BEYOND AMBIENT EXPERIENCE: AN AUDITORY DISPLAY DESIGN FRAMEWORK FOR UBIQUITOUS COMPUTING ENVIRONMENTS

by

Milena Droumeva
Bachelor of Communication, Simon Fraser University

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

In the
School of Interactive Arts and Technologies

© Milena Droumeva 2007

SIMON FRASER UNIVERSITY

Summer 2007

All rights reserved. This work may not be reproduced in whole or in part, by photocopy or other means, without permission of the author.
## APPROVAL

<table>
<thead>
<tr>
<th>Name:</th>
<th>Milena Droumeva</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree:</td>
<td>Master of Arts</td>
</tr>
<tr>
<td>Title of Thesis:</td>
<td>Beyond Ambient Experience: An Auditory Display Design Framework for Ubiquitous Computing Environments</td>
</tr>
</tbody>
</table>

**Examining Committee:**

**Chair:**

Maia Engeli  
Assistant Professor, School of Interactive Art and Technology

Ron Wakkary  
Senior Supervisor  
Associate Professor, School of Interactive Art and Technology

Barry Truax  
Supervisor  
Professor, School of Communication

Martin Gotfrit  
External Examiner  
Associate Professor, School of Contemporary Arts

**Date Defended/Approved:**  
July 30, 2007
DECLARATION OF PARTIAL COPYRIGHT LICENCE

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the “Institutional Repository” link of the SFU Library website <www.lib.sfu.ca> at: <http://ir.lib.sfu.ca/handle/1892/112>) and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author’s written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, BC, Canada
STATEMENT OF ETHICS APPROVAL

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

(a) Human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

(b) Advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

(c) as a co-investigator, in a research project approved in advance,

or

(d) as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, BC, Canada
ABSTRACT

Ubiquitous computing environments are becoming commonplace yet sound displays for them often lack consideration for contextual and embodied aspects. In addition, the audio displays have not been systematically examined as meaningful feedback. This thesis proposes a new framework for auditory display design for ubiquitous computing. The framework is created by synthesizing theories of acoustic communication and auditory display design; by designing and testing a ubiquitous computing prototype named socio-ec(h)o; and through a set of experimental studies focused on specific sound issues. Particular attention is given to an intensity-based gradient approach to feedback. This approach includes complex, environmental sound, and utilizes sonification principles that convey information, as well as support embodiment, sociality, and mediation in a ubiquitous computing setting. The study also examines the role of participatory design workshops and modified experimental studies informed by design methods as methodological innovations in the research of design and audio displays.

Keywords: Sound Design, Design Methods, Ubiquitous Computing, Auditory Display, Design-based Research, Sound Feedback, Sound Ecology, Acoustic Communication

ACKNOWLEDGEMENTS

I am deeply thankful to those who have supported and encouraged me through my undergraduate and graduate degrees, especially in my intellectual endeavours and this work, in particular. Some of these good people are: Barry Truax, for his much appreciated mentorship, learning and support for years and counting; Ron Wakkary, my senior supervisor, whose calm and expert guidance has always reassured me; Norbert Ruebsaat for friendship and stimulating conversation; Suzanne de Castell for much help with talking about and reading over parts of my thesis; Kenneth Newby for his initial guidance and friendship, and bringing me to SIAT; and many others who I feel have made a positive difference during my time at SIAT: Alissa Antle, Jim Bizzocchi, Marek Hatala, Tom Calvert, Rob Woodbury and Tamara Smyth, to name a few.

Outside school, I hold deep gratitude for all the folks at ICAD who have generously offered their friendship, support, help and mentorship over the last three years, especially Derek Brock, Alberto del Campo, Terri Bonebright, Tony Stockman, Bill Martens, Bruce Walker, Emma Murphy, Gregory Leplâtre, Thomas Hermann and many others.

On a personal note I’d like to acknowledge the friendship and support of many school friends who have made my time and work at SIAT a great experience. Some of them are Nis Bojin, Lorna Boschman, Gilly Mah, Rob Lovell, Caitlin Akai, Nima Motamedi, Greg Corness and countless others.

Last, but not least, many thanks go to my partner in crime/life Anne O’Shea for always being there for me; to my sister Radina for being a moral support, and most of all to my parents, Stefka and Vladimir, who have helped me accomplish the things I’ve set my mind to, and whose coming here allowed me to explore the things I am interested in.
# TABLE OF CONTENTS

- Approval .......................................................................................................................... ii
- Abstract .............................................................................................................................. iii
- Acknowledgements ............................................................................................................. iv
- Table of Contents ............................................................................................................... v
- List of Figures ..................................................................................................................... vii
- List of Tables ....................................................................................................................... ix
- Glossary ................................................................................................................................ x

## Chapter 1: Introduction ...................................................................................................... 1
- Research Problem Area ........................................................................................................ 1
- Research Questions .............................................................................................................. 3
- The Role of Sound in the Environment .................................................................................. 3
- Background of Existing Frameworks for Sound Design ....................................................... 6
- New Framework Justification ................................................................................................ 8
- Thesis Outline ....................................................................................................................... 9

## Chapter 2: Literature Review .......................................................................................... 11
- Historical Background of Listening .................................................................................... 11
- Listening before Electroacoustic Reproduction ................................................................... 12
- Listening in the Electroacoustic Era ..................................................................................... 13
- Listening in the Age of Media ............................................................................................. 15
- Media, Technology and Space ............................................................................................. 18
- Sound in Cinema .................................................................................................................. 18
- Listening and Virtual Audio ................................................................................................ 19
- Interactive - Adaptive Audio in Games ............................................................................... 20
- Interactive Auditory Displays in Ubiquitous Computing Environments ............................. 22

- Sound Design Examples in Ubiquitous Computing .............................................................. 24
- The Gap ............................................................................................................................... 27
- Acoustic Communication Framework ................................................................................ 29
  - Acoustic Ecology Model .................................................................................................... 29
  - Acoustic Communication Model ..................................................................................... 31
- Auditory Display Framework ............................................................................................... 34
  - Historical Perspective of Auditory Displays ..................................................................... 37
  - Sonification ....................................................................................................................... 39
  - New Framework Conceptualization .................................................................................. 43

## Chapter 4: Methodological Discussion .......................................................................... 46
- Research Methodology ........................................................................................................ 46
- Theory Building .................................................................................................................. 48
LIST OF FIGURES

Figure 1.1: Areas of significance that characterize ubiquitous computing technologies and environments ......................................................... 5

Figure 1.2: Structure of the research inquiry of the thesis, including relevant areas of consideration and/or investigation ........................................... 10

Figure 3.1: The big circles signify areas of keynote soundscapes, while smaller disk shapes signify soundmarks representing the content of the exhibit area ......................................................................................... 25

Figure 3.2: An overview of different sonic environments that bear relevance to the framework of auditory display design for ubiquitous computing .......... 28

Figure 3.3: A diagram showing the main areas that the two existing frameworks contribute to a conceptualization of sound design in a ubiquitous computing setting ........................................................................ 44

Figure 5.1: Participant on individual exploration in Participatory Workshop I ............................................. 59

Figure 5.2: Illustration of the speaker arrangement in PD Workshop I. The four dotted circles represent acoustic space ranges, and there were two areas of projected light on each side of the scrim (wavy line) ...................... 60

Figure 5.3: Illustration of the speaker arrangement and acoustic space ranges in Participatory Workshop II ........................................................................ 67

Figure 5.4: Participants creating ecologies in PD Workshop II ................................................................. 68

Figure 5.5: Participants playing socio-eco(h)o ...................................................................................... 74

Figure 5.6: A model of our sound and perception approach .................................................................. 75

Figure 5.7: Mapping of the intensity gradient and audio display to sample game play, © 2006, Ron Wakkary, by permission ....................................................... 77

Figure 5.8: Reiterating the structure of the research inquiry in this thesis. The highlighted areas are the ones that will be addressed here ........................................................................ 87

Figure 6.1: A schematic showing the interplay between mediation, perception and action, as nested within their context and physical environment .......... 92

Figure 6.2: Model of the world map indicating areas of low, medium and high intensity, as well as three sample trajectories ......................................................... 101

Figure 6.3: Mapping of audio parameters to the 2-D map .................................................................. 105

Figure 6.4: Screenshots of each iteration of the map game with its corresponding audio waveform track .................................................................................. 108

Figure 6.5: Illustration of the intensity-based auditory feedback model ............................................. 110
Figure 7.1: Comparative observation results regarding accurate perception of sound intensity gradient sound feedback in the different conditions of embodiment. ................................................................. 119

Figure 7.2: A schema showing the progression that the research presented here took, identifying the different emerging concepts leading up to the proposed design guidelines. ................................................................. 124

Figure 7.3: A representation of the relationships between main design areas identified here as central to the framework for auditory display design for ubiquitous computing. ................................................................. 130
**LIST OF TABLES**

<table>
<thead>
<tr>
<th>Table 2.1:</th>
<th>Key concepts from the music-as-environment paradigm</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 2.2:</td>
<td>Key Concepts from the media-as-environment paradigm</td>
<td>17</td>
</tr>
<tr>
<td>Table 2.3:</td>
<td>Sonic components that provide contextual information in game soundscapes</td>
<td>22</td>
</tr>
<tr>
<td>Table 3.1:</td>
<td>Key concepts from the acoustic ecology and acoustic communication models</td>
<td>32</td>
</tr>
<tr>
<td>Table 3.2:</td>
<td>Overview of the main areas of auditory display systems</td>
<td>35</td>
</tr>
<tr>
<td>Table 3.3:</td>
<td>Types and distinctions of auditory tasks for data comprehension</td>
<td>36</td>
</tr>
<tr>
<td>Table 3.4:</td>
<td>Types of auditory displays and their conceptual mapping to meaning and information</td>
<td>36</td>
</tr>
<tr>
<td>Table 5.1:</td>
<td>Sound avatars and their representations, which include three levels of increasing intensity</td>
<td>66</td>
</tr>
<tr>
<td>Table 5.2:</td>
<td>Audio display methods used in levels</td>
<td>76</td>
</tr>
<tr>
<td>Table 6.1:</td>
<td>This table shows the three instruments used in the workshop and their respective intended approaches to intensity</td>
<td>99</td>
</tr>
<tr>
<td>Table 6.2:</td>
<td>The table shows a few examples from the sound plotting section of the workshop. On the models to the left the X axis represents time in seconds, while the Y axis represents intensity from 0% to 100%</td>
<td>103</td>
</tr>
<tr>
<td>Table 6.3:</td>
<td>Coded results outlining participants' observed and/or reported use of the audio feedback devices and approach to representing intensity</td>
<td>107</td>
</tr>
<tr>
<td>Table 6.4:</td>
<td>Breakdown of participant comments and observations related to each sonification parameter in the game</td>
<td>114</td>
</tr>
<tr>
<td>Table 7.1:</td>
<td>Table comparing accuracy of perception results from the experimental studies, across several different settings of embodiment</td>
<td>118</td>
</tr>
<tr>
<td>Table 7.2:</td>
<td>A breakdown of the comparison between main features of the two existing frameworks and the proposed new framework for auditory display design</td>
<td>132</td>
</tr>
</tbody>
</table>
GLOSSARY

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory Displays (ADs)</td>
<td>Designed sound signals or whole soundscapes, which have assigned meaning and convey information.</td>
</tr>
<tr>
<td>Ubiquitous Computing</td>
<td>A field of human-computer interaction that involves the embedding of computing into a physical environment</td>
</tr>
<tr>
<td>Ambient Intelligence (AmI)</td>
<td>An artificial intelligence system that is embedded in space and persistent in time and takes input and provides output continuously</td>
</tr>
<tr>
<td>Embodied Interaction</td>
<td>A term that I use here with a specific meaning coming from the field of interaction design, and describing types of technologies and human interactions that utilize physical movement and full-bodied being in space in a fundamental way.</td>
</tr>
<tr>
<td>Responsive Environments</td>
<td>A space that is technology-augmented in such a way that it responds to human action with different modes of display and feedback</td>
</tr>
</tbody>
</table>
CHAPTER 1: INTRODUCTION

Ubiquitous computing has in recent years become commonplace in many everyday spaces and contexts of life. As an umbrella term, ubiquitous computing encompasses a number of application domains including handheld portable and networked devices, mobile phones, art and museum installations, responsive environments, ambient intelligent spaces and other contexts of embedded technology. One operational definition of ubiquitous computing that will be taken in this thesis describes any physical space where computing is embedded in an environment naturally inhabited by people; artificial intelligence reasons about human activity; a set of display systems also embedded in the environment provide feedback so as to communicate with users; and people interact with the ubiquitous computing space in an embodied, social manner.

While certain aspects of ubiquitous computing technologies such as advances in sensing and motion tracking, networking, artificial intelligence and video displays, have been given deserving attention from industry and academic research, developments and research into auditory displays – sound feedback – for ubiquitous computing contexts, have lagged behind. Sound feedback is vital in situations where interaction is physically engaging, contextual and embodied, as it is in ubiquitous computing contexts. Sound is important there precisely because such environments are bridging the gap between technology with all its historical baggage of digital soundscapes, and the natural sonic environment where sound is a rich, influential modality of our interaction with the world. As a hybrid entity comprised of technical and natural elements, ubiquitous computing clearly requires a system of sound feedback that is functional within a computational framework, yet ecologically nested within a physical and social setting.

Research Problem Area

So how then is sound design approached from the standpoint of ubiquitous computing technologies? The reality is, as a domain within the fields of human-computer interaction and interaction design that are themselves relatively new, sound design
principles for ubiquitous computing environments are still at their early stages of formalization and methodical investigation. Sound design standards from computer-mediated communication and scientific auditory displays have gradually made their way into ubiquitous computing applications for contexts of work, training and productivity. Standards of media production and digital sound composition have, on the other hand, influenced art and museum audio-augmented installation design. However, both design paradigms are insufficient for addressing the complex requirements of ubiquitous computing as a distinctly separate application domain, characterized by embodied interaction, sociality and a responsive environment. Ubiquitous computing environments persist in time and space, and thus have to be designed accordingly to be part of their specific local ecology. Such environments require sound to take a more active, communicative role within a system, where it reliably and persistently conveys information, and dynamically reflects changes that occur in the interaction between people and environment (see Figure 1.1). As embodied, multi-modal interactive technologies ubiquitous computing spaces must encompass perceptual, cognitive, usability-related, cultural and ecological design considerations.

Embodied interaction employs perception, cognition and actions that are qualitatively different from traditional print and digital applications (Dourish, 2001). Contextual, physically involved activities require the support of an equally dynamic, cohesive auditory display system. As a growing domain of embodied interaction, ubiquitous computing environments offer different affordances for perception, cognition and engagement with sonic feedback than do other technological applications. To borrow Gibson's conceptualization of affordances, they are not simply physical properties of objects (in this case - sonic vibrations in the air); rather, they are latent characteristics of the environment relative to the actor/listener (Gibson, 1966) and encompassing perceptual, physical, cognitive, embodied and socio-cultural characteristics. What is required then is a new investigation into all of these areas, reflecting a new framework of auditory display design specifically for ubiquitous computing contexts. Further, such investigation must necessarily be conducted using a methodology sensitive to the contextual nature of this type of auditory feedback.
Research Questions

The overarching goal of this research is to explore and build a theoretically informed and empirically grounded model for using sound as an informative and contextually sensitive feedback mechanism in environments of embodied interaction such as ubiquitous computing. In such exploration, it is important to identify not only the relevant aspects of a novel framework for sound design, but also to articulate a justification for a methodological approach that would be effective in tackling situated design issues. This thesis attempts to propose, explore and refine a novel framework for auditory display design for ubiquitous computing. It will do so by making a theoretical argument for it; then fleshing it out using the experience gained from interpreting a set of multiple case studies; and finally formulating some preliminary design guidelines. The specific research questions I intend to explore are:

1. What are the requirements for a framework for auditory display design for applications of ubiquitous computing, which involves embodied, social interaction in a physical space, and effectively communicates relevant information?
2. What are effective methods of examining and improving such a framework in a manner following contextual inquiry and situated design?
3. What can be learned from design-based research methods, specifically participatory workshops, about designing sound feedback for ubiquitous computing environments?

The Role of Sound in the Environment

In the rich audible world there are a multitude of sounds at any given time that are being generated from different directions and distances, with different qualities and amplitudes. Our ears, bodies and minds are extremely good at interpreting these signals – paying attention to the important ones when needed; ignoring the insignificant ones; hearing where things are, what state they are in; recognizing familiar sounds with ease and sometimes even determining intricate detail about objects and people based on sound alone. Both in the foreground and in the background of our attention sounds constantly provide meaning to us.

To make matters more complex, alongside our natural acoustic environment, where sounds are produced by real physical objects and beings, the development of
electroacoustic reproduction and subsequently media has brought with it a kind of reconstructing of reality through bits and pieces of sound, text and image. We have gradually become accustomed to experiencing fully designed soundscapes in the context of cinema, radio, television and games, where sounds comprising these synthetic sonic environments have become objects, abstracted from their counterparts in the natural acoustic realm. Some of these sound objects represent ‘real’ sounds, while other, more abstract ‘sound effects’ have gained their meaning over time via consistent use in the language of media. While the soundscapes design in these domains is of great importance in achieving a rich sensorial affect, the primary function of these media experiences has traditionally been aesthetic, immersing us in a world of suspended disbelief and hyperrealism. Sound immersion has also developed as an integral part of commercial environments such as malls, elevators, office buildings, hotels and dentists’ offices, in the form of Muzak – an ambient, mood-setting continuous stream of compressed music tracks. Rather than building meaning or relaying information, this form of soundscape manipulation has instead consistently fostered distracted and background listening habits, creating a normalized music ambience in semi-public spaces.

In addition to media soundscapes and background music environments, both the acoustic soundscape and our places of work have over time become inhabited by an increasing number of mechanical and electronic sounds. With technological advances in computing, especially mobile and other portable devices, there is a whole new multitude of auditory signals being introduced in our everyday life. Many of these sounds are short tonal sequences, in either flat synthetic or richer MIDI formats. Unlike media soundscapes, they do not pretend to reconstruct reality; rather, they require being heard as intrusions because they are designed to attract our attention, inform us of things. These ‘auditory displays’ – sound-based representations of information – are the basis of most sonic feedback in computer-mediated communication, and as an extension, in ubiquitous computing applications currently.

The meaning of designed sounds is highly important within the systems of information in which they function. Yet, all of the sounds’ meanings are completely learned, and as such – arbitrary and interchangeable. An electronic device does not have a ‘sound’ intrinsically, like a dog has a bark, and a person has a voice. An electronic device could theoretically reproduce any sound. So, if the meaning of
designed sound signals is highly important (that is, it is important that they are interpreted correctly, and sometimes with speed), and any sound could typically be reproduced electronically – the question is, which sounds are better, for which purposes, when, and why; how do they fit in their intended or unintended environments; and what methods can we employ to best explore these issues? While we have already become relatively good at accepting and interpreting arbitrary electronic sound signals both in our everyday lives and in our places of work, this habituation does not prove that those sound designs are good, appropriate or efficient. Because, as mentioned above, our cognitive-perceptual sensitivities to sound are very well suited to our natural habitat, it is this thesis’s proposition that any framework that speaks to the design of auditory displays within contexts of ubiquitous computing must learn from studying the natural sound environment.

Figure 1.1: Areas of significance that characterize ubiquitous computing technologies and environments
Ubiquitous computing technologies combine elements from the ambient nature of background music environments, the sensory-rich craft of media soundscapes, and the highly informative functionality of electronic signals and computer sounds. Ubiquitous computing technologies are rooted in space and persistent in time and as such form novel, hybrid, everyday surrogate environments in their own right. At the same time they strive to convey different types of information, so unlike Muzak (which is a persistent ‘atmospheric’ auditory display primarily designed to be in the background and serve an aesthetic and commercial function) they may have to fluidly manage different types of listening attentions from their users. In order to attempt to construct a framework for designing sonic feedback within ubiquitous computing contexts, we could first look at two already available paradigms of research and design inquiry – acoustic communication, and auditory display design.

**Background of Existing Frameworks for Sound Design**

Focused on studying designed sounds in their intended environments of work, science and training, the field of auditory displays attempts to build onto psychophysics research of auditory perception and optimize the conceptual mapping between sound and its meaning within an informational context. Some of the priorities in this domain are speed of response on the part of the human user; greater clarity of large sets of data or information; minimizing the fatigue and annoyance factor in the auditory display; identifying a particular numerical value from a sound; identifying a specific message and applying a specific action based on an auditory display; and many others. If the field of auditory displays were to be categorized by specific goals, presumptions and associated methods of investigation, it would be segregated into:

1. Alerts/Alarms (anything form medical equipment to the military, highly based in psychoacoustics and psychophysical research)
2. Earcon/Auditory icons research (desktop-based and increasingly mobile/portable computing domains – based on HCI and mixed psychology methods)
3. Sonification and Auditory Graphs (science and research monitoring, expanding into HCI, based on HCI and investigated through mixed psychology methods)
4. Virtual Audio (heavily reliant on computationally modelling spatial segregation of sounds, for training applications, based on quantitative psychology metrics, engineering, psychophysics)
There are many other domains that use designed sound and place importance on it – art and museum installations, film and games, mass media. For example, art and museum installations typically use an artistic, compositional approach to design, rather than a usability, perception-based model that is propagated by the auditory display paradigm. This is not to say that the design of sound in art installations and as part of museum and gallery exhibits is any less effective in terms of experience, or sophisticated in terms of design. In fact, sound in such contexts of use is often designed (or crafted) with a much greater understanding of its role within the physical and spatial configuration of the environment, and its role as a social entity within the realm of shared and individual human experience. Quite often, however, the design process in this domain is not made explicit, making it hard to derive guidelines and formalize.

Another powerful paradigm of understanding sound could be found in the social sciences. Acoustic communication and acoustic ecology are only two incarnations of a larger focus on sound and listening that is rooted in socio-cultural and historic perspectives of listening, sound-making, music production, oral cultures, architecture and acoustic space. Researchers coming from a variety of home disciplines, including artists, media producers, sociologists, natural and social scientists, have approached the study of sound and its meanings, by attempting to look back to the natural environment and both categorize sounds according to their role and meaning within a community, as well as discover patterns in how people actually listen to, and experience sounds. Approaches in acoustic ecology and acoustic communication are based largely on ethnographic field research, including contextual observations, interviews, as well as recordings and personal reflections. Other methods, seeking to examine contemporary media soundscapes also rely on media analysis and sound measurements. Some of these explorations are motivated by a commitment to the ecology of acoustic environments and to fighting the growing noise pollution problem. As my main starting point within this paradigm, the acoustic communication model necessarily sees sound as an equal participant in the socio-cultural and communicational exchange within a community of people, inside a particular geographic space. Ultimately, the framework of acoustic communication relates to design in its proposition that only by understanding (and if possible, deconstructing) what sound means in people’s natural environments and how they experience it, can we begin to make informed decisions about designing sound in technological contexts. This knowledge would inform both individual auditory displays to be nested in pre-existing environments; as well as complete auditory
environments as part of multi-modal interactive systems and spaces, which is where ubiquitous computing resides.

**New Framework Justification**

Here's where the problem lies. The acoustic communication framework offers leading ideas about thinking of sound as an integral part of a communicational loop between sound, listener and environment. It also categorizes sounds in terms of keynotes, soundmarks, signals (among others), and it categorizes (Truax, 2001, p.21-27) different listening 'positions' in terms of attention they demand from a user (listening-in-search versus listening-in-readiness, analytical listening, media listening and background listening). Further, the aspect of acoustic ecology in this model speaks to the interplay of sounds within a sonic environment in terms of balance with regard to acoustic profiles, frequency spectra and masking (an auditory occlusion of a sound by another), as well as balance of sounds belonging to an acoustic community. While all these propositions are useful and specific, they are more a lens through which to see and organize design, rather than concrete guidelines that could structurally help with designing sound for rich, physical ubiquitous computing spaces. Acoustic communication sees meaning-creation through sound in terms of personal experience, memory and association; as well as membership within a particular acoustic community. Meaning is mediated between sound and listener through the soundscape. Yet, how could sound convey specific information? A more structured set of guidelines needs to be generated here before the acoustic communication perspective could become instrumental to design. Relaying information through sound is indeed the domain of auditory display design, which relies (as mentioned above) on disciplines such as cognitive science, psychology and engineering, among others. Frameworks and advances in auditory display design tend to be quite concrete and generated by small-scale, narrowly focused experiments of specific approaches for conceptual mappings between data/information and sound. They are less able to consider holistic implications in design, and when they do, those are directly tied to the cognitive activity at hand. Ubiquitous computing includes a rich array of environments; takes place in physical spaces, in social contexts; and encompasses a wide variety of interactive activities, from work, telecommunication, training and learning, to play and entertainment.
I propose that combining perspectives from the acoustic communication framework as well as design directions from the auditory display (AD) frameworks associated with different strands of AD research (earcons, sonification, virtual audio, etc.), could provide a good starting ground for building a framework of auditory display design for interactive, ubiquitous computing contexts; and help establish sets of requirements, characteristics, guidelines, propositions and solutions that speak to the design of sound in such contexts. What I hope to do in this thesis is create a conceptual argument for this framework, then explore it by actually designing an auditory display system for a ubiquitous computing prototype, and on the basis of that experience, build a set of experimental studies to investigate issues of sound design further. In its entirety, this research undertaking aims to help refine and build theory around the requirements, issues and core concepts of a new framework of auditory display design for ubiquitous computing. The research is organized in two phases, with a theoretical discussion following each phase. Sources of evidence will include participant observations, video capture of embodied interactions, user-generated artifacts, co-investigator interviews and group discussions. The first phase represents the design process of an auditory display system for a ubiquitous computing prototype named socio-ec(h)o. In keeping with the design-oriented nature of the research, this phase includes a set of participatory design workshops exploring the potential of sound display for ubiquitous computing spaces. They are followed by a description of the complete prototype, in addition to a report of some preliminary user testing. This phase is completed by a theoretical discussion that helps elucidate a set of propositions and research questions regarding the new framework of auditory display that are to be addressed in the next phase. The second phase deals with designing a set of naturalistic experiments as exploratory studies that tackle more specific situated design questions regarding auditory display. Finally, the results from these exploratory studies are analysed in order to redefine and flesh out the new framework for auditory display design for ubiquitous computing environments.

**Thesis Outline**

I will begin by providing a literature review of the two frameworks for soundscape design, as well as a historical background on listening practices in the natural and technological environments. Then I will propose a hybrid framework for auditory display design in ubiquitous computing contexts, informed by the two models introduced. The
The next chapter will describe the methodological approach taken to investigating the research questions set forth here, as well as introduce design methods as an instrumental basis of the experimental studies dealing with auditory display issues. Further, I will describe the two phases of research, and using the experience gained from them, will refine the new auditory display design framework. I will follow with a discussion of its saliency as an approach – conceptual, practical and methodological - to investigating contextual sound design issues for ubiquitous computing environments, and make an argument for why it may be more effective than the two frameworks initially presented. I will conclude by recognizing and discussing the limitations of the research inquiry and the new framework, and will point to future work, as well as future applications of auditory displays in ubiquitous computing. Figure 1.2 below shows the conceptual structure that my inquiry will consider.

![Figure 1.2: Structure of the research inquiry of the thesis, including relevant areas of consideration and/or investigation.](image-url)
CHAPTER 2: LITERATURE REVIEW

“Everywhere we are, we are surrounded by noise. When we ignore it, it annoys us. When we listen to it, we find it fascinating.” (John Cage, 1973)

“If a tree falls in the middle of the forest and nobody is around, does it make a sound?” Common saying

There is sound, and there is listening. It could be said that listening is the social experience of hearing. Hearing then, could be understood to refer to our psychophysical abilities to detect sonic vibrations and experience them as sound. There is no argument that any study about designing sound has to begin by understanding how we hear. Less talked about, but even more important for design I believe is the understanding of what it is to listen. Listening is subjective, yet listening is also universal, a product of its time, society and culture. In this chapter I will try to provide a historical background of the general shifts in the activity of listening, which, as suggested in the acoustic communication framework, has changed dramatically through the last century. In tracing these changes I’ll take specific notice of how sound and music have come to occupy different shared physical environments, as I believe the lessons that could be learned there are especially important for the context of ubiquitous computing, seeing that it strives to put technology back into the environment, rather than take the environment out of computing. Following this theoretical background, I hope to then shed some light on how the two different frameworks for conceptualizing sound – acoustic communication and the domain of auditory displays, understand, theorize and research listening. From there, I will connect the two perspectives into a preliminary framework for designing auditory displays for ubiquitous computing environments.

