests a greater amount of training is necessary.

In the vision condition, although attempts were made to limit the amount of available light illumination by requiring the participants to wear a headlamp, the reflections from the floor did allow greater illumination of the passage. The optic array created by light was more extensive than that created by ultrasound providing information that was not available in the acoustic array. While this did not provide an ideal comparison, some similarities were noted among the vision condition and the ultrasound conditions in various aspects of this study.

### ***7.6.1 Time to pass through aperture***

As expected, the time to pass through the aperture in the vision condition was found to be significantly shorter than the other two conditions but a difference did not exist between the two ultrasound conditions. However, the variability measures showed increased precision of timing in the outward direction by those who had prior experience with the device whereas those with no experience showed greater repeatability in the forward direction. This finding suggests that there was a carry-over effect between the earlier study and the present study. As there was a month between the two tests and the tasks were quite different, effects of learning were not expected. Individuals who had participated in the earlier experiment indicated a preference for the orientation of the receivers in the outward direction. This orientation provides a lower intensity signal as the transmit signal and received signal are perpendicular to each other (Figure 7.12). These participants may have fine-tuned their ability to detect more subtle cues in the signal that enabled them to pass through the centre without a need to consciously orient themselves in their surroundings, or they may have simply had more confidence in a system that was familiar to them. The orientation of the receivers in the forward direction provided more information about what was immediately in front than at the sides. The intensity of the signal

was higher in this condition as the transmit signal and receiver directions were similar. More distinct cues defining the edge of the aperture were available as the signal would completely disappear while passing through the aperture. This seemed to be easier for novice participants to interpret.

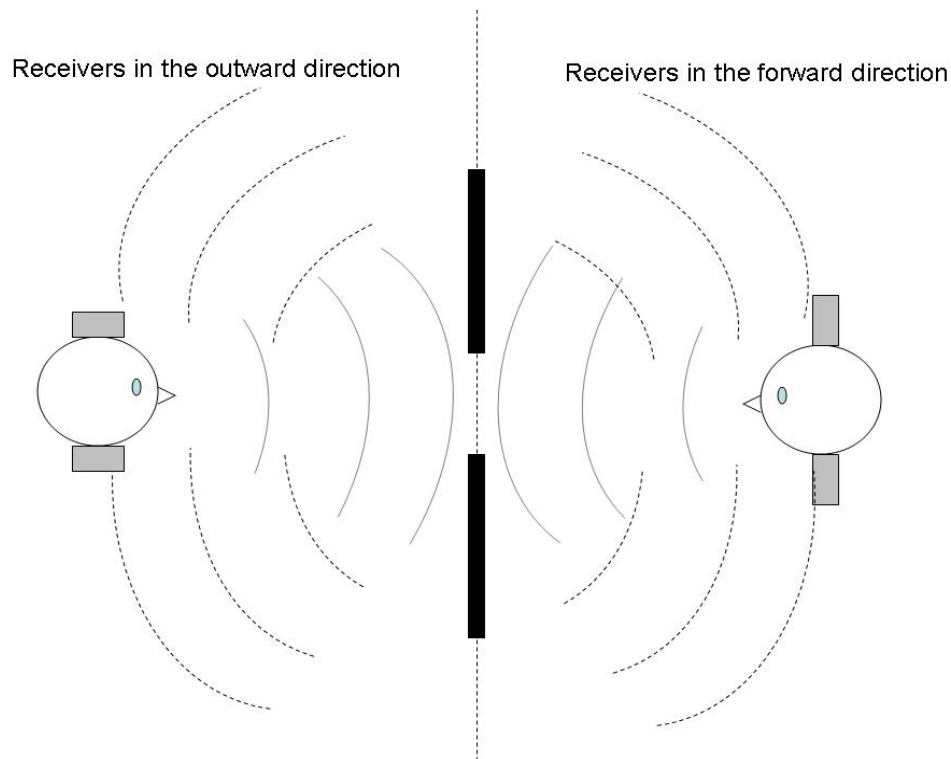


Figure 7.12 Direction of ultrasound signal upon reflection in each ultrasound condition. The solid lines represent the transmitted ultrasound and the dashed, the reflected.

There was no evidence that more time was required to pass through smaller apertures than larger ones. The hypothesis stated that the smaller apertures would require more time to negotiate, but this was not the case. The time to negotiate the gap was independent of the gap size. The acoustic information provided sufficient time to judge the gap, either through a perception of a



gap or perception of the walls and deduction of a gap. Audification in this sense is providing sufficient information to respond actively in response to a stimulus. The individual does not take additional time to evaluate the gap, rather he responds directly to the stimulus and passes through the aperture possibly without the need to create a full spatial map. There is a theory that an individual uses path integration through proprioceptive information to evaluate distances (Loomis et al., 2002). Loomis et al. have studied this concept in detail and believe that an internal representation is determined at the outset of the task and proprioception during self motion allows an individual to achieve this task. These studies are predicated on providing information to a participant and requiring them to complete a task with no additional information. Since the present study involves a perception action task with continual information from the acoustic array being presented, it is difficult to determine whether path integration is involved in this process. It is known however that if path integration is used in creating an internal representation, the additional feedback of acoustic stimuli on approach of the aperture enables a change in the internal representation to be undertaken. Thus, even if a map exists, stimulus response due to direct perceptive evaluation may enable continuous updating.

### ***7.6.2 Ability to pass through centreline***

There were no significant differences in the ability to pass through the centre of the aperture in any of the conditions. The feed forward control resulting from the auditory array was arguably effective in providing information about the location of the two sides of the aperture as the information from the visual array.

There was evidence that the individuals who had previously participated had greater accuracy when passing through the midline of the aperture than novice participants. Upon detection of the

aperture they were able to adjust their approach paths to enable passage along the midline.

Better judgement as a result of increased exposure to the device appears to have resulted in more accurate performance than that of the novice participants.

The second hypothesis expected the audification resulting from orientation of the receivers in the outward direction to be more accurate than that of the forward direction. Although not significant ( $F_{(2,825)} = 2.14$ ,  $p = 0.11$ ), the orientation of the receivers in the outward direction for the ultrasound condition did appear to enable greater accuracy in the passage of midline than the orientation in the forward direction. The field of view in this orientation was larger and allowed for better detection of the aperture sides. While traversing through the aperture, information about the sides was still received making it sound as though the sides were close. It is possible that the orientation of the receivers in this direction gave the perception of the aperture being small requiring increased accuracy through the midline.

Since the receivers in the forward direction collected information from a limited field of view, the binaural signals were very similar. This likely made judgement of distance from the sides of the aperture more difficult as the participant approached. To be able to detect both sides, one would require more head movement to scan the full environment than with the receivers in the outward orientation.

The mean value of the deviation from the midline showed a preference to the left hand side for both the vision condition and the forward receiver orientation. Arguably, similar information about the field of view is received for both conditions. With vision, it is important to view the

area beyond the obstacle or aperture, or as Gibson suggests (1979), allow the optic array to guide away from an obstacle. While approaching the aperture, the relative size of the obstacle increases, but once confident that clearance is inevitable, the individual looks beyond the aperture rather than fixating on the edges. This allows for forward planning of future movements but may reduce the accuracy through the centre of the aperture. In this sense, when the receivers are pointed forward, a similar phenomenon is achieved such that the sides of the aperture are no longer observed and the individual must “look beyond” the aperture. The perceived information is similar to that of vision and can be processed the same way.

Had the vision condition not been tested to allow for a “gold standard” comparison, the preference to the left of the centreline in the aperture would likely have been explained by an increased intensity level in one receiver relative to the other that was more evident in the forward orientation. Since a similar observation was made regarding the vision condition, the veering toward the left side while passing through may be a result of the testing being performed in New Zealand where individuals drive and pass each other on the left side. There was one European participant who showed deviations to the right, so this may be a viable hypothesis.

As hypothesis 3 predicted, the smaller gap sizes allowed more accurate centreline judgements than the larger gap sizes. The variability of the aperture size showed that increased precision was evident for the apertures 55 cm and 65 cm. Interestingly, there was evidence that even though there was less precision in the larger apertures, the accuracy through the midline improved such that the mean midline position at 95 cm was similar to that of the 55 cm gap.

This indicates a preference for midline such that the binaural information was minimised equally (in both ears) as passage through the aperture occurred.

### ***7.6.3 Angle of passage through the aperture***

The notable effect of shoulder rotation is that of condition by aperture. In the vision condition the body rotated significantly to pass through the aperture at 55 cm, but by only a few degrees through the other apertures. In the vision condition, individuals could effectively combine proprioceptive information about their visual array and shoulder location. This allowed for increased rotation of the body through the smaller apertures. The overall shoulder rotation through the apertures in the ultrasound condition with receivers in the outward direction was similar to that of vision (as expected in hypothesis 4). Thus, the outward condition allowed for detection of both sides of the aperture and individuals could proprioceptively orient themselves to ensure safe passage through the aperture. One participant commented after several trials “I know this is a narrow aperture, so I should turn my body” but overall the angle through the smallest aperture was not as great as in the vision condition. Since the shoulder rotation in the ultrasound condition of the outward direction is greater than that of the eye direction the individuals are likely detecting the side of the aperture and increasing the shoulder angle to reduce risk of hitting the side. In this way, they were more cautious with the receivers in the outward orientation than in the forward orientation.

There was evidence that the group who had previously participated did not rotate their shoulders as much as the novices to pass through the aperture. Perhaps they were more confident in their evaluation of aperture size and responded accordingly. With additional training, individuals may respond more closely to that of vision if they have detected a narrow aperture.

#### **7.6.4 General Discussion**

Static monocular vision has been found to be sufficient for the judgement of aperture passability (Warren and Whang, 1987). If it is assumed that minimal acoustic information is required for judgement of aperture passability, it is not surprising that there were few differences between the orientations of the receivers in the “ears” (outward) and “eyes” (forward) directions. The information provided by the receivers in the forward direction is limited to a very specific field of view which could be compared to the limited information available with monocular vision. Perhaps the ability to pass through apertures is a task that is not complex enough to effectively judge orientation of receivers for an ultrasound device for locomotion.

Following Gibson’s approach, optic flow surrounds the locomoting animal such that humans can perceive either outflow or inflow depending on the direction that they are going. Animals with laterally placed eyes, can experience inflow and outflow simultaneously. The experience of inflow and outflow led to a more accurate understanding of the location of the door while passage through was occurring, enabling more accurate centreline passage.

On the other hand, this task does provide insight into the possibility that audification to evaluate aperture passability may provide more information than sonification. Hughes (2001) used another sonar device that relays information through sonification (the KASPA) to judge the size of aperture such that participants commented on their ability to pass through. Although this task required perception and conscious report instead of a more ecological approach as this perception-action event, participants were unable to effectively perceive the “meaning for action” that resulted from the acoustic array provided by the device (Hughes, 2001). The pattern

of sonified reflections from this ultrasound device was insufficient to allow for effective judgement of passage size. Instead, Hughes notes that subtle changes in frequency at the edge of the aperture were difficult for the novice user to interpret as it was not defined by an abrupt change, rather a subtle frequency change. The audification provided by this device enabled very precise determination of the edge of the aperture allowing these participants to pass through the aperture along the centre line safely and efficiently.

## **7.7 Conclusions**

This experiment involved perception-action events of audified echoes that resulted in successful passage through the centre with shoulder rotation nearing that of vision in apertures larger than 1.3 times the shoulder width. Individuals using the receivers oriented outwards were able to pass through the aperture closer to the centreline, in the same amount of time as the forward direction and simulate the shoulder rotation of vision for the apertures from 65 cm to 95 cm. The experienced users were more effective than the novice ones. This indicates that perhaps the receiver orientation in the outward direction may be better than that of the forward direction, but additional research is required with respect to the orientation of receivers. Both the distance from the centre and the amount of shoulder rotation were similar to the vision condition indicating that signals from the acoustic array may provide sufficient information to proprioceptively orient oneself in the spatial environment. Perhaps with additional practice, the speed of the passage through the aperture may more closely reach that of vision.

## **8 Comparison of audible signals to downconverted ultrasound signals for localisation**

Auditory localisation by humans is a technique that is learned from a very young age allowing individuals to determine where a sound is coming from and respond accordingly. For example, if an infant's mother calls his name, the child turns toward the voice. The main purpose of this experiment was to determine the effectiveness with which localisation could be achieved using direct auditory signals and audification of ultrasound signals. At distances greater than 1 m, the duplex theory suggests that sound pressure level is the main method of localising. The first part of this experiment involved using the receivers to collect the ultrasound information from various speakers around the room and analysing the sound pressure levels at each position. Following quantification of sound pressure levels at the various positions, human participants performed a localising task using the same ultrasound receiver orientations as before and compared to the auditory sense with no receivers.

### **8.1 Test Environment**

To reduce the effect of external environmental sounds and uncontrolled reflections that might influence the results, a 5-m x 5-m anechoic chamber was used. The supporting floor of the chamber was an acoustically transparent mesh below which the structural floor was covered with absorbing foam wedges. Ten loud speakers were mounted at two heights (0.6 m and 1.2 m) in five azimuth locations on a circle of 1.9 m radius round the participant 35° apart (Figure 8.1). The two sets of loud speakers were positioned such that one was 15 cm higher than the base of the participants' chair and the other at 60 cm above the first. Since this localisation device is intended for detection of obstacles above waist height these two positions can give an indication of the relative localisation ability in both azimuth and elevation with and without the device.

The two heights were selected to be on the same cone of confusion to allow evaluation of the ability to resolve cone of confusion ambiguities. Each loud speaker was paired with an ultrasound transmitter and an LED (light emitting diode) indicator (Figure 8.2). A central control box selected which loud speaker or transmitter to radiate the test signal played from a digital recorder (IPOD with LINUX installed) at 96 kHz. The signal will be discussed in more detail in the following section.

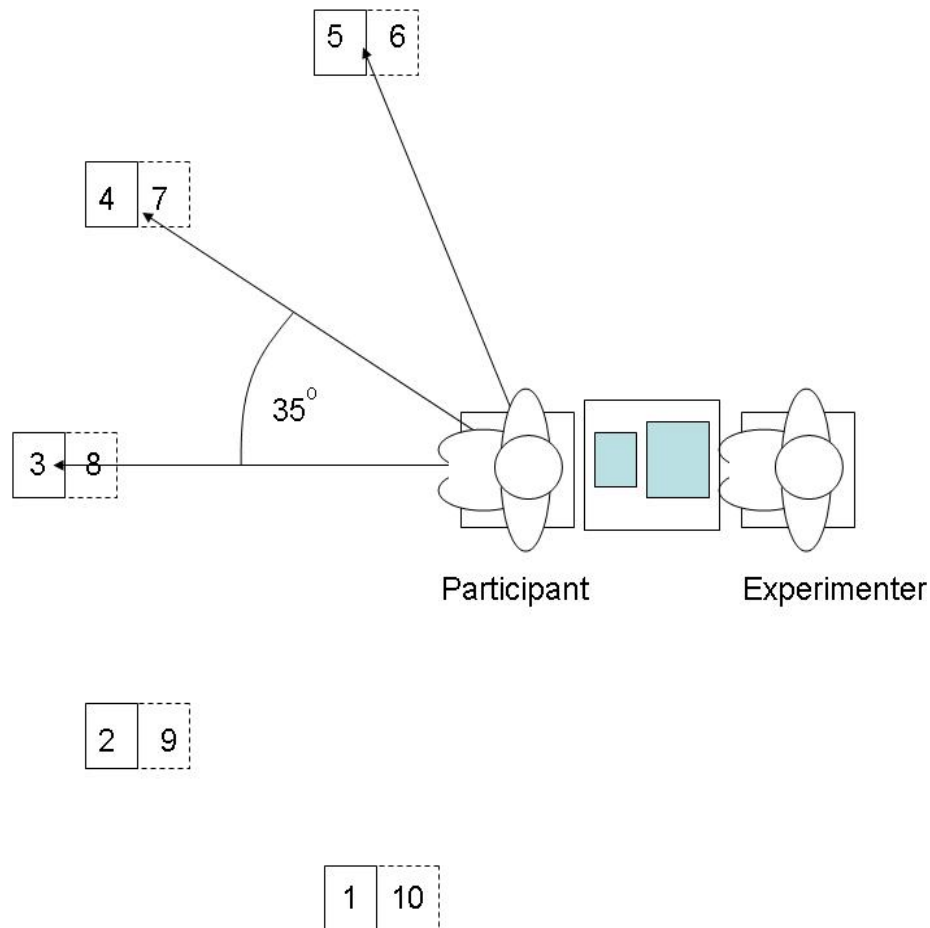


Figure 8.1 Experimental apparatus for localisation experiment performed within an anechoic chamber.