Historical Background of Listening

Interactive technologies, especially ones that involve sound feedback range from personal computing and communication devices to complex virtual simulations and ubiquitous physical interactive environments. In each case, sound design plays an important role in generating a sense of place, movement and interaction, providing aesthetic affect, and giving a tangible cultural context for the activity at hand. Yet such
recent developments in auditory display (AD) systems and immersive audio have not occurred in isolation from other social, cultural and technological changes over the last century. Advancements in sound reproduction, electroacoustic sound and multimedia are only one part of a larger equation. Scientific developments in acoustics, psychoacoustics and sound engineering, together with a greater understanding of audition and perception, comprise another part. The ensuing socio-cultural changes in the patterns of listening, modes of perception, and interaction with our technologically-extended world, all comprise the foundation of the development, popularity and ubiquity of modern auditory displays. These interlinked processes not only affect the practice of AD design, but also need to be considered within its very framework and context, as they embody its implicit history.

Every historical shift in the sonic environment helps contextualize and determine the directions in which subsequent acoustic environments are heading. In this section, I discuss the critical transition from sound-as-environment to music-as-environment, where music, especially background music, becomes the ‘normal’ ambient environment, and new listening patterns emerge. From there follows the transition to design-as-environment, where fully designed systems of auditory displays take on the functionality and experiential modalities of all previous sound environments.

**Listening before Electroacoustic Reproduction**

Walter Ong’s extensive work on orality and communication points to many of the characteristics of listening prior to print literacy and the recording and stockpiling of speech/sound (Ong, 1982). The ephemeral quality of sound - the fact that nothing could be repeated or reproduced in the exact same way again - contributed to an active, feedback-oriented ‘everyday listening’ in oral societies (Truax, 2001, p.23), where sounds mediate the communication between people and the environment that surrounds them. Albert Bregman, James Ballas and other researchers of sound, have focused precisely on this ‘everyday listening’ of environmental sounds, and their work points to an understanding of auditory perception in terms of sound events, emphasizing the functional ecology of this interaction (Bregman, 1994, p.xix; Ballas, 1993). This balance is reinforced by another important condition of the period before electricity and amplification – that sounds can only be as ‘loud’ as it is physically possible for the sound
source to transmit energy, so problems of masking and obstructed communication are minimized (Truax, 2001).

In terms of musical expression and soundmaking prior to electroacoustic sound, Jacques Attali gives us his account of the different relationships that music mediates between people and society before and after the mass reproduction of sound. In the stage of what he calls 'sacrifice,' music is simply a means of "channelling society's violence" (Attali, 1985, p.25). In other words, music reaffirms order and counteracts the chaos that is noise. In this regard, music, much like sound in the acoustic communication model, is a mediator of social relationships in the context of the environment and community. In the period of 'representation,' starting in the 17th century in Europe, music begins to embody more than just sounds. It becomes a symbol of power, status, and cultural expression (Attali, 1985, p.83). It starts exhibiting use/exchange value as an object, and with the emergence of virtuosity, concerts and opera performances for a paying audience, it becomes situated in make-believe surrogate commercial environments removed from the street, where music used to be. In the era of individualized representation, as Attali argues, “music can no longer affirm that society is possible. It repeats the memory of another society” (p.120). This new status of music as an object, and not an everyday activity, helps shape a new type of listening – a more attentive, analytical listening, focused on the individual elements of sound, rather than on a holistic perception. In addition, scientific developments in subsequent eras – specifically, the reverberation formula developed by Wallace Sabine in the early 1900s, contribute to a newly created interest in architectural acoustics for concert halls and thus, to a new attitude towards listening to music (Thompson, 2004, p.81). Once again, this 'expert' preoccupation with the quality of musical sound in built spaces changes the dynamics of the listening experience. As described in detail by Emily Thompson, the control of reverberation in the concert hall, and its subsequent removal in the recording studio, in turn create a different kind of surrogate environment, one that excludes traces of the space itself and promotes a 'clean' dry sound for an anonymous mass audience.

Listening in the Electroacoustic Era

Attali’s concept of mass music, based on Adorno and Horkheimer’s texts on mass culture, proclaims the destruction of cultural meaning. “The absence of meaning is
the necessary condition for the legitimacy of a technocracy's power," he announces grimly (Attali, 1985, p.112), describing the era of "repetition," the stockpiling of music and the emergence of copyright - music as intellectual property. In addition, the era of mass music turns the idea of silence into a "death in the heart of life" (p.120). Schafer (1977) also comments on the vilification of silence in the contemporary soundscape as a 'dead space' and a negative force in society. Repetition symbolizes uniformity, compliance and 'programmed events', as Ursula Franklin (2000, p.14) argues in "Silence and the Notion of the Commons". She describes the disappearance of silence as an "enabling environment" and its replacement by "the silencing that comes when there is the megaphone, the boom box, the PA system" (Franklin, 2000, p.15).

Mass reproduction of music exists long before electroacoustic sound, with opera and instrumental concerts; however, electroacoustic technology makes possible the exact replication of any sound and its independent recollection thereafter. This schizophrenic sound (Schafer, 1977) is disembodied from its source, context and time of occurrence, and becomes an abstract 'aural object' of representation (Metz, 1985, p.158). Sound comes to symbolize power, control, use and exchange value, and private ownership (Attali, 1985, p.101). Electroacoustic technology also puts sound under unprecedented scrutiny as well as aesthetic appreciation, fostering analytical listening (Truax, 2001, p.163). Alongside the move towards clean, non-reverberant sound in architectural acoustics (Thompson, 2004), digital sound further removes any perceptual reference to space by eliminating acoustic colouration. The transition facilitates the easy transfer of sound/music objects into a variety of different surrogate environments – restaurants, concert halls, stadiums, malls, schools, and other shared spaces. This fluidity in turn results in the acceptance of music-as-environment in both the private and public spheres of life, and leads to the blurring of the lines between the two. Background music, defined by Satie as music "like furniture," (Cage, 1973, p.76) not only becomes part of the environment, but is the environment. It builds invisible surrogate relationships between people sparing them from obligatory interaction and "filling up heavy silences between friends" (p.76).

Similar to the Telharmonium of 1906, which provides 'atmosphere music' in restaurants using the telephone line (Weidenaar, 1995), the Muzak Corporation was created in 1922 to first provide music over the telephone (Attali, 1985, p111), and later to 'program' music in various public spaces by use of market research. Muzak is perhaps
the first materialization of complex soundscape design. Hildegard Westerkamp’s research in background music solidifies the views put forward by Attali and others – that music in the era of schizophrenic mass reproduction becomes a "soundtrack for consumerism" (Westerkamp, 1990, p.227). Sound “becomes associated in our memories with environments and products. In essence it becomes the ‘ambiance’ of the media environment” (Truax, 1992, p.374). However, the new fluidity and flexibility of sound freed of its source does not result in endless diversity of spaces and sounds, but in the emergence of archetypal surrogate environments. These most often commercial spaces come with pre-packaged sound quality – compressed, narrowband dynamic range and dry digital sound, as well as content standards – slow light rock for the department store aisles and fast pop at the restaurants. These emerging sound environments foster "passive listening" (Westerkamp, 1990, p. 227) and superficial disengagement from the social environment. As Westerkamp argues, the phenomenon of background music also results in the inevitable silencing of spontaneous human soundmaking, and with it our active interaction with place and time ([p.227).

Table 2.1: Key concepts from the music-as-environment paradigm.

<table>
<thead>
<tr>
<th>Key concepts from the Music-As-Environment paradigm:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Music as an Object of Representation</td>
<td>Music has use and exchange value, and it symbolizes power, status, etc.</td>
</tr>
<tr>
<td>Background Music</td>
<td>Ambient music programming such as Muzak, radio, TV flow, etc.</td>
</tr>
<tr>
<td>Schizophrenic Sound</td>
<td>Electroacoustic sound that is separated from its source and played at another space/time</td>
</tr>
<tr>
<td>Passive Listening</td>
<td>Inattentive, background listening, lack of active interaction with sound</td>
</tr>
<tr>
<td>Surrogate Environment</td>
<td>A re-created environment, into which a chosen sound or an auditory display system is inserted</td>
</tr>
</tbody>
</table>

Listening in the Age of Media

Archetypal sonic spaces such as concert halls, malls, and restaurants, initially defined as public environments, gradually move into people’s private lives with the emergence of the phonograph, the telephone, and especially the radio. The private space becomes another sound-programmed surrogate environment, usually designed for passive, background listening. This shift is not unrelated to the trend of development
of mass media and mass production/consumption. In a society where urban alienation and post-industrial capitalism shatter traditional forms of community and social interaction, media becomes a surrogate social milieu for the 'masses,' and binds people together in imaginary relationships. Radio is a particularly important cornerstone in mass media development in that it plants the roots of contemporary media language, including temporal flow, structural density, foregrounding-backgrounding cues, advertising language, audience as product and media’s overall relation to consumerism. When wireless radio technology first emerged it was praised as a web of interconnection, universal communication and utopian democracy – “house of our dreams” as named by Gaston Bachelard (Dyson, 1996, p.78).

Mendelsohn’s work in the 1960s focuses specifically on the characteristics of radio programming, conceptualising the cultural changes ensuing from its introduction in the private and public social worlds. Radio literally ‘educates the consumer’ and creates a new language of media consumption, aural sensitivity, listening and cognition. One of the major functions that radio fulfils is “bracketing the day” through program flow – that is, providing structure to daily activities by setting predictable patterns – news at noon, followed by music, announcer, advertising, then more news, etc (Mendelsohn, 1964, p.242). In addition, as with background music in malls and restaurants, radio ‘sets the mood’ for the day and “lubricates” social relations (p.243) in alienated urban settings. As such, radio functions as a shared aural environment and implied shared physical and mental space. Similar to background music, radio settles as a predictable ‘accompaniment media’ to daily life, first confined to the home, and soon invading the streets and offices. This transgression of public-private boundaries leads to portable, personalized sound accompaniment in the age of the ‘Sony Walkman,’ as characterized by cultural theorist Paul du Gay (1999) and later in Michael Bull’s work (Bull, 2003). Accompaniment media becomes a standard sonic companion, and listening habits adjust accordingly, as we become the composers of our public and private electroacoustic experience.

Combined with structured programming, radio changes the listening experience from background listening to media listening. An important aspect of media listening, also characteristic of TV flow as an extension of radio flow, is the pattern of amplitude flow. That is, the majority of normal programming is broadcast in narrow bandwidth at relatively constant amplitude, designed to blend with the background of daily activities,
while advertising and other special elements are broadcast with a broader spectrum and
greater dynamic range, demanding foreground attention from the listener (Truax, 2001,
p.187).

In effect, media sound does not merely foster a dependent kind of listening, but it
tells us how to listen. It trains us to increase or decrease our auditory attention by use of
carefully crafted cues, until they become second nature. These gestalts of auditory
perception then seamlessly integrate in cinema sound, carrying the promise of total
immersion, suspension of disbelief, and realistic experience design. Ultimately, as our
environment changes and we become more saturated in media flow we start to
experience sonic phenomena such as radio, TV, portable audio and Muzak as
environmental sound. This dramatically changes our relationship with the acoustic
environment, as electroacoustic and acoustic sounds become intertwined and blend into
each other, rendering the modes of listening that we use for these two sound milieus
interchangeable. As a result, we begin relying less on active, engaged, information-
processing kind of listening, and more on habitual background and media listening in all
of our surroundings. It is perhaps the emergence of ubiquitous computing and interactive
sound design that may finally shift the attention back to a more holistic perception of
sound and active engagement with it, relying on a more communicational ‘everyday
listening’ mode of experience.

Table 2.2: Key Concepts from the media-as-environment paradigm.

<table>
<thead>
<tr>
<th>Key Concepts from the Media-As-Environment paradigm:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structured Flow</strong></td>
</tr>
<tr>
<td><strong>Media Listening</strong></td>
</tr>
<tr>
<td><strong>Accompaniment Media</strong></td>
</tr>
</tbody>
</table>
Media, Technology and Space

Background music is the earliest fully designed and controlled surrogate sonic environment, which, through interaction with other phenomena in sound technologies and media, continues to shape the contemporary listening experience. Other modern examples of immersive audio are film sound, game sound, interactive auditory display installations for art and museum contexts, and audio-augmented ubiquitous computing environments. These interactive soundscapes can be either virtual – where sound is entirely engineered and usually delivered via personal headphones to the exclusion of other acoustic sounds; or physical – realized through computer programming but embedded in a shared physical and cultural space. These alternate sound environments are made possible by advances in digital audio, virtual reality, the Internet, networking, sensing technologies and artificial intelligence.

Sound in Cinema

In its basic definition, an immersive sound space is a carefully and intuitively designed surrogate environment that creates a more full-bodied experience involving the senses to a fuller capacity than traditional media. Sound in cinema is an important cornerstone in soundscape design because it simultaneously builds upon and draws on media and background sound, as well as influences and feeds into more sophisticated virtual and multimedia environments. Using Barry Truax's acoustic communication model for sound design, an immersive environment contains three major elements: speech, music and soundscape (2001, p.50). In cinema, these break down to vocal material, composed musical score and sound effects database, including speaker diffusion patterns. Hearing these sounds in a darkened auditorium and in combination with larger-than-life visuals promotes the experience of immersion. However, it is important to note that the idea of immersion as virtual or 'augmented' reality is not only an extension of earlier electroacoustic technologies but also capitalizes on media listening and decoding. As Truax comments, "once background listening becomes a habit, it is ready for exploitation by the media" (2001, p.27). The cinema soundscape is reduced to "easily recognizable sound objects" (Truax, 2001, p.83) and even though realism is increased through high quality audio, "this realism is not born of the 'real' [but is] constructed through other media" (Dyson, 1996, p.86) - radio and television. Sound, is thus hyper-real rather than real. This view is also echoed by Christian Metz in his
Aural Objects essay, where he describes film sound as being based in sound objects – individual representations of real life sounds, and elements from a larger conceptual sonic media language (Metz, 1985, p.159).

Listening and Virtual Audio

Virtual Reality (VR) is a special type of surrogate environment based in digital audio, 3-D animation, rich graphics and audio coding. In VR spaces, sound is a virtual ‘aural object’ of representation, carefully designed to elicit recognition, action and response, or create an emotional mood. As such, virtual audio builds on the patterns of listening and recognition created by previously existing media and aural phenomena, yet it also results in new sensitivities and cognitive modes of interaction. Following from McLuhan’s conceptualization of the electronic world as oral (Schafer, 1993, p.21), virtual audio creates a new kind of aurality in technological environments. Using this virtual “aural medium,” one can “enter a space of no space” and be immersed in a compelling, increasingly tangible experience (Dyson, 1996, p.73).

There are two predecessors to virtual audio that shape and influence its authenticity and perceptual-listening framework. One is the microphone, which, similar to radio, provides surrogate intimacy and “a spiritual and atmospheric nearness of broadcaster and listener” (Dyson, 1996, p.77). The microphone eradicates implied physical distance in the demolition of metaphorical distance between real and representation. Another predecessor of virtual reality is headphone technology. The possibility of bringing the outer world into the inner world and creating a personalized surrogate sound environment has many cultural critics fascinated. Paul du Gay looks specifically at the phenomenon of the Walkman and terms the experience of listening to portable audio a “soundtrack to life” (du Gay, 1997, p.92). Headphones literally immerse us in a designed soundtrack of imaginary space, and mediate relationships between the listener and that space, establishing a new VR phenomenology. While this relationship is rich and enjoyable, it is individualistic and exclusionary, and based on commercial music culture.

Audio technology creates a ‘sound field’ designed to appear as if it occurred naturally in the environment, however, the very existence of digital media has implications for the role of virtual audio. Sound is divorced completely from a physical source, and is controlled entirely by code. It is technically abstracted – sound attributes,
distances and behaviours are de-constructed, mapped and coded, and then reconstructed to fit a virtual space. This way, algorithms represent sound and sound behaviours, rather than expressing its physical characteristics, similar to Attali’s argument that musical code is a “language without meaning” (1985, p.25). Ultimately, the abstract nature of virtual audio and the limitations and opportunities brought by the technology of headphones and program code create a “new space of perception and embodiment” (Dyson, 1996, p.79) that is only now beginning to be explored.

Interactive - Adaptive Audio in Games

As suggested in the acoustic communication and cultural studies perspective, there is a strong connection between sound design in modern commercial computer games and the standard cinematic sound model of speech – music – soundscape, where speech/vocal information is the focal point of the listening experience, music provides the affective characteristic of the scene, and soundscapes consist of carefully put together sound effects, which provide relevant contextual information (Truax, 2001, p.50). Music scores for games arrive from their counterparts in cinema, where sound serves a primarily aesthetic and affective purpose by providing mood, atmosphere and background. Sound effects in cinema provide sonic snippets to aurally enhance the representation of physical objects, events and actions (Metz, 1985). Metz terms them “aural objects,” and points out how they solidify sometimes arbitrary associations in the audience’s aural memory between film sounds and real sound situations (such as fist-fight sounds in movies, which have no counterpart in reality). Sound effects in games serve a derivative function of event-triggered audio icons that provide minimal contextual and/or environmental information about the imagined world. Sometimes aural objects from cinema directly transfer to game sound (again, fist-fight sounds are a good example). While information-rich, they still have a primarily supportive role, reinforcing visual representations and/or game shifts and events, rather than serving an essential function to the game. The listening that develops with this type of sound environments is partly derivative from media, background and analytical listening, but partly influenced by the inherent interactive nature of gaming. This positioning brings more focus on sound as signal, as an alert to something happening in the game, thus positioning listening-in-readiness for the next chime. It also creates a kind of dependency on sound feedback as confirmation of actions, an immediate gratification reassuring the user that they are doing the right thing and the system is working.
Contemporary approaches to sonic feedback in games come closer to designing it in a more holistic and ecological way – where sound mediates the relationship between listener/player and sound environment/game environment. They do so by providing layered contextual sonic information, which attempts to create a realistic ‘sound and feel’ of the game world (see Table 2.3). Environmental sounds, as well as material/object impact sounds must work together with the visual display to give the player a sense of “place” and personalized control – for example, if their avatar is running on grass, they should hear the sound of footsteps on a soft grassy surface – a direct auditory feedback that confirms their actions within the game. Yet in order for sounds to be arranged in a truly holistic and ecological way, then issues of masking, timbral qualities and acoustic balance must also be considered. For example, many contemporary commercial games contain loud music tracks that often overpower and mask other relevant sonic information about the environment, thus causing an imbalance in the soundscape of the game. This contemporary sound design model synthesizes dynamic soundscapes in real-time using large banks of sound samples, effects and filters in order to create soundscapes that are essentially unique to each player. Ultimately this approach renders vocal, musical and contextual elements of sound in real time resulting in a complex multi-layered informational soundscape. Yet, even this approach leaves sound outside the set of core mechanics of the game. The feedback, while rich, is still confirmatory and contextual in nature. A player can switch off the sound and still play the game.

Games are part of popular culture and as such their audio-visual components are experienced in much the same way as their counterparts in cinema, TV, popular music, malls, shops and advertising. The result of this blending of everyday perceptual experience is that listeners become accustomed to “backgrouding” most aural information they are bombarded with in daily life with the exception of ‘instant feedback’ beeps and alerts. We spend most of our lives trying to block out traffic noise, TV commercials, mall music, other people’s conversations. When using interactive digital media applications we have to actively reverse engineer sound back into the foreground of our attention. Understandably, the task is hard. In the breakdown of speech – music – soundscape even well-designed sonic components are often experienced only half-attentively by the player.
Table 2.3: Sonic components that provide contextual information in game soundscapes

<table>
<thead>
<tr>
<th>Types of Sonic Components in Game Soundscapes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acoustic Space</td>
</tr>
<tr>
<td>The acoustic properties of a specific visual environment (e.g. a cave should sound different than a forest)</td>
</tr>
<tr>
<td>Object/Character Interaction</td>
</tr>
<tr>
<td>When visual game objects come in contact with other objects or characters this should be reflected aurally as well (banging on wood should sound different than banging on metal)</td>
</tr>
<tr>
<td>Localization</td>
</tr>
<tr>
<td>Systems that utilize multi-channel spatialization formats could emulate discrete sound source locations by distributing different sounds in different channels</td>
</tr>
<tr>
<td>Sound Effects</td>
</tr>
<tr>
<td>Standard sound effects (sounds that don’t necessarily have a counterpart in reality, such as fist-fight sounds) are compiled in real-time</td>
</tr>
</tbody>
</table>

Interactive Auditory Displays in Ubiquitous Computing Environments

Sound feedback is vital in systems where interaction is physically engaging, contextual and embodied, as it is in the context of ubiquitous computing. Hands-free and eyes-free feedback is often especially important, and best achieved with a well-designed auditory display system. There, sound takes on a more prominent role in communicating and supporting human activity, compared to its role in traditional computer-mediated communication.

In traditional auditory display design, concerned with alerts and notifications, there are important considerations about the environment or context in which sounds are heard and perceived. Aside from psychoacoustic issues, there are challenges with masking in busy sound environments. With physical interaction systems, specifically responsive environments, where almost every parameter of the experience is technology-driven, mediated and controlled, the auditory display system is the environment. Not only is the sound, and hence – the experience – designed in minute detail, but also users most often experience the space via headphones, further minimizing acoustic sounds and colourations.

With the development of 3-D audio the restrictions of stereophony are finally transgressed. By applying digital filters and multidimensional streaming techniques, sound can effectively be reproduced anywhere in virtual space simulating a tangible
sound source (Brugnart, 2002, p.93-4). In addition, multi-channel speaker systems, commercially known as ‘surround sound,’ provide a shared immersion experience, where, ideally, each speaker functions as a sound source on its own. In effect, AD design has finally shifted back to a more ‘acoustic’ model of aurality, sound distribution and perception. As McLuhan argues, quoted by Schafer, “the electric world” is aural, and it moves us back to an acoustic space of preliterate cultures (Schafer, 1993, p.29). In this type of approach, sound has a strong connection to the environment and mediates a relationship between the listener and space, reinforcing acoustic communication.

Ubiquitous computing presents a special case to Interactive sound design in that it is typically nested in a real physical environment, rather than constructed as a ready-made technology to immerse the user in. Thus, feedback would presumably be delivered via speakers in a shared aural space, rather than by individual headphones. As such, it attempts to use our senses in the more physical, embodied act of listening, engagement and interaction, rather than in the energy-transfer model of hearing and tone perception. It extends the existing soundscape within the surrogate sound environment by drawing attention to itself, rather than passively ‘filling up’ a space to create a mood for consumption. It does so by asking us to relate to, orient ourselves through sound, locate other objects within a space and give and receive feedback about actions and events. Interactive audio also provides variety and colouration by creating adaptive artificial intelligence for sound behaviour and allowing active modulation and interaction with its listeners.

The next section will provide some current examples of ubiquitous computing environments and technologies using sound and from this basis begin to build theory about a new framework of auditory display design for contexts of ubiquitous computing.
Another step before attempting to synthesize an initial conceptualization of a new auditory display framework for ubiquitous computing, is surveying current practices, application domains and design approaches to creating meaningful sound feedback for existing ubiquitous computing spaces. As mentioned earlier, there are many instances and various levels of sophistication of sound designs for different ubiquitous computing applications. Audio installations in the art and museum domains are one such setting. In them, sound is essential to the affective experience of the artwork or exhibit, and to the activity of interacting with it. The human-computer interaction field, both in its industry and academic incarnations, has also been attempting with more frequency to incorporate sound feedback and auditory displays in interactive technologies. Again, there are far too many case studies to include here, so instead I have chosen to describe a few representative examples. In addition, I deliberately picked projects that use sound with a distinct informational or otherwise system-essential agenda. What I hope to demonstrate with them is that while the potential of sound feedback as an integral, important part of ubiquitous computing technologies is clear, the level of formalization of design principles, the methodological discussion into situated design issues, and as a result – the prototyping of auditory displays, which tap into sound’s full potential is as yet under-developed.

Sound Design Examples in Ubiquitous Computing

As an example of an ubiquitous computing environment, one of the projects preceding and informing the current research exploration – ec(h)o - is a multi-modal ambient intelligent exhibit installation for the Museum of Nature in Ottawa, Canada. It is “an audio augmented reality interface utilizing spatialized soundscapes and a semantic Web approach to information” (Wakkary, Hatala, Newby, Evernden & Droumeva, 2004). In this sound-enhanced exhibit the visitor navigates through several layers of audio – ambient soundscapes, location-based soundmarks, hierarchical audio icons for selection and audio narratives relating to the artifacts at hand. This interactive museum installation
is an example of design-as-environment providing a multi-modal shared surrogate environment in a physical space. Individual motion-tracked headphones deliver the immersive sound experience, and the interaction takes place via an ergonomic colour-coded cube, connected to visual recognition software (Wakkary et al., 2004). The soundscape elements are modelled upon the acoustic ecology and acoustic communication models, where sound, environment and listener create a ‘feedback loop’ of interaction (Truax, 2001, p.23) - see Figure 3.1 for a floor plan schematic – small circles represent soundmarks and big circles represent keynotes.

The acoustic functionality of this system is defined by sound behaviour that is “adapted to its environment and understands beauty as a value expressed through people’s attitudes” (Truax, 2001, p.110). As shown in the diagram below, there are several thematic soundscapes, localized in a radius around a relevant exhibit and broadcast over FM frequencies. These soundscapes are modulated in real time by the visitor’s movements, fading in and out of each other. In addition, soundmarks, chosen specifically to reflect important aspects of a given set of artifacts are incorporated into the soundscapes. Another element to the audio experience is a number of recorded narratives, that can be recalled via three aural ‘prefaces’ appearing in the left, centre and right speakers of the headphones, utilizing sound localization techniques. All of these elements: speech, soundmarks and keynotes interact with each other creating an ‘acoustic space’ that characterizes the museum community.

Figure 3.1: The big circles signify areas of keynote soundscapes, while smaller disk shapes signify soundmarks representing the content of the exhibit area.
Another growing domain of application for auditory displays within the human-computer interaction and design field is technologies and interactive environments for the blind and visually impaired. These kinds of interactive environments must use alternative modalities to represent elements that are normally experienced through vision, thus the resulting environments are sensory-rich. They may use audio displays, tangibles and textures to represent space, feel, pictures, maps, characters, numbers, combinations, motion and navigation. For instance, a project called Monster Hospital attempts to sonify furry toy monsters for blind children by creating a synesthetic experience from different kinds of fabric being matched up with different sound timbre qualities (McEliggot, 2004, p.10). The result is a tactile sonic experience. Many other projects use sound displays and feedback in prototypes for blind children (McElligott, 2004; Parente & Bishop, 2004; Sanchez & Flores, 2004; Sharlin et al., 2004) in a more overtly informational way. AudioMath, specifically, uses sonic feedback to represent basic mathematics concepts to blind children (Sanchez & Flores, 2004). The AudioMath interface has virtual, tangible and spatial components. Children are asked to identify numbers and perform mathematical actions based on short-term auditory memory. While the interface is information-rich, the sound feedback is limited in its quality and array of representations, and does not guide children to the right answer, only confirms when they arrive at it.

The creators of BAT: the Blind Audio Tactile Mapping System (Parente & Bishop, 2004) have taken the auditory interface route further by creating a rich, narrative, audio interface for exploration of geographical maps. When a participant scrolls over the map, a dynamic system of sonification guides them through the map. Movement triggers the system to play local soundmarks and recorded environmental sounds in conjunction with abstract auditory icons that signify major cities, distances between locations and other contextual information. This sonic interface comes much closer to delivering a system of information and acoustic ecology by balancing elements in an intuitive and communicative way. However, since it is a browsing interface, it is passively experiential in nature and lacks directive feedback elements – such that could assist a user in making choices or solving problems.

Sound spatialization is also a popular technique used in the design of interfaces and tangibles especially for blind and visually impaired users. Spatialization is one of the only aural cues that can be used to spatially construct a world that is physical and can
be navigated meaningfully (McElligott, 2004; Sharlin et al., 2004). A prototype project for children called *World Aloud!* situates a child user within a space and provides localized audio feedback which guides and helps orient them in space through various sound effects, signals and diffusion techniques (McElligott, 2004, p.66). Another project, the *Tangible Pathfinder* uses a similar approach of localized cues to aid path finding navigation for the blind (Sharlin et al. 2004). A large number of tangible tools for children use sound for recording and playback of the child’s voice (McElligott, 2004, p. 69) as well.

Moving into situations of increasing embodied interaction, a project where sound is used as feedback for karate training is another example of an auditory display system for a ubiquitous computing system (Takahata, Sakane & Takebayashi, 2004). Continuous and gradually changing musical tones respond to minute wrist and body movements of karate students, giving them hands-free and eyes-free feedback on their Kata postures. This project uses simple musical tones designed with some consideration for their perceptual qualities (based on tone perception) and clarity of feedback, however, less attention is given to sound’s role in the context of the space, activity, and information ecology. An example of a more complete multi-sensory ubiquitous computing environment is MEDIATE – a fully controlled space designed for low-functioning autistic children to gradually experience and become comfortable with different types of sensory feedback (Péres et al. 2005). The system includes tactile, auditory and visual displays and a complex artificial intelligence engine, which learns and dynamically adapts the environment’s responses to the user’s behaviour.