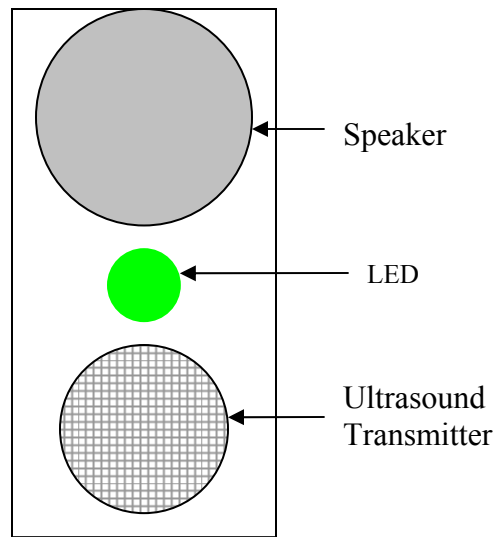


Figure 8.2 On each speaker mount, a speaker, and LED and an ultrasound transmitter.

For the first part of the study that involved quantifying the sounds from the speakers in the two orientations, a tripod was positioned in the same position as the participant at a height of 1 m. The receivers were placed on the tripod in two orientations (forward and outward direction) with a distance of 18.5 cm between the two.

## 8.2 Localisation signals

A broadband signal was transmitted by each of the ten speakers. The stimulus for the ultrasound condition was a five second burst of broadband noise in the frequency range of 26 kHz to 40 kHz which, when sampled at 40 kHz would downconvert to pink noise in the range of 0 Hz to 14 kHz. The stimulus for the auditory condition was a five second burst of pink noise with cut-off frequencies of 0 Hz and 14 kHz (white noise that is band limited to being within the range of hearing). Since individuals were required to move their heads to determine the location of the

sound, the sound had to be long enough to allow for head motion (Thurlow et al., 1967).

Although research involving head movement to localise has been limited, most studies in the field use short signals (less than 800 ms to ensure no head movement). Thurlow, Runge and Mangels (1967) allowed 5 seconds of transmit signal to allow for effective localisation with head movement. Since participants in this study required individuals to localise with a prototype device that likely provided interference with head motion, 5 seconds was deemed appropriate to ensure adequate “gaze” fixation in all conditions.

### **8.3 PART A: Characterising speaker levels to judge direction**

It was important to ensure that the receivers were able to provide sufficient information to determine localisation based on sound pressure levels from transmitters distributed around the room.

#### **8.3.1 Purpose**

This evaluation study quantified the difference in sound pressure level among ten different speakers to ensure that humans would be able to localise using the ultrasound device in different orientations. As discussed, at distances greater than 1 m, sound pressure level is used for localisation. According to the duplex theory, interaural intensity differences (or sound pressure levels) are important for localisation of sounds greater than 1 m away. Here the sound pressure level was quantified at each of ten locations to ensure that there were sufficient intensity differences from each receiver to enable sound source localisation of sounds transmitted at ultrasound.

### **8.3.2 Method**

The sound was played at each position from the ultrasound transmitter, and recorded from the device after downconversion had occurred. Receiver positions were evaluated as in the previous experiments:

- a) with the receivers oriented face forward, 17.5 cm apart (the distance between the ears), and;
- b) with the receivers oriented outward, 17.5 cm apart.

### **8.3.3 Analysis**

The average sound pressure level at each position and in each orientation was determined from a digital recording (Figure 8.3). The 0 dB reference level is arbitrary, as it could be adjusted through the intensity of the transmitted signal, or the amplification of the received signal. In this analysis, 0 dB corresponds to the maximum sound pressure level attainable at 1 m separation between transmitter and receiver.

### **8.3.4 Discussion**

Starting with the receivers facing forward, when the ultrasound signal is transmitted from positions 1 and 10 or 5 and 6, the intensity of the signal is very low in both receivers. When the signal is immediately in front (positions 3 and 8), maximum intensity levels are achieved by both the left and the right receivers. Maximum intensity of both receivers is obtained when the “look direction” is the same direction as the speaker signal. On the other hand, the intensity drops off at the sides suggesting that peripheral information is very limited.

When the receivers are oriented in the outward orientation, the intensity of the sound was largest in the left receiver for positions 1 and 10 and loudest in the right receiver for positions 5 and 6.

As the sound from one receiver falls in intensity, the intensity of the receiver on the other side grows, such that a truly binaural signal is evident. At the centre positions (3 and 8), the sound pressure level was more than 40 dB below the sound pressure level at the outer positions. In this case the drop-off suggests that there is a null in intensity in the “look direction” due to the receiver gain pattern, but peripheral information is readily available.

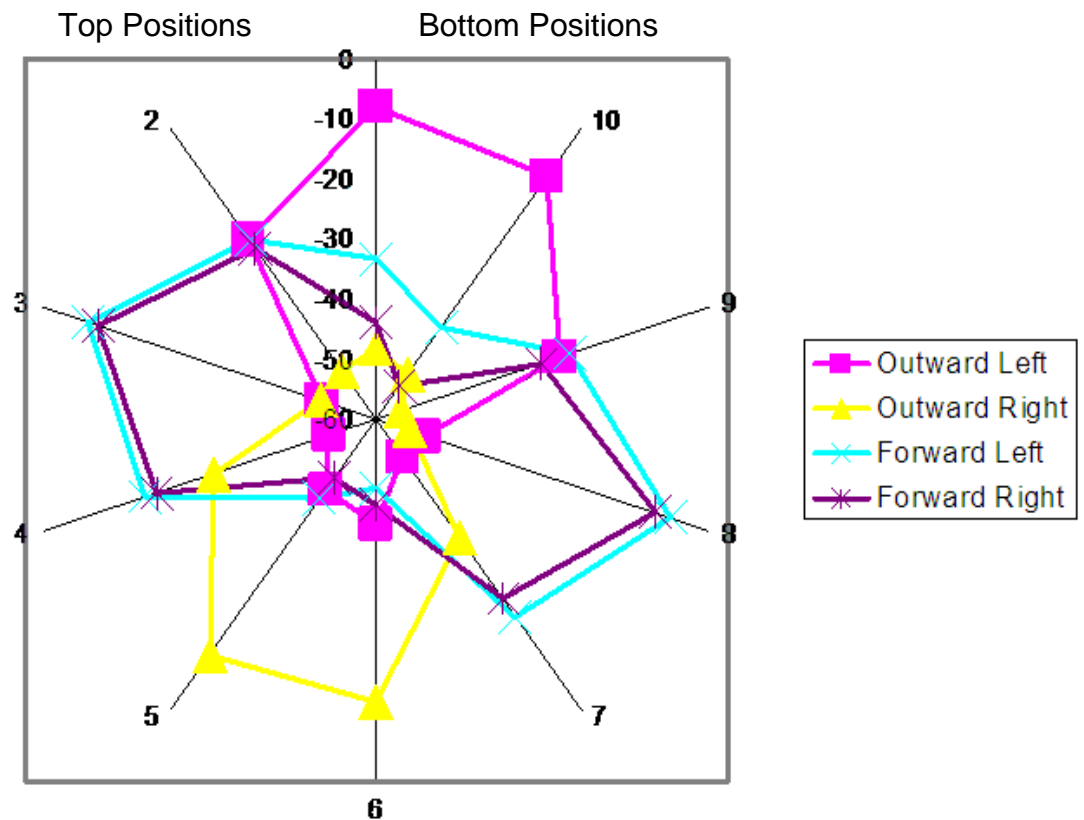


Figure 8.3 Sound pressure levels at each of the ten positions.

There are two reasons for the drop off that occurs at the periphery for the forward orientation and at the centre for the outward orientation. First of all, the angle of incidence is 70 degrees when the signal comes from these positions rather than 0 degrees when the signal is received directly.

Secondly, the receivers themselves have a limited gain pattern that restricts reception at this angle (Figure 8.4). At  $70^\circ$  and  $290^\circ$  there are nulls in the gain pattern of the receivers making it more difficult to receive signals at these great angles.

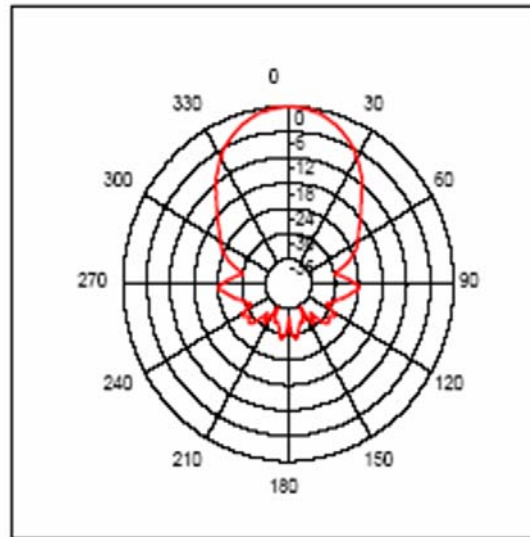


Figure 8.4 Gain pattern for ultrasonic receiver and transmitter (Air ultrasonic ceramic transducer 400 ST/R 160 from Farnell-in-One datasheet 65326)

Similar results were observed for both the bottom and the top transmitters. Based on intensity levels alone, it was not possible to judge the two different elevation levels of the different speakers.

### 8.3.5 Conclusion

The orientation of the receivers facing forward allowed for determination of those transmitters that were located immediately in front and within  $35^\circ$ . The orientation of the transmitters in the outward direction enabled peripheral information to be collected, but limited information about the look direction was available. It was not possible to differentiate between the top and the bottom receivers based on intensity alone.

## **8.4 PART B: Human testing of localisation**

### **8.4.1 Purpose**

This study identified similarities and differences in individuals with no training in echolocation of localisation performance of audible signals relative to audified signals.

### **8.4.2 Hypotheses**

H1: The localisation accuracy between direct audible and audified ultrasound signals will be similar.

P1: Since the sound will be a direct sound as opposed to echoed sounds, the only differences between the two signals will be the type of receiver used to collect the signals. The ear will collect the audible signals, but the ultrasound receiver will collect and audify the ultrasound signals. In the event that no differences are evident, this would support the argument that audification of ultrasound can provide similar information to that of hearing. Any differences that do result between ultrasound and auditory signals will be due to the sensor differences in collecting the signal.

H2: Receiver positions for localisation will be most effective when oriented outward.

P2: The contribution of this work versus other systems is that the user will be able to directly localise from audible information provided with minimal signal processing. To localise as effectively as one would using auditory signals, the placement of the receivers would have to be placed in the same positions as the ears. Since auditory echolocation was being compared directly to audified localisation by the same users, the results should be comparable.

### 8.4.3 Method

Fifteen individuals with no known deficiencies of hearing volunteered to participate (Figure 8.5). Four of these participants had been involved in one of the previous studies (all male) and three had been involved in both (1 male). Each participant read and signed an informed consent approved by the University of Waterloo Human Ethics Committee and reviewed by the University of Auckland Human Ethics Committee (Appendix 3). As this test involved precise localisation in all three conditions, it was important to characterise the hearing in both ears. An audiogram was conducted with each individual prior to the start of the test to determine hearing ability. Each ear was tested with an audiometer at pure tone frequencies of 4000 Hz, 2000 Hz, 1000 Hz, 500 Hz, 250 Hz, and 125 Hz. Each tone was played starting at 30 dB HL (hearing

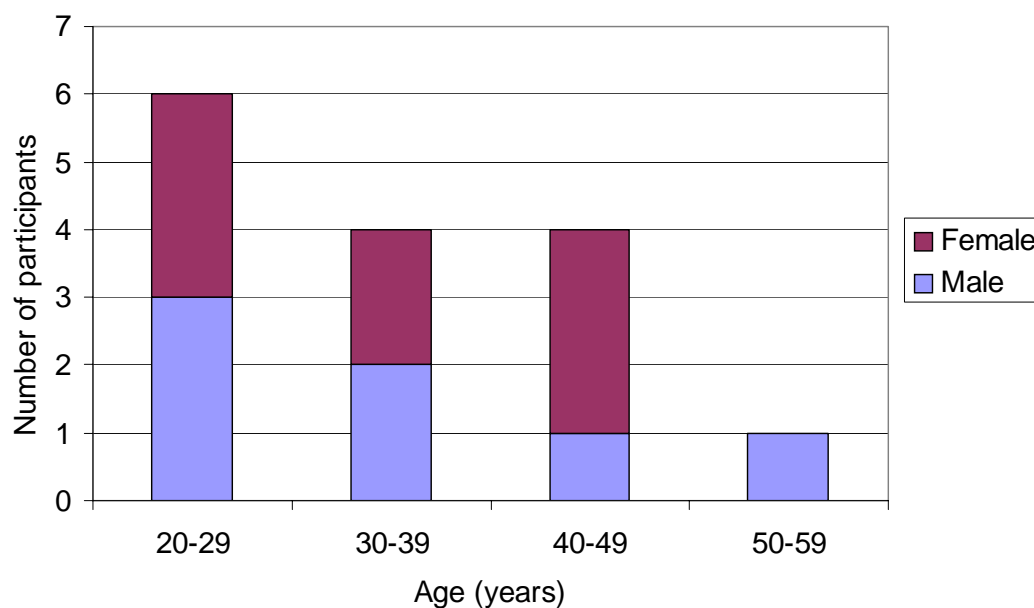


Figure 8.5 Demographics of participants

level) and decreased by 5 dB HL until the threshold was found. A plot of frequency versus hearing level for each ear was developed and reviewed. A collapsed audiogram of all participants is shown in Figure 8.6. All values are within the normal range of hearing which is 10 dB to 25 dB (1996).

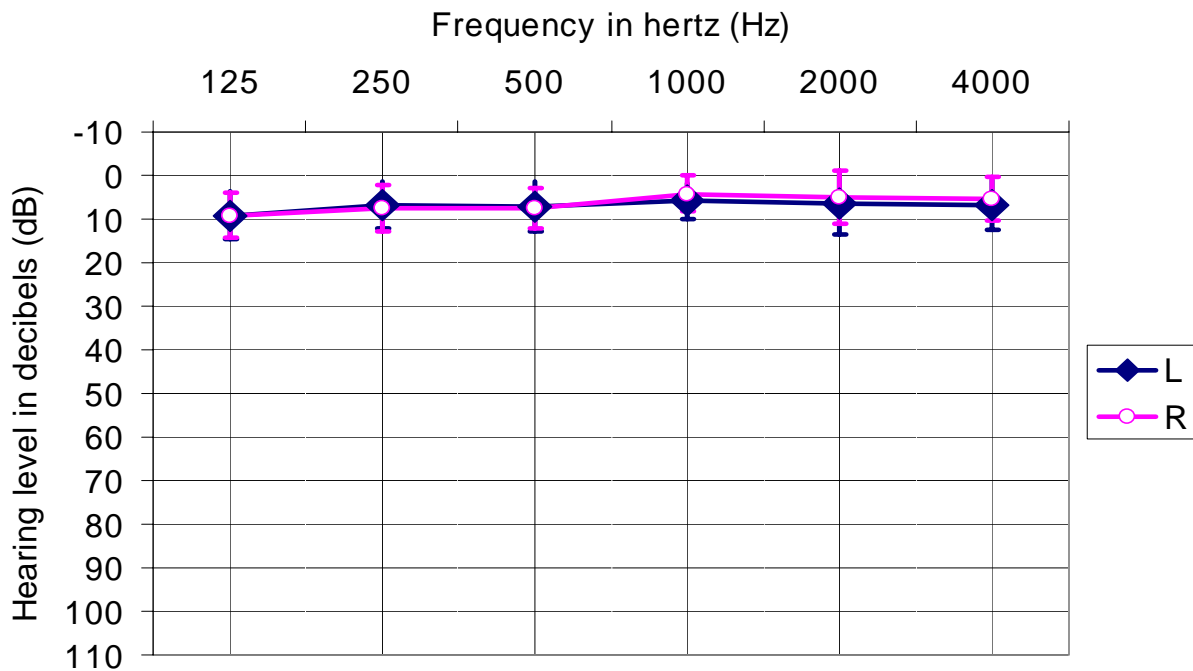


Figure 8.6 Audiogram of the mean values of hearing level for all participants.