**The Gap**

It is important to reiterate that there are too many examples of auditory displays in different interactive technologies, tangibles and ubiquitous computing environments to survey here. The selection I chose to include was specifically favouring systems where sound has a more essential, informative function that could potentially convey abstract information through non-speech sound, aside from the experiential meaning-creation that occurs naturally in sensory-rich settings, or the compositional, performative and aesthetic role that sound takes on within art installations. In all of the above-described cases, I was interested in how decisions about sound design were made and what general patterns emerged regarding the role of the auditory display within a system.
was looking for design conceptualizations of sound ecology, listening patterns, embodiment; explicit understanding of human auditory perception, evidence of formative design research into sound as feedback; and any formalization of observations into a usable framework for sound design for future use. There seemed to be a gap there, in that if the auditory display was designed for effective information retrieval, other contextual considerations often lacked; while soundscapes composed well and fit the context rarely aimed or achieved a more informative, epistemic function. Without delving into phenomenology and theories of knowing, I use the term “epistemic function” in its narrower definition borrowed from interaction design, as an action, a mental or physical transformation of perception, space, environment, artifact, state or physical manifestation, which results in understanding, meaning-building and/or information transfer. When it comes to sound, there hasn’t been as yet a systemic collection of concepts, principles and guidelines of auditory display design specifically for ubiquitous computing environments.

<table>
<thead>
<tr>
<th>Type of sound displays in relation to environment</th>
<th>Design priorities that dominate sound display domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physically embedded in space</td>
<td>Ambient in space, fits in the soundscape but for commercial purposes</td>
</tr>
<tr>
<td>Muzak, Background Music</td>
<td>Ambient or localized, passive or interactive, affective and spatial qualities, thematically tied to an exhibition</td>
</tr>
<tr>
<td>Art and Museum Audio-augmented Installations</td>
<td>Alerts and signals, localized to person, clarity of feedback, aesthetics are important, arbitrary relationship to the natural environment</td>
</tr>
<tr>
<td>Portable/Mobile Devices Alerts and Tones</td>
<td>Realistic, highly affective soundscapes</td>
</tr>
<tr>
<td>Media Soundscapes - Cinema, Games</td>
<td>Highly programmed sound flows, managing listener attention</td>
</tr>
<tr>
<td>Media Soundscapes - TV, radio</td>
<td>Hierarchical systems of audio signals, meaning and perception is key, immediate feedback</td>
</tr>
<tr>
<td>Sound-enabled Computer-Mediated Communication (Desktop)</td>
<td>Framework for Auditory Display Design for Ubiquitous Computing</td>
</tr>
</tbody>
</table>

Figure 3.2: An overview of different sonic environments that bear relevance to the framework of auditory display design for ubiquitous computing.
Ubiquitous computing contexts, being interactive, embodied, informative, immersive and social entail many of the features of the above-mentioned examples of sound technologies and spaces (see Figure 3.2), and yet the lessons, requirements and design principles from these domains don’t seem to have found themselves together in one framework, capable of supporting auditory display design for ubiquitous computing. To extend the kind of experiences, which are possible in ubiquitous computing environments, an auditory display design framework is needed where sound is immersive and ambient as well as perceptually and cognitively effective for the activity at hand. To achieve this, requires a feedback system that supports the acoustic and information ecology of a physical space and provides directive, informative feedback. To develop such a framework, I suggest looking at two existing approaches to theorizing about sound, investigating sound, designing and describing sound – the acoustic communication framework, and the auditory display model of design.

**Acoustic Communication Framework**

“There are simply no sounds in nature that will destroy your hearing. I have often thought of this as I listen to the interactive messages of the natural soundscape, where there is always a time for sounding and a time for listening.” R.M. Schafer “Voices of Tyranny, Temples of Silence”

**Acoustic Ecology Model**

A pioneer in the field of Acoustic Ecology, R. Murray Schafer first defined the notion of a soundscape to mean a holistic system of sound events constituting an acoustic environment and functioning in an ecologically balanced, sustainable way (Schafer, 1977). Born out of the threat of urban noise pollution, Schafer’s ideas cantered on conceptualizing an ecological balance in the acoustic realm. He developed the terms ‘hi-fi’ and ‘lo-fi’ to describe different states of aural stasis in the environment. A hi-fi soundscape, exemplified in Schafer’s work by the natural environmental soundscape, is one where frequencies occupy their own “spectral niches” (Wrightson, 2000, p.11) and are heard distinctly, thus creating a high signal-to-noise ratio. A lo-fi soundscape, on the other hand, often exemplified by modern urban city noise, is one where amplified sound, traffic and white noise, mask sound signals and obstruct clear aural communication, creating a low signal-to-noise ratio (Truax, 2001, p.23).
Schafer’s answer to noise pollution and the un-balanced urban soundscape is a combination of aural education, sound awareness and a new public approach to understanding and designing the sound environment through soundscape composition (Schafer, 1977; Schafer, 1993). His central thesis is that the acoustic environment could and should be heard as a musical composition and we must acknowledge our own responsibility for its composition (Wrightson, 2000, p.10). This responsibility has both positive and negative manifestations – soundmaking and music, or noise and signal masking. As Schafer’s colleague and acoustic communication theorist, Barry Truax points out, “the necessity of the ecological concept springs from the context of loss, or at least from the present threat to survival. The question for us now is whether a new balance can be regained. Can we – with consciousness – be part of a new eco-system?” (Truax, 1992, p.364). Within the paradigm of soundscape design, we have to be informed by the past, and maintain an ecological balance of sound components. To start, we have to understand design as a system that “comprises the knowledge and the techniques that we understand and can put into practice,” (Truax, 1992, p.365) and that it involves everyone as listeners and soundmakers, not just the designer/composer.

Based on the World Soundscape Project (WSP), a long-term ethnography and fieldwork initiative, which began in the 1970s, Schafer and the WSP team developed, using a grounded theory method, a classification system of core terms to describe soundscape components in the natural environment. Within this ontology, keynotes are ambient, background sounds that are present a large portion of the time in a given space; sound signals are foreground sound events that transmit information about happenings in the environment, and soundmarks, similarly to landmarks, are unique sounds typical for local community soundscapes and often characteristic of them (Truax 2001, pp.23-24). Yet these terms are not set categories, but characteristics of sound, as each individual listener’s perception will determine the class to which the sound belongs at any given time, and in any given context.

Another important idea that Schafer developed in his 1993 book *Voices of Tyranny, Temples of Silence*, is the concept of ‘acoustic space’ (Shafer, 1993, p.26). His conceptualization has many implications for soundscape design and auditory perception studies, because it emphasizes the multi-directionality of the sound field and the complex, almost unpredictable nature of sound behaviour. Acoustic space consists of many sounds and is coloured by all physical properties of the environment, including the
listener, thus creating a unique atmosphere of being and place (Schafer, 1993, p.32). Along with acoustic colouration - the audible properties of sound produced by their interaction with the physical environment - these soundscape components characterize an acoustic community (Wrightson, 2000, p.12).

**Acoustic Communication Model**

Following Schafer’s work, Barry Truax developed a multi-disciplinary framework for understanding sound based in notions of acoustic ecology as well as communication theory. This framework – acoustic communication – moves away from the energy-transfer view of sound, where the focus is on the mechanical transmission and reception of sound vibrations. Instead, it models sound, listener and environment in a holistic interconnected system, where sound mediates a two-way relationship between listener and environment (Truax, 2001, p.12). It also places importance on the role of context to the process of listening, which Truax defines as “the processing of sonic information that is usable and potentially meaningful to the brain” (p.11). This notion emphasizes the listener’s ability to extract meaningful information from the content, qualities and structure of the sound precisely by situating this process in their knowledge and familiarity with the context and environment (p.12). This new understanding of sound allows us to bring considerations of culture and political economy into the soundscape paradigm alongside auditory perception and cognition. Drawing on scientific and psychoacoustic notions of sound, cultural and sociological histories of listening, as well as major communication theories, Truax creates a rich perspective of the role of sound in the environment before and after the emergence of electroacoustic technology, as well as a critical account of large-scale shifts in listening patterns. With the acoustic communication model, the soundscape can be seen as a multi-faceted point of reference representing the many relationships that sound mediates between environment, society, listeners, culture, public and private domains, class, status and politics.

The acoustic communication model also extends to a new understanding of psychoacoustics, listening and perception. Traditional models of auditory perception conceptualize listening as a process of neural transmission of incoming vibrations to the brain (Cook, 1999; Neuhoff, 2004, McAdams, 1993) that, shaped by our physiology, allows us to experience sound qualities. In fact, as pointed out by Truax (1992) and
others, listening is a complex activity involving multi-level and dynamically shifting attention, as well as higher cognitive functions such as memory associations, template matching, foregrounding and backgrounding of sound; and inevitably dependent on context (Truax, 2001, p.11). Again, this model points to the importance of understanding listening as a physiological as well as a cultural and social practice. From a design perspective, it is also imperative to understand that listening is a dynamic and fluid activity that in turn affects the perception and experience of sounds in the acoustic or electroacoustic environment, and helps mediate the relationship between actor, activity, context and environment. Two major classifications of listening are ‘everyday listening’ – (Bregman, 1990; Ballas, 1993; Gaver, 1994; Kramer, 1994) – omni-directional, semi-distracted, adaptive-interactive listening that focuses on immediate information-processing of sound; and ‘analytic listening’ (Truax, 2001, p.163) – attention to detail, ‘expert’ activity focused on an aesthetic or analytical experience of sound, and rooted in context as its frame of reference for extracting information from sound characteristics. Based on the idea of different classifications of listening, Truax developed a number of categories exemplifying major listening modes and processes (2001, pp.21-27). Listening-in-search involves a determined seeking of a particular sound template in an aurally busy environment. Listening-in-readiness involves background listening with an underlying expectation for a particular sound or set of sound signals (such as a baby’s cry). The ‘cocktail party effect,’ for example, is a special mode of listening-in-search, which involves a ‘zooming in’ on a particular sound source – often semantic-based (speech) and familiar in an environment of competing sound information in the same spectrum (p.22).

Table 3.1: Key concepts from the acoustic ecology and acoustic communication models.

<table>
<thead>
<tr>
<th>Key Concepts from the Acoustic Ecology and Acoustic Communication models:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hi-fi Soundscape</td>
</tr>
<tr>
<td>A high-information environment with a high signal-to-noise ratio</td>
</tr>
<tr>
<td>Lo-fi Soundscape</td>
</tr>
<tr>
<td>An environment with a high degree of masking and low signal-to-noise ratio</td>
</tr>
<tr>
<td>Keynote</td>
</tr>
<tr>
<td>An ambient sound, present in an acoustic community most of the time and cognitively backgrounded by listeners</td>
</tr>
<tr>
<td>Sound signal</td>
</tr>
<tr>
<td>Foreground sound events, providing information</td>
</tr>
<tr>
<td>Soundmark</td>
</tr>
<tr>
<td>Unique sounds, characterizing a community</td>
</tr>
</tbody>
</table>
Key Concepts from the Acoustic Ecology and Acoustic Communication models:

<table>
<thead>
<tr>
<th>Soundscape Composition</th>
<th>The process of recreating a soundscape using electroacoustic techniques of sound manipulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Everyday Listening</td>
<td>Omni-directional, semi-distracted, adaptive-interactive listening that we engage in on a daily basis with the goal of immediate information processing (Ballas, Gaver)</td>
</tr>
<tr>
<td>Analytical Listening</td>
<td>Attentive, foreground listening, usually to the finer details of sound</td>
</tr>
<tr>
<td>Cocktail-party effect</td>
<td>An aural 'zooming in' in a busy acoustic environment</td>
</tr>
<tr>
<td>Masking</td>
<td>The result in perceptual loss due to white, traffic, or other broadband noise prominently present in the environment</td>
</tr>
<tr>
<td>Acoustic Space</td>
<td>A sound field of propagation and interaction between sound and environment</td>
</tr>
</tbody>
</table>

Media listening and distracted listening are two positions of listening that Truax argues are a direct result of the transition to electroacoustic sound and especially the way in which sound has evolved in its use in media. Since much of media is experienced as a 'background to life,' often in the visual background, programming flow has developed sophisticated and strong aural cues in order to manage and direct listeners' attention to the next item on the media program.

The most important elements of this framework to be noted in relation to the domain of ubiquitous computing, are: the concept of sound as a mediating tool between listener and environment (or user and system); the patterns of sound categories (keynotes, soundmarks, signals) as patterns of everyday listening affordances; the role of context and familiarity in the way listeners extract detailed information from sound; and the idea of listening positions as dependent on a variety of factors including context, place, type of sound, belonging to an acoustic community, cognitive and physical activity. Lastly, it is even more important to note that classifications of sounds and of listening positions are not exclusive, firm categories, but rather fluid degrees of perception that are constantly being re-directed and re-constructed by people's interaction with their environment.
"At three o’clock in the morning, the wailing siren of an ambulance outside my bedroom window is noise. Unless I am lying unconscious on the floor.”

(Austin, K.)

As articulated in the acoustic communication framework, in the natural acoustic environment aural cues provide us with important ‘semiotic shortcuts’ - quick recognition, orientation and other relevant information, often bypassing visual aides. Using auditory displays (ADs) – designed electroacoustic sounds - to ‘sonify events, objects or systems, is becoming an increasingly popular and effective way in the field of human-computer interaction of communicating data and information to an end user, especially in environments of cognitive load. As well, the rich variety of auditory displays already present in everyday life has ensured that electroacoustic sound is one of the primary ways in which people interact with their public and private environments. Yet some of the traditional design practices associated with auditory display systems seem to have departed from the holistic model of aural cues in the natural environment. It is only in recent years that the design of ADs is conducted with consideration to its intended environment, and multiple dimensions of sound qualities are being considered in designing ADs.

Before I begin to explore the core areas of development in auditory display design, brief word is needed to situate the term “auditory displays.” Even though this term is widely accepted in industry and academia, albeit mostly in the fields of natural and applied sciences, it is not entirely unproblematic. Seemingly, the use of “auditory” rather than “audio” reflects the traditional relationship between classic tone perception and AD design. As Kramer points out, AD is a “form of applied auditory perception” (2004 p.154) and an “engineering art” (2004, p.157). In other words, the name itself emphasizes the mechanism of reception (auditory) rather than the medium of delivery (audio) or the mode of perception (aural). Furthermore, “display” reveals the traditional bias towards visual perception, appropriating the language of visual media. Many use “earcons” and “audio icons” interchangeably with ADs in order to solve this semantic problem, however the rich variety of functionality and perceptual models of sound displays require the use of specific classes within a more general umbrella term.

There are several major groups of auditory displays that currently exist and are being researched and developed. They can be categorized by functionality type (what
type of information they communicate) - Table 3.2; conceptual mapping systems (what is
the relationship between sound and assigned meaning, including their internal hierarchy
and organization) - Table 3.4; and perceptual task required (what type of aural-cognitive
attention is involved and intended) - Table 3.3 (Kramer, 1994).

Table 3.2: Overview of the main areas of auditory display systems

<table>
<thead>
<tr>
<th>Functional Characteristics</th>
<th>Description, Distinction and Main Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warnings, alarms, status indicators</td>
<td>Traditionally one-dimensional, simple, periodic, synthesized sounds, predominantly non-vocal. Newer studies point to the usefulness of voice-based alarms and warnings and 3-D spatialization of ADs in pilot training (Brungart, 2002).</td>
</tr>
<tr>
<td>Alerts and notifications</td>
<td>Also traditionally non-vocal. This class of sounds includes everything from e-mail and other computer functional notifications, to wireless mobile devices notifications such as ring tones. There is a wide range of complexity in these types of auditory displays — from very simple (e-mail) to very rich and culturally-coded (ring tones)</td>
</tr>
<tr>
<td>Audification</td>
<td>The direct translation of data streams into waveform information and sound display (Kramer, 1994). In the 1994 book on ADs, Kramer is the first to make a distinction between audification and sonification.</td>
</tr>
<tr>
<td>Sonification</td>
<td>Representation of data streams through abstract non-vocal sound, and transformation of data relations into perceived sound relations (if temperature goes up, the perceived pitch goes up) (Walker &amp; Kramer, 1996)</td>
</tr>
</tbody>
</table>
Table 3.3: Types and distinctions of auditory tasks for data comprehension

<table>
<thead>
<tr>
<th>Perceptual Task</th>
<th>Description, Distinction and Main Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Template Matching</td>
<td>A (prompt) recognition of an aural image from our working or long-term memory. Cognition, memory, as well as the context of listening all play a role in template matching</td>
</tr>
<tr>
<td>Data Exploration</td>
<td>The ability to monitor an audio stream of data and be able to discern stretches of information from the flow (Kramer, 1994)</td>
</tr>
<tr>
<td>Trend Analysis</td>
<td>The task of determining patterns in the data set over the course of several data points (e.g. is the price of stock rising or falling), as per Walker et al.</td>
</tr>
<tr>
<td>Point (Value) Estimation</td>
<td>The task of determining as near as possible the exact value of a data stream at a specific point of time or space (Walker &amp; Kramer, 2004)</td>
</tr>
</tbody>
</table>

Table 3.4: Types of auditory displays and their conceptual mapping to meaning and information

<table>
<thead>
<tr>
<th>Conceptual Mapping Systems</th>
<th>Description, Distinction and Main Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auditory icons</td>
<td>The counterpart to visual icons, auditory icons represent events, functions, objects or actions in an &quot;analogical&quot; relationship to the original sound source (Kramer, 1994), i.e. they keep the intrinsic connotation encoded in the object or event (waste basket sound in Mac Finder resembling dustbin paper trash sound).</td>
</tr>
<tr>
<td>Earcons</td>
<td>They are part of a hierarchical system of sound representations (also traditionally non-vocal), which create a metaphorical language of sounds. That is, earcons are interrelated within a system, recognizable as being of the same system and users have to learn the ‘language’ of that system. Earcons have mostly a metaphorical or “symbolic” (Kramer, 1994) relationship to the event or object they represent.</td>
</tr>
<tr>
<td>Voice-activated telephone directories</td>
<td>Voice-activated telephone directories are perhaps the one example of earcons that use entirely vocal (sampled or synthesized) material to deliver information. They are used to manage high-traffic telephone depots through automated generation and assembly of speech. Notably, the new generation of phone directories use all sampled sound organized in natural ‘flows’ and the aural image of a fictional character thus conveying a more fluid, humanistic and less ‘automated’ communication.</td>
</tr>
</tbody>
</table>
Historical Perspective of Auditory Displays

The use of alarms, warnings, as well as other non-speech or musical sounds to communicate signals or information to people has been used throughout history – some notable cases that Kramer (1994) mentions are the Roman use of drum and horn patterns in battle to signal combat strategies, and the Morse code communication system (Kramer, 1994, p.34). Other natural acoustic signals as examples of 'auditory displays' include dog barking as a warning sign, steam whistles in mines and ferry boats to signal work shifts or locate the shore, respectively. However, in terms of 'scientific' attempts at using sound to represent data or discern information, some examples date back to the first decade of the 20th century. Echo-ranging sonar became popular in 1912 after an iceberg sank the Titanic. This sonification system was used all during WWI and covered a frequency range between 500Hz and 1kHz (Kramer, 1994, p.30). Later sonar systems include the supersonic sonar and the passive listening sonar, with a range well beyond 20kHz.

A different type of AD – an interface for blind users, was developed as early as 1914. It utilized a series of simple tone beeps and signals (Kramer, 1994, p.36) and could be seen as the predecessor of a personal computer audio interface. Stemming from research for the blind, the next few decades set the stage for research into situations where the user is not blind but their visual faculties are fully occupied, and additional information can be related through sound. Factory control rooms, planes and other data monitoring jobs were becoming increasingly saturated visually and at the same time audio technologies were experiencing an enormous development in the first three decades of the 20th century. The possibilities of using auditory displays in aircrafts surfaced in 1945 with T. Forbes’ work “Auditory Signals for Instrument Flying.” This work is also perhaps the first to consider the context and conditions under which the ADs will operate, rather than just the mechanics of the sound. Forbes articulated several principles for the design of successful aircraft ADs, including taking into considerations pilots' habitual style of thinking and operating, as well as masking patterns of audio and other multimodal displays in the cockpit (Kramer, 1994, p.35).

Data sonification efforts also began around the same time, in 1954, with Ficks’ research outlining the use of abstract sound to illustrate data. Eight binary variables were mapped to eight sound parameters such as pitch, loudness, sound/noise alternation rate, etc. What Ficks and Pollack found, which also resonates with
contemporary AD findings, is that multidimensional displays, i.e. ones that use mappings of multiple parameters of sound outperformed uni-dimensional ones (Kramer, 1994, p.37). In 1961 Speeth conducted related experiments with audification of seismic data, by transposing it to audible range and seeing if respondents could discern any informative patterns in the audio streams (Kramer, 1994, p.37). Into the 1970s, German inventors Kaiser and Greiner developed and patented the first multidimensional auditory monitoring of printing press activity. Within this system, multiple printing press operations controlled multiple parameters of sound that outputted into a single auditory display. Around the same time, engineers in Bell Laboratories developed a simple 3-dimensional auditory display (pitch, timbre, amplitude modulation) to assist in the classification of data (Kramer 1994, p.38).

Historically, AD research and AD design have been heavily influenced by traditional psychoacoustics, music theory, cognition and psychology. More recently, the field has also incorporated some 'ecological' approaches to sound starting with Gibson's original proposition of ecological validity (Gibson, 1977). It includes the research of Vanderveer in environmental sound recognition in 1979, followed by the works of James Ballas and Steven McAdams in "everyday sound" perception in the early 1990s, and subsequently Gaver and Kramer's notions of "everyday listening" (Gaver, 1994, p. 418). Another crucial work that has been highly influential to AD design is Albert Bregman's research into auditory stream perception articulating human patterns of perception and auditory grouping within a system of multiple sound displays (Bregman, 1990).

For the purposes of this thesis I won't go into detail describing many established forms of auditory display, such as auditory icons and earcons, virtual audio, or auditory graphs. Rather, I will focus on sonification, as I believe it holds great potential for ubiquitous computing environments, with its qualities of being ambient, persistent, (ideally) ecologically nested within the surrounding acoustic environment, and highly informative at the same time. The following section should illustrate why I think this field may have a lot to offer to sound feedback design for ubiquitous computing. However, this is not to say that research into the requirements, qualities and characteristics of alerts, earcons, and especially auditory icons, are not applicable to future research into auditory display design for ubiquitous computing.
Sonification

Sonification is not only the oldest form of organized, designed auditory display, but it also incorporates the elements and characteristics of many of the types of ADs discussed thus far. It could be argued that it also holds more possibilities, as well as challenges for aesthetic considerations than other ADs, because, unlike them, the primary task required of sonification is dynamic monitoring patterns, not discrete template matching. Sonification operates on the fundamental principles of primary audition, streaming and habituation, rather than high-level abstract recognition. Major concerns in sonification, then, are masking in the surrounding environment, spatial segregation, fatigue and annoyance, as well as interaction (particularly in multimodal displays). Sara Bly is the pioneer in the field of sonification and the first to provide a focused analysis of representing data through sound in her 1982 doctoral thesis (Gaver, 1994, p.419). She developed the first methodologies for creating meaningful and effective representation of logarithmic, time varying data using sound synthesis techniques (Kramer, 2004, p.153). In short, sonification allows data of any kind to “change the parameters of a synthesized tone, usually in a highly metaphorical or symbolic manner...[and] to allow the listener to extract meaning from the sound” (p.153).

In the 1994 anthology called *Auditory Displays* Gregory Kramer provides an overview of advancements in the different fields where sonification has been applied quite successfully. These include sonar; streaming seismic data; monitoring of medical parameters of the human body for diagnostics; sonifying graphical representation in order to provide alternative perspectives into underlying data; spatial sonification of financial trading and stock market data, sonifying debugging programming scripts, and multidimensional chaotic data, among others. Kramer also suggests thinking of sonification as “auditory graphs” (Kramer, 1994, p.153-4) as a counterpart to visual graphic representation, though since then auditory graphs have developed as their own field of inquiry. And as with the development of visual aesthetics, Kramer advises AD designers to apply more aesthetical and contextual considerations when mapping synthesis parameters to data streams. Mara Helmuth is an example of an artist who has used sonification of glacier events and mapped it to physical parameters of speed, size, volume, etc. to create the most aesthetically appealing and most compositionally organized self-generating soundscape (Helmuth & Davis, 2004). Her work is, in a sense,
edging on the opposite spectrum of science, where data sonification becomes art for art’s sake.

With the exception of some attempts in vocal formant synthesis – modelling of human vowel sounds (Lee, 2007), sonification has utilized largely non-vocal (often MIDI-based) synthetic musical sound. With aesthetics in mind, and influenced by advances in tone perception and musicology, Diane Sonnenwald et al. created the system InfoSound in 1990. InfoSound is "an interface toolkit that provides [the user with] a mechanism for design both musical and everyday sounds, and to have real-time, continuous control from a parallel or sequential application" (Sonnenwald et al., p. 541). The authors of InfoSound believe that musical patterns could be used to represent complex sets of data, specifically, program debugging scripts, for rapid recognition of errors that would be difficult to discern visually. InfoSound also uses an external MIDI device and the sonification stream is mapped to occurrences of events (errors) in the system (Sonnenwald et al., p.544). Another extension of what DiGiano and Baecker call “program auralization” is LogoMedia developed in 1992 – a program that allows the programmer to map program events to musical sounds and sound effects (Vickers, 2004, p.2). In the Sonnet system – another debugging auralization tool from 1994 – “the code to be debugged is tagged with auralization agent that define how specific sections of code will sound” (Vickers, 2004, p.2). This allows the user to develop analytical ‘gestalt’ recognition patterns, where deviations from correct script will really stand out aurally. The next generation of debugging sonification is the ADSL language, which utilizes “tracks” of sound instead of assigned mappings to program parameters. This system also allows the script to be sounded out as it runs (Vickers, 2004, p.3). ADSL can use MIDI, composed musical sequences, sound effects and recorded speech. A similar product, LISTEN from 1994, is designed to monitor automobile controlling activities and uses both pre-processing techniques and mapping sound to events. In a discussion on aesthetics in computer sound, Vickers analyses all of the above-mentioned sonification systems in search of a balance between aesthetics or “musicality,” and efficiency in terms of information transfer. He concludes that the contour and metric structure of the sound is much more effective for communicating information and fast recognition than harmonic, tonal structure (Vickers, 2004, p.5-6).

One other aspect of sonification that has been researched more recently is user interaction. As early as 1994, Kramer poses the question of interactive sonification
displays – how much and what kind and quality of interaction is needed in sonifying data, and how does that change the nature and effectiveness of perception-in-context (Kramer, 1994). Since we are obviously well-equipped to understand our environment through interacting with it, it is only logical that “dynamic human interaction to explore datasets while they are being transformed into sound” is a desirable research area (Hunt and Herman, 2004, p.1). As Hunt and Hermann point out, our “perception of objects builds over time; it itself is an interactive process...for instance, to understand a 3D object, different views are needed” (2004, p.1). From these fundamental concepts of human perception, interaction and learning, Hunt and Hermann develop approaches to integrating high-dimensional, continuous control in data sonification, including hand-motion hardware/software for 3D data representations (p.6). What they suggest is the introduction of “model-based sonification as a framework that integrates interaction as one of its defining constituents” (Hunt and Hermann, 2004, p.7).

Scientific education and equipment are increasingly becoming computerized and while the majority of interfaces are visually based, the need for complementary multimodal interaction and access for users with visual impairments becomes more urgent (Walker & Kramer, 1996; Flowers et al., 2001). In addition, advances in mobile computing and its customary use in dynamic competitive industries such as stock trading (Nesbitt & Barrass, 2002; Smith & Walker, 2002) calls on new demands for sonification even for sighted users. This presents opportunities for aiding or even replacing visual displays with auditory ones in divided-attention tasks, environments of sensory and cognitive overload, or situations where screen space is limited (for example in portable mobile devices). More importantly, auditory displays and data sonification are the only way for blind and visually-impaired users to interact with computerized interfaces, as well as gain intuitive, quick understanding of large sets of information, discern patterns or trends of data, and perform exact point estimation of numerical values.

With visualization, data parameters are mapped to visual concepts, simply put, in a way that ‘makes most sense’ – that is, a way that is most effective, given our innate visual abilities and preferences, as well as assumed training and habituation. Similarly, in sonification of data, parameters are mapped to sound characteristics, such as pitch, amplitude, timbre, granularity, tempo/phase, etc. It is important to note that these are all sound characteristics that lend themselves well to continuous, dynamic scaling driven by fluctuations in data. The alternative would be to sonify data events conceptually by
designing unique auditory icons for each data point, and requiring a user to essentially
learn all these icons and perform template matching in order to determine an exact
value. Research in sonification has shown that using data-driven sonic parameters of a
single sound (rather than a multitude of different short sounds) is a much more intuitive
way of representing data and enables “quicker, more informed decisions through
perception of higher level information in patterns across a wider range of data and over
longer time periods” (Nesbitt & Barrass, 2002). Thus, the most critical design decisions
in data sonifications are the saliency and effectiveness of parameter mapping, scaling
and polarity of sound (Walker & Kramer, 1996). Effectiveness here is an umbrella term
encompassing ease-of-use and comprehension, perceivability (Smith & Walker, 2002) of
auditory output, satisfactory response time for making decisions based on sonified data,
aesthetic considerations (especially for long-term exposure), and [relatively] accurate
point estimation.