The participant was then taken to the anechoic chamber where the overhead lights were turned off to ensure that the individual did not see the test apparatus. Two low-level floor lights were used to guide the participant to the participant chair. The individual was blindfolded and a head-mounted camera was secured on the forehead. A head-pointing technique rather than an arm-pointing technique for direction determination was used. Head pointing has been found to yield better performance than arm pointing (Makous & Middlebrooks, 1990). The camera was used to collect position information based on the look direction by locating the LED mounted with each



speaker and transmitter would be bright in the darkened image of the look direction. The LED was 5 mm in diameter which resulted in a cross three pixels by three pixels in the snapshot image. From pictures taken when the participant was looking toward the sound, the deviation from the look direction relative to the actual location of the LED could be determined. Three conditions were examined: an auditory condition without the device, an audified ultrasound condition with the receivers in the forward direction, and an audified ultrasound condition with the receivers facing outward. For the auditory condition, only the camera was mounted on the participant. For the ultrasound conditions, the receivers were mounted on a stereo headset which was placed on the individual and checked to ensure that the sounds could be heard in each ear. The order of each condition was counterbalanced among individuals to reduce the effect of crossover information.

For each of the audified ultrasound conditions, participants were permitted three practice trials (of five seconds each). These three trials allowed the individual to hear the audified signal at the extreme left, the centre and the extreme right. After these three practice trials, the testing began. The system was designed to allow for two output devices to be plugged into the processor to hear the signal. The experimenter sat behind the participant and listened to all the audified sound signals at the same time as the individual. This allowed the experimenter to make general qualitative observations about the various techniques used in processing the information.

When the subject was ready, the experimenter played the stimulus and the subject oriented the head in the direction of the sound. The experimenter took a snapshot image of the look

direction. There were a total of 5 measures taken at each of the ten locations for each condition, thus 150 trials. The procedure took approximately 1.5 hours to complete for each subject.

#### **8.4.4 Analysis**

Each image was read into Matlab and the position of the LED in each frame was calculated. Since a head-mounted camera was used to determine look direction, calibration was required for each subject. In each condition, the position information from all the trials was averaged and the mean in both the azimuth direction and the elevation direction was used as the “look direction”, also referred to here as the “centre” of the image. This step allowed for normalisation of the camera positions based on the individual and the actual orientation of the receivers (relative to different head sizes, heights, orientation of the camera). Analysis of the data was performed in units of pixels rather than converting to an angle as these units allowed for comparisons of very small differences. Each pixel represented a distance of 0.19 degrees.

The distance from this calibrated centre in both the azimuth and elevation directions were computed for each trial and used in a repeated measures ANOVA examining a Group  $\times$  Trial order  $\times$  Condition  $\times$  Position ( $2 \times 5 \times 3 \times 10$ ) for each of the relative distances to the centre of the image in the two directions. Group 1 consisted of eight novice users of the device and group 2 were seven individuals who had participated in at least one of the earlier experiments. In addition, the precision for each subject was calculated and an ANOVA conducted on the Group  $\times$  Condition  $\times$  Position ( $2 \times 3 \times 10$ ) for the variability in each of the azimuth and elevation directions. A Tukey post-hoc test was performed on any differences identified by the ANOVA as having a p-value of less than 0.05.

### 8.4.5 Results

#### *Azimuthal Error*

There was a three way interaction effect of Group  $\times$  Condition  $\times$  Position ( $F_{(18,1640)} = 2.26$ ,  $p = 0.002$ , Figure 8.7) of azimuthal error. Also in azimuth, there was a statistically significant Group  $\times$  Position effect of distance ( $F_{(9,1640)} = 2.36$ ,  $p = 0.01$ , Figure 8.8) and an interaction effect of Condition  $\times$  Position ( $F_{(18,1640)} = 3.05$ ,  $p = 0.001$ , Figure 8.9). Finally, there were main effects of both group ( $F_{(1,1640)} = 30.48$ ,  $p = 0.0001$ , Table 8.1) and position ( $F_{(9,1640)} = 53.63$ ,  $p = 0.0001$ , Figure 8.10).

Statistical significance was only found in the main effect of condition for the variability ( $F_{(2,343)} = 22.16$ ,  $p = 0.0001$ , Table 9.2) such that the vision condition showed considerably less variability ( $p=0.0001$ ) or increased precision relative to the other two conditions. All other factors tested were non-significant.

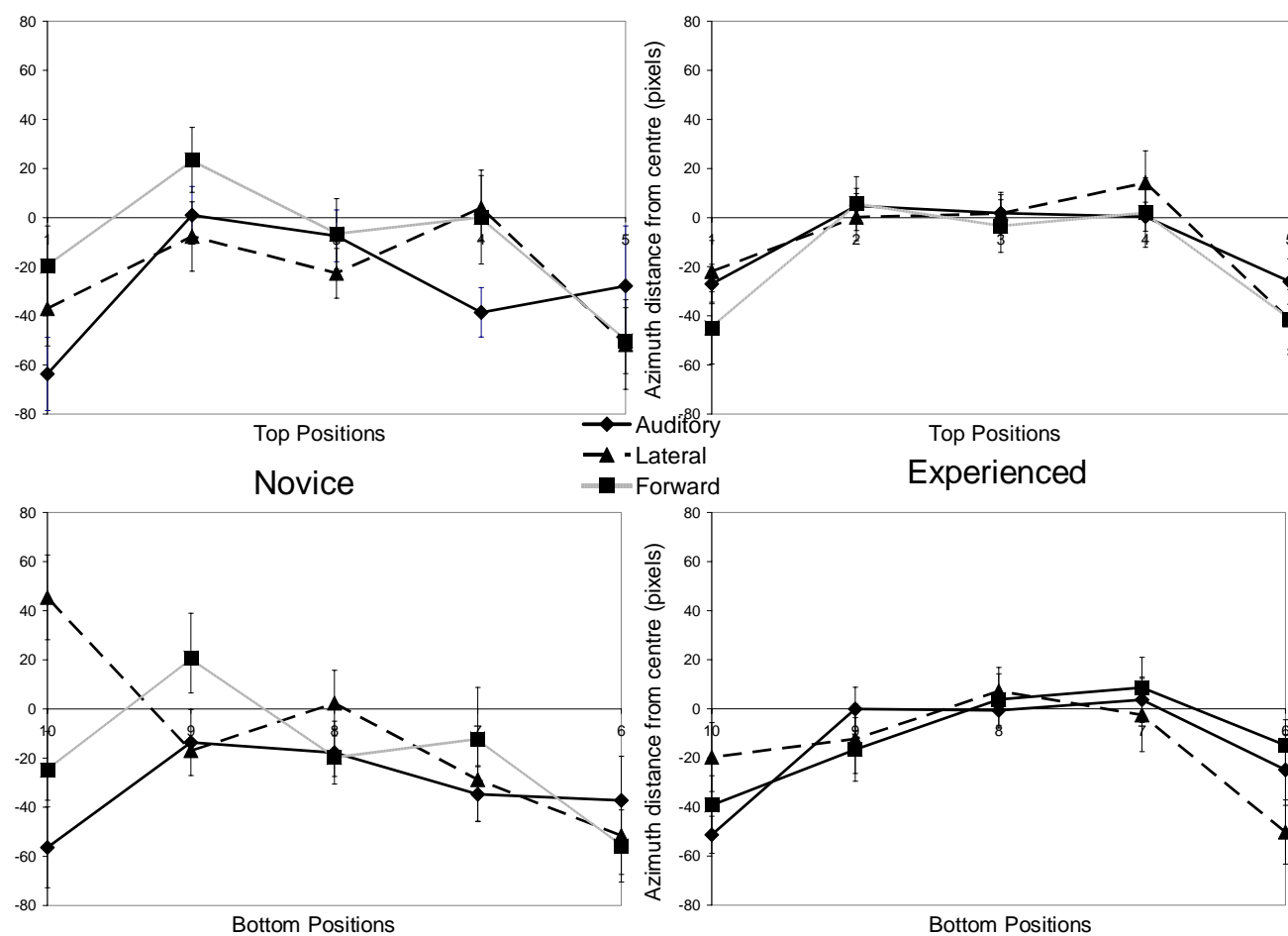


Figure 8.7 This figure represents the three way interaction of Group  $\times$  Condition  $\times$  Position of the distance from the calibrated centre of the image in the direction of azimuth. Error bars represent standard error of the mean.

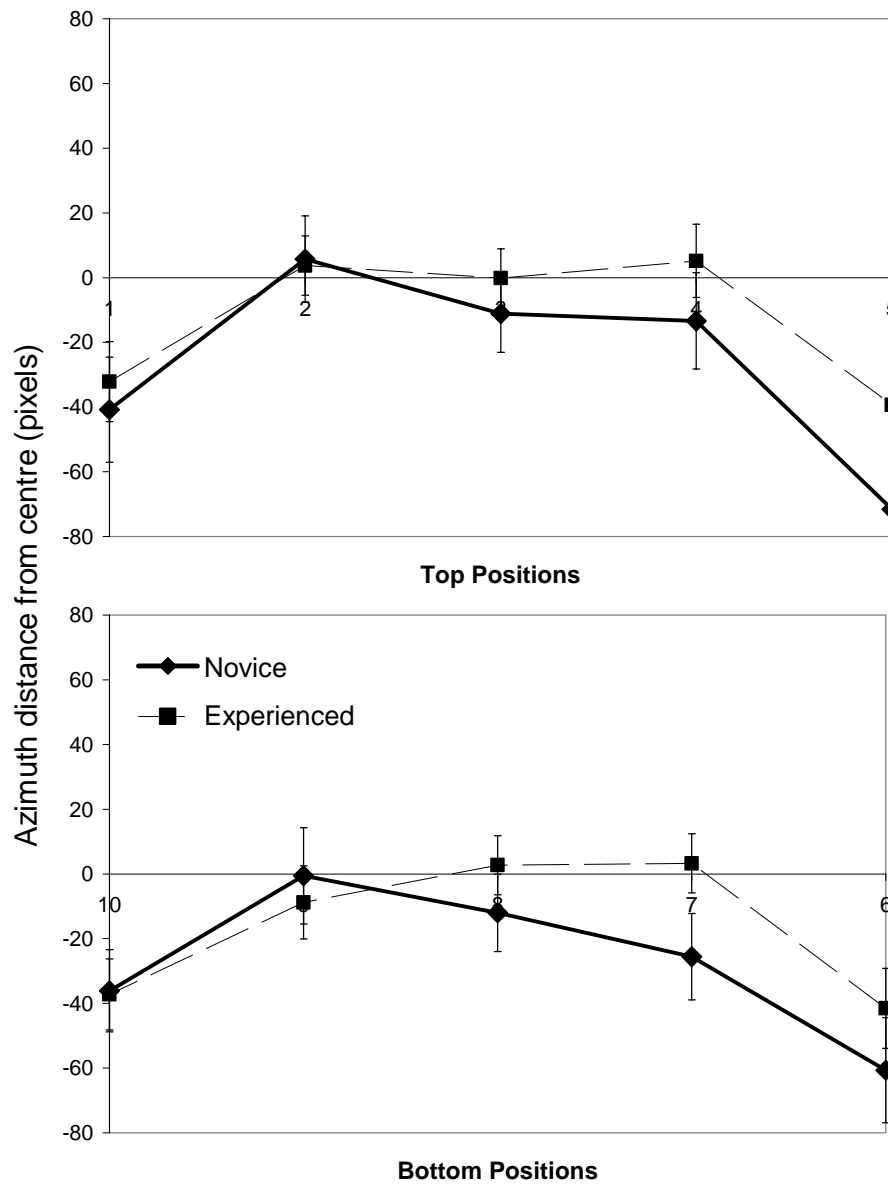


Figure 8.8 Interaction effect of group by position.

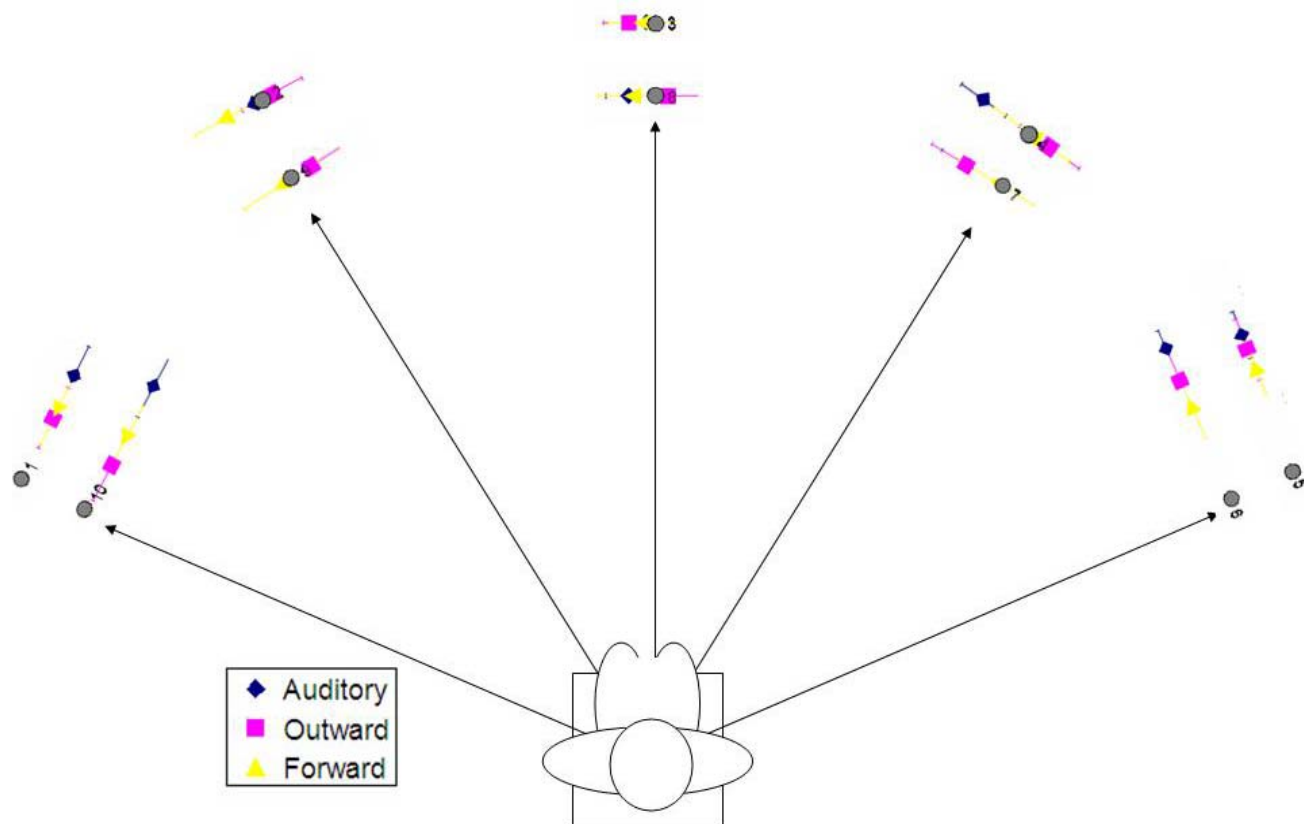


Figure 8.9 Effect of condition by position of the look direction. The grey circles are numbered in the speaker locations. The coloured symbols represent the mean look direction for all participants in each condition with standard error bars.

GROUP	MEAN DIFFERENCE IN AZIMUTH DIRECTION	STANDARD ERROR
Novice (Group 1)	11.01	16.46
Experienced (Group 2)	-1.56	12.34

Table 8.1 Mean difference in pixels from the calibrated centre to the LED in the look direction in azimuth.

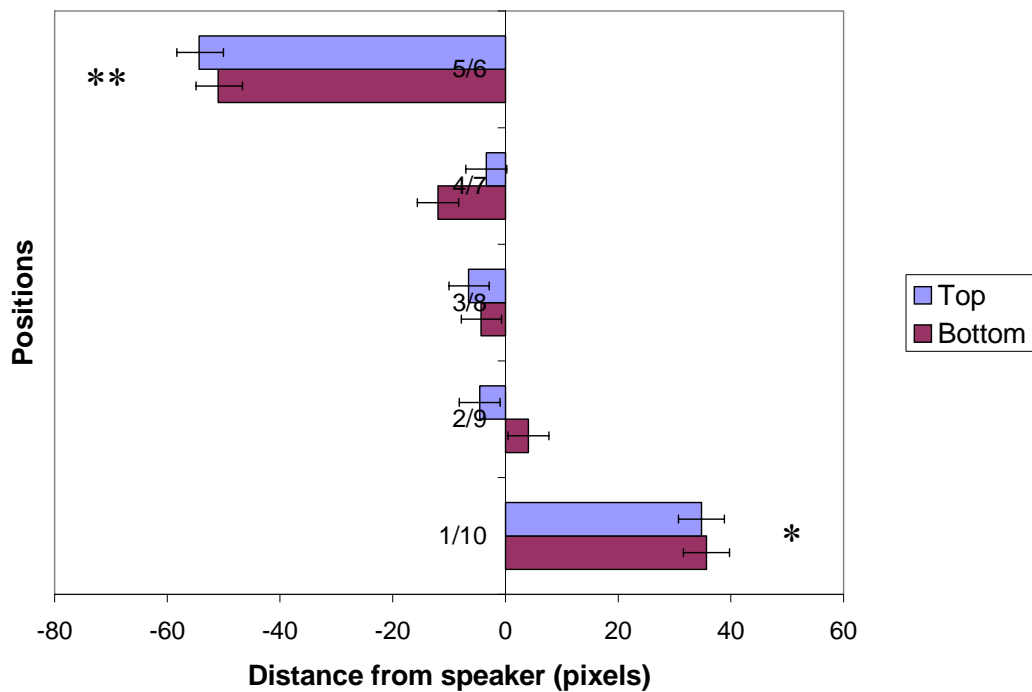


Figure 8.10 This graph shows the main effect of position. Positions 1 and 10 were significantly different from all others ( $p=0.0001$ ) as were positions 5 and 6 ( $p=0.0001$ ).