Data Mapping, Scaling and Polarity

Sonification is a way of representing data using a continuous stream of sound
driven by changes in values that result in an audible difference in the sound. Sonification
constitutes an auditory feedback model that is both ambient and informative and could
inform and direct actions. It is used in environments where large sets of information need
to be analyzed hands-free or vision-free. Many researchers in sonification have
examined best practices in mapping of data to sound (Walker & Kramer, 1996).
However, variations in the context and complexity of human activity preclude the
development of a universal prescriptive system of design guidelines for data sonification.
Context and complexity of human activity significantly impact data-sound mappings,
which, in turn, affect perceivability of sound and task performance. Research does
provide us with specific aspects of sonification, which must be considered in developing
a model for sound feedback. There are several major sound parameters that can be
dynamically varied in a single sound. These are: amplitude (i.e., loudness); frequency
(i.e., pitch); timbre/spectrum (e.g., soft/harsh); and phase (i.e., timbre/rhythm). Lessons
from sonification suggest that the most intuitive mappings of information to sound rank
as follows: amplitude; pitch; tempo and finally timbre (Walker & Kramer, 2004, p.158;
Data-to-parameter mapping refers to the choice of which data parameter is mapped to which sound variable. For example, we could map temperature to pitch, or to tempo. We could represent volume with timbre, or with amplitude. These design decisions should attempt to balance conceptual and perceptual associations of data and sound parameters as described below. Scaling refers to the minimum and maximum value that a sound parameter will gradate between, driven by incoming data. This is also a significant decision. Even though humans can perceive relationships between harmonic tones (i.e., we can discern that one tone is an approximate amount higher than other), there isn’t an inherent sense of a scale in any a particular sound. Scaling is tied to a particular design situation and varies for context of an activity and type of sound. Polarity refers to the direction of gradient of change mapped between data variable and sound parameter. An example of positive polarity is when an increase in temperature is mapped to an increase in pitch. An example of negative polarity is when an increase in volume is mapped to a decrease in tempo. Decisions about polarity are important. Non-intuitive mappings may confuse users and result in inaccurate comprehension of information. Positive polarity is considered to be more intuitive than negative polarity (Walker & Kramer, 1996).

Hopefully evidenced in the description above is the reason why I believe sonification, as a design framework, could be a great starting point for sound design in ubiquitous computing environments. It is precisely its continuous, ambient nature which persists in space and time, just as ubiquitous computing technologies persist in space and time, that make sonification, its principles and perceptual models, most appropriate for the domain of ubiquitous computing. Further, its dynamically changing display may indeed enable the kind of fluid model of listening that an embodied, interactive, social context of ubiquitous technology requires. Further, advances in researching and generating design guidelines for sonification within the auditory display domain present the type of systematic, interdisciplinary, mixed-method approach to inquiry that I feel is needed in the area of sound design for ubiquitous computing.

New Framework Conceptualization

So what should a framework for auditory display design look like in the context and physical environment of ubiquitous computing? How could such a framework be informed by and incorporate core ideas, methods and concepts from the acoustic
communication framework as well as from the auditory display design field? As illustrated in Figure 3.3, the acoustic communication framework contributes a valuable focus on listening as a meaningful, dynamic and rich cultural activity – what we listen to and how we listen is coloured by where we are, what we are doing, who are we with and what else is transpiring in the sonic environment.

![Diagram of acoustic communication and auditory display design frameworks](image)

**Figure 3.3:** A diagram showing the main areas that the two existing frameworks contribute to a conceptualization of sound design in a ubiquitous computing setting.

The concepts of acoustic community, acoustic space and sonic balance – acoustic ecology – are also important additions to a potential framework of auditory display design for ubiquitous computing environments. Acoustic space refers to the way sounds propagate in the physical environment, forming an acoustic horizon and creating a specific type of soundscape balance – whether it is a high fidelity (sounds are heard clearly) or a low-fidelity (lots of masking, louder sounds overpower quieter sonic cues) one. This balance in turn determines the listening positions that listeners take within an environment. The concept of acoustic community encompasses the idea that in any given soundscape there are certain keynotes (constant sounds), soundmarks (significant, characteristic cues) and signals (incidental sounds, carrying meaning and attracting attention). Truax defines it as a bounded system of information exchange.
where "community means that acoustic cues and signals constantly keep the community in touch with what is going on from day to day within it." (2001, p.66) Because ubiquitous computing fits into a real physical environment and cultural context, I believe the idea of acoustic community with its categories of sounds complements this technology quite well.

The field of auditory display design, on the other hand, contributes a vast body of research on absolute and relative measurements of auditory perception, as well as a focus on sound as a channel of conveying meaning, data and information. Concentrated on cognition, information-transfer and sound-mediated activity, it contributes many practical models and guidelines for how to build and design auditory displays for different contexts, from cockpit simulations to mobile technologies. Sonification, in particular, contributes a growing number of models and frameworks for different applications of mapping continuous sound to information. The sonification model of constant, ambient sound complements the idea of a keynote, with the added conceptualization of conveying specific information via the purposeful manipulation of one or more sound parameters. The study of alerts and auditory icons further connects with the concept of signals and soundmarks in the environment – characteristic, incidental sounds that allow listeners/users to form lasting associations and respond with active attention and action. Thus, a framework of auditory display design for ubiquitous computing, borrowing concepts from the two above-described fields, would be one where sound is essential and meaningful, situated in a balanced acoustic space and specifically designed to allow listeners to gain information from it and make sense of their environment through it.
CHAPTER 4: METHODOLOGICAL DISCUSSION

“In focusing not on measuring but on understanding people’s or subjects’ behaviour, qualitative psychologists claim to produce what they believe psychology should produce: a better understanding of human thinking and acting which could lead to human beings being able to understand themselves better and to improve their thinking and acting.” (Zazie p.5)

Research Methodology

The research challenge of building a new framework could of course be addressed in a number of different ways. I could follow the acoustic communication approach and use a grounded theory method where I categorize, classify and describe the approaches taken to auditory display in a number of different ubiquitous computing cases, in an attempt to synthesize a framework. Alternately, I could attempt to conduct controlled experimental studies attempting to isolate and establish the relationship between auditory perception, type of activity, level of embodiment, and other important factors; in order to establish ways of effectively relaying information using sound. What precludes me from attempting the former is the fact that, as illustrated earlier, there are relatively few examples of complex auditory displays where sound is made meaningful, essential, communicative, engaging and social within a ubiquitous computing setting. What keeps me from attempting the latter approach is similar, in that the research into actual ubiquitous computing environments is still in the process of bridging knowledge between psychology, cognitive science, human-computer interaction, and issues of embodiment, engagement and socio-cultural aspects. Controlled experiments into auditory display issues are simply insufficient at this point in time in addressing these complex requirements of ubiquitous computing, as we don’t yet understand enough their foundational components.

In traditional psychology experiments on auditory perception, as well as in quantitative human-computer interaction methods, there is rigid stimulus-response formula. Such studies are good for determining absolute, static thresholds and values of human auditory perception. Some auditory display studies, particularly Kramer and Gaver’s work (Gaver, Smith & O’Shea, 1991) in auditory icons, and more currently the
works of Kramer, Walker, Brewster and Barrass in earcons and sonification (Brewster, Wright & Edwards, 1992; Nesbitt & Barrass, 2002; Kramer et al., 1999), have gone farther by situating the stimulus-response formula in a real task situation with a meaningful activity that uses sound feedback. Among others, McAdams' and Ballas' work on environmental sound perception (Ballas, 1993; McAdams & Begand, 1993) also takes the inquiry further by examining how people's experiences with everyday, environmental sounds happens within different models of perception, cognition and action, than is suggested by laboratory tests with synthetic tones. Sonification studies, as a subset of auditory display research, focus on determining measures and thresholds of auditory perception responses of continuous, changing sound in terms of scaling, polarity, data-to-parameter mapping and spatialization (Adcock & Barrass, 2004; Walker & Kramer, 1996). Contemporary sonification studies take context and other user-centred design issues into account in the research inquiry, as well as emphasize considerations such as time-prolonged exposure to sonifications; auditory and cognitive fatigue; and aesthetic aspects as an element of efficiency of the sonification (Vickers, 2004; Leplâtre & McGregor, 2004). The use of non-traditional sound categories for display - using vocal or other non-synthetic content (Lee, 2007; Butz & Jung, 2005; Hunt & Hermann, 2004) has also been investigated, though by few. In any case, the unit of analysis is always the human response, measured as a threshold, a just noticeable difference, a correct or incorrect response – all in an attempt to capture generalizeable human perceptual and cognitive abilities and predict reactions.

On the other end, while the fields of human-computer interaction and interaction design place emphasis on user-centred design in developing and researching technology, they rarely deal with investigating sound in a significant, systematic manner. When they do, it is often through traditional evaluation and usability methods, which still may not provide a setting that is ecological and holistic enough to allow for building and validating contextual, situated knowledge about sound in ubiquitous computing environments. As Gibson points out, “...awareness is rooted in meaningful experience of the environment: thus ecological validity results from studying subjects/people in their own natural environment, in motion, in active exploration. For people this environment is social, cultural, systemic, economic, political, etc.” (Gibson, 1966, p. 55). Examining social contexts, group interactions, and embodied experiences and their interplay with sound, listening and soundscapes, could not therefore be achieved through traditional psychology or usability methods alone, and requires, I...
argue, adopting the situated, activity-based approach of participatory design methods in order to achieve this task better.

As proponents of using activity theory as a research framework in HCI, Nardi and Kuutti suggest that this necessitates “enlarging the research object” and making activities with all their complexities the unit of analysis (Nardi, 1996, p.26). With regard to the present investigation, I propose that a framework that attempts to describe the role of sound feedback in ubiquitous computing environments, as well as investigate its design principles, has to also employ a larger unit of analysis and contextualize activity. It should be one where human abilities are a point of interest, on the same plane as contextual considerations (social, embodied activity, engagement), as well as patterns of mediation practices - how people develop ways to derive meaning through sound (epistemic activities).

**Theory Building**

For the reasons outlined above, in the present research investigation I will attempt to build theory about auditory display systems in ubiquitous computing in a way that not only pays respect to ecological validity (is conducted in a manner that preserves the authenticity of the end context) but also addresses design directly by generating design guidelines and tangible concepts, rather than simply offering a higher level theoretical discussion. In order to do so, I first have presented a literature review of two existing frameworks that deal with sound design – auditory display design and acoustic communication. From there, I offer a way of combining elements from both, in order to build a theory of auditory display design specifically for ubiquitous computing. In the chapters to follow, I describe the design of an actual ubiquitous computing prototype, as well as the design workshops used in the process. Based on the experience of designing, as well as from testing the prototype, I generate issues, questions and propositions adding to my theory about auditory display design for ubiquitous computing. If my aim is to build theory around ecological validity and design implications of sound, it is crucial to raise design issues from within design, from within the experience of designing. The motivation for this phase of research is concerned with what can be learned from the experience of designing an auditory display for an actual ubiquitous computing space. What can be learned from testing it with users? Here it is important to acknowledge the distinction between design methods and research methods. While
design methods facilitate, further and ultimately improve the end design goals for a prototype (in my case an auditory display system), they do not intrinsically answer research questions, uncover relationships or explain phenomena. However, design workshops could be seen from a research perspective as cases - instances of activity, interactions or experience that, when analyzed, could offer salient concepts towards building a theory about the role of sound in ubiquitous computing, and elucidate significant relationships between auditory display and contextual elements.

Such concepts, precisely because they are generated from within design experience, take me to the next phase of research. Their importance to fleshing out the new framework for auditory display design is in helping refine and generate research questions and propositions to be explored further through research. The next phase attempts to tackle these questions via a set of naturalistic experiments – exploratory studies – informed by design methods. This research phase not only furthers the theory-building process by generating and refining new knowledge about sound in ubiquitous computing, but also demonstrates how modified design methods could help preserve the ecological validity of the inquiry. These experimental studies allow me to re-examine and refine the new framework further by addressing specific auditory display issues with more rigour and building possibilities for analytical generalization. The research instrument in my investigation is the collection of various data points from the different experimental study iterations, including video and audio interaction analysis, observation and discussion logs as well as user-generated artifacts. Taken together, these data points form the basis for a research discussion and analysis contributing towards the proposed new framework for auditory display design for ubiquitous computing environments.

The research investigation ends with a final refinement of the theoretical concepts relating to sound in ubiquitous computing environments, and a set of guidelines, addressing design implications and demonstrating the relevance of the proposed framework to the domains of ubiquitous computing and auditory display design. For this final stage of theory synthesis I use both phases of inquiry including the ubiquitous computing prototype case and the participatory design workshops preceding it, as well as the experimental studies focusing on auditory display issues. To reiterate, just as design experience is central to generating questions that could then be explored from a research perspective, the final outcome of my inquiry relates directly back to offer
design implications for auditory display design in ubiquitous computing. Because design methods are foundational to this research the next section offers a background and rationale for their use, purpose and structure, with a particular focus on participatory workshops.

Design Methods Rationale

The discussion to follow refers largely to the principles and theory of design methods, as a paradigm of approaching design and research, that I argue fits well within the proposed framework for auditory display design for ubiquitous computing environments. Such discussion is necessary in that the conceptual work of building a new framework cannot be separated from a discussion of the methods that have to address the unique challenges and requirements it proposes. A framework may contain an explicit or implicit commitment to a methodology – reasoning behind using particular methods – as part of its paradigm of inquiry and analytical approach. I strongly believe that a discussion of an auditory display design framework for ubiquitous computing environments has to involve an understanding not only of the characteristics of sound design in contexts of embodied interaction, but also of the specific requirements for a methodological approach that could best address that.

In the previous chapter I presented the two frameworks for designing and conceptualizing sound – the more descriptive paradigm of acoustic communication, and the more practically and scientifically-oriented paradigm of auditory display research. Nascent in each of them, are the methods of investigation that they use to organize and construct notions of sound design, listening and perception. The acoustic communication framework seeks to qualify and describe sounds within the acoustic and electroacoustic environments, using qualitative and quantitative methods of media analysis, as well as fieldwork and ethnography. The auditory display paradigm, on the other hand, combines a range of investigative methods suited for its different application domains, and spanning from controlled psychology and psychophysics experiments, to empirical usability methods and human-computer interaction case studies. Neither of the two frameworks’ methodological approaches, however, has as of yet been applied specifically to the context of ubiquitous computing, with a focus on sound. Further, upon examination, both sets of methodologies appear inadequate in addressing the situated, contextual nature of auditory displays for ubiquitous computing environments on their
own. For this reason, I propose building a framework for auditory display design for ubiquitous computing applications complete with a methodology that is capable of addressing the particular requirements of ubiquitous computing, and a set of methods that could be used and modified to aid the design of such feedback and/or build conceptual knowledge around it. As ubiquitous computing falls within the paradigm of interaction design, it seems fitting to first look at design methods.

**Design Methods Foundations**

"Because there is so much benefit in the physical world, we should take great care before unreflectively replacing it. More precisely, from a design perspective, solutions that carefully integrate the physical and digital worlds — leaving the physical world alone to the extent possible — are likely to be more successful by admitting the improvisations of practice that the physical world offers." (Klemmer, Hartmann & Takayama, 2006, p.140)

Design methods are part of a situated, user-centred framework for conducting inquiry in the areas of interaction design and human-computer interaction, where design issues are investigated in-context, through the use of system or artifact prototypes (Sanders, 2001; Svanes & Seland, 2004; Simsarian, 2003; Suchman, 1987; Winograd, 1997; Iaccuci, Iaccuci & Kuuti, 2002). In other words, design is central to the inquiry but advancing the design towards the solution of a particular problem is the ultimate goal. Design methods have emerged out of a need to investigate not simply user reactions to presented stimuli, but also to better understand user interactions with real objects and environments, especially interactions that are not pre-determined by the designers. One prominent paradigm of design methods could be found in the Scandinavian tradition of participatory design (PD). In his cornerstone book on work-oriented design, Pelle Ehn proposes that, "human practice and understanding in everyday life should be taken as the ontological and epistemological point of departure in inquiries into design and use [of computer artifacts]." (Ehn, 1989, p.28). Based in Marxist philosophy and emancipatory practice, PD emerged to address social, technical and power relation issues in designing within organizations. Traditionally, the method involves lengthy involvement with stakeholders within the users’ settings that result in an empowered stakeholder and informed designer, co-designing solutions. As Christopher Jones suggests, design of computing technologies and environments is no longer "...in the making of products and systems and bureaucracies, but in composing the contexts that include everyone, designers too (Jones, 1977).
Moving slightly away from the emancipatory goals of the Scandinavian participatory design movement, the user-centred interaction design field has instead adopted participatory workshops - small-scale situated design inquiries. Participatory workshops preserve the principles of genuine user participation, design within end-user settings, and enabling participants to co-design, however in a severely shortened time period and without the goals of in-depth contextual design, transformation of users into designers, or systemic sustainability (Muller & Kuhn, 1993; Taxén, 2004). Rather, the workshops are a quick, flexible and powerful tool that allows designers to investigate specific activities, situations and environments. As one of the pioneers of using participatory workshops in design research, Liz Sanders emphasizes the importance of "make-say-do" in workshop-based research – not only listening to what users have to say, but accessing their internalized "know-how" by watching what they make, and how they do it (Sanders, 2001). The aim of such workshops is typically to move beyond traditional user-centred design and to harness participants' creativity, utilize their expertise for hands-on concept development (Taxén, 2004), as well as model and manipulate simulated environments through role-playing (Svanaes & Seland, 2004).

While industry uses participatory workshops, and design methods in general, for product development especially of novel technologies, design theorists and academics have been attempting to formalize design as knowledge-creation in and of itself, and design methods as vehicles for research inquiry. To think of design in this way is to recognize that knowledge is contained within practice, and that this type of knowledge is qualitatively different from analytical, intellectual knowledge. As Ehn stresses, the role of practice in the everyday, the interconnected cycles of action, reflection and knowledge, are fundamental to both designing and using artifacts (Ehn, 1989, p.60). This conceptualization is echoed and developed by others, including Donald Schön, who sees designing as a reflective, iterative process of fluidly moving between the stages of analysis, synthesis and ideation. Knowledge is generated through the process of reflection and iteration and is cyclical as its aim is as much about generating answers as it is about generating problems, framing and re-framing them, and gaining greater understandings (Schön, 1985). Activity, doing, practice is then clearly central to using design methods as inquiry (Ehn, 1989; Nardi, 1996; Dourish, 2001).
CHAPTER 5: RESEARCH DESCRIPTION: PHASE I

Before approaching the question of what can be learned from designing an auditory display for a ubiquitous computing prototype, we have to first ask how do we design such an auditory display, and what kinds of considerations are important for the type of system we want to have — one that is embodied and social, and where sound has a prominent and epistemic role? In this chapter I describe the first phase of my research investigation, which contains a set of two participatory design workshops leading up to the development of a ubiquitous computing prototype named socio-ec(h)o. I will describe both the workshops, and the preliminary user testing sessions, including a discussion of the findings. Before I begin, I want to acknowledge the context within which some of the research took place. The design and evaluation process for socio-ec(h)o was led by professor Ron Wakkary and professor Marek Hatala with the research contributions of Kenneth Newby, Jim Bizzocchi, and a team of graduate research assistants that I was a member of. This phase of my research is explicitly part of the socio-ec(h)o design process and conducted collaboratively with the research team, while the second phase represents my own contributions. As socio-ec(h)o was going to be a physical ambient intelligent system — a game played by a team with a significant responsive environment element - designing an appropriate and effective auditory display was a priority. This is why we concentrated some effort on conducting exploratory workshops dealing with issues of sound as part of our design process.

In this section I will describe the structure and results of the workshops, synthesize some core observations about the role of sound within an environment of embodied interaction; and demonstrate how the lessons learned from the workshops helped shape socio-ec(h)o. Then I will describe the design approach taken to the auditory display system of socio-ec(h)o and discuss the experience from our preliminary user testing of the prototype. Following that I’ll attempt to again re-synthesize core emergent concepts relating to the new framework for auditory display design for ubiquitous computing, and articulate a set of propositions and questions brought about in the experience.
Two Participatory Design Workshops

The two workshops that are described in this phase came midway through the design process for our ambient intelligent (Ami) game socio-ec(h)o. We had previously hosted several other participatory workshops and conducted concept development meetings where we developed the conceptual foundations of socio-ec(h)o, including core game mechanics, game progression and structure, and narrative development. We had yet to build a working prototype. Our main concern at this stage was the design of a compelling environment based on user engagement, movements in physical space, immersion, and narrative game progression. We knew at this point that we needed to investigate specifics in the role that the audio display would have. We had determined that the technical preconditions included location tracking, and an ambient interface that might involve body and object movement, location, and gestures. Given the Ami nature of the project we ruled out a graphical user interface of any kind. Thus what we wanted to explore was whether giving sound and ambient light a prominent role in the space would afford for different kinds of interactions, activity and experience, and what would they be?

Physical Setup

Both workshops were set within the physical game space intended for the final prototype: a black box environment (enclosed theatrical space with black floor, ceiling, walls and curtains so as to allow complete darkness) with controlled light and sound displays. All system feedback in the workshops was delivered via Wizard of Oz techniques. A Wizard of Oz experiment is one that simulates the functionality of a technological system without actually building an automated prototype. While participants interact with the “system” as if it were autonomous, human researchers provide behind-the-scenes functionality in response to user actions (e.g. bringing up light or sounds, triggering events). This method allows for exploratory testing of user interactions and experience patterns. It focuses on the effectiveness and possible uses of the simulated prototype, rather than on the usability of an entire system.

Our conceptual starting point for the audio feedback was the use of sound avatars that would allow game characters to be used in the game mechanics and would provide a vehicle for narrative progression. Both workshops were organized around this conceptual starting point. The underlying focus was on the interaction patterns between
players and system, and the role of the ambient response in audio and light. After each workshop, we invited participants to suggest changes in the response and interaction rules based on their experience of the environment and their avatars.

**Foundational Sound Design Concepts**

In designing the workshops, we reviewed some auditory perception and AD design literature and based on it, as well as on our formative conceptual game development up to that point, we identified several sound design concepts that became operational in carrying out the participatory workshops. These concepts were selected on the predication that they might be useful in communicating game information between players and system, signalling progress and changes in the game, and creating an immersive atmosphere. As mentioned, these concepts have a theoretical and practical underpinning in applied psychoacoustics, as well as a commitment to the acoustic communication framework of soundscape design. Below is a breakdown of the operational design ideas expressed through sound feedback concepts, and brief discussion of their importance, reason for selection, the role that they may take in the system’s auditory responsive environment, and the perceptual models of listening they embody.

**Keynote Sound as “Ground”**

A keynote sound, as proposed by Schafer, is the ‘ground’ from the figure-ground paradigm of listening (Schafer, 1977). It is any sound that is present in a certain space for most of the time. Sound is an extremely powerful tool in creating a sense of ambience in a space, as well as fostering an evocative cultural experience for the users. This is exemplified in various media, especially cinema. In both our workshops, we felt it was important to create an atmospheric keynote sound that would serve as a ‘ground’ for localization of the game’s acoustic space, and situate the rest of the auditory display within a context.

**Musical Expression as Avatar**

As already determined, a core mechanic in our game was going to be player identity. Through previous design sessions we had come to the idea of using unique sounds within a system of display to represent game characters. One way to do that
through sound is by using musical sound. This “musical expression avatar” approach would utilize a discrete musical (MIDI) phrase to sonify players’ identities, their actions and spatial locations. This model relies on perception research of periodic sounds, auditory streaming and template matching (Cook, 1999; Bregman, 1990; Neuhoff, 2004). We used this approach in Participatory Workshop I, in the form of four individual parts of a counterpoint MIDI composition. Given that the phrases could combine in a number of different ways, we wanted to know how well this approach could work in terms of recognition and identification, and in the formation of sound ecologies, supporting narrative and play.

Environmental Metaphor as Avatar

Perceptually, this approach also relies on template matching of discrete sounds, unique to each player. However their quality as ‘everyday’ environmental sounds presupposes a listening position that is strongly coupled with auditory memory, prior experience and associations. We felt that the richness of environmental sound alone deserves exploration as a vehicle for facilitating recognition, identification with character, and narrative possibilities. In this concept we were informed both by Schafer and Truax’s work in acoustic ecology, and their classification of the natural sound environment in terms of keynotes, soundmarks and signals (Truax, 2001). We also consulted Ballas’ and Gaver’s work in recognition and perception of environmental sounds (Ballas, 1993; Gaver, 1993).

Timbre-Based Sonification of Game Events

Besides using avatars to sonify player identities, we wanted the auditory display system to represent subtleties in the play and game shifts that reflected players’ groupings, their level of activity, their proximity to one another and to the sound sources. We felt that this could be achieved by using timbre changes as a model of sound colouration in the real world. Timbre could be affected by applying simple reverberation to a given sound source. Players will have to listen for complexity, colour and quality of the sound in order to derive meaning. In simple perception terms, this technique could be categorized as a gradient of “muffled” to “bright” or “distant” to “close-up” sound quality. This approach is based on holistic everyday listening, in which we detect small changes in sound quality when we are required to extract information from sound (Truax,
Since this approach is the most intuitive and qualitative, we anticipated that it would be hard to gauge its effectiveness from the workshop.

**Sound Spatialization**

This consideration comes from existing literature on sonification of information and attention management (Adcock & Barrass, 2004; Wenzel, 1994; Butz & Jung, 2005, Hunt & Hermann, 2004), which suggest that separating sound in different spatial locations helps in recognition and effective interpretation of its significance within a rich responsive system of information. For our intended context of an ambient intelligent game, we hypothesized that this approach would work well with our attempt to introduce sound avatars as a game mechanic. We thought that spatializing sound and mapping its virtual location to the physical location of participants would reinforce their connection with an individual sound avatar. We used this approach specifically in Participatory Workshop I.

**Hierarchy of Auditory Displays**

Internal consistency of the system of auditory display was another design consideration. While using a single sound avatar proved promising in Workshop 1, there was a need to provide more coherency and richness of the sonic characters. As research in auditory icons design suggests (Gaver, 1994; Adcock & Barrass, 2004), a hierarchy of internally and semantically consistent audio icons creates better recognition and facilitate navigation and utilization of a system than randomly designed signals. In Workshop II we developed a set of three semantically related sounds that increase in perceived intensity in order to represent each game avatar. Essentially, we constructed a hierarchy of sound avatar representation related to gameplay, movement and character interaction (see Table 5.1).

**Delayed Feedback as Core Game Mechanic**

This approach was developed and tested in Participatory Workshop II, in contrast to Participatory Workshop I, where the feedback was continuous and immediate. Instead of instant system response, in the second workshop we provided no auditory feedback unless players achieved a specific configuration of spatial positions. Further, the feedback was delayed in that players had to hold their position for at least 3 seconds.
before they were rewarded with an auditory response from the environment. This model specifically explored the idea of subverting sound’s traditional role as auditory reward in computer games and auditory signal in electronic devices, by delaying the sonic gratification in order to establish clearer yet subtler gameplay.

**Participatory Workshop I**

Following is a description of the first participatory workshop and a discussion of the sound issues we addressed as design questions, as well as an outline of the design concepts we started out with, workshop structure and protocol. A technical discussion of the sound set-ups and Wizard of Oz techniques is also provided. Finally I discuss what did and did not work in conducting these exploratory workshops, and the lessons we learned. Participants were students at the university and included both males and females aged between 22 to 34 years old.