CONDITION	MEAN VARIABILITY	STANDARD ERROR
Audition	25.55	1.76
Outward	41.76	1.87
Forward	37.78	1.81

Table 8.2 Decreased variability in the audition condition relative to the ultrasound conditions when detecting azimuth.

### *Elevation error*

With respect to elevation as defined by the distance from the calibrated centre of the image to the y position of the LED, there was a significant interaction effect of Condition  $\times$  Position ( $F_{(18,1640)} = 1.89$ ,  $p = 0.01$ , Figure 8.11). The distance from the centre of the image was significantly different between the top and bottom positions. The estimated distance at position 6 in the auditory condition was significantly more accurate than the other conditions. There was also an interaction of Group  $\times$  Condition ( $F_{(2,1640)} = 2.99$ ,  $p = 0.05$ , Figure 8.12) such that group 2 exhibited higher estimates of position for both the forward and outward oriented positions than the group 1 estimates for the forward orientation, but neither of these were significantly different from the auditory condition. There were also significant main effects of group ( $F_{(1,1640)} = 14.03$ ,  $p = 0.0002$ , Table 3) and position ( $F_{(9,1640)} = 389.1$ ,  $p = 0.0001$ , Figure 8.13). The ANOVA model of variability in the elevation differences in distance showed no significant differences ( $F_{(59,343)} = 1.06$ ,  $p = 0.375$ ).



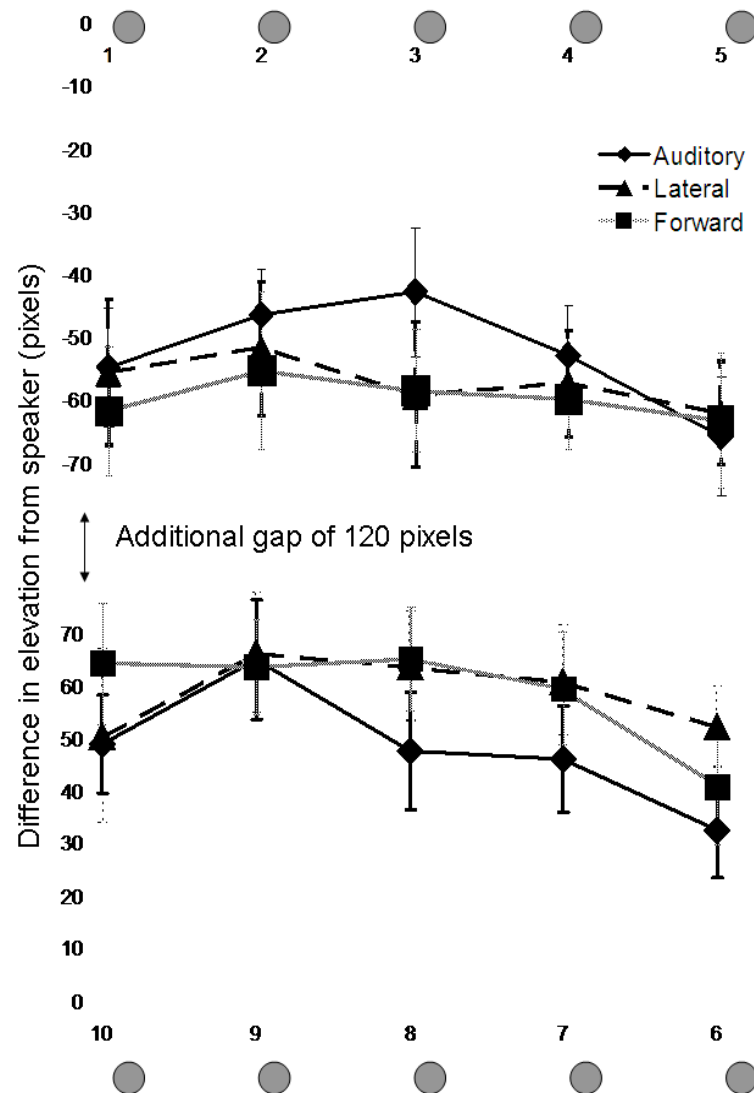


Figure 8.11 This figure represents the two way interaction effect of position and condition when the distance from the calibrated centre of the image was determined. Error bars are standard error of the mean.

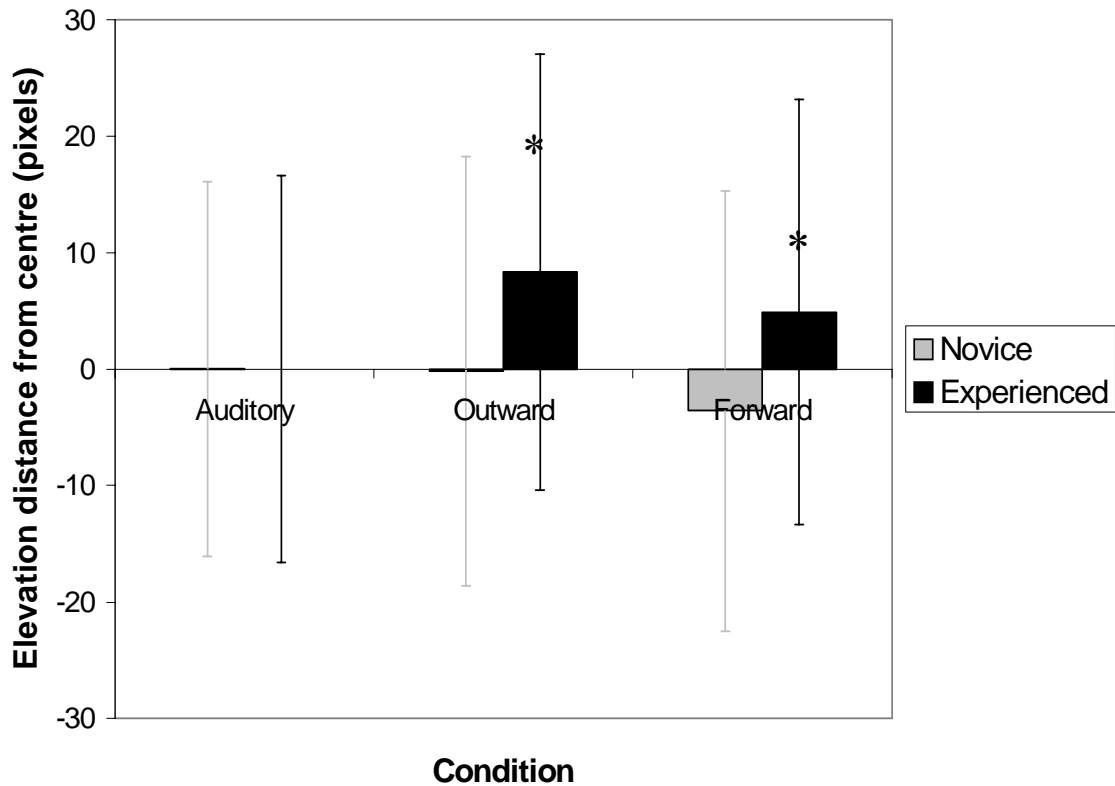


Figure 8.12 Effect of Group by Condition. The mean distance from the centre is represented by the bar accompanied by lines of standard error of the mean. The \* represents significantly different distances from group 1 in the forward orientation.

GROUP	MEAN DIFFERENCE IN ELEVATION DIRECTION	STANDARD ERROR
Novice (Group 1)	1.30	17.76
Experienced (Group 2)	4.19	17.82

Table 8.3 Mean difference in pixels from the calibrated centre to the LED in the look direction in azimuth.

Although not quantified, there appeared to be significantly more environmental scanning with the audification of ultrasound than with the auditory condition. In the auditory condition, individuals appeared to turn and fixate within the first couple seconds of the signal, whereas in the audified conditions they appeared to scan for the duration of the sound source signal and

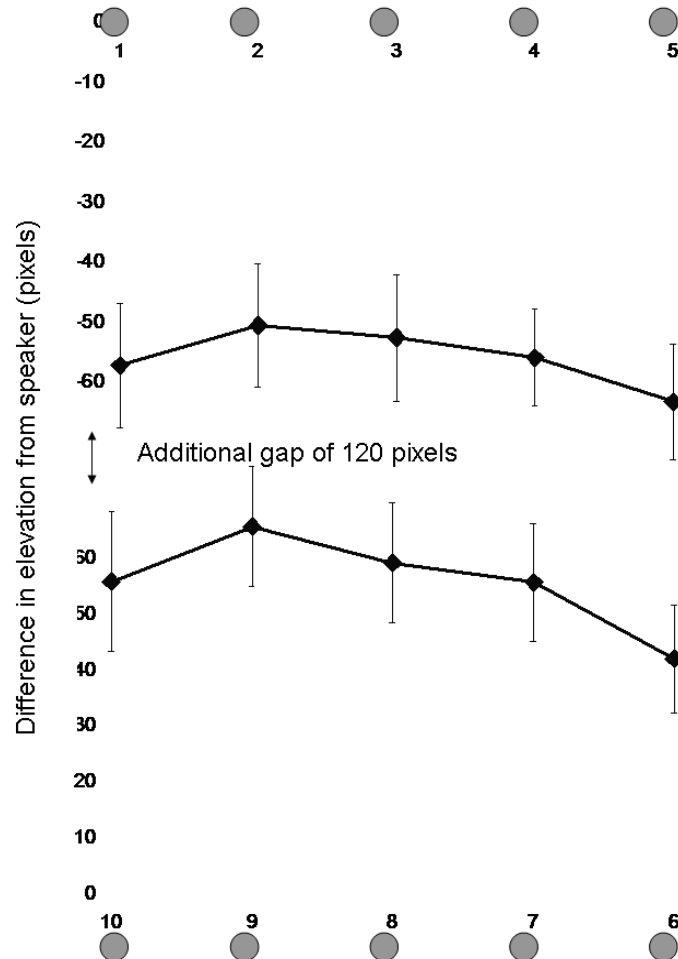


Figure 8.13 Main effect of position on the distance from the calibrated centre of the image to the LED in the elevation of the look direction.

settled on a look direction as the signal stopped. Different techniques were used for scanning dependent on individual and condition. In the condition with the receivers facing outward, some individuals passed the sound from ear to ear several times, whereas others would slowly rotate the head until the sound was balanced. Many commented that this orientation was not intuitive and they had difficulty tracking the sound (it would take a couple seconds to determine whether the sound was coming from the left or the right and the experimenter noted that they would quickly scan the full 140 degree azimuth while searching for the sound source). For the most part, participants appeared to rotate the head (left to right) rather than tilt (up and down) while conducting the testing with the receivers oriented in the outward direction. In the condition with the receivers facing forward, there was also rotational scanning, but to enable accurate elevation localisation (to get the maximum intensity), many realised that tilting the head up and down allowed for increased intensity.

It did appear that individuals developed a technique during the first half of the experiment in each condition and used that technique for the rest of the testing in that condition. This observed improvement in ability to localise was not evident in the statistical analysis of trial order ( $F_{(4,1640)} = 0.90, p = 0.46$ ). Although these different techniques were used, at the end of the testing when the participants were permitted to view the experimental setup many commented that they were not aware of two different elevations.

#### **8.4.6 Discussion**

##### *Localisation in the direction of azimuth*

The interaction of Group  $\times$  Condition  $\times$  Position showed that there was a considerable learning effect for those who had used the system before (Figure 8.7). In all conditions (Figure 8.8), the

participants in group 2 (the experienced group) exhibited more consistent results at the positions immediately in front (speaker positions 3 and 8) and 35° to each of the left and right (speaker positions 2,4,7, and 9). This type of learning was unexpected as this task was very different from the previous tasks. The novice participants had more difficulty in estimating the position of the sound in the direction of azimuth, but the precision (as judged by the variability for each subject) was not significantly different between the groups for either azimuth or elevation. Previous studies of localisation have provided many hours of training with feedback performing localisation tasks to achieve consistent results (Makous & Middlebrooks, 1990; Oldfield & Parker, 1984). Oldfield and Parker (1984) gave at least two hours of training very specific to their localisation task. Makous and Middlebrooks (1990) provided at least ten training sessions with visual feedback of an LED to the sound source location. The present experiment provided fifteen seconds of training. Those who had participated in previous experiments had a maximum of 2 hours practice in conditions that were not similar to these. Arguably, a more ecological approach to localisation testing was used for this study than for other studies, but also showed that previous use of the device did enable more accurate localisation.

For those sound sources that came from the peripheral directions (positions 1, 10, 5 and 6), the accuracy in the azimuth direction was better with the ultrasound audification than with auditory signals (Figure 8.9). In order to fixate on these sound sources, the head was required to rotate seventy degrees. Sometimes it is easier to rotate the torso as well as the head to pinpoint a source that far from the midline. Although possible in these and ideal conditions, individuals did not appear to attempt to rotate fully in the auditory condition. This is not an unusual observation. The study by Thurlow et al. (1967) showed that participants did not turn sufficiently to face the

sound source during the first rotation of the head. By more than 50 percent of their participants, there was scanning back and forth around the sound source location several times. Since the head motion during the auditory condition was not monitored as closely as the audified condition in this experiment (as an additional headset was not being used to eavesdrop on the received signal), it is difficult to comment on the technique used in the auditory condition. While performing the experiment with audification, there appeared to be additional scanning after rotation toward the sound source which required the individual to move back and forth past the source before determining a look direction. By first rotating past the source, a more accurate measure of the location may have been determined. Although the accuracy was improved in the audified conditions, it was not as good at the peripheral positions as it was in the central ones (Figure 8.10). Makous and Middlebrooks (1990) also found that accuracy diminished at more peripheral locations even after considerable training sessions with experimenter feedback.

#### *Localisation in the direction of elevation*

Individuals commented that they were not aware of two elevations of speakers. This was also apparent in the results (Figure 8.9 and 8.10). Participants had a tendency to underestimate the elevation of the speakers that were mounted on the top level and overestimate the elevation of the speakers located on the bottom level. They did not however, fixate on the midpoint between the two levels. For those that were on the top level, the head was tilted slightly upward and for those on the bottom level, the head was tilted down.

In the auditory condition, individuals would overestimate or underestimate the elevation such that the mean results showed no difference (Figure 8.12) between the groups. The individual

was just as likely to underestimate the position of the speaker as she was to overestimate its position. In the audified condition with the receivers pointing outward and the receivers forward facing, participants in the experienced group were more likely to estimate the source of the sound as coming from higher than the actual position. For the novice group, there was a tendency to slightly underestimate the sound source elevation with the receivers facing forward, but neared zero with the receivers facing outward. Prior experience with locomotion and avoidance of obstacles may have encouraged the individuals in group 2 to be more comfortable with scanning up and down and settling on a position slightly higher out of an abundance of caution. Important to note when characterising this finding is that all results were within 10 pixels of the mean of elevation. This corresponds to 1.9 degrees of accuracy with respect to elevation which in an ecological environment would likely not change the course of action.

The peripheral locations (Figure 8.13) were judged as being lower than the actual position for the speakers located at the top level and higher than the actual position for the speakers located on the bottom level than the other positions in front of the participant. These positions also seemed to take longer to localise as individuals often did not have sufficient time to complete their scans and remarked that they were not sure of the location. Once the head is rotated, the heterosensory component of motional theories may affect peripheral judgements. Feedback control guiding head orientation also requires input from the vestibular sense as well as position, tension, and posture sensors from the back of the head (Blauert, 1983). If these are strained to detect peripheral sound source locations, proprioceptive information about head orientation may not be as accurate as for those locations to which a participant is normally responding.

In the case of the forward-facing receivers, the expected movements to determine the sound source would be to rotate the head to the sound source location during which the sound source would increase to its maximum value in the azimuth direction (as with the auditory sense, Blauert, 1987), then to tilt the head up and down to maximize the intensity for the elevation determination. With the receivers facing outward, determination of azimuth would require the rotation of the head such that a minimum intensity existed between the two receivers. To be able to determine elevation in this condition, pivoting of the head from side to side had to occur. This is an unnatural movement that is rarely used for localisation purposes but individuals in this test, especially novices, appeared to gain slightly better results in elevation using this technique than the tilting of the head.

The first hypothesis stated that the localisation accuracy between auditory signals and audified ultrasound signals would be similar. In the direction of azimuth, those who had prior experience with the device showed similar accuracy in all three conditions for the positions that were immediately in front of the participant and 35° to each side. Those with no experience also showed comparable results, but the mean estimates were not as accurate as the experienced group (refer back to Figure 8.7). There was evidence that the peripheral locations were not as effectively localised in any of the three conditions. The measures of precision showed that the auditory condition was significantly better than that of the ultrasound conditions. While localisation accuracy among all three conditions was similar in the plane of azimuth, the precision was lacking in the ultrasound audification.



The results with respect to elevation showed that only at position 6 (the lowest, most peripheral speaker location) there was better accuracy in the auditory condition, otherwise the accuracy among all top positions and among all bottom positions were similar across condition.

Localisation in elevation did not differ among conditions in the forward facing orientations.

Based on these observations of both azimuth and elevation the first hypothesis was found to be true for localisation ability immediately in front of the participants and with a field of view of 30 degrees on either side. Increased accuracy across conditions was found for “experienced individuals” but the precision was not as good in the audified ultrasound signals as compared to the auditory condition.

The second hypothesis predicted that the receivers oriented in the outward direction would be more accurate than those in the forward direction. There was no conclusive evidence that one orientation was more accurate than the other. A spontaneous movement of the head toward the position of the audible sound enables increased precision as the source is brought into the region of sharpest hearing (Blauert, 1983). The area of sharpest hearing is that oriented in front of the head such that the two ears have similar input. This is also the look direction that was examined in this experiment. The hypothesis suggested there would be better performance of the receivers oriented outwards which was presumed to be the same direction of the ears. The prediction was based on an assumption that the ears face outward to collect sound. The shape of the pinna allows for collection of sound at the periphery as well as in front. In audible localisation, the head shift to look toward the sound source allows for finer resolution of the sound. When the ultrasound receivers were oriented outward, the participant had to determine the position of the

sound based on a null rather than increased intensity of the sound. In the forward direction of the receivers, an increased intensity was evident allowing the individual to more precisely locate the sound. The orientation of the receivers in the forward direction appears to represent a more similar response to that of localisation in the auditory domain than that of the receivers pointing outwards.

Before completely discounting this hypothesis in the design of systems for individuals who are legally blind, a realisation must be made that there is a difference in spatialisation ability of these individuals compared to those with vision. Sighted individuals use vision to calibrate their hearing. Individuals who have never had vision develop a sophisticated understanding of their relative spatial environment through proprioceptive information (Lewald, 2002) and a better understanding of personal audiomotor control. This results in localisation performance equal to that of sighted individuals in the frontal domain and better peripheral hearing than the sighted (Roder et al., 1999). Those individuals who have functional vision (such as light perception) have actually been found to have decreased accuracy in localisation than totally blind or sighted individuals (Lessard et al., 1998). These individuals have neither been able to calibrate sounds through vision or developed a finer tuning of their audiomotor control. Although the audification information received by the receivers pointing in the outward direction seemed more difficult for blindfolded, sighted participants, perhaps a blind population would show a preference for this orientation. Further work must be done with respect to orientation of receivers before accepting or discounting this hypothesis.

## **8.5 Conclusions**

Localisation ability was tested using auditory signals and audified ultrasound signals. No difference was evident in the azimuth or elevation accuracy of these signals. Previous practice with the device allowed for increased accuracy even though the training was not specific to this task. Although precision was better for the auditory condition, it is possible that with practice, precision will be increased with audified signals as well. This research supports the notion that audification of ultrasound can be effective at localising point source sounds.



## 9 General Discussion

The present day secondary mobility devices developed for individuals with blindness are difficult to use and require considerable training (Blasch et al., 1987). They lack easy to interpret, decipherable interfaces. This work sought to improve the current interface designs by evaluating the needs requirements in such a design. Development of an audified design followed. A system was constructed to provide audified ultrasound information. Testing of the audified system involved evaluating human participants' ability to determine distances, pass through apertures, and localise point sources. Did this research develop an auditory interface that could provide sufficient information for safe and efficient travel for those who lack visual input? To answer this question, the questions introduced at the beginning of this work must be revisited:

- 1) What critical information must be displayed to a visually impaired traveller to provide sufficient information for obstacle avoidance while attempting to minimise cognitive load?
- 2) How can auditory information be displayed to allow for direct perception of an environment?
- 3) Can audification in the form of intentionally aliased ultrasound provide adequate information for detection and localisation of environmental obstacles?
- 4) Does audification allow for direct perception of environmental obstacles without training?
- 5) How is audification influenced by the manner in which the information is retrieved from the environment?

This chapter discusses the five research questions in detail. It is followed by a discussion of the limitations of this work and provides some possible directions for future research.

## **9.1 Answering the research questions**

This part is presented in five sections that address each of above research questions and is followed by a discussion of limitations of this work and future design directions.

### **9.1.1 Information requirements**

*What critical information must be displayed to a visually impaired traveller to provide sufficient information for obstacle avoidance while attempting to minimise cognitive load?*

A design framework was applied to the development of an interface for auditory display which was based on cognitive work analysis. This required an understanding of the specific tasks involved in performing the task and applied ecological interface design to determine informational requirements for a system that must enable safe and efficient travel by individuals with loss of vision. The hierarchical approach involved the definition of two conflicting functional purposes and determining the abstract functions leading to generalised function, physical function, and physical form. This led to a set of informational requirements that were necessary for the design of an auditory interface to allow individuals to access all information that could enable effective locomotion in an obstacle-rich environment. The elements of context that would provide the most information at a perceptive level of behaviour were visited and the decision was made to develop an interface that uses audification.

In the design of a system that provides spatial information to an individual without sight, it is important to display all information in the field of view. What constitutes this field of view? Looking to Gibson's definition, it is that solid angle of ambient light that can be registered by the ocular system (1979). If instead, a similar field of view is applied to an acoustic array, it would include those solid angles of sound which can be registered by the auditory system. Sound, like

light is omni-directional, but is limited in the power of illumination. Occlusion of sound, as with the nose in vision, results from the head and shoulders. The central area of acuity in front of the head in a visual system is clear in the centre and more variable toward the periphery. The same is true of an auditory system such that sounds immediately in front are more easily discernable as compared to those at the edges. The ear itself also provides occlusion to certain sounds, specifically those behind. The intended field of view in this experiment included the field in front of the individual in which the individual was most likely to encounter obstacles. With the receivers oriented outward, the field of view was extended to include areas at the side of the head.

To allow auditory information to be collected from echoes of the surroundings, a continuous transmission ultrasound system that illuminated all obstacles in the direct field of view of the ultrasound transmitter was used. Reflections of the ultrasound from environmental obstacles were collected by receivers. These reflections were audified using intentional aliasing by a direct conversion receiver and relayed binaurally through headphones to the traveller. As the traveller was moving toward a stationary obstacle, the rate of approach could be judged by Doppler shift. Increased intensity of the sound resulted from decreased distance or increased obstacle size, the same way it would with auditory echolocation. From the aperture test, it was evident that the edges of doors could readily be determined. From the localisation test, it was found that scanning up and down with the device permitted characteristic properties of the obstacle such as height to be determined. All three studies contributed to understanding the effectiveness of the device to meet those informational requirements identified by the work domain analysis.

This design process incorporated the SRK taxonomy in an attempt to minimise cognitive load. Although audification is believed to minimize cognitive load (Sanderson, 2005) relative to other auditory elements of context, this was not specifically tested. Future testing could involve the use of n-back testing (Klatzky et al., 2008) which involves vibrotactile stimulation while assessing action controlled by auditory display. When compared to other systems, this would give a better indication of whether cognitive load is minimised using audification.

### **9.1.2 Auditory displays for direct perception**

*How can auditory information be displayed to allow for direct perception of an environment?*

The auditory displays of past devices have included sonification (for example, KASPA, vOICe) requiring knowledge-based behaviour to interpret the interface and thus a significant amount of training, and earcons (for example, Sonic Pathfinder) which are rule-based in nature but provide limited information about what is located immediately in front of the individual. To be able to better display this information, the behavioural nature should require less knowledge and instead be more intuitive. Since sonification is a knowledge-based task, this project attempted to exploit those that were more skill- or rule-based in nature. A more skill-based approach was applied such that audification could provide localisation information. Since ultrasound is conducive to audification due to its physical nature, a device was developed to allow audification of ultrasound to be presented.

Audification is a skill-based element of context that capitalises on the physical properties of ultrasound to provide perceptual auditory information. Aliasing the Doppler shift of ultrasound signals allowed the reflected ultrasound to be converted to the auditory domain. Since the sound



was directly converted from ultrasound, there was no additional interference by the experimenter in “determining the best sound”, instead it was provided directly. This allows perceptual processes to occur naturally, without additional signal processing. Since this interface could not be applied to the current systems, a system had to be developed to enable audification.

Direct perception results from an active response to environmental stimulus such that the individual perceives the relationship between himself and the environment and using this perceptive information performs action (Gibson,1979). Attempts have been made to compare the acoustic array provided by echolocation to that of optic flow (Lee 1997, Rosenblum 2001) which results in direct perception. Hughes (2001) argues that the stimulus for active echolocation consists of discrete events, clicking or snapping, and thus is not comparable to a continuous event like optic flow. This novel system uses audification of Doppler signals to create a true acoustic flow that changes based on motion of the participant, from the participant’s perspective, similar to that of the optic flow. If there is movement either by the participant or by another person in the room, the Doppler signals can be observed. Since a true acoustic flow is presented to the individual, an argument for direct perception can be made.

In his book, Gibson refers to three sets of experiments from which direct perception can be deduced. These three include experiments relating to surface layout, those relating to changing surface layout and those that involve self movement. This thesis has sought to investigate all three of these components suggested by Gibson as being indicative of direct perception.

First, direct perception of surface layout was studied by the first experiment of distance perception. In vision, optical contact within the surface of support is necessary to evaluate surfaciness. Since the device requires reflections from obstacles, the sound source must also create acoustical contact with the surface before being perceived. Gibson argues that perception of an edge is an ability that develops as affordance perception. A similar result can be argued with the experiment of door passage in that the edge must be perceived and a null in the sound offers affordance of passage. These indicators are suggestive of direct perception in vision and parallels can be drawn to sound based on these experiments.

Secondly, changing surface layout is discussed by Gibson. Invariants exist in vision allowing time to collision to be determined perceptually. These same invariants appear to exist in sound such that time to collision enables individuals to effectively pass through an aperture or approach a wall (after giving the distance judgement). The aperture experiment provided a good first step in evaluating this ability to perceive the changing surface layout upon approach and effectively respond to it.

Finally, Gibson suggests that co-perception of own motion is necessary for direct perception. The centre of optical expansion is not a sensory cue, but an optical invariant. He uses the example that expansion of a body towards a person initially illicit a blinking response, but after several trials, individuals no longer blink but are still perceiving the changing light cues. This thesis examined the centre of acoustic expansion (or an invariant of increased intensity) throughout both the tests of distance perception and aperture avoidance. As an individual approached the wall, the intensity of the reflections increased. A further study could involve

impact assessment as a sound moves towards the individual who must determine whether it is on an impact path. Further studies would show a better understanding of the human co-perception of own motion and perceived environmental stimulus in the form of sound.

Although a claim for direct perception based on auditory stimulus would be premature at this stage of the testing, these tests cover some of the basic assumptions by Gibson as required to achieve direct perception. Further testing would be required to better substantiate these claims in the same manner that Gibson undertook a series of tests to provide a claim of direct perception of a changing optic array.

An ecological approach task involving apertures of different widths was used to evaluate the ability to directly perceive and avoid obstacles in the environment. Individuals were able to successfully turn and walk through an aperture with centreline passage and rotation angles similar to those of vision. Although the time to proceed through the aperture was not matched to that of vision, there was evidence that individuals with practice were better able to evaluate the environment. There appeared to be two methods of approach. The first method specifically involved determining the edges of the aperture by approaching both sides independently and relying on memory information from the previous scan to guide them through the aperture. The other method was possibly indicative of direct perception in that the participants could detect and respond without the requirement for defining the edges of the aperture. Acoustic flow was initiated when the individual started to move, steering toward the centre of the outflow by minimising the signal to each ear. Successful passage was signalled (as with the optic array) with a gain in the structure within the closed contour of the gap (Gibson, 1979). It appeared that

most participants undertaking this task were able to perceive and avoid apertures suggesting that there was evidence of co-perception of self-motion relative to the acoustic array.

### **9.1.3 Audification to avoid environmental obstacles**

*Can audification in the form of intentionally aliased ultrasound provide adequate information for detection and localisation of environmental obstacles?*

Environmental obstacles can be detected upon approach with the use of an audified system. Two tasks were undertaken to evaluate the ability of individuals to judge distance of obstacles and proceed to avoid obstacles with audification. The first task compared the ability of individuals to judge distance based on audification of ultrasound as compared to auditory distance judgement with echolocation. Echolocation usually requires a pulse-echo event such that both the direct and the indirect signal are heard and can be used for distance judgement subconsciously (Stoffregen and Pittenger, 1995). When individuals are tasked with providing distance information, the task no longer becomes subconscious as cognitive action is required to report. Although this is not an ideal task to evaluate distance judgement, it did serve to provide comparative judgements between information from auditory echolocation and those of audified signals. The audification provided significantly better results than those of auditory reflections especially at distances greater than 0.5 m. This confirms that the audification of ultrasound reflections is better at distance determination than that of auditory echolocation.

The detection of an environmental obstacle in this condition was only determined with acoustic flow. Without motion, no environmental information was available. This acoustic information flowed outwards while the individual was travelling toward the obstacle. When required to make a judgement of distance, individuals had to cognitively put a number on their perceived

responses. Although they likely perceived the distance from the obstacle, it was difficult to provide a specific number. To enable increased observation, the participants found they could obtain additional information by rocking. Both outflow and inflow occurred allowing them to more accurately perceive the exact location relative to themselves.

The second study involved a perception-action task that required individuals to detect two sides of a gap and pass through the middle of it. This task relied on more subconscious judgement of aperture size as the individual perceived the environmental information through the acoustic array and responded accordingly. This task was compared to a visual setting in which the individuals wore a head-mounted light to perform the same task. Although the time through the aperture was better with vision, similar responses to vision were observed for shoulder rotation angle through the aperture and distance from the midline of the aperture.

Gibson argues that perception and visual kinesthesia are coupled allowing for tasks like this to be an act of attention rather than a learned response. The results of this study agree in general that individuals directly responded to environmental obstacles by angling their bodies (with outward facing receivers) and passing through the centre of the aperture to avoid possible collision. On the other hand, the first trial took longer than the others to perform. Perhaps feedback from the brain provides instructions to the body. This may suggest a form of spatial updating or path integration (Loomis et al., 2002) such that the stimulus is encoded through direct perception, but a spatial image is updated during movement which elicits a response. The visual system contributes to proprioceptive information through learning. Infants and primates spend considerable time looking at their hands and the actions that can be achieved when moving them

(Gibson,1979). Individuals who participated in this study were preconditioned through years of practice to perceive an aperture with vision and rotate through it. They had not experienced a similar conditioning through sound. It is possible that individuals with visual impairments may have already been preconditioned to sound as sighted individuals have to vision. With participants who are visually impaired, the aperture size may be directly perceived and the individual will proprioceptively respond similarly to those with vision. There was evidence that after the first trial, the individuals directly perceived the aperture and continued through successfully at a speed that they felt was comfortable.

Based on the results from these two experiments, the conclusion can be drawn that audification can be used effectively for judgement of distance to an obstacle location as well as providing avoidance information to individuals passing through a gap between obstacles.

#### **9.1.4 Audification virtually eliminates the need to train participants**

*Does audification allow for direct perception of environmental obstacles without training?*

Neuhoff suggests that considerable training does not provide an ecological environment for testing, rather one that can only be applied to novel conditions (2004). One purpose of this work was to show that audification provides sufficient information for localisation and obstacle avoidance without training. The maximum amount of training given to each individual at the start of the test procedure was five minutes per task per orientation. No learning effects were evident on a trial by trial basis except in the case of aperture passage. In the aperture passage experiment, there was a significant difference in the amount of time that an individual required to

evaluate the environment and pass through for the first trial at each aperture, after which the passage time showed no discernable differences.