**Workshop Setting and Structure**

In this first workshop, the specific sound issues that were explored included: the introduction of personalized sound avatars (including their spatialization patterns); the effectiveness of using different sound categories – music, voice, abstract, or environmental sound - with regard to observed recognition, perception and identification; interaction and play. As well we were interested in the use of audio process (reverberation) in sonifying player location, and amplitude in sonifying team activity. Within the black box space, a circular area was designated as the interaction/play space. Four speakers were placed on the floor back to back, forming four semi-distinct acoustic spaces, or zones (see Figure 5.2). As the sound avatars were not fixed in space, but rather, meant to follow their character’s movements, the auditory zones and sound ecologies were dynamically created by players’ interactions and groupings. The four participants were first engaged in a pre-discussion during which they were given basic information about the workshop. The participants were told that they would have an individual “sound avatar” but were not told what it would be; and that the avatar would “follow” them in space.
The workshop consisted of three parts. The first stage was an exploration in which each player individually listened to their “sound avatar” and its behaviour in the physical game space (see Figure 5.1). In the second stage, all four players explore the space together and discover relationships and audio combinations that they can create with their sound avatar and other participants’ sound avatars. The third stage included discussions, suggestion for changes, and real-time implementation of some of the suggested changes. While we had spent most of our conceptual design work on developing game and narrative progression, this workshop was open-ended in terms of narrative and game mechanics so that the role of sound and light could be explored in more depth. The workshop investigated the following game event and narrative components relating to the sound feedback mechanism:

- Discovery of sound avatar (who you are in the game);
- Discovery of audio combinations with other participants (exploration and manipulation of collective identity);
- Sound ecologies challenge, addressing players’ ability to affect the environment by forming new ecologies (their movement stimulates the dynamic soundscape – in periods of inactivity the environment decays);
- Ecology as metaphor, discovering the right configuration of players and/or activity that will result in a prominent sound ecology. Conversely, the “wrong” combination of players and/or activity could result in a negative ecology or complete decay.
Each part of the workshop lasted approximately an hour, with a 20-minute group interview in the end. The entire workshop was video taped for further video interaction analysis. Several researchers were informally observing the participants’ activity and one researcher was taking notes. As an exploratory workshop, though it was focused on interactions with sound, additional points of interest and observation for us were participants’ movements in space – particularly how they used their bodies, their range of movements and speed; participants’ patterns of play, communication and socialization with each other, and mediation strategies that they used to attempt to interpret and manipulate the system’s rules and set of responses.

**Figure 5.2:** Illustration of the speaker arrangement in PD Workshop I. The four dotted circles represent acoustic space ranges, and there were two areas of projected light on each side of the scrim (wavy line)

**Sound Settings in Part 1**

In the first stage of the workshop, we used the Musical Expression as Avatar approach to audio display. A Bach counterpoint piece (in MIDI) was deconstructed into four parts and each was assigned to each of the four participants. The sound avatar (a subset of the Goldberg variations) physically followed each participant, thus reinforcing a sense of association with a sound-based avatar. The sound ecologies that players constructed by movement formed different musical orchestrations. Here we also tested
the use of reverberation to affect timbre of sound and mapped this to distance (proximity to sound source). Amplitude levels were mapped to intensity/level of group activities and movements in the play space. Our objectives were to observe the perception of musical parts, on their own and in different combinations, as well as the perception of distance through timbre in the form of reverberation, and amplitude levels (volume) as a response to levels of activity. Since we were manually driving the system’s response, we were able to adjust those elements fluidly throughout the session. In an exploratory way, we also wanted to test notions of emergent play, free-form play, movement, gesture, and social interactions as those are mediated through sound and embodied interaction.

Sound Settings in Part 2

In the second part of the workshop we used the Environmental Metaphor approach to sonic display. Four environmental sounds signifying earth, fire, water and wind were created. We again used the idea of creating combinations and ecologies with the sound metaphors. Also, amplitude levels and reverberation were used to make subtle changes in the sound, as a response to activity, movement, proximity to sound source, and types of participant groupings. Two known issues we were aware of in working with environmental sound, were the semiotic mappings between avatar and its sound representation (will the sound be recognized as representing what we intended it to), and masking (environmental sounds are typically broadband, thus potentially overlapping with each other). Naturally, water and fire are concepts that could be translated in direct representations, while earth and wind are more metaphoric representations. For example we used footsteps to represent earth, and a processed soft whooshing sound for wind.

The workshop objectives here were in large part similar to those of Part 1. The idea was to explore how environmental sound, as a different sound category from music/MIDI, would influence participants’ experience of exploration, discovery, interaction and play. We also wanted to see whether having environmental rather than musical sound would encourage formation of sound ecologies or game narrative in any specific ways. The display was once again restricted to only one sound for each category. We used amplitude, panning (moving sound around to different speakers), and reverberation to respond to patterns of game events and/or player actions.
Wizard of Oz Techniques

For both parts of this workshop we created a custom Wizard of Oz sound display tool in Max/MSP (www.cycling74.com). The Wizard of Oz method is the manual simulation of unimplemented technology. Besides playing the four parts of the Goldberg Variations (Part 1) and the four environmental sounds (Part 2), the audio tool allowed us to spatialize sound, add reverberation, vary sound levels, and apply granulation. The tool was operated via a UC-33 MIDI controller for faster response time, and operated by two people at once for more accuracy. We had an audio station set to the side of the interaction space enabling a clear view of the participants’ location, actions and group configurations. In stage one, we only played the sound avatar of the individual who was in the space. We attempted to localize the sound wherever the player was, as well as increase or decrease the sound’s amplitude depending on how close the player was to a speaker. In the second stage, we monitored and responded to the formation of particular ecologies, by mixing in the different musical parts, or different environmental sounds. As well, we introduced reverberation as an indication of distance of player from sound source (a speaker). Amplitude, on the other hand was increased when engaging activity and play occurred, and decayed over time if there was little or no movement.

What Worked and Didn’t

For the most part we were able to respond instantly to relevant participant interactions, and reward behaviours that we wanted to encourage. The value of using a Wizard of Oz approach, is that in a workshop, which is highly exploratory and quite loosely structured, it allows us the freedom to examine what kind of activities and interactions we want to support, encourage and reward. It also allowed us to spontaneously and dynamically adjust the sound display to match the players’ ease-of-use of the system, as they became more familiar with its features, thus facilitating their engagement and playfulness. The technique also allowed us the ability to improvise and bring in a special sound reward (a granulated vocal composition) if we felt the players achieved a particularly creative configuration.

However, in Part 1, we could not make the personalized sound parts follow the individual as well as we had intended due to a programming flaw, thus everyone had difficulty identifying with their sound in a spatial sense, and they felt as if the ecologies they formed as a group were somewhat random and arbitrary. In Part 2, we were able to
make the sound follow the participants across two directions, which dramatically improved the reported individual experience of avatar discovery. However, the sound couldn’t follow participants everywhere, which limited the directionality and acoustic space size of the ecologies created. Also the sounds of water and wind proved to be too broadband and partially masked the other two sounds in the ecologies. This resulted in some players’ inability to identify their sound, and as a result be able to develop strong affinities with it.

Lessons Learned

After the first workshop, based on our informal and formal observations, as well as the group discussions held after each part, we refined and distilled our formulation of the core game states in relation to sound feedback for the development of our ambient intelligent prototype into the following conceptual categories as follows:

• Evolution (interactions and game state shifts)
• Relationship between play and mastery (skill acquisition)
• Discovery and exploration (as core game mechanics)
• Game types (how different players affect the game)
• Characters and identity (sound avatars/characters)
• Narrative represented by environment (sound, light)
• Sustainability (engagement and generative play)

The ideas of ecology that we started out with related more to testing the development of a narrative-rich world. Ecology was understood to encompass both environmental ecology – the internal consistency of the ambient immersive world, and social ecology – the sustainability and engagement in social group models that were formed. Again, even though the structure of the workshop was not formalized, the design had a built-in internal consistency of representation through sound – sound avatars did not mutate or change during the play session, rather they created multiple ‘sound mixes’ based on groupings and activity.

Even though players were not always able to identify their sound part within the group musical composition, their movement across space indicated that they didn’t feel constrained by this. Their actions showed that they kept trying to affect the system in some way and get a response or a clear idea of their sound part – what it is and where in the acoustic space it is. We are unable to say whether participants consciously
registered the changes in timbre/reverberation and how it may have affected their interactions or movements. It was clear from the after-discussion that changes in amplitude levels had a much stronger perceptual connection to the sense of responsiveness of the system, as those were more readily registered and remembered. We believe that spatialization of the sound was beneficial, as some participants commented positively on it especially in the individual sessions, however, since we could not fully simulate it again, it is unclear how this would have affected the interactions observed.

While participants seemed much more playful with the music components in the workshop they reported a high satisfaction with having environmental sound characters. Participants felt they were able to identify with them. The players with fire and water avatars reported feeling particularly attached to their “character” and feeling a strong sense of projected identity. In Part 2 of the workshop, during the group experience all participants expressed a desire for sound to better support the “narrative” formations that they were trying to construct. For instance, the fire avatar employed gesture and composition-like movements, and acted “threatening” to other sound avatars (except water, since the participant declared that water quenches fire). All participants seemed frustrated with not having strong enough response and requested a more “clear feedback” system response. In terms of sound narrative with environmental sounds, participants suggested that a greater array of sounds should be used to represent each character, rather than have only one sound.

**Participatory Workshop II**

Because the two workshops were conducted within a larger design process with an overarching design goal, it should be noted that the experience and reflections from the first workshop directly influenced the structure of the second workshop. Similar to the previous section, I will describe the second participatory workshop by discussing the sound issues we addressed, the structure of the workshop, protocol and design concepts. Technical discussion of the sound set-ups and Wizard of Oz techniques will again be provided. A discussion of what did and did not work in the workshop, as well as lessons learned will follow. Participants were students at the university and included one male and three females aged between 22 to 36 years old. Two of the participants had already taken part in Participatory Workshop I as well.
Workshop Setting and Structure

After having established, from Participatory Workshop I, that participants found environmental sounds more meaningful and rich in narrative potential than synthetic musical melodies, we focused on utilizing only the four environmental sounds signifying wind, earth, water and fire in the second workshop. In it, however, we attempted to create a narrative based on the sound characters by creating three accents for each sound with increasing levels of intensity (see Table 5.1). In the previous workshop we learned that participants wanted more “direct” feedback: a clearer connection between their actions and the perceived system response. As a result, we developed a response pattern that had a delayed, yet clearly defined response (our Delayed Feedback as Core Game Mechanic approach). The objective for this workshop was to better understand, 1) the effectiveness and recognition of richer sound avatars through consistent, yet varied (gradually intensified) content; and 2) the role of a delayed sonic feedback for gameplay.

Participatory Workshop II was also set in the black box space. A circular play area was marked out on the floor and a white curtain tied in the middle, hung from the ceiling (see Figure 5.3). Overhead lighting projectors were used to create four circles of light within the darkened space. A 4-channel sound system was implemented in the space. Similar to PD Workshop I, participants were assigned sound avatars and the first stage allowed for individual exploration of the avatar. Again, this second workshop consisted of two parts, each containing individual as well as group exploration.

In Participatory Workshop II we decided to employ constraints in order to encourage specific types of interactions related to desired game mechanics, namely, exploration, discovery and achievement. We aimed to build on the free play results of the previous workshop, but see if a more explicit audio and visual display could shape the play and ultimately encourage skill acquisition in the game, as well as conscious exploration of the game rules with more intentionality than before. We focused on clearly distinguishing different aspects of gameplay by explicitly directing what players can and cannot do. In the individual and group sessions, the four circles of light were the only places where players would get auditory feedback, which might be their sound avatar or combinations of sound ecologies (see Figure 5.4). Other areas of the space were non-interactive. The group play was more restrictive, since we had decided to only encourage even groupings of two and four participants in the lit zones but not of configurations of one or three in lit zones (see diamonds in Figure 5.3).
Table 5.1: Sound avatars and their representations, which include three levels of increasing intensity

<table>
<thead>
<tr>
<th>Sound Avatar</th>
<th>Representations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind</td>
<td>1- Processed wind sound – soft</td>
</tr>
<tr>
<td></td>
<td>2- Stronger wind sound</td>
</tr>
<tr>
<td></td>
<td>3- Soft musical sound - marimba</td>
</tr>
<tr>
<td>Earth</td>
<td>1- Footsteps in forest</td>
</tr>
<tr>
<td></td>
<td>2- Eagle call – mountain ambience</td>
</tr>
<tr>
<td></td>
<td>3- Wolf sound</td>
</tr>
<tr>
<td>Fire</td>
<td>1- Light crackle</td>
</tr>
<tr>
<td></td>
<td>2- Heavier flame</td>
</tr>
<tr>
<td></td>
<td>3- Blowtorch gust</td>
</tr>
<tr>
<td>Water</td>
<td>1- Rumbling of small stream</td>
</tr>
<tr>
<td></td>
<td>2- Outdoor waterwheel</td>
</tr>
<tr>
<td></td>
<td>3- Splash of an ocean wave</td>
</tr>
</tbody>
</table>

Thus, whenever only one participant occupied the spotlights, or two participants were joined by a third, the system decayed and stopped auditory feedback completely. An additional nuance to this workshop was system response time as a variable. We designed the sound and light feedback to be delayed, requiring a player or group of players to linger inside a circle of light for a fixed amount of time (between two and four seconds) before the system responded. Again, this was an attempt to shift user’s active attention to the relationship between their actions and system response. The game mechanics that were operational in our design concept development in this workshop were as follows:

1. Skill acquisition of game rules (delayed feedback rule – player have to be in a circle for at least 3 seconds)
2. Discovery and exploration of individual avatar (players discover their sound avatar and its three levels of intensity, see table 2)
3. Forming ecologies – skill acquisition (players learn that combinations of two and four people create an auditory and visual response, while other combinations cause the system to decay and stop)
4. Forming ecologies of sound (players explore sound ecologies and learn other players’ sound avatar by entering into combinations with them)
As with PD workshop I each part of the workshop lasted approximately an hour, with a 20-minute group interview in the end. The entire workshop was video taped for further video interaction analysis. Several researchers were informally observing the participants’ activity and one researcher was taking notes. As another exploratory workshop building on the first one, it was focused on users’ interactions with sound, specifically their experience with sound layering and gradual intensity related to duration, both in the individual and in the group exploration portions. As well, we were interested in their experience with a delayed feedback model of system response. Additional points of interest for observation were participants’ movements in space – particularly how they use their bodies, their range of movements and speed; their patterns of play, communication and socialization with each other, and mediation strategies that they used to attempt to interpret and manipulate the system’s rules and set of responses. In this last set of questions relating to activity, we were specifically interested in the differences that we might observe in this iteration of the workshop, which is much more constrained, compared with the free-flowing playful nature of the first one.
Sound Settings in Part 1

In PD Workshop II we based our auditory design approach entirely in the Environmental Metaphor as Avatar concept. We used environmental sounds to create a system of auditory icons again using the four different avatars of earth, wind, fire and water. We enriched the complexity of the sound characters by adding layers of increasing intensity to the base sound (Table 5.1). This approach aimed to support participants’ request from the first workshop for more narrative dimensions to the sounds and stronger and richer qualities in the sound avatars.

Sound Settings in Part 2

This setup was identical to the first sound setting in this workshop, with the exception of a light ambient sound of frogs’ chorus, which was present at all times. We also introduced a wildcard auditory event triggered when a single player spent more than 3 seconds in a circle of light. In this case, a composed musical sound (light marimba musical phrase) was played, rewarding this discovery.
Wizard of Oz Techniques

Our Wizard of Oz set up was virtually identical to that of Participatory Workshop I. Only in this workshop we included four sets of three environmental sounds and several “wildcard” pre-composed auditory rewards to the Max/MSP control system. In terms of system response, the environment was “silent” unless one of the four circles of light was “activated” by participants. Only one circle of light could be active at a time, and the first player or group to achieve a desirable configuration and hold it, would determine the audio and visual response. As mentioned above, participants had to stay inside the circle for at least 3 seconds in order to hear a sound. The longer they stayed in the circle, the more “intense” the sound would become; the sound would cross-fade through its three intensity levels and would increase in amplitude over time.

What Worked and Didn’t

This time our sound set up was relatively easy and we managed to respond through audio display precisely as we had intended. In addition, we were able to reward a few moments of play/exploration with a pre-composed vocal/musical sequence. For example, one player started interacting with the scrim hanging in the middle of the dark space (a white transparent curtain) and we were able to respond to her touching gesture with great precision. Soon other players joined too and explored this newly discovered system feature, visibly enjoying the auditory feedback. Because this iteration of the participatory workshop had a tight constraint structure and system rules, the only form of feedback we provided were sound avatars in various combinations, and two additional layer accent sounds for each avatar. While we still varied amplitude depending on duration of player actions and movement, this session resulted in a rather dull soundscape. Players later reported being quickly bored with it, after they discovered their avatar in all its dimensions.

Lessons Learned

Individual explorations with sound avatars went quite well in this workshop. Players typically spent an average of five minutes getting to know their avatar and learning how to elicit system responses. Participants reported being frustrated with having to restrict their movement to the lit circular areas and felt disappointed at the absence of immediate sound triggers. They were also disappointed at having to discover
the system's behaviours, rather than the system responding to their free flow of behaviours. However, they also reported that the response patterns of the system were very clear and the learning/exploration and discovery were greatly facilitated by introducing clear restrictions and consistent feedback. Players did a lot of lingering within a light circle and listening (or what appeared to be listening) to the sounds in great attentiveness.

Some players exhibited more compositional objectives, and others more exploratory activity; it became clear that both should be supported by our system. It was also clear that people were thinking about narrative and different groupings of sound avatars and expected more of a reward in terms of sonic development in response to different groupings. They expected the system to respond with something more than just the mix of the present sound characters. The challenge to us then seems in how to give them more, yet not confuse them by creating conceptual incoherence in the sound content choices. A number of gesture movements were also observed bringing attention to the possibility of gesture mapping to sound being able to afford players the sense that they are "making something happen." Yet since gesture is a compositional tool, it provides a challenge for motion tracking, clear mapping to sound and effective user recognition of cause and effect.

Participatory Design Workshops: General Discussion

To summarize the key observations, the relationship between narrative elements and sound ecologies repeated itself throughout the two workshops. Later, we able to draw on the distinct aspects of the sound avatars and game narrative in creating an internal narrative coherence, thus supporting skill acquisitions (learning behaviours) and communication resulting from awareness of manipulation and creation of sound ecologies. Specifically, at the points of ecology creation, narrative associations were especially evident with environmental sound. In discussion, a participant commented that the "fire" avatar seemed to keep running away from the "water" avatar because water would put out their fire. We had not anticipated these narrative developments, and therefore had no way of supporting them through the system's auditory display response. Yet these results became extremely useful in future explorations of the use of sound in fostering and developing narrative constructs and an immersive story world. Both of these workshops are an example of using representational sound (whether
music or environmental) by tapping into evocative individual memory, from where narrative structures are bound to emerge.

The workshops were also useful in identifying problems to be avoided in designing auditory displays for physical interactive shared spaces. Not unexpectedly, masking of sounds contributes to a lack of clarity of feedback and its relationship to system events, so it is to be avoided. Some issues to think about when attempting to explore sound’s narrative qualities in gameplay situations, are improving on the internal consistency of sound character ‘databases’ (as exemplified in table 2) and their clear delivery in the game; as well as improving on the mappings between player interactions and system response (when they do this, that happens); working on play/game/event time and state shifts in relation to feedback; and supporting different ecological activities inside the play space through sound feedback shifts.

The participatory workshops described above allowed us to make some preliminary conclusions about auditory display design for responsive environments and its relationship to the concepts of narrative, identity and sound ecologies. What was valuable to us at the time was that the workshops helped move our design process forward in a structured and meaningful way. As a result, the availability, manipulation and experience of a sound avatar were major emergent game elements. Sound characters or unique sound characteristics enabled players to assume a story world identity that fostered communication, exploration, skill acquisition, socializing with others and a sense of progression. This approach encouraged specific exploration of the system and its responses based on one’s character, making participants more aware of the subtleties and narrative aspects of their experience. This enabled participants to form internal representations, associations and expectations about how a sound could or should act in the game.

**Participatory Design Workshops: Emerging Concepts**

While the discussion above emphasized the importance of the two participatory workshops to our design process by providing useful design ideas relating to sound feedback, in this section I’d like to focus on the more theoretical implications that these workshops present. Aside from uncovering the significance of narrative, ecology and physical play in relation to sound, certain general ideas about the role of auditory feedback to embodied interaction emerged as well. These concepts, while simple and
perhaps self-evident, help build theory around the new framework of auditory display design, because they recognize and incorporate the specific characteristics and requirements for sound in ubiquitous computing spaces. In listing them, I will also demonstrate how they relate back to the two originally presented frameworks of soundscape design — acoustic communication and auditory display.

**Sound is meaningful** - Sound avatars gave users a chance to assign meaning to a particular sound within a mixture of other sounds, follow it and care about its development. This relates to the concept of soundmarks, which become characteristic sonic cues for a certain acoustic community, gaining associative and cultural significance.

**Sound is social** - This mediating function of sound contributed to participants socializing together through forming sound mixes. If given an opportunity, people use sounds in various ways to play, socialize and communicate with each other, speaking to the performative aspects of interaction through sound in the natural environment, whether it is voice, media or technological sound.

**Sound shifts** - Even though users most likely weren’t able to detect many of the subtle changes in sound, they perceived some of them and paid great attention to the shifts of the sounds around the space. This concept relates to the idea of keynotes, however it departs from their original definition in that specific, minute change and shifts in the ambient soundscape are highly meaningful and essential to the interaction. This epistemic quality begins to come closer to the conceptualizations of sound-based information transfer within the field of auditory display design, but also reiterates the idea within acoustic communication of extracting information from the very temporal and spectral structure of sound, depending on the context in which it is heard.

**Sound is spatial** - When sound localization was technically possible, participants reported satisfaction and exhibited active attention in following and being followed by sound in space. This concept relates both to the notion of acoustic space from the acoustic communication paradigm, but also to a vast body of research on spatial segregation (Bregman, 1990), and localized auditory perception (Neuhoff, 2004).

**Sound is active** - Sound levels and types of sound mixes often seemed tied up with physical activity and engagement in the play space. The fact that sound facilitates action, especially in situations of full body-engagement has long been recognized in the
fields of alarms and alerts, and subsequently in human-computer interaction domains. Psychoacoustically, the mapping between amplitude (loudness) and activity, fun and engagement seems to be one of the most intuitive ones.

**Sound is ambient** - Sound was also often left as ambience, lingering in space while participants focused on interacting, moving and playing with each other. The ideas of background music, background listening and music-as-environment (Westerkamp, 1990) come to mind immediately, speaking to the proliferation of spaces and context where designed sound lingers in the air, filling the gaps (Cage, 1973). While this concept seems to be at odds with sound’s epistemic qualities within ubiquitous computing environments, I believe this contradiction could be bridged with the idea of managing listening positions – directing users’ auditory attention to different moments of significance.

Observing and reiterating these simple patterns resonates strongly with the concept of mediation between listener and soundscape that the acoustic communication framework refers to (Truax, 2001, p.12). Mediation describes not only how people determine the qualities, characteristics and meaning of the sound feedback afforded to them, but also how sound enables them to socialize, as well as how sound relates to their physical, embodied presence and activity in a space. These workshops also reinforce the idea of listening positions, as defined in the acoustic communication framework. They demonstrate the fact that listening attentions, and as a result interpretations, change dynamically over time, constructing and de-constructing the meanings of sound, and people’s relationship to it; and as a result, different parts of the auditory display take on different significance. This is a very important point to consider, because if we want to design auditory feedback that is informative for ubiquitous computing settings – that is, it requires users’ attention in order to communicate meaning or information – shifting listening positions need to be explicitly and systematically taken into consideration. To that end, it’s important to note when, why and how they shift, in order to attempt to design auditory displays that support these functional, epistemic requirements.

**The socio-ec(h)o Prototype**

To exemplify the preliminary design ideas generated from the first phase of design workshops, I want to introduce the ambient intelligent (AmI) prototype that we
ultimately developed, named socio-ec(h)o. It could be seen as an example of how the approach to sound feedback, inspired both by practices in sonification, auditory display research and concepts from acoustic communication, actually played out within a ubiquitous computing context. The auditory responsive environment was also largely influenced by the outcomes in our design workshops, described in the previous section. The auditory display in socio-ec(h)o was delivering information and direction, yet in an ambient, gradient manner, with a delayed response allowing for dynamic shifts in listener positions while users are physically engaged in playing the game.

Figure 5.5: Participants playing socio-ec(h)o

The socio-ec(h)o game is played by a team of four players and features six levels of increasingly difficult word puzzles solved by coordinated body positions and movements. The environment is responsive to players’ actions through abstract light and sound (see Figure 5.5). Players’ movements are tracked using a Vicon motion capture system (www.vicon.com). A custom reasoning engine was developed to track the game state and infer players’ actions. The goal was to create an ambient feel and function that is in keeping with AmI and supports group interaction. Based on our design workshop results, we felt that the interface required a gradient rather than direct response to players in order to best represent game states and anticipate players’ actions. This approach relates to the strong emergent pattern of sound’s narrative potentials in our exploratory design workshops. This intensity gradient of sound would shift smoothly
between layers of sound and provide not only narrative complexity but also have epistemic qualities in relaying information and guiding player actions through ambient real-time sonification. In its simplest form, the approach can be described as dynamic soundscapes that are not only recognized by players, but in turn interpreted to determine the state of the game (see Figure 5.6) -- how 'close' players are to solving the puzzle. The soundscapes are made of sounds that can be categorized as musical, abstract and environmental and there is a single base soundscape type associated with each level of the game. The idea is that meaningful sound aids in understanding of the game state and the effect of player actions. Just as we employed sound avatar layering in PD Workshop II in the first phase of this research, we felt we could modulate this understanding by increasing or decreasing the rate of intensity or changes in the sound along a gradient. An example of this is the game of "hot and cold", where one player uses words along the continuum of hot to cold to signify the proximity of another player to a solution.

Figure 5.6: A model of our sound and perception approach.

The question then was what sound techniques would work best to signal increased intensity and change along a gradient? This is where existing sonification frameworks came in. While we had realized we needed a rich narrative, ambient sonic display, we also needed it to convey active, dynamic information to the users. Sonification deals precisely with communicating information about fluidly shifting states
or values, in a background, ambient way. The basic idea is to have a sound that doesn’t change in its essence, but one or more of its parameters (pitch, tempo, spectrum/timbre, or other) change noticeably. Usually sonification for data uses simple, synthetic sounds to preclude confusion and ambiguity. We, on the other hand, wanted to not only produce a rich immersive world of sound in our ubiquitous computing prototype, but also to use sound as a game mechanic — thus, ambiguity was a positive, rather than an unwanted aspect. Outlined below are the specific approaches we took and choices we made about the sound feedback in socio-ec(h)o.

**Audio Display Schema**

The audio display for socio-ec(h)o consists of three components: a real-time ambient sonification that has a different soundscape for each level of the game; an anticipatory feedback sound to signal when all participants are working together towards the goal; and a confirmatory feedback sound, which signals the completion of the goal and the progress to the next level (see Table 5.2 for detailed auditory display approaches). Using the acoustic communication classification system, the ambient sonification was designed as a keynote sound that is present all the time; while the anticipatory feedback is a signal, alerting players that the solution is near; and the confirmatory feedback is a soundmark, marking the end of a level. The latter two feedback signals are consistent through all levels, so as to create a coherent expectation of success. These responses work in conjunction with each other. The gradient intensity is applied to the real-time ambient sonification component and builds up to the two types of feedback.

**Table 5.2: Audio display methods used in levels**

<table>
<thead>
<tr>
<th>Level</th>
<th>Method</th>
<th>Description of Effect</th>
<th>Modality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Amplitude</td>
<td>Sound levels slowly come up</td>
<td>Volume</td>
</tr>
<tr>
<td>2</td>
<td>Phaser, Layer fader</td>
<td>Tempo goes up, cross-fade of five sounds</td>
<td>Tempo, Timbre</td>
</tr>
<tr>
<td>3</td>
<td>Layer fader, Pitch shift</td>
<td>Cross-fade of five sounds</td>
<td>Timbre, Pitch</td>
</tr>
<tr>
<td>4</td>
<td>Layer fader, Low-pass filter</td>
<td>Sound cross-fades and becomes muffled-to-bright</td>
<td>Timbre, Association</td>
</tr>
<tr>
<td>5</td>
<td>Phaser, Pitch shift</td>
<td>Sound goes up in tempo and pitch</td>
<td>Tempo, Pitch</td>
</tr>
<tr>
<td>6</td>
<td>Layer fader</td>
<td>Cross-fade of five sounds</td>
<td>Timbre, Association</td>
</tr>
</tbody>
</table>
For example, players in socio-ec(h)o must achieve a specific goal at each level, which is a coordinated configuration of movements and body positions. For example, an answer to the word puzzle "lo and behold" was for all players to crouch low and be still. Further, they must hold this configuration for a short period of time, typically 4-6 seconds. This ensures that the actions were purposeful and knowingly completed rather than an accidental formation. In Figure 5.7 we illustrate players' actions and the mapping of intensity levels and the audio display components, in a test scenario. Reading from left to right, the players explore different combinations of movements and actions. As some players discover a combination, the intensity of the real-time sonification increases, in this case to a value of 2 out of a maximum of 4. As more players discover the solution, the intensity level briefly drops but then increases to the maximum level of 4. At this point, an anticipatory feedback sound occurs signaling to the players that they've done something right and if they keep at it they will complete the solution. However, we see that the players are unable to hold the combination and the intensity level rapidly decays to 1 (in our observations players often stop after an increase in intensity or anticipatory feedback in order consider what they just did). The players quickly resume the configuration and the anticipatory feedback sound occurs again. This time the players complete the solution by maintaining the configuration for the required duration, and the confirmatory feedback sound occurs. The real-time sonification fades away then back as it transitions to the next level. Here, a new soundscape is composed and the intensity level is reset to a minimum.

Figure 5.7: Mapping of the intensity gradient and audio display to sample game play. ©2006, Ron Wakkary, by permission
Discussion of Preliminary Testing Observations

In this section I will talk about our preliminary user testing, focusing on user interactions with a functioning ubiquitous computing environment. The testing included two three-hour sessions with eight participants, and an additional two-hour session with four other participants. The testing included a total of three females and nine males ranging in ages from twenty-one to fifty-nine. Each session began with a warm-up to introduce the concept of puzzles solved through physical action and support through implicit responses. Each team of four played two levels followed by discussions. After all levels were completed, or a total of two hours of interaction elapsed (60 minutes in the short session), the game was stopped and an open-ended group interview was conducted. We have organized our results by first discussing what combination of sound categories and audio processing techniques created recognition, followed by a discussion of techniques that best achieved an intensity of sound, gradient of change, and effective feedback.

Sound Category

The first two game levels featured a baseline of an abstract musical soundscape. Level 3 had an abstract non-musical sound (crackling of rocks), and the last three levels had environmental soundscapes. Level 4 and 5 were represented by fire and water respectively, while level 6 featured a more complex soundscape of a forest that gradually turned from a calm night to a full-fledged storm. The socio-ec(h)o soundscapes evolved from one-dimensional, internally coherent sounds, to complex ecologies. In the course of our user testing, it became clear that the more abstract sounds (such as a crackling of rocks, or highly processed musical abstractions) were hard for participants to recognize. As a result, these sounds posed difficulties in determining the level of change in the gradient response. In the discussion after levels 2 and 3, players commented that they could tell the sound display was intensifying, but found it hard to tell "by how much." In contrast, after levels 4 and 5 (featuring water and fire sounds) players reported a noticeable positive difference in the "sharpness" of the ambient response. The environmental sounds were more recognizable, and players reported an easier time interpreting changes in them. Players also found environmental sound more contextualized and richer in narrative. As one participant commented: "the sound of fire didn’t just get louder, it was different, more intense, and I automatically associate it with
a positive thing, with success — building a big fire!” It seems, if players have prior experience with a sound, they have a better sense of its inherent scale of intensity, i.e. they know what full-blown fire sounds like, as opposed to a faint crackle. This accounts for the difficulty with unfamiliar abstract sound. The players lacked a sense of scale and therefore found it harder to interpret the gradient of change.

Perception of Intensity

Just as in data sonification, socio-ec(h)o uses sound that is coherent in its basic characteristics of pitch, rhythm and timbre within the same game level, and represents success rate in the game by intensifying these basic sound characteristics. In Table 5.2 we detail the range of audio design approaches we experimented with to achieve gradient sound display in each game level.

The Layer fader is a simple cross-fader between 5 layers of sound which could be specifically chosen or pre-composed to represent increased intensity of the same group of sounds (i.e. dripping water gradually changing into a fast river stream, over 5 steps). The Low-pass filter affects the perception of intensity through adding and subtracting frequencies from the sound’s spectrum. The result is a sense of brightness or muffled sound that is quite recognizable. The Phaser approach relies on running a select sound source through a simple phaser, which varies the sound’s phase — i.e. its perceived tempo heard as a pulsation. This approach was made even more complex with culturally associative environmental sounds such as water. Intensity here was related inversely to tempo — shifting from a pulsating beat to full, continuous sound at the completion of the goal state. As research in classical and contemporary psychoacoustics suggest (Neuhoff, pp. 255-259; Walker & Kramer, pp. 154-161), amplitude change, followed by pitch change and tempo change are the most readily and easily perceived sound variance characteristics. Thus game levels are based on these approaches (and combinations thereof) in order of their “ease of perception”. For example, the bi-quad filter’s quality was effective as an approach to gradual change in our design workshops, yet didn’t work as well in preliminary testing. One possible reason for this is the abstract nature of the sound category used with this effect. Low-pass filtering could still be effective in representing a gradient, only with less abstract sounds. The Phaser effect (slowing down or speeding up of a pulsating beat of the main soundscape) produced a perception of change, yet participants reported it hard to the gauge on the gradient
scale. When prompted during discussion, participants suggested pitch shifting from low to high as a clear way to represent rising intensity. This reinforced one of the approaches we had already incorporated in the system.

**Perception of Change**

8-channel spatialization was used to aid and amplify the perception of change in the soundscape. This was done by gradually moving the sound in space in a circular configuration. The audio display cross-faded sounds of different intensity in and out, giving the players the impression that sound is “going away” or “getting closer” to their relative position in space. Participants commented positively on the sound’s spatial movements, saying it made the feedback appear “more crisp” and helped them interpret change easier. Participants responded the most positively to a combination of environmental sound (whether intensifying fire or going deeper into the forest) and a multi-channel diffusion. All participants commented positively on the immersion quality of the play space.

**Anticipatory and Confirmatory Feedback**

An issue that became clear from participants’ comments was the importance of the anticipatory and confirmatory feedback sounds, or ‘reward’ sounds, as we termed them. As players put it, they wanted to know when they were “on the right track,” but they also wanted an indication of when they “got it,” in order to “celebrate after all the hard work.” Recognition of the sound, as well as overcoming masking, i.e. sound has to be heard clearly over the ambience were key issues. The sounds were random pitch variation of two soft abstract sounds (granulated tapping of glass). We played the anticipatory and confirmatory feedback sounds to all participants before the game began. We found that prior listening was sufficient (nevertheless crucial) in creating recognition of sounds when heard again later in the game.

The evaluation of socio-ec(h)o as an immersive multi-modal ubiquitous computing environment clearly has to be tested in a more rigorous manner and with a larger sample. An evaluation was indeed recently completed, however, the vast amounts of rich data have not yet been fully analyzed, and so I restrict my current discussion to the results from the three preliminary user-testing sessions. Still, this preliminary user tests are already a valuable example of observing people interacting with an ambient
intelligent sound display in a realistic setting. There are several salient observations and
discussion points gained here, and while not indicative or generalizable, they are
important indicators of what worked and how, and which aspects of the auditory display
need refinement, in what ways. The experience gathered from this user testing ultimately
led me to attempt to separate out and explore specific individual issues of perception,
recognition and embodied interactions with regard to the intensity-based gradient
approach to sound we developed.

socio-ec(h)o Prototype: Emerging Concepts

Ambient intelligent environments rely heavily on a meaningful and clear, yet
ambient feedback. In socio-ec(h)o we introduced the idea of providing intensity-based
audio display using a gradient approach. In terms of its relative effectiveness and fit
within the system we were encouraged by the results of our preliminary tests. Even in
game levels where the auditory display was less clear, participants were still able to use
the audio as a reliable and informative feedback system throughout the game, and
reported enjoying the soundscapes and feeling immersed. In addition, specific
approaches, such as using environmental sound and certain audio processing
techniques proved very effective in contextualizing the game and providing game state
information to players. Future work would have to include further study of more specified
relationships between the audio processing of different sound categories, and the
auditory perception of recognition, intensity and change within contexts of embodied
interaction.

The anecdotal results from user testing the ubiquitous computing prototype
clearly help inform and improve a future design of such an environment, even though
these lessons are, in this preliminary incarnation, still far from being a formal evaluation
of the saliency of our approach to sound feedback. Nevertheless, they offer a valuable
experience both in designing an auditory display system for ubiquitous computing, and
from observing users interact with that prototype. These two points of knowledge
building once again serve to advance my inquiry into the requirements for an auditory
display framework for ubiquitous computing environments. Analyzing the results from
that perspective makes evident several emergent concepts that were common through
all three user-testing sessions.
Sound is ambiguous (and it is not necessarily a bad thing!) – All three groups reported being on many occasions confused or unsure about the sound feedback, however it was often a matter of degrees of ‘clarity’ rather than complete confusion; and teams did not seem to want to quit because of it, in fact it seemed to invite more attentive listening and feeling of challenge. Again, in terms of relation to the acoustic communication as well as the auditory display paradigms, this concept speaks to the type of everyday listening that people are proficient in utilizing already – making sense of confusing, unclear, complex sonic situations by selectively focusing or shifting attention on different aural elements, and fine-tuning their ear to particular sound structures and qualities.

Sound is associative (carries prior associations) - This is true even for electronic, synthetic sounds, because they are so omnipresent in any environment, however, here associations took on a special importance because sounds were ‘real’ and even recognizable sampled from the natural environment (fire, water). Prior associations and familiarity with sound definitely coloured user experiences with the sonic feedback and their interpretation of its meaning, range, connotations, etc. In fact, interestingly, recognizable sounds seemed, in our preliminary tests, to present a more intuitive range and scaling (to use sonification terms) to users than abstract sounds. This concept also relates to the qualitative psychoacoustic notions in the acoustic communication framework, were associations of sounds determine how they are experienced in the soundscape (as keynotes, soundmarks, sound phobias or sound romances) (Truax, 2001, p.28).

Sound is the voice of the system (sound as a means of accessing and understanding the system’s intelligence, rules and behaviour) - Sound feedback appeared as something that users relied on heavily in order to understand how the system works and responds to them. In a way it showed a conversational style of experimentation and gauging response, in order to learn to manipulate, interpret and be in the environment. Again, this concept speaks to the idea of everyday listening (Gaver, 1993; Truax, 2001) – selectively paying attention to minute sound shifts to gain information about the soundmaker’s state.

Sound as reward – Having feedback at all times, in this case continuous soundscape feedback (a keynote sound), reinforces and rewards user efforts of interacting with the system. This idea of system response or display is at the heart of the
field of human-computer interaction, and auditory display design. Our case of intensity-based gradient soundscape most resembles a sonification scenario where the response is constant and dynamic, and it requires constant attention. The anticipatory reward sound – signaling that users are close to completion served, I would argue, as a soundmark, which users became quickly familiar with, and could instantly identify so as to adjust their actions accordingly. Upon completion of each game level the auditory reward sound – a signal in the soundscape - seemed to be really enjoyed by the players, and important to their leaving the system confident and satisfied, rather than frustrated and defeated.

Redefining Questions and Propositions

“Embodiment is the property of our engagement with the world that allows us to make it meaningful.”

(Dourish, p.126)

socio-ec(h)o exemplifies a design-based research approach, in that it through the experience of designing, as well as from its finished system we are able to investigate the effectiveness, utility and implications of its components with ecological coherency. However, by the time the socio-ec(h)o system was completed, the sound feedback incorporated so many intertwined approaches to change and types of soundscapes in addition to an equally complex theatrical lighting feedback system, that it was virtually impossible to separate out and examine the core issues of interest with respect to auditory display design, so as to build generalizable knowledge for future use about how well it functioned. So, what was learned from the experience of designing an auditory display system for a ubiquitous computing prototype? Following the theoretical discussions pertaining to both the participatory design workshops and the socio-ec(h)o user study, I can now begin to identify important concepts describing the role of sound within ubiquitous computing environments, and form a growing paradigm of interconnections between these concepts. This section deals with redefining the initial propositions about how different notions of embodiment, sound perception, listening and mediation relate to and interplay with one another, and I will now attempt to re-synthesize research questions by directly addressing the approach to auditory display developed in Phase I.

What is valuable at this point in the research inquiry is the ability to articulate and confirm the importance of certain concepts and relationships between sound, context,
embodiment and interactivity, so that they could be further examined. More specifically, the ideas of directive auditory feedback, ambient response, anticipatory and confirmatory sound cues clearly established their significance in socio-ec(h)o, and so deserve further and more focused attention. The umbrella term we used for our auditory display model – an intensity-based gradient approach to sound – rests upon these concepts and relationships, however, before undertaking the next research or design step, there are some gaps of knowledge that need to be recognized and examined. The very idea of sound intensity gradients deserves some consideration first, as well as the aspects of auditory perception pertaining to it. The outlining of these questions and gaps below helps continue the process of building theory and fleshing out a new auditory display framework for ubiquitous computing.

**Sound Intensity Gradients**

Despite the vast proliferation of audio-augmented technologies and spaces, relatively little work has been done to date in designing multi-layered informational auditory displays for responsive environments that actively guide human activity towards solving a problem or achieving a goal (Droumeva & Wakkary, 2006; Takahata et al., 2004, Péres et al., 2005). As discussed already, confirmatory feedback signals when a desired state (or sub-state) is achieved, while anticipatory feedback confirms that users are very close to the solution, but not there yet. Directive feedback, the object of my investigations here, provides a constant, seamless ambient soundscape in the context of ubiquitous computing, which intensifies or de-intensifies to reflect the rate of success of users in a given activity. This type of approach provides a consistent and reliable source of communication with users instead of just responding to actions with positive or negative feedback. If interpreted correctly, users would know not only if they are on a right track within a certain task, but also, how far along they are to completing it. In socio-ec(h)o, this feedback was termed intensity-based gradient approach to sound, since it dynamically moves along a gradient of game progress mapped to sound parameter change.

While intensity-based sound feedback occurs normally in particular everyday circumstances (e.g. paging your cordless phone at home and going from room to room listening to its sound intensifying), there are few studies of these everyday phenomena, and fewer still of their possible translation into design guidelines for sound feedback.
Furthermore, few studies focus on perception of complex everyday changing sound, while taking into account context and purpose of activity, level of embodiment, or familiarity and associations with the sound. A methodological investigation into all these different components of the model is needed in order to understand more fully how to better use and design such auditory displays.

Re-defining the Areas of Inquiry

Auditory perception of complex (everyday) changing sound: This is the central approach of using sound to convey information in the proposed intensity-based gradient model. Existing knowledge comes mainly from sonification, which has a growing body of guidelines and frameworks, adding to the general field of auditory perception (of changing sound), however, its soundscapes are typically synthesized, one-dimensional and abstract, rather than complex and representational (they could be recognized, and have connotations). Issues to be considered here are polarity, scaling, possible spatialization and data-to-parameter (in my case intensity-to-activity) mapping of sound.

The relationship between intensity-based (dynamically changing) sound feedback and the level of embodied interaction: Following the discussion after Phase I, it is clear that embodiment has a salient relationship to the experience and perception of sound within ubiquitous computing. It is crucially important what type of interaction the system supports or requires - whether it is passive (visual or auditory); desktop - involving hands-eyes coordination; tangible - involving haptic interaction in space; or ubiquitous (whole body involvement). Again, sonification has traditionally been investigated for inside spaces of limited embodiment, for places of work, training or monitoring. It should be noted that some contemporary applications of sonification take it outside into mobile devices, albeit still for individual, work or business-oriented use. (Nesbitt & Barras, 2002)

The relationship between intensity-based sound and type of activity: Just to further the "context"-sensitive argument in relation to sound feedback in ubiquitous computing, this point serves to illuminate the fact that a different locus of activity – work, training, play, or learning – yields a different expectation, experience and perceptual-cognitive aptitude to sonic feedback. While my current investigation has been very much constrained to activities of fun and playful learning, research in sonification and other
auditory display domains quickly demonstrates that type of activity is an important contextual consideration to designing auditory displays.

The relationship between intensity-based sound feedback and cognition: Cognitive theory, especially the “hill climbing” problem-solving approach discussed by Newell and Simon (1972) are particularly applicable here in complementing the idea of gradient progress feedback and the notion of experiential cognition (learning-by-doing) as it relates to interpretation of sound feedback. While I won’t be able within the scope of this thesis to really address this relationship, or even detail a larger theory of cognition as an experiential as well as mental activity, this is nevertheless an important concept that has to be acknowledged in a discussion on sound in ubiquitous computing.

The relationship between intensity-based sound feedback, level of embodied interaction and social setting (shared space, where groups rather than individuals interact) – Museums and art installations are becoming increasingly conscious of such issues and take effort to create affordances for public, shared interaction, however, the function of such auditory environments is typically affective, immersive and experiential, less focused on relaying meaningful, informative (or directive) feedback.

The relationship between intensity-based sound and the bigger soundscape: This question includes the use of complex, everyday or environmental, representational sound as a sound base, as well as the overlaying of discrete sonic elements such as soundmark or signal cues on top of the ambient intensity gradient soundscape.

The relationship between sound and ubiquitous computing space (architectural, physical, cultural): This question relates to what the acoustic communication framework refers to as acoustic community, acoustic field and acoustic space (Truax, 2001). Each of these terms describes space either from a socio-cultural perspective (acoustic community) or as a combination of physical characteristics of sound behaviours (acoustic field). From Phase I the experience of space or “place” already emerged as a really important one.

Researching these problems using contextual, ecologically valid methods: Making sure the methods used to investigate these problems and relationships reasonably reflect the complexity and core features of the ultimate intended technology and environment. Again, sonification uses some situated methods, however, it rarely delves into issues of embodiment and sociality, since they aren’t typically part of its
intended application. Research in interaction design affords situated methods of investigations, but rarely conducts systematic investigations of auditory displays specifically for ubiquitous computing contexts.

Narrowing the Scope

It is precisely because of the experience of designing an auditory display system for an ubiquitous computing prototype that I am now able to redefine and re-articulate design concepts of auditory display that would have to be integral to a new framework (see Figure 5.8).

As well, this experience allows me to crystallize and formulate more specific research questions that would have to be answered in response to the uncovered gaps in research into auditory display for ubiquitous computing. It is important to note that being able to point to these gaps is only possible because of having explored, designed
and tested an auditory display system for an actual ubiquitous computing prototype. It is important to recognize that questions resulting from this first phase of research necessarily relate and speak to some of the particular issues of our approach to sound design in socio-ec(h)o – the intensity-based gradient approach to sound. This approach is far from being the only one appropriate or potentially effective for contexts of ubiquitous computing or even ambient intelligent play. However, in the interest of delving deeper into one perspective of sound design for ubiquitous computing rather than attempting to generalize across application domains with their different requirements and approaches, I will hereby focus on exploring further this particular auditory display model – the one of intensity-based gradient sound. What is more, I will not, in the confines of the present inquiry, be able to address all the issues that I have identified above with regard to intensity-based gradient approach to sound. Rather, in the next phase I will focus on examining the use of complex changing sound for conveying information, its relationship to level of embodiment, and the application of contextual inquiry methods informed by design workshops (see Figure 5.8). In order to do so, I pose several key questions that have to be addressed in order to more fully understand, describe and conceptualize a framework based on this model.

1. **Context:** How does level of embodiment influence the perception of different approaches to intensity-based sonic feedback for ubiquitous computing applications?

2. **Auditory Perception:** What is the significance (if any) of using different categories of sound (vocal, musical, abstract, representational) as a basis for an ambient soundscape within a ubiquitous computing space? How do different sound categories influence the perception of scaling and polarity, as well as intensity-to-activity mapping? Are there differences in the perception, comprehension and action with regard to varying different sound parameters in ubiquitous computing environments?

3. **Mediation:** What mediating tactics do people employ to deal with ambiguity of sonic response? How does listening attention to sonic feedback shift in different contexts of embodied interaction?

While these inquiries are general and complex in their current state, the goal in the next phase of my research is to sift through and pose questions in a more focused way, and address them through research by finding and modifying suitable methods.
Design Methods as Basis for Experimental Study Design

The main challenge in investigating intensity-based auditory feedback lies in isolating variables that influence the way people perceive sound in the performance of a given task. Based upon the already noted guidelines for sonification (scaling, polarity, data-perimeter mapping, etc.), as well as upon the aforementioned models of acoustic communication and auditory display design, there are two main points of interest in this next investigation, relating specifically to sound. The first is in the type of sound used to represent or communicate information. Recognizable sounds carry pre-existing associations (Truax, 2001, p.27) and are easier to identify than unfamiliar, abstract sounds. Sounds perceived as annoying may affect perceivability in one way, while sounds that are deemed pleasant may be influential in another (Vickers, 2004; Leplâtre & McGregor, 2004). Further, dimensions of complex sound such as timbre or spectrum (quality of sound, spectral makeup), pitch, amplitude (loudness) and envelope (temporal signature) may also have different effects on users' perception.

The second major area of investigation is type of change representing sound intensity – that is, the data-to-parameter or (in my case) intensity-to-activity mapping. How should sound relate to specific meanings within a system of information? There are a number of characteristics of sound that could be dynamically varied, while still preserving a core quality of a sound or soundscape. Pitch, amplitude, timbre, envelope and rhythm are some of them, and, to complicate the issue, they can often be manipulated in more than one way, or have a compound effect. For example, timbre is affected both by changes in pitch, changes in envelope and effects such as spectrum filtering. Varying the speed of playback of a sound could change pitch, however, filtering of sound (attenuating or de-attenuating certain frequencies) also affects the perceived pitch, as well as the perceived amplitude (Cook, 1999). The best way to fully isolate pitch from timbre and from amplitude is to only use single sine tones for investigation. However, this would preclude harnessing the rich and potentially more effective pool of everyday sounds for feedback in experiential, embodied situations.
As an alternative, I decided to look at design methods, to scaffold my modified experimental study design, in pursuit of situated yet constrained, empirical inquiry. The game activity approach is one frequently used in HCI research as a platform for studying various aspects of user experience, cognition, perception, system usability or interface design (Muller & Kuhn, 1993). One of the reasons is its tight structure and quantifiable outcome – that is, a game activity could be designed in such a way that the only way to "win" is to utilize particular system elements as core mechanics (Salen & Zimmerman, 2003). Investigating the intensity-based auditory feedback model in particular, requires an activity that involves incremental progress towards a goal, which could be mapped to a sound intensity gradient. Participatory workshops often call for low-tech design games, where users brainstorm concepts and build models, usually in groups or teams, relating to the mechanics, affordances and requirements of the technology under investigation.

One of the biggest challenges with design-based research, especially open-ended participatory workshops and design games, is interpreting the "results" from them in ways that are meaningful to the design process and/or research question, and directly address the issues under investigation. Indeed, in the area of user-centred design there is a certain fascination and enthusiasm about participatory design, because for the first time it affords the user the tools and opportunity to express and ideate design, rather than having these imposed upon (Schuler & Namioka, 1993; Taxén, 2004). However, it takes careful planning and moderating to produce rich yet useful data. For my modified experimental studies, I decided to use a participatory workshop format that offers a situated setting but has a more rigid structure and more built-in constraints to help with my analytical generalization goals. Rather than using the Liz Sanders model of providing only raw materials and having users ideate and build from scratch (2001), I implemented the design game approach (Muller and Kuhn, 1993) where I situate users in a pre-determined activity, but allow for informal interactions. This design game approach could allow me to isolate sound feedback as the only type of guiding response within a game, if I make it essential to attaining the goal. However, challenges to data interpretation still remain, and the analysis process would require creative approaches to representing and deriving useful patterns from the rich qualitative data.

Video analysis, audio analysis, conversation analysis, in addition to observational notes and semi-structured discussions make up some of the common methods in the area of interaction analysis (Jordan & Henderson, 1995). Usually, the goal of conducting
such workshops is to further design of a novel artifact or system. In this case, the study attempts to achieve a bit more – answer research questions about sound intensity perception by trying to access tacit perceptual experiences related to intensity-based sound feedback. Only further down the road of investigation would this research inform the design of ubiquitous computing environments. In the current experimental study, the main focus is on analyzing interactions through video annotation (looking at pointer movements in response to the sonic feedback); interpreting the informal discussions transcripts; organizing and comparing sonic graphs; and analyzing and comparing intensity-based audio feedback (the actual sound tracks from user-generated sonic responses).

**Action – Mind – Mediation**

"One of the basic ways in which socio-cultural setting shapes mental functioning is through the cultural tools employed. Mediation provides a formulation of how this shaping occurs, in order to specify how cultural tools exist and have their effects; it is essential to focus on human action as a unit of analysis." (Wertsch, 1991, p.256)

In order to tie this investigation back to the interaction design conceptualization of activity theory mentioned earlier in the thesis, I bring in cognitive psychologists Michael Cole and James Wertsch's mediational theory of the mind, where mind is created and must be studied in relation to context, activity and artifacts. According to this line of thinking, "the mediation of activity through artifacts is the fundamental characteristic of human psychological process and the human environment." (Wertsch, 1991 p.190). Activity, mind and mediation together, as a unit of analysis, form a better point of departure for attempting to analyze complex interactions, instead of as separate variables (see Figure 6.1). The significance of perception in an environment of rich sensory feedback and high cognitive demand cannot be separated from the transformations that occur within activity through embodied interaction, nor can they be divorced from the mediating factor that is the context with its socio-cultural characteristics.
Figure 6.1: A schematic showing the interplay between mediation, perception and action, as nested within their context and physical environment

**MIND: Perception + Cognition**: This aspect of the unit of analysis describes the role of our physical and intellectual abilities to partake in, function, use and transform technological systems of interaction, including arrays of feedback. With regard to sound, this concept rests upon psychoacoustic research, limitations and boundaries of human hearing and auditory perception at a general level (masking, pitch discrimination, just noticeable differences, thresholds of hearing and of difference, auditory scene analysis, spatial segregation). In addition, this concept describes cognitive theories about meaning-creation through sound. Thus it is the analytical approach of the auditory display research paradigm that informs and contributes to this aspect of analysis.

**ACTION: Activity + Interaction + Embodiment**: Human activity as a subject of inquiry and as a focal point of understanding consciousness, knowledge and experience has been discussed by a vast number of philosophers and theorists, so it is beyond the scope of this work to address the concept in its deserving depth and entirety. Even the theorists mentioned above whose work addresses the interconnections between human activity, cognition and context, derive some of their foundational ideas from earlier works.
by Vygotski and Leont’ev in activity theory; and Heidegger, Marx and Merleau-Ponty, among others, in phenomenology. For the purposes of my inquiry relating to design within ubiquitous computing environments, I’ll narrow the lens through which activity is seen by looking at the framework of embodied interaction put forth by Paul Dourish (2001). The main reason for choosing to look there is that his conceptualization directly implicates interaction design, and situates embodiment in the way that is significant in human-computer interaction, including ubiquitous computing. Embodiment, he defines as the quality of being and acting, in a real place in real time. Embodied interaction then “is not simply…an interaction that is embodied, but rather it is an approach to the design and analysis of interaction that takes embodiment to be central to, even constitutive of, the whole phenomenon.” (Dourish, 2001, p.102). Further, using the concept of embodied interaction, rather than simply action or activity, emphasizes the fact that action is not the same as interaction. While action is mediated by context and in turn mediates the meaning creation in a given environment through use of artifacts or other means, interaction suggests a further mediating relationship between human actors and a technological system – in my case, a system that is physically embedded in the real world, a ubiquitous computing environment.

**MEDIATION:** In the context of sound, mediation describes how people interpret feedback, what tools and activities they use to make decisions and perform actions based on sound. Exploration, experimentation, transformation and (re)-mediation of conceptual understandings are some of the mechanisms characterizing activity in a setting with mediated means. Mediation thus requires recognizing and examining the influence that context has on perception and interpretation of sound feedback – what is the task at hand, what type of activity is it, is it social, how long does it last, what type of attention does it require over time? As Benford, Beaver & Caver suggest, well-designed ambiguity is one of the most powerful features of a technological system, since “by impelling people to interpret situations for themselves, it encourages them to start grappling conceptually with systems and their contexts, and thus to establish deeper and more personal relations with the meaning offered by those systems.” (2003, p.233). The idea of mediation also fits extremely well within the paradigm of acoustic communication, enriching the idea of a listener’s mediated relationship with a soundscape through the conceptual addition of the aspect of active engagement and embodied interactivity.
These three dimensions of auditory display design can be accommodated by using design methods, because they not only situate users in an engaged activity that is social, shared and physical (just as the requirements for ubiquitous environments call for); but also generate, through the process of design iteration, multiple points of analysis, such as user responses, embodied actions, user-generated artifacts and observable patterns of interaction, communication and socialization in a shared physical space, that provide a richer picture of investigation and analysis.

Two Experimental Studies

In computing science, human-computer interaction and psychology, some see disadvantages to a qualitative way of approaching auditory display design, especially using design methods, as a result of which these methods are yet to gain popularity and become commonplace in these communities. First, a low-tech workshop (what many design methods call for) by definition means a lack of absolute consistency in the feedback – being manual, every time it is provided it is slightly different, because of variation both within and across individuals’ approaches and physical abilities. In addition, because engagement is an issue in the context of an informal embodied activity (game or non-game-based), there is the added complexity of designing a suitable engaging workshop activity. This adds enormous value to the quality of the interaction, however, it also creates more considerations – such as, was the activity fun or engaging enough? There is rarely time for repetitive trials of the same type of feedback being provided to each participant to ensure statistical generalizeability – a problem in the natural sciences but not a goal of the present inquiry, as my concern lies with ensuring ecological coherency. The informal participatory approach can provide surprising depths of information, revealing important and at times salient patterns of user interactions in relation to directive sonic feedback. The idea of allowing users to represent intensity themselves, instead of imposing feedback on them and recording responses, is that this forces users to externalize their tacit, latent ideas about sound intensity with regard to all of the relevant elements from the emerging framework – scaling, polarity, approach to intensity gradient, spatialization and embodied interaction. The units of analysis then become user interactions as they represent intensity, offering a glimpse into their understanding of both sound parameter change and intensity-to-activity mapping. User actions in response to a manually-generated sound intensity feedback could also be observed as an indication of how embodied interaction plays out as a factor in the
activity's ecology. An exploratory, experimental study informed by design methods then, would likely both provide a participatory, situated setting as well as some critical experimental constraints that allow deriving empirical evidence with some potential for analytical (in the absence of statistical) generalization and validation.