Direct perception of motion involves a response to optical expansion. Gibson (1979) evaluated optical expansion using a screen with a small silhouette that increased in size. He found that the participant would blink in response to the stimulus, but this would eventually cease. Although the response changed, the perception did not change. The perception was not dependent on a learned conditioning, but rather the information from the optic array. In the case of all three experiments presented within this thesis, the only effect of trial order was evident for the first trial passing through each aperture. Thus human behavioural response was unchanged over the course of the experiment. These individuals did not learn how to use the device within the time of the test procedure, yet could accurately perceive distance, avoid apertures and localise within their acoustic environments. This supports the notion that direct perception may have been achieved without significant learning.

That said, previous use of the device did allow for more precision in following tests. Although learning was not evident within each experiment, seven participants were involved in more than one of the experimental procedures. There were significant differences between the two groups (novices and experienced) in the ability to process the information. There was a considerable carryover effect for those individuals who had participated in at least one previous experiment. This was unexpected as the experiments were completely different in task requirements and performed at least a month apart. For the aperture study, there was evidence that the experienced group took longer to pass through the aperture but passed closer to the centreline. This group did

not angle their bodies through the aperture as much as the novice, perhaps more confident in their judgement of the aperture size. The localisation task also showed better performance by the experienced group with respect to an ability to localise effectively using the audified information to precisely identify the azimuth direction. Individuals likely developed an ability to hear information and subconsciously process it after limited experience with audification. This suggests that on a task specific basis, learning may not be evident, but when integrated with other tasks trained responses might be carried over allowing future tasks to be easier than the initial task. While not expected, the age old saying “practice makes perfect” holds true when performing tasks with new devices.

Although precision was improved by the experienced group with respect to the tasks performed, there was evidence that novice users could effectively use audification to pass through apertures with similar accuracy. Accuracy of point-source localisation by novices was also found to be similar among all three conditions. Thus, indicators of improvement for specific tasks do not show a learning effect, but increased exposure to different tasks does appear to increase the effectiveness of individuals (based on increased precision) in their ability to use the device.

#### **9.1.5 Achieving effective directional audification**

*How is audification influenced by the manner in which the information is retrieved from the environment?*

Gibson has argued that any information required for obstacle avoidance is present in the optic array and distance is not judged through “depth perception” per se, but through invariant environmental features. In support of this claim, it has been shown that monocular vision allows



for effective passability of obstacles (Warren, Jr. & Whang, 1987). In all three experiments for this work, two orientations of the receivers were used to judge the location of obstacles by direct ultrasound audification from an acoustic array. In two of these studies, there was a comparison with auditory sound perception and in one of the studies a comparison to vision was performed.

In judging distance using audified ultrasound in both the outward and forward directions, the distance estimates were significantly better than auditory echolocation of blindfolded sighted individuals. A significant main effect of receiver orientation was observed such that the outward direction was better than that of the forward direction. For this task, individuals also expressed a preference for the outward facing receivers and there was less variability in these measures. This observation would support the concept that sufficient information for depth perception is present in an acoustic array rather than requiring differential binaural input through each ear. The receivers placed in the forward direction would have enabled differential input between the two receivers, but this did not enable better depth perception. Further research into the ability to evaluate obstacles closer and farther away is required to evaluate this claim.

For the task that required obstacle avoidance through an aperture, the results of the two orientations of the receivers showed similar results with respect to time through the passage. For the group that had previously used the device, there was slightly better accuracy with the receivers facing outward, and the opposite was observed for the novice users. Overall, the distance to midline was closer for the condition of the receivers facing outward than those facing forward as more acoustic information was available during the actual passage through the

aperture. On the other hand, the distance from midline using the receivers in the forward direction more closely simulated that of vision.

In the final localisation experiment, similar azimuth results were obtained for the frontal positions for the two receiver orientations. Elevation appeared slightly more accurate when using the outward facing receivers, but only for the novice group.

Direct and reflected ultrasound signals were collected, audified, and provided binaurally to participants of these studies such that two different orientations of receivers could be evaluated for obstacle avoidance tasks. The sounds provided by each orientation were characteristically different (collecting reflections from different orientations), yet both provided sufficient information for individuals to judge distance, avoid apertures, and localise point sources. Based on the observations for effective travel using this device, the orientation of the receivers facing laterally appeared to offer the most information about distance and aperture size. Localisation of point sources is not so critical during locomotion tasks and this was the only experiment that did not support orientation of receivers in this direction. The ears are complex structures with an ability to hear sounds in front, peripherally, and behind. An orientation of the receivers or a structure that better simulates the ear may provide additional information not characterised in this study. The research here can form a good basis for future work in determining the best receiver orientation for audified information.

### **9.1.6 Conclusion: Can audified ultrasound be used for safe and efficient travel?**

Audified ultrasound has never been studied as a means to provide information to individuals about their environment. Audification in general can only be applied to those signals that have similar physical characteristics of audible sound and in the past has had very limited application. As ultrasound is a high-frequency signal, intentional aliasing can bring the ultrasound signal directly into the auditory domain. Audification as an element of context to auditory interfacing requires skill-based behavioural processing and is thus intuitive to interpret.

This thesis discusses the planning, development, and testing of an auditory interface that allows for detection of obstacles, passage through an aperture or between two obstacles, and localisation of point source sounds using an audified system. The audified interface has been shown to exhibit characteristic human responses similar to both the natural auditory sense and the sense of vision. In conclusion, audification provides enough information to allow for safe and efficient travel by individuals with minimal training who are somehow limited in their sense of sight.

## **9.2 Limitations and future research**

This section discusses the realised limitations of this research and provides ideas for future work. There are many directions that this research could continue along, but this audified ultrasound system provides a unique evaluation tool for examining perception-action behavioural responses in acoustic environments.

These studies have examined the abilities of blindfolded sighted individuals to perform common obstacle avoidance tasks and localisation tasks. Although individuals with vision have similar

abilities to localise, they have not developed the same spatial abilities and calibration techniques as those who do not or who have never had the sense of vision. The results from these studies can be extrapolated to the general population but may not provide similar results that would be observed by individuals with either functional or total blindness. Future work will involve discussion and testing of audification and the developed system with individuals who are visually impaired.

This device is limited to those individuals who have functional or total blindness, but have effective hearing. Approximately 11 per 100 000 individuals within Canada are deaf-blind which is defined as having "a condition, that combines any degree of hearing loss with any degree of vision loss that interferes with communicating and acquiring information; even though Deaf-Blind persons may still have varying levels of useful vision and hearing" p. 16 (Watters et al., 2004). An individual who is not able to effectively localise sound would likely not benefit from this device.

The orientations of receivers were limited to two directions for this study. These two orientations were supposed to be representative of the same direction as the eyes and the same direction as the ears. Although both seemed to work fairly effectively in the judgement of distance and the ability to avoid apertures, it became readily apparent during the testing of localisation that in fact the ears are not only oriented outwards. A baby turns toward its mother's voice to be able to hear it more clearly. Looking toward a sound not only allows an individual to use vision to more precisely locate it, but also brings that sound into the area which can most easily be interpreted, that which allows similar information in both ears, or the area immediately

in front of the individual. In the localisation experiment with the receivers “oriented in the direction of the ears”, the individuals were searching for a null in the sound. This appears counterintuitive as compared to the actual auditory strategies developed for listening with ears. The natural hearing technique more closely simulates the condition in which the receivers were oriented “in the direction of the eyes”. Although audification was effective for the tasks presented, perhaps more efficiency could be obtained by evaluating other orientations of the receivers.

The tasks chosen for this experiment were representative tasks of those that require “observation” of obstacles, a perception-action response to obstacles and the ability to pinpoint the source of a directly transmitted sound. Although good starting points for evaluation of an individual’s ability to effectively avoid obstacles, there are many more conditions that an individual may encounter. For instance, the aperture structure was made of individual sheets of 5-mm depth medium-density fibre board that were hung up. These did not give the same information as a door frame which would also have a depth component to it of at least several centimetres. Also, what if there was a door hanging in the frame? The door could be opened to a variety of apertures requiring that individuals perceive the angle of the door before either pushing it open or squeezing through.

In these studies, only stationary obstacles existed, there was no evaluation of the ability of participants to avoid other moving obstacles. One of the first participants did comment “what is all the pretty music that I hear?” at the very beginning of his training session. The experimenter did not realise that her rushing around to set things up while the individual was self-training was

being heard by the participant. The evaluation of Doppler shift with approaching and receding obstacles was not examined in any of these experiments. Since avoidance of obstacles often requires avoiding other pedestrians, a good future experiment would be to evaluate self motion relative to other motion in the same environment.

The device used for the purpose of this experimentation was a prototype that was large and bulky requiring the individual to be carrying around significantly more weight than would be intended in a final design. Although these studies did show an ability to interpret the information, it is possible that different, or more accurate results might be obtained if the individuals were not constrained by wiring and boxes strapped to their heads. The next model of the ultrasound system is currently under development and is not expected to be much larger than a hearing aid. This will facilitate studies in a more ecological setting.

All testing was performed in an anechoic chamber. This allowed for differentiation between the different conditions by eliminating any other confounding factors of outside noise. This is an ideal environment for testing but cannot readily be applied to the “outside world”. Information about other noises in the environment must also be examined. The basic circuitry for the receivers followed from the design of bat detectors. What other organisms might be communicating at these ultrasound frequencies? This device was designed with a centre frequency of 40 kHz due to the choice of transducer, but is this the best signal to use? Future research must take into account more “noisy” areas to evaluate the effectiveness under more ecological conditions.

Considerable effort has gone into the design of a new system for providing spatial information in an intuitive manner and the methods to best display this information through hearing, but there is still a lot of work to be done. This research has shown that audification is an effective method of display for obstacle avoidance strategies but has not been exhaustive in the possible research environments to test the display. It has not been tested with visually impaired individuals. Other applications in which vision is occluded have not been examined (fire fighting, night-time surveillance). The one area of research that may also benefit considerably from this design is that of bat detection. Chiroptologists are always looking to new ways of hearing bats and this binaural method of audification will allow for effective detection and localisation of these mammals. Not only can this system be used in research for specific applications, but the concept of using this system to evaluate an acoustic tau that examines “time-to-collision” can provide insight into the basic science of perceptual processes.

### **9.3 Summary**

This work suggests that audification of ultrasound signals can enable safe and efficient travel by people with visual impairments. Direct perception through an acoustic array can allow for locomotion, suggesting that in fact, senses other than light reflection can be used for control of locomotion. Audification was shown to be effective for localisation of direct environmental sounds in both the azimuth and elevation directions. Audification through intentionally aliased signals creates an acoustic flow through which individuals can travel. There is evidence that depth perception occurs more effectively with laterally placed receivers that provide concurrent inflow and outflow cues. Direct perception through information in the acoustic array allowed for detection of obstacles and passage between them. The presence of acoustic flow and the ability

to navigate through an obstacle-rich environment has been shown. This novel approach will allow individuals with loss of vision to better “see” the world through their own ears.



## References

- ANSI (1996). Washington, DC: 3.22 American National Standards Institute.
- Allard, F. (2001). *Information Processing in Human Perceptual Motor Performance*. University of Waterloo.
- Amedi, A., Stern, W. M., Camprodon, J. A., Bermpohl, F., Merabet, L., Rotman, S. et al. (2007). Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. *Nat.Neurosci.*, 10, 687-689.
- Ashmead, D. H., Hill, E. W., & Talor, C. R. (1989). Obstacle perception by congenitally blind children. *Percept.Psychophys.*, 46, 425-433.
- Ashmead, D. H. & Wall, R. S. (1999). Auditory perception of walls via spectral variations in the ambient sound field. *J.Rehabil.Res.Dev.*, 36, 313-322.
- Ashmead, D. H., Davis, D. L., & Northington, A. (1995). Contribution of listeners' approaching motion to auditory distance perception. *Journal of Experimental Psychology: Human Perception and Performance.*, 21, 239-256.
- Au, W. W. L. (2004). Echolocation signals of wild dolphins. *Acoustical Physics*, 50, 454-462.
- Au, W. W. L., Carder, D. A., Penner, R. H., & Scronce, B. L. (1985). Demonstration of adaptation in beluga whale echolocation signals. *Journal of the Acoustical Society of America*, 77, 726-730.
- Auvray, M., Hannetono S., & O'Regan J.K. (2003). Localisation and form recognition in sensory substitution. In Proceedings of the 26 European Conference on Perception .

- Bach-Y-Rita, P., Collins, C. C., Scadden, L. A., Holmlund, G. W., & Hart, B. K. (1970). Display techniques in a tactile vision-substitution system. *Med Biol.Illus.*, 20, 6-12.
- Ballora, M., Pennycook, B., Ch Ivanov, P., Goldberger, A., & Glass, L. (2000). Detection of obstructive sleep apnea through auditory display of heart rate variability. In *Computers in Cardiology 2000, Sep 24-Sep 27 2000* (pp. 739-740). Cambridge, MA, USA: Institute of Electrical and Electronics Engineers Computer Society, Los Alamitos, CA, USA.
- Batteau DW (1967). The role of the pinna in human localization. *Proceedings of the Royal Society of London B Biological Sciences.*, 168, 158-180.
- Beall, A. C. & Loomis, J. M. (1996). Visual control of steering without course information. *Perception*, 25, 481-494.
- Bigel, M. G. & Ellard, C. G. (2000). The contribution of nonvisual information to simple place navigation and distance estimation: an examination of path integration. *Can.J.Exp.Psychol.*, 54, 172-185.
- Blasch, B. B., Long, R. G., & Griffinshirley, N. (1987). Evaluation of Electronic Travel Aids (Etas) for Visually Impaired Individuals. *International Journal of Rehabilitation Research*, 10, 458-459.
- Blasch, B. B., Long, R. G., & Griffinshirley, N. (1989). Results of A National Survey of Electronic Travel Aid Use. *Journal of Visual Impairment & Blindness*, 83, 449-453.
- Blauert, J. (1983). *Spatial Hearing, the psychophysics of human sound*. Cambridge, Massachusetts: The MIT Press.

- Bongers, R. M., Schellingerhout, R., van Grinsven, R., & Smitsman, A. W. (2002). Variables of the touch technique that influence the safety of cane walkers. *Journal of Visual Impairment & Blindness*, 96, 516-531.
- Brewster, S. A. (1997). Using non-speech sound to overcome information overload. *Displays*, 17, 179-189.
- Brewster, S. A. & Crease, M. G. (1999). Correcting menu usability problems with sound. *Behaviour and Information Technology*, 18, 165-177.
- Brewster, S. A., Wright, P. C., Dix, A. J., & Edwards, A. D. N. (1995). The sonic enhancement of graphical buttons. In *Proceedings of 5th International Conference on Human-Computer Interaction (INTERACT'95)*, 27-29 June 1995 (pp. 43-48). Lillehammer, Norway: Chapman & Hall.
- Brewster, S. A., Wright, P. C., & Edwards, A. D. N. (1994). The design and evaluation of an auditory-enhanced scrollbar. In *Proceedings of ACM Conference on Human Factors in Computer Systems*, 24-28 April 1994 (pp. 173-179). Boston, MA, USA: ACM.
- Brisben, A. J., Hsiao, S. S., & Johnson, K. O. (1999). Detection of vibration transmitted through an object grasped in the hand. *J.Neurophysiol.*, 81, 1548-1558.
- Brown, C. P. & Duda, R. O. (1998). Structural model for binaural sound synthesis. *IEEE Transactions on Speech and Audio Processing*, 6, 476-488.
- Burns, C. M. & Hajdukiewicz, J. R. (2004). *Ecological Interface Design*. Boca Raton, FL: CRC Press.

- Caclin, A., Soto-Faraco, S., Kingstone, A., & Spence, C. (2002). Tactile "capture" of audition. *Percept. Psychophys.*, 64, 616-630.
- Coward, S. W. & Stevens, C. J. (2004). Extracting Meaning from Sound: Nomic Mappings, Everyday Listening, and Perceiving Object Size from Frequency. *The Psychological Record*, 54, 349-364.
- Davies, T. C., Burns, C. M., & Pinder, S. D. (2007). Testing a novel auditory interface display to enable visually impaired travelers to use sonar mobility devices effectively. In *PROCEEDINGS of the HUMAN FACTORS AND ERGONOMICS SOCIETY 51st ANNUAL MEETING—2007* (pp. 278-282).
- Davies, T. C., Burns, C. M., & Pinder, S. D. (2006). Using ecological interface design to develop an auditory interface for visually impaired. In *OZCHI 2006* (pp. 309-312). ACM Press.
- Davies, T. C. & Patla, A. E. (2004). Obstacle Avoidance Strategies Using a Sonic Pathfinder. In *Proceedings of the Canadian Society of Biomechanics, Halifax, Nova Scotia, 2004*.
- DeLong, C. M., Au, W. W., Harley, H. E., Roitblat, H. L., & Pytko, L. (2007a). Human listeners provide insights into echo features used by dolphins (*Tursiops truncatus*) to discriminate among objects. *J Comp Psychol.*, 121, 306-319.
- DeLong, C. M., Au, W. W. L., & Stamper, S. A. (2007b). Echo features used by human listeners to discriminate among objects that vary in material or wall thickness: implications for echolocating dolphins. *Journal of the Acoustical Society of America*, 121, 605-617.

- Dombois, F. (2002). Auditory Seismology, on free oscillations, focal mechanisms, explosions and synthetic seismograms. In *Proceedings of the International Conference on Auditory Display in Kyoto, Japan, 2002*.
- Easton, R. D. (1992). Inherent problems of attempts to apply sonar and vibrotactile sensory aid technology to the perceptual needs of the blind. *Optom.Vis.Sci.*, 69, 3-14.
- Eldridge, A. (2006). Issues in Auditory Display. *Artificial Life*, 12, 259-274.
- Feinstein, R. (2001). Hearing is Believing.  
[http://www.bentvoices.org/bentvoices/feinstein\\_click.htm](http://www.bentvoices.org/bentvoices/feinstein_click.htm) Date Accessed: 12/10/2007  
 [On-line].
- Gibson, J. J. (1958). Visually controlled locomotion and visual orientation in animals.  
*Br.J.Psychol.*, 49, 182-194.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Hillsdale, New Jersey: Lawrence Erlbaum and Associates.
- Goldreich, D. & Kanics, I. M. (2003). Tactile acuity is enhanced in blindness. *J.Neurosci.*, 23, 3439-3445.
- Graham, R. (1999). Use of auditory icons as emergency warnings: evaluation within a vehicle collision avoidance application. *Ergonomics*, 42, 1233-1248.
- Heyes, A. D. (1980). Use of musical scales to display distance to objects in an electronic travel aid for the blind. *Percept.Mot.Skills*, 51, 1015-1020.
- Heyes, A. D. (1983). Human navigation by sound. *Physics in Technology*, 14, 68-75.

- Heyes, T. (1984). Sonic Pathfinder. *Electronics & Wireless World*, 90, 26-29.
- Hofman, M. & Van, O. J. (2003). Binaural weighting of pinna cues in human sound localization. *Exp.Brain Res.*, 148, 458-470.
- Hofman, P. M., Vlaming, M. S., Termeer, P. J., & Van Opstal, A. J. (2002). A method to induce swapped binaural hearing. *J.Neurosci.Methods*, 113, 167-179.
- Hughes, B. (2001). Active artificial echolocation and the nonvisual perception of aperture passability. *Hum.Mov Sci.*, 20, 371-400.
- Johannsen, G. (2004). Auditory displays in human-machine interfaces. *Proceedings of the IEEE*, 92, 742-758.
- Johnson, M., Hickmott, L. S., Aguilar, S. N., & Madsen, P. T. (2008). Echolocation behaviour adapted to prey in foraging Blainville's beaked whale (*Mesoplodon densirostris*). *Proc.Biol.Sci.*, 275, 133-139.
- Johnson, M., Madsen, P. T., Zimmer, W. M., de Soto, N. A., & Tyack, P. L. (2004). Beaked whales echolocate on prey. *Proc.Biol.Sci.*, 271 Suppl 6, S383-S386.
- Jones, W. G. (2004, February). Sight for sore ears. *IEEE Spectrum*, 13-14.
- Kay, L. (2000). Auditory perception of objects by blind persons, using a bioacoustic high resolution air sonar. *J.Acoust.Soc.Am.*, 107, 3266-3275.
- Kay, L. (2001). Bioacoustic spatial perception by humans: a controlled laboratory measurement of spatial resolution without distal cues. *J.Acoust.Soc.Am.*, 109, 803-808.

- Kish, D. (1995). *Echolocation: How humans can "see" without sight*.
- Klatzky, R. L., Giudice, N. A., Marston, J. R., Tietz, J., Golledge, R. G., & Loomis, J. M. (2008).  
An n-back task using vibrotactile stimulation with comparison to an auditory analogue.  
*Behav.Res.Methods*, 40, 367-372.
- Kramer, G., Walker, B. N., Bonebright, T., Cook, P., Flowers, J., Miner, N. et al. (1997).  
*Sonification Report: Status of the Field and Research Agenda* National Science  
Foundation: **International Community for Auditory Display**.
- Kramer, G. E. (1994). Auditory display: Sonification, audification and auditory interfaces. In *An introduction to auditory display* (pp. 1-78). Reading, MA: Addison-Wesley.
- Kuc, R. (2002). Binaural sonar electronic travel aid provides vibrotactile cues for landmark, reflector motion and surface texture classification. *IEEE Trans.Biomed.Eng*, 49, 1173-1180.
- Kulkarni, A. & Colburn, H. S. (1998). Role of spectral detail in sound-source localization.  
*Nature*, 396, 747-749.
- Kunkler-Peck, A. J. & Turvey, M. T. (2000). Hearing shape.  
*J.Exp.Psychol.Hum.Percept.Perform.*, 26, 279-294.
- Lagrow, S. J., Blasch, B. B., & Delaune, W. (1997). Efficacy of the touch technique for surface and foot-placement preview. *Journal of Visual Impairment & Blindness*, 91, 47-52.
- Lee, D. N. (1976). A theory of visual control of braking based on information about time-to-collision. *Perception*, 437-459.

- Lee, D. N., Simmons, J. A., Saillant, P. A., & Bouffard, F. (1995). Steering by echolocation: a paradigm of ecological acoustics. *J.Comp Physiol [A]*, 176, 347-354.
- Lee, D. N., van der Weel, F. R., Hitchcock, T., Matejowsky, E., & Pettigrew, J. D. (1992). Common principle of guidance by echolocation and vision. *J.Comp Physiol [A]*, 171, 563-571.
- Lessard, N., Pare, M., Lepore, F., & Lassonde, M. (1998). Early-blind human subjects localize sound sources better than sighted subjects. *Nature*, 395, 278-280.
- Lewald, J. (2002). Vertical sound localization in blind humans. *Neuropsychologia*, 40, 1868-1872.
- Loomis, J. M., Klatzky, R. L., & Golledge, R. G. (2001). Navigating without vision: basic and applied research. *Optom.Vis.Sci.*, 78, 282-289.
- Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W., & Fry, P. A. (1993). Nonvisual navigation by blind and sighted: assessment of path integration ability. *J.Exp.Psychol.Gen.*, 122, 73-91.
- Loomis, J. M., Lipka, Y., Golledge, R. G., & Klatzky, R. L. (2002). Spatial updating of locations specified by 3-d sound and spatial language. *J.Exp.Psychol.Learn.Mem.Cogn*, 28, 335-345.
- Loomis, J. M., Marston, J. R., Golledge, R. G., & Klatzky, R. L. (2005). Personal guidance system for people with visual impairment: A comparison of spatial displays for route guidance. *Journal of Visual Impairment & Blindness*, 99, 219-232.



- Maberley, D. A., Hollands, H., Chuo, J., Tam, G., Konkal, J., Roesch, M. et al. (2006). The prevalence of low vision and blindness in Canada. *Eye*, 20, 341-346.
- Makous, J. C. & Middlebrooks, J. C. (1990). Two-dimensional sound localization by human listeners. *J Acoust.Soc.Am*, 87, 2188-2200.
- Marr, D. (1976). Early Processing of Visual Information. *Philosophical Transactions of the Royal Society of London.Series B, Biological Sciences*, 275, 483-519.
- Marr, D. & Hildreth, E. (1980). Theory of Edge Detection. *Proceedings of the Royal Society of London.Series B, Biological Sciences*, 207, 187-217.
- Meijer, P. B. (1992). An experimental system for auditory image representations. *IEEE Trans.Biomed.Eng*, 39, 112-121.
- Mershon, D. H. & King, L. E. (1975). Intensity and reverberation as factors in the auditory perception of egocentric distance. *Perception & Psychophysics*, 18, 409-415.
- Middlebrooks, J. C. & Green, D. M. (1991). Sound Localization by Human Listeners. *Annual Review of Psychology*, 42, 135-159.
- Munhall, K. G., Jones, J. A., Callan, D. E., Kuratate, T., & Vatikiotis-Bateson, E. (2004). Visual prosody and speech intelligibility: head movement improves auditory speech perception. *Psychol.Sci.*, 15, 133-137.
- Neuhoff J.(Ed) (2004). *Ecological Psychoacoustics*. San Diego, California: Elsevier Academic Press.

- Neuman, A. C. (2005). Central auditory system plasticity and aural rehabilitation of adults. *J Rehabil.Res.Dev.*, 42, 169-186.
- Oldfield, S. R. & Parker, S. P. (1984). Acuity of sound localisation: a topography of auditory space. I. Normal hearing conditions. *Perception*, 13, 581-600.
- Patla, A. E. (1997). Understanding the roles of vision in the control of human locomotion. *Gait & Posture*, 5, 54-69.
- Patla, A. E. (2004). Adaptive human locomotion: influence of neural, biological and mechanical factors on control mechanisms. In A.Bronstein (Ed.), *Clinical Disorders of Balance, Posture, and Gait* (5 ed., pp. 20-38). Oxford University Press.
- Patla, A. E., Davies, T. C., & Niechwiej, E. (2004). Obstacle avoidance during locomotion using haptic information in normally sighted humans. *Exp.Brain Res.*, 155, 173-185.
- Patla, A. E., Niechwiej, E., Racco, V., & Goodale, M. A. (2002). Understanding the contribution of binocular vision to the control of adaptive locomotion. *Exp.Brain Res.*, 142, 551-561.
- Philbeck, J. W., Klatzky, R. L., Behrmann, M., Loomis, J. M., & Goodridge, J. (2001). Active control of locomotion facilitates nonvisual navigation. *J.Exp.Psychol.Hum.Percept.Perform.*, 27, 141-153.
- Philbeck, J. W., Loomis, J. M., & Beall, A. C. (1997). Visually perceived location is an invariant in the control of action. *Percept.Psychophys.*, 59, 601-612.

- Philbeck, J. W. & Mershon, D. H. (2002). Knowledge about typical source output influences perceived auditory distance (L). *Journal of the Acoustical Society of America*, 111, 1980-1983.
- Pinder, S. D. & Davies, T. C. (2007). Exploring direct downconversion of ultrasound for human echolocation. In *Proceedings of the 7th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction: design centered HCI* (pp. 49-54).
- Ramsey, V. K., Blasch, B. B., Kita, A., & Johnson, B. F. (1999). A biomechanical evaluation of visually impaired persons' gait and long-cane mechanics. *J.Rehabil.Res.Dev.*, 36, 323-332.
- Rice, C. E., Feinstein, S. H., & Schusterman, R. J. (1965). Echo Detection Ability of the Blind: Size and Distance Factors. *Journal of Experimental Psychology.*, 70, 246-251.
- Roder, B., Teder-Salejarvi, W., Sterr, A., Rosler, F., Hillyard, S. A., & Neville, H. J. (1999). Improved auditory spatial tuning in blind humans. *Nature*, 400, 162-166.
- Rosenblum, L. D., Gordon, M. S., & Jarquin, L. (2000). Echolocating distance by moving and stationary listeners. *Ecological Psychology*, 12, 181-206.
- Sanderson, P., Anderson, J., & Watson, M. (2000). Extending ecological interface design to auditory displays. In *Proceedings of the 2000 Annual Conference of the Computer Human Interaction Special Interest Group of the Ergonomics Society of Australia (OZCHI2000)* (pp. 259-256).

- Sanderson, P. & Watson (2005). From Information Content to Auditory Display with Ecological Interface Design:Prospects and Challenges. In *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting* (pp. 259-263).
- Sanderson, P. M. (2005). Beyond visual display design:Prospects for auditory and cross-modal information representation in complex socio-technical systems. In *7<sup>th</sup> Cognitive Systems Engineering Consortium webcast*.
- Schellingerhout, R., Bongers, R. M., van, G. R., Smitsman, A. W., & Van Galen, G. P. (2001). Improving obstacle detection by redesign of walking canes for blind persons. *Ergonomics*, 44, 513-526.
- Schenkman, B. N. (1986). Identification of ground materials with the aid of tapping sounds and vibrations of long canes for the blind. *Ergonomics*, 29, 985-998.
- Schenkman, B. N. & Jansson, G. (1986). Detection and localisation of objects by the blind with the aid of long-cane tapping sounds. *Human Factors*, 28, 607-618.
- Sebkova, J. & Bamford, J. M. (1981). Evaluation of binaural hearing aids in children using localization and speech intelligibility tasks. *Br.J Audiol.*, 15, 125-132.
- Speigle, J. M. & Loomis, J. M. (1993). Auditory distance perception by translating observers. In *Virtual Reality, 1993. Proceedings., IEEE 1993 Symposium on Research Frontiers in* (pp. 92-99).
- Stoffregen, T. A. & Pittenger, J. B. (1995). Human echolocation as a basic form of perception and action. *Ecological Psychology*, 12, 181-206.

- Strelow, E. R. (1985). What is needed for a theory of mobility: direct perception and cognitive maps--lessons from the blind. *Psychol.Rev.*, 92, 226-248.
- Strelow, E. R. & Brabyn, J. A. (1982). Locomotion of the blind controlled by natural sound cues. *Perception*, 11, 635-640.
- Strutt.L.Rayleigh (1907). On our perception of sound direction. 13:214-232, 1907. *Philosophy Magazine*, 13, 214-232.
- Supa, M., Cotzin, M., & Dallenbach, K. M. (1944a). "Facial Vision": The Perception of Obstacles by the Blind. *The American Journal of Psychology*, 57, 133-183.
- Supa, M., Cotzin, M., & Dallenbach, K. M. (1944b). "Facial vision"; the perception of obstacles by the blind.
- Thurlow, W. R., Mangels, J. W., & Runge, P. S. (1967). Head movements during sound localization. *J Acoust.Soc.Am*, 42, 489-493.
- Thurlow, W. R. & Runge, P. S. (1967). Effect of Induced Head Movements on Localization of Direction of Sounds. *The Journal of the Acoustical Society of America*, 42, 480-488.
- Uslan, M. M. (1978). Cane Technique - Modifying Touch Technique for Full Path Coverage. *Journal of Visual Impairment & Blindness*, 72, 10-14.
- Vicente, K. J. (1999). *Cognitive work analysis : toward safe, productive, and healthy computer-based work*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Vicente, K. J. & Rasmussen, J. (1992). Ecological interface design: Theoretical foundations. *IEEE Transactions on Systems, Man and Cybernetics*, 22, 589-606.

- Walker, B. N. & Kramer, G. (2004). Ecological psychoacoustics and auditory displays. Hearing, grouping and meaning making. In (.
- Walker, B. N. & Lindsay J. (2006). Navigation Performance With a Virtual Auditory Display: Effects of Beacon Sound, Capture Radius, and Practice. *Human Factors*, 48, 265-278.
- Walker, V. A., Peremans, H., & Hallam, J. C. (1998). One tone, two ears, three dimensions: a robotic investigation of pinnae movements used by rhinolophid and hipposiderid bats. *J Acoust.Soc.Am*, 104, 569-579.
- Wall, R. S. (2001). An exploratory study of how travelers with visual impairments modify travel techniques in winter. *Journal of Visual Impairment & Blindness*, 95, 752-756.
- Wall, R. S. & Ashmead, D. H. (2002). Changes in biomechanical features of the two-point touch technique as it is learned. *Journal of Visual Impairment & Blindness*, 96, 829-841.
- Walraven, J. (1982). (Rep. No. Report 1ZF).
- Warren, W. H., Jr. & Whang, S. (1987). Visual guidance of walking through apertures: body-scaled information for affordances. *J.Exp.Psychol.Hum.Percept.Perform.*, 13, 371-383.
- Waters, D. A. & Abulula, H. H. (2007). Using bat-modelled sonar as a navigational tool in virtual environments. *International Journal of Human-Computer Studies*, 65, 873-886.
- Watson, M. & Sanderson, P. (1998). Work domain analysis for the evaluation of human interaction with anaesthesia alarm systems. In *Proceedings 1998 Australasian Computer Human Interaction Conference. OzCHI'98, 30 Nov.-4 Dec. 1998* (pp. 228-235). Adelaide, SA, Australia: IEEE Comput. Soc.