Unlike the first research phase, which involved situating users and design/research activity right into a simulated or realistic ubiquitous computing environment, the workshops presented in this phase occur in simpler, everyday settings, (almost) free of any technology. The idea is that despite moving the research inquiry farther away from the intended application setting these low-tech workshops would still preserve and emulate the type of physical and social engagement that is at the core of interactions with ubiquitous computing technologies. Further, low-tech workshops have the advantage of being cheap and easy to run (even Wizard of Oz studies require a good amount of technical setup and preparation, while fully functioning prototypes are a result of months, even years, of work). Low-tech workshops require users to generate or interact with everyday artifacts, thus reducing preparation, as well as the learning curve and competence connotations associated with technology.

Summary

There are two workshops presented in this section, as well as a pilot test informing the experimental design of Workshop I. In order to address some of the issues identified above, I used the format of participatory workshops (Schuler & Namioka, 1993; Muller & Kuhn, 1993) to engage a group of participants in a fun, informal activity where sound plays a major role, so that I can observe emergent patterns in their interaction with sound, with each other, with the space and with the activity. However, because of my interest in level of embodiment and social dimensions of interaction, I introduced several distinct iterations of the inquiry, as exemplified in workshops I and II. The first workshop, involved a group of people who were physically (full-bodied) engaged in the activity, working in teams and freely interacting in a physical space together. The activity was set in a real-life everyday setting and the objects used, including the sound-producing tools, were all physical, tangible objects. Workshop II, on the other hand, was delivered through a desktop interface, on an individual basis. The reason for introducing these differences in the study protocols was twofold. One, I was interested in seeing whether the user experiences with auditory perception of complex changing sound will...
be consistent despite the pronounced contextual differences. Two, I wanted to explore precisely how user interactions with sound, and with the activity, may be affected by the differing levels of embodiment and social engagement.

**Pilot Study and Lessons Learned**

I started out wanting to design a low-tech workshop where I can examine people’s perception of intensity-based sound, roughly the way we had conceptualized it in socio-ec(h)o, within the confines of a fun, informal activity. Even though I was not organizing a controlled experiment, I still wanted to attempt to isolate interactions with sound as an observable research objective. I wanted to try and explain how intensity-based sound is perceived and works in situations of embodied interaction. The idea of a game came to mind, because games have a tight structure and a "quantifiable outcome" (Salen & Zimmerman, 2003) so that if sound were made to be essential to winning a game, I could use this type of success in gauging how well sound worked as feedback. Because of the gradient, persistent, ambient nature of the auditory feedback I was interested in, the activity couldn’t be just any game. It had to be informal and engaging that at the same time is goal-oriented in an incremental fashion, suitable for a gradient sound feedback; and one whose rules are formalized enough to allow systematic observation and analysis of interactions. After contemplating a few ideas, with the help of friends, I came up with a geography trivia game, where secret locations on a physical World Map have to be discovered using only sound as guidance. A player would start by tracing a pointer over the map, and the researcher (me) would provide incremental sonic feedback, using a simple sound-making instrument, signifying how close the player is to the goal location. I had a chance to informally test out this activity with a group of children, and I treated that as a pilot study that would tell me if the activity was appropriate and if anything can indeed be learned from using sound feedback in such a fashion.

In the beginning of the pilot study I was the one providing feedback using one of three instruments – a kazoo, egg shakers and clave sticks. To me these instruments signified pitch-based, timbre/tempo-based and clean tempo-based approach to intensity, respectively. Ideally I wanted to test out each instrument with each child, but soon realized this task will be incredibly time consuming and children’s attention spans would not allow it. The way the game worked is I had made up secret trivia questions, referring
to finding the homes of non-existent animal species. This was to preclude users from finding the goal based on prior knowledge. Children had to find the secret location by tracing a pointer over the map and listening to the sound feedback. When they found the location I had another helper play a chime note on a musical triangle. The idea of the 'reward' sound had proved quite important in our socio-ec(h)o prototype testing already, so I wanted to see if it would elicit any comment or response here.

The experience from the workshop was very encouraging to me, as children were engaged and seemed to easily grasp and effectively use the intensity-based sound feedback to play the game. They were very positive about the gradient feedback, reported it being more fun than if someone was giving them verbal feedback, and liked the game as well. Interesting to note are also the rich embodied interactions and socialization patterns that I observed here. While the game was meant to be turn-taking, children were easily becoming involved with each other's turns, giving advice, pointing, laughing, walking over the map and talking over each other. Two things that came to mind were that the physical set-up afforded for these fluid bodily and gestural interactions (pointing is not a motion that can be done in a desktop situation, for example). The ambient, yet consistent sound, on the other hand, seemed to afford for competing verbal interaction without obliterating its meaning and significance as directive feedback – another feature, vital to a ubiquitous computing context that is shared and social.

Just when I thought the session was over, children asked if they could produce the feedback themselves, using the instruments (almost all used the egg shakers) and take turns playing. I hadn't planned for this, but since this was a pilot, I decided to go along and let them, to see if any interesting observations could be made. Indeed, while watching children provide feedback for each other, I really started noticing emerging patterns of how they interpreted the intensity gradient as related to the spatial layout of the map. As well, there were patterns of the range of feedback they provided, and in the level of discrete gradations they used to signify success. In the end of this portion, I decided that the activity was well-suited for my needs, and that allowing users to generate feedback is a much richer, and worthwhile source of observation for me as a researcher, than providing it myself and logging their responses. Thus I moved on to designing another experimental study, incorporating the lessons I learned from the pilot,
and using a similar protocol, artifacts and setting. Below is a detailed description of the study and its results, intertwined with observations and descriptions from the pilot.

**Workshop 1: Geography Sound Map Game**

**Research Questions:**

As articulated after Phase I of the research, there are a number of specific research questions that need to be addressed with regard to investigating an intensity-based gradient approach to sound within contexts of ubiquitous computing. As shown on Figure 5.8, I am focusing only on a few in the scope of the present inquiry. They are:

1. Perception of complex changing sound of different categories (environmental, abstract, musical, etc.) and using a different approach to intensity (such as tempo-based, pitch-based or amplitude-based sonic changes): Can users correctly and consistently judge intensity change, and how do different approaches affect perceivability and effective interpretation of the feedback? (Efficiency here is defined as both an ease of perception, and an intuitive translation to activity requirements.)

2. Using (participatory) design methods - What can be gleaned about participants' tacit knowledge of intensity-based sonic feedback from allowing them to provide that feedback themselves? Is there a way of establishing a basis for comparison of intensity/change auditory perception across experimental study iterations?

**Workshop Setting and Structure**

To create a situated, contextual study for investigating issues of intensity-based gradient sound feedback, I organized a participatory workshop in two parts. In the first, I collected written responses of a passive sound listening exercise, representing sound intensity changes (explained in more detail below). In this part I wanted to examine passive perception of various approaches to sound intensity, as another condition of level of embodied interaction) and attempt to find patterns of commonality or difference among participants. For the second part I used the same geography guessing game from the pilot workshop, where the puzzles are secret locations hidden on a physical poster-size map of the world. In this turn-taking game, one player decides upon a secret location and helps their team-mate find it. A tangible pointer device was used for tracing across the map and seeking the hidden goal. The only form of feedback that one team-
mate uses to guide the other is generated by hand (by the first player) using one of three provided instruments – kazoo, egg shakers, and clave sticks (see Table 6.1). In this part I wanted to find out what can be gleaned about participants' tacit knowledge of intensity-based sonic feedback from allowing them to provide that feedback themselves?

Table 6.1: This table shows the three instruments used in the workshop and their respective intended approaches to intensity

<table>
<thead>
<tr>
<th>Audio Display</th>
<th>Approach to Intensity</th>
<th>Polarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kazoo</td>
<td>Pitch Shift (Complex Tone) + Amplitude</td>
<td>Positive</td>
</tr>
<tr>
<td>Clave sticks</td>
<td>Tempo Shift + Amplitude Clean Timbre</td>
<td>Positive</td>
</tr>
<tr>
<td>Egg Shakers</td>
<td>Tempo Shift + Amplitude Rough Timbre</td>
<td>Positive</td>
</tr>
<tr>
<td>Triangle</td>
<td>Confirmatory Feedback</td>
<td>N/A</td>
</tr>
</tbody>
</table>

I decided to introduce the passive graph plotting exercise, as a preamble to the physical geography game activity, in order to explore intensity-based sound, especially sound that is complex and/or environmental, and involves intensity gradients with several types of changes (pitch, as well as amplitude and spectrum). Participants were given forms with 16 blank graphs (see Table 6.2.) and asked to listen to 16 sound excerpts of 10 to 30 seconds each that represent some kind of intensity change of sounds ranging from simple sine tones, to complex environmental or abstract sampled sounds. They were asked to both graphically plot the change they hear, and write a short one-sentence description of the excerpt. The scaling, polarity and sound content of the complex sounds in this section of the workshop were directly borrowed from the auditory feedback model in socio-ec(h)o, for consistency reasons (Droumeva & Wakkary, 2006). Prior to starting the game activity participants were reminded of the children’s “hot and cold” game analogy as a metaphor for the mechanics of the activity. They were then told to use only sound generated by one of the instruments of their choosing to help their team-mate find the secret location. The instruments were musically simple to mitigate skill and experience factors (virtually no learning curve), and each had a distinctly different sound quality and intensity constraints (e.g. while a kazoo could perform obvious variations in pitch, tempo and timbre, claves and shakers have
virtually no variation of pitch, only of amplitude and tempo). All instruments were briefly modelled just to ensure that players knew how to use them.

This low-tech, participatory format was crucial to my core research question in this study, which was – how would people both generate and interpret sound intensity gradients in relation to a goal-oriented embodied activity? Would I see the same patterns of representing and using intensity-based sound feedback emerge here as they did in the pilot study? Would I see similar patterns of mediation and embodied interaction as I saw in the workshop with children? Methods of data collection for the workshop included observation, semi-structured discussion, paper sheets with plotted sound intensity graphs, and video/audio capture. There were a total of 4 users, all female, ages ranging from 30 to 51, and a total of six iterations of the game activity, including two blindfold turns, initiated voluntarily by two of the participants. Only one user had previous musical training, and all reported normal hearing.

In order to make explicit my own concepts and expectations of sound intensity feedback with regard to the geography map game, I created a framing schema that could then be used for analysis of the workshop results. The schema was partially borrowed from the pilot study. From it, I took the interaction patterns that proved most important and organized them in order to see if they surface again in this iteration of the workshop. First the sound types (instruments) and sound parameters (pitch, amplitude, tempo) were identified with regard to the musical instruments provided (Table 6.1.). Of course, users were free to deviate from the intended approaches and descriptions, but I thought that having a pre-existing schema would facilitate identifying these instances and analyzing them later. In addition, Figure 6.3 shows the expected mappings between the 2-D geography map artifact and sound intensity feedback. Specifically, physical distance from the pointer device to the secret location at any one time of the game is naturally mapped to the gradient sound intensity. Theoretically, if one is far away from the goal, intensity is low, if they are close to the goal intensity is higher, changing gradually following the trajectory of the pointer device. The concept of trajectories (pointer leading towards or away from the goal) was identified from the pilot study as another useful concept. The basic idea of it is demonstrated in Figure 6.2. For example, Trajectory 1 starts far from the goal, yet moves towards it, so feedback should intensify, while Trajectory 3 starts closer, yet moves away, so feedback should de-intensify, even though it has started from a higher position on the gradient.
Sonic Graph Plotting Analysis

The passive graphic plotting task did appear to be an effective way to isolate and test out a number of key sound characteristics, intensity gradients, sound categories and ranges in a way that still allowed users to externalize and represent their perception and understanding in an intuitive way, albeit less embodied and interactive. 16 sound excerpts were tested. Because this section was not structured as a tightly controlled experiment, there wasn’t sufficient variable separation and validated sequence modelling of the excerpts presented. However, it is still important to identify the theoretical underpinnings of this section of the study. Aside from the concepts of scaling, polarity and data-to-parameter mapping, coming from the sonification framework, I also grounded my expectations in some of the traditional auditory perception research. Typically, I expected participants to have no problem identifying gradient changes in sine tone excerpts, and to have a much harder time identifying changes in complex everyday and abstract sounds (McAdams & Bigand, 1993; Cook, 1999). Also, I expected amplitude changes to be the easiest to identify, followed by pitch, them tempo and finally spectrum-based timbre shifts (Walker & Kramer, 2004, pp.154-161).

Sounds started from simple/sine tones with straightforward parameter changes and moved towards complex, everyday sounds with more subtle and compound parameter changes. The first two excerpts were sine tones, and they were partially used...
as a way to establish 'sound intensity competence' – which all users seemed to possess (see Table 6.2, line 1) as all correctly identified a rising and falling sine tone. It was also a test to see if user-generated sonic graphs could be compared to one another with any degree of consistency. Once again, the answer appeared to be yes. The rest of the sounds were more ambiguous, involving degrees of greater or less change, as well as featuring sound categories that varied from musical, abstract to environmental. There is still so much to be learned about perception of complex, everyday changing sound that investigating all of its variations via design methods may be extremely time consuming and difficult. This graph plotting sound identification portion of the workshop was conducted as a way of exploring perception of complex changing sound, especially to see what kind of analysis the sonic graphs and captions may lend themselves to. It is an efficient, quick way of testing out a large number of sound intensity samples using standardized reporting instruments. As a passive identification exercise, this section employed minimal embodied interaction, in contrast to the game workshop activity, and so presented an opportunity for comparison on the basis of level of embodiment.

I saw graphing, rather than choosing a range or number, as a more direct way of getting at users' perceptual experiences. I also encouraged them to freely draw sound as they 'hear' it, hoping to observe some interesting representations reflecting the different timbres and qualities of sounds presented. The rigid structure of a graph makes it possible for different graphs to be compared directly. While I did not know what kind of results I might get, the examples presented in Table 6.2 demonstrate that even with a small sample and noticeable variation in drawing styles, definite similarities and common patterns can be discerned when comparing user representations of changing intensity-based sound. Full breakdown of all 16 entries could be found in Appendix 1. However, it was still important that users had some space to write a short descriptor for each sound, as some sounds might have been too complex to graph, and verbal description could better represent what participants heard. This turned out to be a wise decision, because when reading and comparing the graph-plots and transcripts, it was discovered that users often drew different-looking graphs, but expressed similar experiences through words, and the other way around (see Appendix 2). Specifically, the sonic graphs both confirmed and contradicted some of my underlying expectations of auditory perception of changing sound. While sine-tone excerpts were indeed easily and consistently identified by all users, whether they used pitch or phase (tempo) as the parameter of change, when it came to complex sound, tempo and spectrum-based changes elicited
the most accurate and consistent representations, while pitch seemed more elusive to interpret. Abstract sounds fared worse than recognizable complex sounds, in all aspects of parameter change. Curiously, while users were sometimes slightly inaccurate regarding gradient range and even direction change (whether intensity is increasing or decreasing) in simple sounds, they were surprisingly consistent and perceptive when it came to some complex sounds with more than one intensity change per excerpt.

Because of the seeming gap between psychology studies of auditory perception and their application into context-sensitive perceptual frameworks for sound, making sense of the data has been a ground-up, iterative process. Following the methods of interaction analysis I have constructed a visual comparison schema that includes the actual (intended) intensity of the sounds, and a composite of each participant’s plotted graph of it (Table 6.2). My provisional suggestion is that this representation is in itself a valuable contribution towards developing non-traditional forms of data analysis in exploring changing everyday sound that often reveals aspects and perspectives of research that could otherwise be missed.

Table 6.2: The table shows a few examples from the sound plotting section of the workshop. On the models to the left the X axis represents time in seconds, while the Y axis represents intensity from 0% to 100%

<table>
<thead>
<tr>
<th>Type and Degree of Change (intended)</th>
<th>Participant-Generated Graphs (interpreted)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sine tone at 100Hz rising steadily to 440Hz</td>
<td><img src="image" alt="Graph for Sine tone" /></td>
</tr>
<tr>
<td>Pulsating sine tone of 220Hz increasing in tempo from 12% to 87%</td>
<td><img src="image" alt="Graph for Pulsating Sine tone" /></td>
</tr>
<tr>
<td>Type and Degree of Change (intended)</td>
<td>Participant-Generated Graphs (interpreted)</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>Fire sound through a low-pass filter. Starts muffled and quickly is unfiltered (0% to 100%)</td>
<td><img src="image1.png" alt="Graph" /></td>
</tr>
<tr>
<td>Ticking clock sound through a low-pass filter – starts crisp, then becomes muffled then crisp again.</td>
<td><img src="image2.png" alt="Graph" /></td>
</tr>
</tbody>
</table>

**Geography Map Game Analysis**

In the geography guessing game, several categories of interest were identified for analysis – the user-generated data-to-parameter mapping, scaling, polarity, and chosen sound type. The distinction between location/distance versus trajectory-based approach to feedback was also a point of investigation. This pattern had surfaced already in the pilot study, so I was wondering if it would be present again. It illuminates whether users understand the concept of a progress gradient, the way I had intended it, as per the reasoning built into socio-ec(h)o. Basically, my operational assumption was that there always is a gradient present, and it just intensifies or de-intensifies according to the physical location of the cursor in relation to the desired end. The gradient should also start at a sound intensity reflective of the initial position of each trajectory (again relative to the goal). Instead, what I started noticing in the pilot study was that users provided mostly minimal feedback in the areas of the map that was far from the goal, and all of a sudden intensified the feedback a lot within the vicinity of the goal. So if someone was in a small vicinity of the secret location they really could tell, but if they were far away there was not much indication of which way to head. The ideas of location and trajectory did indeed describe a pattern that I observed here in the user
performances in this workshop iteration. Specifically, they tended to provide relatively unchanging, medium-level feedback while the other player was far from the goal, and rapidly intensifying feedback within a small vicinity of the goal (see Table 6.2). In my interpretation from both the pilot study and the current one, users tended to focus more on the physical location, rather than on the more abstract idea of adhering to a progress gradient – how close the player is from the goal, versus, where the player is on the map.

<table>
<thead>
<tr>
<th>Mapping of sonification parameters to a 2-D world map</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intensity → Trajectory path on map</td>
</tr>
<tr>
<td>Polarity → Positive</td>
</tr>
<tr>
<td>Scaling → Content and sound parameter dependent</td>
</tr>
<tr>
<td>Reward → Location-based on map</td>
</tr>
</tbody>
</table>

Figure 6.3: Mapping of audio parameters to the 2-D map.

A new observation made here was that there seemed to be an interesting combination of location and trajectory-based approaches to feedback. This was different from the pilot workshop in that there were more distinct gradient levels of feedback, ranging from complete silence, to low-mid, high-mid, high and extremely high intensity of sound (see Table 2.). Also, all players insisted on starting in the middle of the map and performing the intensity gradient from there. All four iterations of the game finished in under several minutes. Participants appeared to and reported navigating the map while listening to sonic feedback with ease. Everyone got a chance to provide feedback as well as receive it, and again all seemed to easily and intuitively grasp the idea of a guiding intensity, even though they did it in their own way. It is interesting to note that even though silence was never modelled in either workshop as a form of feedback, both sets of participants from the previous and the present workshops intuitively defaulted to it. One user even remarked that “it is really clear when you are not close, the instrument doesn’t make any sound” and so she suggested making a constraint that “silence isn’t a sound that you can do” so that it would push people to provide more specific feedback. The same user remarked during the blindfold session on her experience of having to completely rely on the sound because she didn’t even have the visual orientation of the
map. This, she described as a much more rewarding and sensory-rich experience, than the sighted one. Yet, curiously, she did happen to begin on the right trajectory right away, and quickly found the goal.

The second blindfold session was very interesting, because the player did not start on the right track, in fact she kept getting sidetracked in the wrong direction, and it became evident just how important the feedback was to her activity, and how the feedback provider struggled to support her quest towards the goal. From my perspective, this was arguably the most interesting iteration of the game, both with regard to user performance with intensity-based sound feedback, but also from the perspective of tacit user-driven approaches to intensity-based sound. As I mentioned earlier, in order to analyze the game’s interactions I used a video annotation approach focused on both the trajectory movement in time – pointer traces (video), as well as on the intensity-based sonic feedback in time – the actual sound track (video/audio). However, for this last rich episode of data, I felt that in order to represent the analysis in a way that does justice to the quality of the observation, I needed a more creative method. Again, design analysis methods came to the rescue. There are a number of design methods focused on constructing narratives, performances and role-playing (Caroll et al., 2000; Iaccuci, Iaccuci & Kuuti, 2002; Laurel, 2003) in order to create a fuller, more empathic vision of a user, technology or context. I decided to use a modified scenario-based account of this interaction (Caroll et al., 2000) to generate a story retelling both sides of the exchange so that relevant issues might surface through the narrative (see Appendix 3).

To analyze more specifically the characteristics of the sonic feedback provided (in terms of scaling, polarity and intensity-to-activity mapping), audio files were created from all six instances of the game, with specific attention being given to their graphical waveform patterns, in order to attempt and hear/see how scaling was used by different players and whether it went roughly through the same degree of intensity or not (see Figure 6.4). Similarly to the pilot workshop, scaling varied a bit, but overall seemed to reach a maximum of a certain capacity for all, no doubt related to human abilities – one could only shake a shaker so fast and so loud. And of course, human abilities differ; hence slight differences in the degree of the scaling range were present. Again, partly because of physical effort, but also perhaps because of intuitive perception, the rate of change, whether it was pitch or tempo, was always tied to amplitude as well. This
finding, together with the strong affinity towards silence [no feedback] as a default state aligns with research in psychoacoustics, showing that amplitude is the strongest cue to sound change/intensity. User-generated polarity, as in the pilot, was also always positive – again compliant with research in data sonification and contemporary auditory perception (Walker & Kramer, 1996). It is worth noting, however, that participants here, just as in the pilot study, showed a high affinity for tempo-based change, and commented positively on its qualities as a guiding feedback.

Table 6.3: Coded results outlining participants' observed and/or reported use of the audio feedback devices and approach to representing intensity

<table>
<thead>
<tr>
<th>#</th>
<th>Audio Display Device + Approach to Intensity</th>
<th>Approach to Intensity - Scaling (Feedback Progression)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Shaker (Tempo and Amplitude)</td>
<td>No Sound – Faster, medium Sound – Very fast, loud sound</td>
</tr>
<tr>
<td>2</td>
<td>Clave sticks (Tempo and Amplitude)</td>
<td>No Sound – Medium tempo – Faster, louder – Very Fast/loud</td>
</tr>
<tr>
<td>3</td>
<td>Kazoo (Amplitude reported) – pitch and tempo observed</td>
<td>No Sound – bursts of tonal sounds (high and low pitch) – Loud, droned out high pitched sound</td>
</tr>
<tr>
<td>4</td>
<td>Clave Sticks (Tempo observed)</td>
<td>Slow – Medium – Faster – Very Fast</td>
</tr>
<tr>
<td>5</td>
<td>Shaker (Tempo)</td>
<td>No sound – Slow – Fast – Very Fast</td>
</tr>
<tr>
<td>6</td>
<td>Clave Sticks (Tempo and Amplitude)</td>
<td>No Sound – Very Slow – Medium/louder – Very Fast/loud</td>
</tr>
</tbody>
</table>

Workshop Discussion

Rather than attempting to advance general knowledge about an absolute efficiency hierarchy of timbre-based versus pitch or rhythm-based feedback, I wanted to use this workshop to build upon the lessons from the pilot test in an increasingly nuanced and fine-grained way. What has proved most interesting already has been the process of allowing users to generate intensity-based sonic feedback themselves, in the context of a playful activity. That is where novel insights emerged with respect to the latent perceptual-conceptual mappings between activity, representations and intensity-based sonic feedback. The research challenge was finding analytical approaches that enabled the parameters of interest to be teased out from the record of that activity. As mentioned earlier, analysis of design-based research workshops for interaction design is
largely understudied and lacks stable guidelines, even though, as suggested, it affords significant opportunities for novel results.

Figure 6.4: Screenshots of each iteration of the map game with its corresponding audio waveform track.

I believe this study contributes to the design of innovative ways of facilitating visual analysis of user-generated data (in this case audio and video footage). Seeing things in alternative ways can sometimes reveal or elucidate issues previously overlooked, or that simply could not be discovered using other, more conventional methods of analysis. A viable future direction, for example, could be analyzing user-generated audio captures in more detailed ways, breaking down their temporal and frequency structures and comparing these patterns across cases. The process of examining sound through user-generated graphs also seems promising. As well, there could be considerable potential in developing a modified scenario-based approach with
which to explore retelling of research activities as narratives in order to capture descriptive, yet also intuitive, analytical and observational notes of the interactions. The idea is that involving users in such ways falls within and furthers design-based research methods, and as such, will be able to viably function within the paradigm of ubiquitous computing.

**Workshop Discussion**

Rather than attempting to advance general knowledge about a salient perceptual or cognitive efficiency hierarchy of timbre-based versus pitch or rhythm-based feedback, I wanted to use this workshop to build upon the lessons from the pilot test in an increasingly nuanced and fine-grained way. What has proved most interesting already has been the process of allowing users to generate intensity-based sonic feedback themselves, in the context of a playful activity. That is where novel insights emerged with respect to the latent perceptual-conceptual mappings between activity, representations and intensity-based sonic feedback. The research challenge was finding analytical approaches that enabled the parameters of interest to be teased out from the record of that activity. As mentioned earlier, analysis of design-based research workshops for interaction design is largely understudied and lacks stable guidelines, even though, as suggested, it affords significant opportunities for novel results.

I believe this study contributes to the design of innovative ways of facilitating visual analysis of user-generated data (in this case audio and video footage). Seeing things in alternative ways can sometimes reveal or elucidate issues previously overlooked, or that simply could not be discovered using other, more conventional methods of analysis. A viable future direction, for example, could be analyzing user-generated audio captures in more detailed ways, breaking down their temporal and frequency structures and comparing these patterns across cases. The process of examining sound through user-generated graphs also seems promising. As well, there could be considerable potential in developing a modified scenario-based approach with which to explore retelling of research activities as narratives in order to capture descriptive, yet also intuitive, analytical and observational notes of the interactions. The idea is that involving users in such ways falls within and furthers design-based research methods, and as such, will be able to viably function within the paradigm of ubiquitous computing.
Experimental Study II: AmbientSonic Map

Alongside the low-tech workshops, I also developed a computer-based prototype, which uses the same approaches to auditory display and design game activity, in order to separate and test out different approaches to intensity-based sound feedback. Once again, I was interested in how different categories of sounds (environmental, musical, vocal or abstract) influence perception of the intensity gradient of sound. As well, I wanted to look into shifting different sound parameters, such as pitch, amplitude, phase (experienced as a pulsating rhythm) and spectrum, in order to find out if any of these approaches are more or less intuitive for interpretation. Finally, I was hoping to be able to compare the results from this workshop iteration to the ones from workshop I. In its simplest form, the intensity-based approach requires an understanding of the underlying sound and entails correct interpretation of its changes. The idea is that this understanding is demonstrated through the user’s ability to act accordingly in the game and ultimately solve the puzzles (see Figure 6.5).

Figure 6.5: Illustration of the intensity-based auditory feedback model

The challenge of designing an auditory display that is ecological, ambient and provides rich guiding feedback through sound (alone or as a primary channel) is in its saliency as a perceptual system that users can reasonably interpret. Thus, auditory perception principles must be pushed beyond the limits of classical psychoacoustics or data sonification design patterns, and into interactive, task-based contextual activities.

Study Structure and Protocol

The AmbientSonic Map game consists of five questions about world geography, and the task is to find a discrete location using only intensity-gradient feedback for direction. The questions and answers are fabricated to ensure that a user cannot find the
correct location without directive sound feedback (based on prior knowledge). Answers to the questions are physical locations on the map (e.g., countries, cities, specific ocean areas). The user has to read the question and explore the map to find the right location, tracing the map with the mouse curser while hearing a dynamically changing soundscape. While it is possible that a user might find the right location by accident, it is unlikely since locations are small rather than large areas (e.g., London) and thus hard to stumble on. This way, users are very much dependent on the auditory feedback to find the correct location.

Each question has a unique single sound attached to it, which is thematically related to the geographic area where the answer is found. For example, one question calls for a specific part of the Atlantic Ocean, and the sound associated with it is of ocean water. The ocean sound is then modulated using a low-pass filter to create a perceptible change in the sound’s timbre (spectrum), as the mouse scrolls across the map image. When a player drags the cursor within a small perimeter of the goal area, a reward sound is triggered to indicate that solution is close at hand. Reward sounds were also thematically related to the questions and soundscape sound – for example, the reward for the “ocean area” question, juxtaposed over the ocean water sound was a short recording of a tall wave crashing at shore.