- Watson, M. O. & Sanderson, P. M. (2007). Designing for attention with sound: Challenges and extensions to ecological interface design. *Human Factors*, 49, 331-346.
- Watters, C., Owen, M., & Munroe, S. (2004). *A study of Deaf-Blind Demographics and Services in Canada: A report Prepared for the Canadian National Society of the Deaf-Blind*.
- Wightman, F. L. & Kistler, D. J. (1989). Headphone simulation of free-field listening. I: Stimulus synthesis. *J.Acoust.Soc.Am.*, 85, 858-867.
- Worchel, P. & Dallenbach, K. M. (1947). "Facial Vision": Perception of obstacles by the deaf-blind. *American Journal of Psychology*, 60, 502-553.
- Zahorik, P. (2001). Estimating sound source distance with and without vision. *Optom.Vis.Sci.*, 78, 270-275.
- Zeilek J.S., Bromley S., Asmar D., & Thompson D (2003). A haptic glove as a tactile-vision sensory substitution for wayfinding. *Journal of Visual Impairment and Blindness*, 97, 621-632.
- Zwiers, M. P., Van Opstal, A. J., & Cruysberg, J. R. (2001). Two-dimensional sound-localization behavior of early-blind humans. *Exp.Brain Res.*, 140, 206-222.

**Appendix 1**  
**Information and Consent Form – Test 2**

Department of Systems Design Engineering  
University of Waterloo, CANADA

Study Title: Localising environmental obstacles using ultrasound and audible sound: Comparison of echoed localisation between audible and downconverted ultrasound signals

Supervisor: Dr. Catherine Burns, Systems Design Engineering (1-519-888-4567, ext. 3903), [c4burns@uwaterloo.ca](mailto:c4burns@uwaterloo.ca)

PhD Candidate: Claire Davies, Advanced Interface Design Laboratory, (NZ mobile 021 1025971 or 1-519-888-4567, ext. 4904) [cdavies@uwaterloo.ca](mailto:cdavies@uwaterloo.ca)

This study is being conducted by a researcher from Canada who has received approval from the Office of Research Ethics at the University of Waterloo. Her application has been further reviewed by the University of Auckland Human Participants and Ethics Committee and granted permission to perform this research at the University of Auckland Acoustics Research Centre. You have received two copies of this Information and Consent Form. Should you decide to participate in the study, please return one signed copy and keep the other for personal reference.

Participation in this study is voluntary, and will take approximately two hours of your time.

By volunteering for this study, you will help to identify differences between sounds that are echoed in an audible domain and sounds that are echoed in the ultrasound domain and translated into the audible domain. The specific task that you will undertake will be to identify the location of a “wall” and walk toward its perceived position once it is removed from its location. The wall will consist of a large square of particle board with sides of 2m.

First, your hearing will be tested. Eighteen sounds will be played to you on a headset, and you are asked to report whether or not you were able to hear them. The range of sound pressure levels during this test will be from 10dB to 50dB. This refers to sound levels as quiet as rustling leaves or a quiet whisper to that of a dishwasher or soft conversation.

Following the hearing test, there will be a total of 100 blindfolded trials. For each trial, you will hear a sound reflected from the “wall” through a headset. Your ears will be occluded with ear defectors and you will remain standing until the “wall” is moved. You will then walk to the perceived position of the wall. There will be fifty trials in each of two conditions, stationary and moving. For the stationary trials, you will determine the position of the wall based on either reflected echoes in the auditory domain or from audible signals of downconverted ultrasound. The moving condition will involve the same signals, but you will echolocate by moving toward the wall for ten feet arriving at the same point that you started for the stationary trials. Your ears will be occluded and you will walk to the perceived position of the wall.



Potential benefits from this study include insight into a blind person's spatial impairment. The study may have implications for future research on the design of aids for blind people, as well as design of living/working environment for visually impaired individuals.

The potential risks from this study are minimal. There is a risk that the signals could affect your hearing in the short term (sound induced hearing loss, which is not permanent) if played too loudly. However, the sound level will be limited to 60 dB (normal conversation levels) to prevent temporary sound induced hearing loss. Also there is a risk that you may trip while walking to the location of the wall. A spotter will be beside you to reduce the risk of falling.

You may decide to withdraw from this study at any time by advising the researcher, and may do so. At the end of the study, you will receive a t-shirt.

All information you provide is considered completely confidential; indeed, your name will not be included or in any other way associated, with the data collected in the study. Data collected during this study will be retained indefinitely, in a locked office to which only researchers associated with this study have access.

With your permission, we will videotape you during the session. This will allow us to make some calculations regarding the positions that you estimated as the location of the wall. Sometimes a certain photograph and/or part of a videotape clearly show a particular feature or detail that would be helpful in teaching or when presenting the study results at a scientific presentation or in a publication.

This study has been reviewed and received ethics clearance through the Office of Research Ethics at the University of Waterloo. However, the final decision about participation is yours. If you have any comments or concerns resulting from your participation in this study, please feel free to contact this office at 1-519-888-4567 ext. 6005 or by email at [ssykes@uwaterloo.ca](mailto:ssykes@uwaterloo.ca).

#### Consent of Participant

I have read the information presented above about the procedures and risks involved in this study, and I have received satisfactory answers to my questions related to this study. The specific details of this study have been explained, as well as the potential risks of the study. I am aware that this project was reviewed by, and received ethics clearance through the Office of Research Ethics at the University of Waterloo (519-888-4567, ext. 6005 or by email at [ssykes@uwaterloo.ca](mailto:ssykes@uwaterloo.ca)). I am aware that I may request a break at any point during the experimental procedure, as well, I may withdraw from the study at any time. With full knowledge of all foregoing, I agree, of my own free will to participate in this study.

---

Print Name

---

Signature of Participant

---

Dated at Auckland, New Zealand

---

Signature of Witness

## **Consent to Use Video and/or Photographs in Teaching, Presentations, and Publications**

Sometimes a certain photograph and/or part of a videotape clearly show a particular feature or detail that would be helpful in teaching or when presenting the study results at a scientific presentation or in a publication.

I agree to allow video recordings in which I appear to be used in teaching, scientific presentations and/or publications with the understanding that I will not be identified by name. I am aware that I may withdraw this consent at any time without penalty. I will not be able to be identified in the videotape.

I was informed that if I have any comments or concerns resulting from my participation in this study, I may contact the Director, Office of Research Ethics at the University of Waterloo, at (1- 519) 888-4567 ext. 6005.

---

Print Name

---

Signature of Participant

---

Dated at Auckland, New Zealand

---

Signature of Witness

## **Appendix 2**

### **Information and Consent Form – Test 3**

Department of Systems Design Engineering  
University of Waterloo, CANADA

Study Title: Localising environmental obstacles using ultrasound and audible sound: Comparison of echoed localisation between audible and downconverted ultrasound signals while passing through apertures

Supervisor: Dr. Catherine Burns, Systems Design Engineering (1-519-888-4567, ext. 3903), [c4burns@uwaterloo.ca](mailto:c4burns@uwaterloo.ca)

PhD Candidate: Claire Davies, Advanced Interface Design Laboratory, (NZ mobile 021 1025971 or 1-519-888-4567, ext. 4904) [cdavies@uwaterloo.ca](mailto:cdavies@uwaterloo.ca)

This study is being conducted by a researcher from Canada who has received approval from the Office of Research Ethics at the University of Waterloo. Her application has been further reviewed by the University of Auckland Human Participants and Ethics Committee and granted permission to perform this research at the University of Auckland Acoustics Research Centre. You have received two copies of this Information and Consent Form. Should you decide to participate in the study, please return one signed copy and keep the other for personal reference.

Participation in this study is voluntary, and will take approximately two hours of your time.

By volunteering for this study, you will help to identify differences between sounds that are echoed in an audible domain and sounds that are echoed in the ultrasound domain and translated into the audible domain. The specific task that you will undertake will be to identify and respond to the presence of obstacles, in this case, two sides of a doorway using only your sense of hearing.

First, your hearing will be tested. Eighteen sounds will be played to you on a headset, and you are asked to report whether or not you were able to hear them. The range of sound pressure levels during this test will be from 10dB to 50dB. This refers to sound levels as quiet as rustling leaves or a quiet whisper to that of a dishwasher or soft conversation.

Following the hearing test, there will be a total of 50 blindfolded trials. For each trial, you will hear information from reflected signals through a headset. These will help you to identify a “doorway”. While listening to the signals, you will walk to the doorway and stand within it. There will be twenty five trials in each of two conditions, audible and downconverted ultrasound. The size of the “doorway” will vary from trial to trial between 55 cm and 95 cm.

Potential benefits from this study include insight into a blind person’s spatial impairment. The study may have implications for future research on the design of aids for blind people, as well as design of living/working environment for visually impaired individuals.

The potential risks from this study are minimal. There is a risk that the signals could affect your hearing in the short term (sound induced hearing loss, which is not

permanent) if played too loudly. However, the sound level will be limited to 60 dB (normal conversation levels) to prevent temporary sound induced hearing loss. Also there is a risk that you may trip while walking to the doorway or perhaps run into the sides of the doorway. A spotter will be beside you to reduce the risk of falling or running into the particle board defining the doorway.

You may decide to withdraw from this study at any time by advising the researcher, and may do so. At the end of the study, you will receive a t-shirt.

All information you provide is considered completely confidential; indeed, your name will not be included or in any other way associated, with the data collected in the study. Data collected during this study will be retained indefinitely, in a locked office to which only researchers associated with this study have access.

With your permission, we will videotape you during the session. This will allow us to make some calculations regarding the positions that you estimated as the location of the door. Sometimes a certain photograph and/or part of a videotape clearly show a particular feature or detail that would be helpful in teaching or when presenting the study results at a scientific presentation or in a publication.

This study has been reviewed and received ethics clearance through the Office of Research Ethics at the University of Waterloo. However, the final decision about participation is yours. If you have any comments or concerns resulting from your participation in this study, please feel free to contact this office at 1-519-888-4567 ext. 6005 or by email at [ssykes@uwaterloo.ca](mailto:ssykes@uwaterloo.ca).

#### Consent of Participant

I have read the information presented above about the procedures and risks involved in this study, and I have received satisfactory answers to my questions related to this study. The specific details of this study have been explained, as well as the potential risks of the study. I am aware that this project was reviewed by, and received ethics clearance through the Office of Research Ethics at the University of Waterloo (519-888-4567, ext. 6005 or by email at [ssykes@uwaterloo.ca](mailto:ssykes@uwaterloo.ca)). I am aware that I may request a break at any point during the experimental procedure, as well, I may withdraw from the study at any time. With full knowledge of all foregoing, I agree, of my own free will to participate in this study.

---

Print Name

---

Signature of Participant

---

Dated at Auckland, New Zealand

---

Signature of Witness

## Consent to Use Video and/or Photographs in Teaching, Presentations, and Publications

Sometimes a certain photograph and/or part of a videotape clearly show a particular feature or detail that would be helpful in teaching or when presenting the study results at a scientific presentation or in a publication.

I agree to allow video recordings in which I appear to be used in teaching, scientific presentations and/or publications with the understanding that I will not be identified by name. I am aware that I may withdraw this consent at any time without penalty. I will not be able to be identified in the videotape.

I was informed that if I have any comments or concerns resulting from my participation in this study, I may contact the Director, Office of Research Ethics at the University of Waterloo, at (1- 519) 888-4567 ext. 6005.

---

Print Name

---

Signature of Participant

---

Dated at Auckland, New Zealand

---

Signature of Witness

### **Appendix 3**

#### **Information and Consent Form – Test 1**

Department of Systems Design Engineering  
University of Waterloo, CANADA

Study Title: Localising environmental obstacles using ultrasound and audible sound: Comparison of direct localisation between audible and downconverted ultrasound signals.

Supervisor: Dr. Catherine Burns, Systems Design Engineering (1-519-888-4567, ext. 3903), [c4burns@uwaterloo.ca](mailto:c4burns@uwaterloo.ca)

PhD Candidate: Claire Davies, Advanced Interface Design Laboratory, (NZ mobile 021 1025971 or 1-519-888-4567, ext. 4904) [cdavies@uwaterloo.ca](mailto:cdavies@uwaterloo.ca)

This study is being conducted by a researcher from Canada who has received approval from the Office of Research Ethics at the University of Waterloo. Her application has been further reviewed by the University of Auckland Human Participants and Ethics Committee and granted permission to perform this research at the University of Auckland Acoustics Research Centre. You have received two copies of this Information and Consent Form. Should you decide to participate in the study, please return one signed copy and keep the other for personal reference.

Participation in this study is voluntary, and will take approximately two hours of your time.

By volunteering for this study, you will help to identify differences between sounds that are echoed in an audible domain and sounds that are echoed in the ultrasound domain and translated into the audible domain. The specific task that you will undertake will be to identify the locations of various sounds by pointing to them.

First, your hearing will be tested. Eighteen sounds will be played to you on a headset, and you are asked to report whether or not you were able to hear them. The range of sound pressure levels during this test will be from 10dB to 50dB. This refers to sound levels as quiet as rustling leaves or a quiet whisper to that of a dishwasher or soft conversation.

Following the hearing test, you will wear a blindfold and a stereo headset. In each of four conditions, you will perform 72 localisation trials. The localisation trial will consist of listening to a speaker sound and pointing in the direction of the sound. The direction of pointing will be recorded using a videocamera. One condition will be direct audible signals. The other conditions will be downconverted ultrasound signals with the ultrasound receivers placed at three different orientations relative to your ears. You will hear a sound point toward it. After each condition, you will be given the opportunity to rest before continuing to the next condition.

Potential benefits from this study include insight into a blind person's spatial impairment. The study may have implications for future research on the design of aids for individuals with visual impairments, as well as design of living/working environment for visually impaired individuals.

The potential risks from this study are minimal. There is a risk that the signals could affect your hearing in the short term (sound induced hearing loss, which is not permanent) if played too loudly. However, the sound level will be limited to 60 dB (normal conversation levels) to prevent temporary sound induced hearing loss. A spotter will be beside you to offer support should you become disoriented.

You may decide to withdraw from this study at any time by advising the researcher, and may do so. At the end of the study, you will receive a t-shirt.

All information you provide is considered completely confidential; indeed, your name will not be included or in any other way associated, with the data collected in the study. Data collected during this study will be retained indefinitely, in a locked office to which only researchers associated with this study have access.

With your permission, we will videotape you during the session. This will allow us to make some calculations regarding the positions that you estimated as the location of the sound. Sometimes a certain photograph and/or part of a videotape clearly show a particular feature or detail that would be helpful in teaching or when presenting the study results at a scientific presentation or in a publication.

This study has been reviewed and received ethics clearance through the Office of Research Ethics at the University of Waterloo. However, the final decision about participation is yours. If you have any comments or concerns resulting from your participation in this study, please feel free to contact this office at 1-519-888-4567 ext. 6005 or by email at [ssykes@uwaterloo.ca](mailto:ssykes@uwaterloo.ca).

#### Consent of Participant

I have read the information presented above about the procedures and risks involved in this study, and I have received satisfactory answers to my questions related to this study. The specific details of this study have been explained, as well as the potential risks of the study. I am aware that this project was reviewed by, and received ethics clearance through the Office of Research Ethics at the University of Waterloo (519-888-4567, ext. 6005 or by email at [ssykes@uwaterloo.ca](mailto:ssykes@uwaterloo.ca)). I am aware that I may request a break at any point during the experimental procedure, as well, I may withdraw from the study at any time. With full knowledge of all foregoing, I agree, of my own free will to participate in this study.

---

Print Name

---

Signature of Participant

---

Dated at Auckland, New Zealand

---

Signature of Witness

## Consent to Use Video and/or Photographs in Teaching, Presentations, and Publications

Sometimes a certain photograph and/or part of a videotape clearly show a particular feature or detail that would be helpful in teaching or when presenting the study results at a scientific presentation or in a publication.

I agree to allow video recordings in which I appear to be used in teaching, scientific presentations and/or publications with the understanding that I will not be identified by name. I am aware that I may withdraw this consent at any time without penalty. I will not be able to be identified in the videotape.

I was informed that if I have any comments or concerns resulting from my participation in this study, I may contact the Director, Office of Research Ethics at the University of Waterloo, at (1- 519) 888-4567 ext. 6005.

---

Print Name

---

Signature of Participant

---

Dated at Auckland, New Zealand

---

Signature of Witness