Sounds that were used and precise approaches to scaling and intensity-to-activity mappings were largely borrowed from the socio-ec(h)o prototype. In the whole system there were three environmental sounds: ocean water, river water and fire; one abstract musical sound; and one abstract non-musical sound. All sounds could be considered “complex,” as they have rich spectra, as opposed to sine tone-based computer generated sounds that only contain one single tone or frequency. Thus ambiguity in the presented sonic feedback was a built-in feature, used precisely to contextualize the display with the subject matter. This is important, as the ultimate goal that this research is contributing to could be a ubiquitous computing environment where richness of experience, narrative and community are design features just as important as efficient and clear information retrieval. The question in this whole exploration really is -- can we achieve both? And how?

The sound feedback approach varied several sound parameters including amplitude, pitch, spectrum and tempo, and used several categories of sound – everyday/environmental sound, abstract musical sound and abstract non-musical sound.
Each sound parameter within any type of sound can be varied relative to the progression of the task. The relationship between sound parameter gradation and activity progress (cursor over map) was linear. For example, if a user is moving away from a solution to a problem then the sound intensity will decrease. If they move towards a solution or become close to completing the task, then a gradually rising intensity sound would be displayed. When a player reaches the final goal or completes the task, a "reward" sound is displayed. This form of final confirmatory feedback was found to be very important in conjunction with ambient sound in the responsive environment of socio-ec(h)o (Wakkary et al., 2005). The translation of sonification principles into game mechanics was again identical to the intended framing in Workshop 1 (see Figure 6.3). This was done in order to further constrain game and system parameters so that user performance can be tracked and possibly compared across the case studies. This kind of formalization also allows for using the same scaling, mapping or polarity mechanism with different core sounds.

**Activity Setting and Structure**

This study asked four subjects to play AmbientSonic Map and perform geographic map identification tasks. The purpose of this study was to further investigate contextual intensity-based sound perception, and its effective application in designing audio displays for ubiquitous computing environments. In this iteration, admittedly, embodied interaction was limited in exchange for greater consistency of the sonic feedback model and game structure. However, I used this fact as an advantage as it allows comparison of the user experience across two situations of different levels of embodiment (the low-tech workshop, and this desktop interface game). The present workshop consisted of four participants, two female and two male, with ages ranging from 27 to 64, reported of normal hearing. They were asked to play the AmbientSonic Map game wearing headphones, and attempt to answer each of the five game questions by moving their mouse cursor over the map and listening to the auditory feedback. Since the game interface has no data collection capabilities, participants' general comments were solicited using a loose Think Aloud protocol. That is, they were asked to comment as they played the game so that spontaneous thoughts would not be lost. I thought this approach to collecting data worked well within the context of the study in conjunction with a post-session interview, because auditory memory is quite short and participants would likely simply forget what they heard if not prompted on the spot. The post-session
interview on the other hand allowed participants to express their overall impressions, final thoughts and suggestions. The sessions were not videotaped, again partly because of the lesser role of embodiment in this activity.

**Workshop Results and Discussion**

Some of the highlights of the workshop results were the unanimous acceptance of positive polarity as mapped to a positive progress gradient. Also, judging from user Think Aloud comments, they all understood the type of change used in each puzzle – whether it was pitch or tempo or spectrum/timbre based. One user suggested that sounds need to be more conceptually related to the puzzle question at hand, as that would aid in the gradient feedback interpretation. All users reported not enough 'change' or 'clarity' of the sound gradient in certain areas to truly signify where they have go next on the map. When prompted to elaborate, one user commented that they had no trouble perceiving low and medium intensity, but the difference between medium and high, on a gradient scale, did not seem significant enough. Another interesting result was the overall feeling participants reported that dynamically changing sounds need to be “simpler” so that changes within the sound can be heard more easily, and one player even suggested there should be only one type of change per sound, and sounds should be single/simple tones. All participants seemed to perceive and interpret very well the bottom and middle sections of the gradient of change, yet, they felt a much more dramatic change in the sound should have occurred when they got within a certain closer range of the solution. As it is, they could not tell if they were 50% on the way to the goal, or 75%. This finding was interesting, as the confirmatory “reward” sound was meant precisely for that purpose, yet players found it hard to even get within that range using the sound feedback gradient alone. Table 6.4 contains the full breakdown of observations as they pertain to the system parameters – intensity mapping, scaling, polarity, sound content and reward sounds. Structuring results that way allowed for some comparisons to be made between participants' performance and helped organize the data.
Table 6.4: Breakdown of participant comments and observations related to each sonification parameter in the game

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear on Positive Polarity</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Scaling – Clear or not?</td>
<td>3 thought it was the clearest, 1 thought it was quite good.</td>
<td>All players thought this was the most confusing type of feedback</td>
<td>All players perceived clear from low to medium, but confusing from medium to high.</td>
<td>All players perceived clear from low to medium, but confusing from medium to high.</td>
<td>1 player thought it was clearest, 3 thought it was better than most others</td>
</tr>
<tr>
<td>Content – Theme, Clarity</td>
<td>Uniform enough, fairly easy to understand how it changed</td>
<td>All players found it too complex and confusing. - heard 2 high/low separate sounds - said sound was changing by itself</td>
<td>1 player liked that the sound was thematically related to the question.</td>
<td>1 player said sound was still too complex to provide clear change gradient.</td>
<td>Feedback easier than other approaches 1 player still heard two separate pitches.</td>
</tr>
<tr>
<td>Reward Identified Ok</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
</tbody>
</table>

Another interesting question – whether the gradient of change (base soundscape and approach to intensity) characteristic for each puzzle question facilitated or influenced performance – revealed several interesting results. Participants’ perception of the quality and ease-of-use of the sonic feedback had a lot to do with the inherent qualities of the core soundscape. Sounds such as water, fire and flat-line musical tones – fairly familiar and one-dimensional sounds – had a much more perceivable gradation of change than more complex, abstract, unfamiliar sounds. This was true regardless of approach to intensity change – pitch shift, tempo shift or spectrum filtering. Since participants did not comment much on the semantic relationship between soundscape and question topic, it is hard to establish whether this was a factor in the perception of audio feedback. Perhaps the most significant outcome of this study was that while players on the whole understood and were able to use gradient sound feedback, there was a sharp difference between their perception of low to mid intensity, and their perception of mid to high intensity. Many commented that a more dramatic difference in the sound’s intensity is needed to signify proximity to the goal. This finding illuminates
the fact that a linear mapping (what I used) between intensity/spatial location as a measure of progress rate, and rate of sonic change may be insufficient to guide players towards a game solution. A logarithmic scale – one that advances in ever-increasing increments, may serve better. It is not surprising that logarithmic gradation of sound parameters may work better as it is a well-known fact in psychoacoustic research that sonic variables have to advance in logarithmic jumps in order to be perceived as linearly progressing (Mathews, 1999, p. 72; Shepard, 1999, pp.151-157). However, I feel this finding is still worth discussing as it not only involves parameter changes of complex, recognizable sounds (as opposed to classical tone perception), but there is also a cognitive, interpretive aspect to this notion of logarithmic intensity – perhaps it has to extend to the conceptual mapping between activity and intensity as well.
CHAPTER 7: THEORETICAL ANALYSIS AND DISCUSSION

All three studies presented in this phase are small-scale, exploratory studies taking a naturalistic experiment format. In comparing them, some important guidelines do emerge about auditory perception of intensity-based sound and embodied interaction. These lessons are, I feel, strong trends that keep surfacing in different exploratory study iterations, and extend beyond problematizing the issue of sound feedback, towards generating potential new angles for future investigation. This way, the amount of data collected and cross-referenced allows an analytical generalization of core findings. Following is a list of emerging guidelines that the second phase of research revealed about the narrowed-down examination of sound intensity gradients as a form of guiding directive sound feedback in situations of embodied interactions.

Experimental Studies Results

Intensity-based gradient approach to sound seems to effectively work as an intuitive system of feedback for performing embodied activities and interaction;

Positive polarity is intuitively related to positive performance in an activity (in all workshop iterations positive polarity was instantly recognized and interpreted, while a negative polarity was deemed confusing and contradictory);

Intensity-to-activity mapping likely needs a logarithmic, rather than linear gradation – a finding that coincides with auditory perception research already, as psychoacoustic variables are known to behave more logarithmically rather than linearly. (Mathews, 1999, pp.72-73). However this notion has to apply to the conceptual mapping between intensity and activity as well, not only between intensity and sound parameter change. This illustrates the common thread between all workshops where participants felt they needed the intensity to increase dramatically within a close proximity to the activity’s solution. It is also reinforced by observations in the low-tech workshops where users provided a rapidly intensifying feedback near the secret location – externalizing the same need for a significant rise in intensity;
Reward is important – as demonstrated in socio-ec(h)o, the reward sound continued to play an important role within the intensity-based model, as a confirmatory signal for completing the activity;

Thematic continuity between soundscape and activity purpose and context is desirable. Most participants across all experimental studies were able to readily identify the representational soundscapes and relate them to the puzzle at hand. Some participants even suggested that having that connection helps in listening to and interpreting the sound’s meaning – another observation very much confirming the acoustic communication model’s proposition that context aids in the subtle recognition of sound quality and temporal structure, and in that facilitates information retrieval from sound;

Amplitude is a strong cue to intensity – As expected, demonstrated by many psychoacoustic studies, amplitude appeared to be the strongest and most intuitive cue to intensity; however, I would mention it was more so when it was a redundant cue, combined with another parameter of change. Amplitude was used in one of the socio-ec(h)o levels by itself, and while it performed well, it did not seem to be the most effective type of feedback, as traditional auditory perception research would suggest.

Tempo is also a strong cue to intensity, following amplitude. However, tempo may be a stronger cue compared to amplitude for information retrieval in situations of highly embodied interaction (see Table 7.1), specifically interpreting the degree of gradient change;

Pitch is reported as being a strong cue; however, in my observations performance didn’t match up. Further, pitch-based change perception works best with simple, harmonic sounds, while, in my speculation, tempo seems to work better with the type of complex, environmental sounds I presented in many of the workshop iterations;

Perception of intensity and interpretation of intensity as representing information such as a progress gradient, are not the same thing. Conceptually, there is still more work to be done, tapping into cognitive theory in order to better understand meaning-creation through sound;

Perception of the middle ranges of a gradient are less accurate than perception of the outer edges of the range (any range containing the minimum or maximum value represented) This is really not surprising given research in sonification and evidence that
users require a reference tone of the baseline value (minimum) to compare it to current gradient sound and aid interpretation (Nesbitt & Barrass, 2002).

Table 7.1: Table comparing accuracy of perception results from the experimental studies, across several different settings of embodiment.

<table>
<thead>
<tr>
<th>Accuracy/Embodiment</th>
<th>Perception Accuracy (Pitch)</th>
<th>Perception Accuracy (Tempo)</th>
<th>Perception Accuracy (Spectrum)</th>
<th>Perception Accuracy (Amplitude)</th>
<th>Tolerance for Ambiguity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passive Listening Task</td>
<td>Medium</td>
<td>Medium/High</td>
<td>Medium/High</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>Desktop Interactive</td>
<td>Medium</td>
<td>Medium/High</td>
<td>Low</td>
<td>High</td>
<td>Very low</td>
</tr>
<tr>
<td>Low-tech workshop</td>
<td>Low/Medium</td>
<td>Very high</td>
<td>Low</td>
<td>Very high</td>
<td>Very high</td>
</tr>
<tr>
<td>Ubiquitous Computing prototype</td>
<td>Low/Medium</td>
<td>High</td>
<td>Medium/High</td>
<td>Very high</td>
<td>High</td>
</tr>
</tbody>
</table>

Represented on Figure 7.1 below, are my combined observations and interpretations relating to perception and performance with the intensity-based gradient approach to sound feedback, in the various contexts of embodiment. On one end of the spectrum we have the passive listening task (non interactive and mostly cognitive), followed by the desktop interface game (interactive with limited physical involvement and feedback), the low-tech game workshop (physical and interactive, but lacking a rich technological component), and finally – the ubiquitous computing environment itself – the socio-ec(h)o prototype. The figure represents only relative relationships, and what I really wanted to show is how visualizing my observations this way helps draw out certain patterns that connect level of embodiment to sound feedback perception. One of the interesting points is how while accuracy and analytical interpretation of sonic intensity was lowest in the low-tech workshop (which was social, shared and physically engaged), tolerance for ambiguity and a preference for tempo/rhythm-based feedback were highest and emerged as clear trends. Another interesting, contrasting point was that tolerance of ambiguity was lowest in the desktop-based study, perhaps pointing to the fact that while we expect challenges and confusions in the ‘real world’ we expect the digital world to be flawless. An unexpected result (and possibly not significant within this scope of inquiry) was the higher accuracy of spectrum-based change (filtering) in both the passive listening and the ubiquitous computing prototype iterations. Tempo-based feedback seemed to do best in terms of perception and performance throughout the workshops,
while, surprisingly (and again, possibly not conclusively from these investigations) pitch did not do as well, in fact in the more embodied sections it clearly lagged behind tempo and spectrum. This was perhaps because of my using complex sound rather than one-tone periodic sound. Amplitude-based change performed just as research suggests – it is the strongest cue to intensity, though I would note that in my experience it performed much better combined with other approaches rather than on its own.

Figure 7.1: Comparative observation results regarding accurate perception of sound intensity gradient sound feedback in the different conditions of embodiment.

Emerging Concepts

What I suggest with the results presented here is that the process of iteration, designing, and prototyping affordances for sound and embodied interaction through the different workshops described, form a sort of triangulation of data that begins to solidify common patterns and trends regarding sound’s role as feedback in contexts of
embodied, social interaction; as well as add to the body of situated auditory perception knowledge. The two participatory workshops in the first phase of the research hold relevance and value in generating emergent, iterative design problems and concepts, and uncovering affordances and requirements otherwise inaccessible through more traditional experimental methods. The preliminary user testing of socio-ec(h)o contributes to building the new auditory display design framework by providing a realistic testing ground of a real ubiquitous computing environment, with a real-time dynamic soundscape engine, delivering directive, ambient sound feedback. Again, the points of research consideration here are the emerging patterns of interplay between sound and embodied user interactions as explicated through the Action-Mind-Mediation paradigm suggested earlier. The second phase of research offers two experimental studies informed by design methods that are used to address more specific research questions by way of introducing relevant constraints while still keeping the nature of the inquiry contextual, embodied and social. Here analysis and research contribution is twofold. On one hand I’ve presented some evidence addressing situated auditory perception of intensity-based sound as feedback for ubiquitous computing, in the form of empirical results, lessons and suggestions. Now, I continue to interpret the studies further in order to address my overarching research questions pertaining to building requirements for a new framework of auditory display design for ubiquitous computing.

Using small-scale participatory workshops may be the long road to constructing a functional model of situated human perception of complex, changing sound in contexts of engagement and active, embodied interaction. Yet such small-scale situated explorations may offer optimal conditions for achieving ecological validity in this area of auditory display design. Beyond designing auditory displays for a ubiquitous computing prototype, and investigating auditory perception of intensity-based gradient sound, what larger lessons or emerging concepts could be drawn that inform and add to the new framework of sound design? Further reflection of my observations, analysis and comparison between the three experiment iterations revealed the importance of several principles, some of which reiterate observations already made after the first phase.

**Tolerance of ambiguity** was one concept that surfaced as an important feature of ambient sound feedback in ubiquitous computing spaces. Changing sound, especially complex changing sound, always has degrees of ambiguity as a form of feedback, however, based on the different experiment iterations, level of active interaction and...
embodiment played a great part in how tolerant users were of this inherent ambiguity. In
the more obvious example, in the desktop version of the geography game users were
extremely vocal about wanting clear feedback, while in the low-tech, as well as in the
ubiquitous computing prototype cases, participants seemed to be quite accommodating
of less than clear auditory feedback (See Figure 7.1).

Another principle that became clear was that **perception is not interpretation** –
identification of sound or even sound intensity does not necessarily result in correct
action. From a human cognition point this principle is not surprising and even though I
am not explicitly examining cognitive aspects of auditory feedback perception, it is worth
outlining how this concepts played out. From comparing the observations and results
from the different sound perception tasks I asked participants to perform in Phase II, it
became evident that often their confusion with the task did not have to do with their
inability to identify the sound, the parameter change or the direction of change; rather, it
had to do with a confusion of what it means – how to interpret the changing sound’s
meaning. The questions then are, what conceptual mappings are intuitive, but also, how
to communicate the meaning-transfer model of the system to users?

Again, not surprisingly, as shown in Figure 7.1, auditory **perception itself
depends upon the context and setting**. This point relates back to the tolerance of
ambiguity issue in that it describes an observation I experienced in all iterations of case
studies in Phase II, which was that user experiences with sound in the different settings
reflected different levels of perception of otherwise similar sounds. However, while
tolerance of ambiguity refers to people’s ability to interpret the sound change and
sound’s meaning, this principle addresses the fact that in different settings their very
recognition and perception of sounds and sound qualities differed. This principle also
feeds back into the acoustic communication theory, which discusses people’s ability at
the micro level of perception to naturally derive information from the structure and quality
of sound, depending upon the context. In other words, it is the context that shapes
listening in such a way that it can fine-tune information retrieval from subtle sound
changes by recognizing patterns in the sound. As Truax points out, it is these patterns
that mediate the relationship between people and environment (Truax, 2001, p.30)

Another point worth mentioning describes a recurring phenomenon that I named
**auditory memory flush**. This pattern was observed as early in the process as the
socio-ec(h)o prototype, and it appeared to persist in the more low-tech iterations with
intensity-based sound feedback. It is also a well-known issue in sonification that several researchers have worked to address. Essentially, auditory memory is fairly short and if people are asked to compare two or more sounds, or continuously compare a dynamically shifting sound in order to derive meaning, auditory memory is fragile. Eventually it gets ‘full’ and people start losing confidence in the just noticeable differences between different sounds or sound portions. Having a bottom-line reference tone is one way to solve this problem. In my case, the way people solved this problem for themselves was that they would knowingly go into a state or area of low intensity, in order to re-adjust their ‘listening position’ and start fresh with interpreting the gradient intensity change from there on.

**Action – Mind – Mediation**

Finally I’d like to return once again to the analytical paradigm of Action - Mind - Mediation as a lens for understanding and analyzing modified experimental naturalistic studies informed by design method approaches. Hopefully such a discussion will elucidate the reasoning and effectiveness of adopting these participatory, situated methods into experimental study design specifically for exploring issues of auditory displays for ubiquitous computing environments. In this conceptualization, **Mind** speaks to the relationship between perception of sound and its interpretation, through both cognitive, as well as physical transformative processes. This second phase of my research addresses more explicitly the perceptual side of intensity-based sound feedback, yet still in a contextual manner – in such a way as to intertwine and examine mind with both action and mediation. As a main point contributing to the framework of sound for ubiquitous computing in this phase, perception is not only different depending upon setting and context, and level of embodiment, but it is also an issue separate and even independent from interpretation. It is an experiential cognition issue, leading us to **Action**, which refers not only to interactivity as a qualitatively different way of experiencing a system or technology, but also to the fundamental understanding that ubiquitous computing needs a body-centric view in dealing with cognition, experience and action (Klemmer, Hartmann & Takayama, 2006, p.140). Embodiment, and embodied interaction add a new flavour to interaction design, especially in its relation to sound. Sound in the natural environment is experienced as a full-body sensation, not restricted to the ears and brain. Both the acoustic communication and the auditory display conceptual frameworks already recognize that. In the present research
investigation, it really became clear how important it was for the sound display system to support embodied interaction and learning-through-doing. In the first phase, for example, participants used their bodies to run, sit, squat, gesture, hold hands, move down to the ground, and many of these actions were performed in order to better understand the sonic feedback. In the low-tech workshops in Phase II participants again used their whole bodies to move over the map, gesture and point to each other in a way that appeared to seamlessly integrate within the structure and sound feedback system of the activity. That takes us to the aspect of **Mediation**, which really refers to the question of how do users make sense of the system for themselves? How do they transform and manipulate their environment or themselves, either physically or perceptually, in order to make sense of it? Mediation is discussed by both Dourish (2001) in his formulation of embodied interaction, and by Nardi and Kuutti's accounts of activity theory in human-computer interaction (1996). In both cases, mediation is seen as an active transformation of artifacts in order to better understand a context, task or system (Nardi & Kuutti, 1996, p.26; Dourish, 2001). According to Cole and Griffin, "[re]-mediation means a shift in the way that mediating devices regulate coordination with the environment" (Cole & Griffin, 1983, p.70). While sound isn't a device nor an artifact, it is a tangible, physical presence in the environment, and one that is highly communicative, evocative and transformative of both activity and interaction. So how does the system of auditory display mediate epistemic interactions and support experiential cognition? For example, in my experience with intensity-based sonic feedback in both the full prototype and in the low-tech studies, the idea of a 'default' feedback state – the state of lowest intensity emerged as a strong pattern of users performing a mediating, epistemic transformation that lets them interpret the system better. In order to perform that auditory memory flush, or simply to re-start their dialogue with the system, they would deliberately regress back to a state of 'default' and then start again, learning the system and building experience along the way. Again in the socio-ec(h)o prototype experience (and to some degree in the participatory design workshops) users often attributed anthropomorphic qualities to the sound display in order to understand it and better motivate their actions and configurations based on it, making comments such as "it likes it when we do this" or "now it doesn't like it!." Within this category I would also suggest that because this type of ambient sonification conveys constant information but does not require constant listening or cognitive attention, it facilitated participants' communication with one another and supported fluid shifting between verbal soundmaking and active listening.
Figure 7.2: A schema showing the progression that the research presented here took, identifying the different emerging concepts leading up to the proposed design guidelines.

Guidelines for Sound in Ubiquitous Computing Contexts

To recall the ultimate question, what have I learned and discovered about requirements for auditory display design for ubiquitous computing environments? What qualities are essential to a sound display system of feedback for a ubiquitous computing environment? How can such qualities be uncovered and supported by research – what methods or modified methods are most useful in investigating situated, specific auditory display issues? The process of articulation of these propositions, questions and concepts with regard to auditory display design emerging from the research is what helps establish the framework, as an epistemology of sound in ubiquitous computing spaces. The new model presented here contains a set of core concepts and guidelines, as a way of synthesizing some of the knowledge and theory-building gained here. As a new hybrid entity borrowing from acoustic communication as well as from the analytical
approaches of auditory display design, the proposed framework not only conceptualizes sound feedback for ubiquitous computing but also suggests methodological approaches of investigating it, emerging from design methods and thus nested well with the paradigm of interaction design (see Figure 7.2). The following section discusses further the emergent properties of sound feedback for ubiquitous computing environments. Auditory displays for ubiquitous computing environment have to support the following functions and possess respective qualities:

Support Epistemic Function: We started out with the assumption that people like to create and derive meaning for themselves from sound feedback, rather than be told authoritatively what to do through sound signals and alerts. Our assumption seemed to hold true in all iterations of research as participants time and again reported having an overall positive experience engaging with sound, trying to figure it out, derive meaning from it. This notion was inspired by lessons from the natural acoustic environment, elucidated both through the idea of everyday listening within the field of auditory displays and Schafer and Truax’s (1977; 2001) accounts of people’s experiences with their natural soundscapes – experiences that are often confusing and ambiguous, but also challenging and rewarding to interpret. In short, both the acoustic communication and the auditory display paradigms emphasize sound’s potential to convey meaning. To elaborate on this guideline, it goes further than simply proposing that sound should be meaningful. As discussed in earlier chapters, meaning could be created in many different ways and take many different forms. Meaning could be emotional, affective, generally experiential, evocative (likes, dislikes), associative (drawing on past memories), or knowledge building. Within the category of knowledge building there are still gradients – an auditory display could facilitate transient knowledge, lasting knowledge, general pattern or trend of information (as in standard sonifications), specific point estimation – guessing exact concepts or numerical values (as in auditory graphs) and other dynamic information-transfer functions. Within the context of ubiquitous computing, I argue, based on my experience from conducting the two phases of research into intensity-based gradient approach to sound, that the type of epistemic function that should be supported by auditory display design for contexts of ubiquitous computing is one of dynamic, continuous information-transfer that is still evocatively and conceptually tied with the context and/or activity at hand, so as to allow supporting other forms of meaning. This kind of ambient, continuous soundscape would be constructed using guidelines from perception research, especially sonification, to allow for actual,
somewhat predictable information transfer and not simply at the levels of experiential meaning-creation. Such an approach to conveying information through sound would still be sensitive to all issues of scaling, polarity, data-to-parameter mapping and other perceptual considerations introduced earlier. Its ambient nature, serving in a way as a keynote, or sonic flow, would still allow even more contextual aural cues such as auditory icons or earcons (soundmarks and signals, to borrow from acoustic communication) while other soundmarks and signal sounds are overlaid on top of the baseline soundscape. An example of applying this guideline would be creating a responsive environment with an ambient soundscape that gradually shifts to convey changes in the space (for example, temperature) while other soundmark and signal sounds come on occasionally to sound reminders and other calls to action.

Support embodiment: The sound designs used in the present research all illuminated the notion that it was important for people to be otherwise engaged physically, while at the same time being able to monitor the sound feedback. In this way the feedback would support experiential cognition, as emphasized by Dourish and others (Dourish, 2001; Klemmer, Harmann & Takayama, 2006). Again, at first look, this guideline may seem self-evident. All sound, in a way, supports embodiment, because our ears are already attached, hands-free, to our heads and we already hear all the time. To come back to the distinction made earlier between hearing and listening, I suggest here that sound in ubiquitous computing environments should support participants’ ability to listen attentively while at the same time be engaged in full-bodied interactions with the system and/or each other. Unlike virtual audio or personal headphone-delivered systems, which require users to pause and listen to artificially spatialized cues, ubiquitous computing happens in a real physical environment, so it is important to point out that it supports free movement in space in a significantly different way. Users should be able to do anything and go anywhere in the environment and still be able to listen to and understand the auditory feedback. An example of this could be a path finding ubiquitous computing environment, where users have to navigate physical obstacles and landmarks, while still relying on a meaningful sound feedback for information.

Support social engagement: In most workshops we found users to be socially and informally engaged and as a result able to share in the perception of the sound feedback together. Again, ubiquitous computing, unlike desktop applications and other personal devices, readily assumes that space, technology and display are shared
between a number of people. This predicates verbal communication and socialization – laughter, raising voices, footsteps and other body sounds, which naturally elevate the ambient sound levels in the environment. Sound display, must then be sensitive to allowing users to socialize and engage in such a way, while still occupying an audible acoustic niche so as to continue conveying information. An example of this could be a family outing at a local museum that features a ubiquitous computing installation. While it’s important for the group to be able to communicate and make sounds, ambient sound feedback could continue to provide useful information.

**Support shifting attention:** Users dynamically gradate between different listening positions and it’s important that they don’t lose or miss crucial opportunities for interpretation of the sound display and potential actions. This guideline further supports the previously mentioned concept of listening while engaging in social embodied interaction in a space. The type of listening positions that users employ will likely shift and change dynamically, as demonstrated in some of the workshops presented here, so that at one time they may be listening attentively, or analytically, while in another they may be listening aesthetically, and in yet another – completely passively and distractedly. This became obvious especially in the first phase of the research where participants clearly took ‘breaks’ from listening to the sound display to communicate, play, move around and socialize. A continuous dynamically updating/shifting soundscape that conveys persistent, coherent information, could ensure that even if users miss whole sections of the display, they may at any point ‘tune back in’ to active listening and explore what their status is, as represented through sound. An example of this guideline could be a ubiquitous computing space with an ambient auditory display environment, and smaller, finite interactive activities embedded in it that users can engage in. These mini-activities may contain their own auditory display system in which case they would “steal” user attention away from the ambient display, however, once the activity is completed a person can easily and seamlessly transition their listening attention back.

**Support emerging mediation tactics:** Users come up with and use different strategies to individually, or collectively interpret the sound feedback provided and respond with actions, whether through verbal communication or movement, or epistemic actions and transformations. More research may be needed to focus on and identify other mediational tactics, but what was observed and discussed in these experiments,
demonstrates several approaches, including moving bodies through space to hear sound better; asking and checking with each other verbally about the meaning of sound; experimenting with different interactions to compare differences in sound intensity gradients; performing auditory flushes (starting ‘fresh’ from the lowest intensity point); relating to the sound display as an intelligent entity with likes and dislikes (“it doesn’t like it when we do this” – from user testing transcripts); among others. An even more explicit example of this guideline could be a situation where users actually have the ability to transform and manipulate sound themselves – by recording, re-playing, moving it around or controlling other aspects of its presentation, in order to help make sense of it as feedback.

Support Individual and Collaborative Exploration: Again, just to reiterate points raised within the descriptions of some of the other guidelines, apart from socializing and playing, it was also very important for the system of sonic feedback to be conducive to both individual exploration (as a form of knowledge creation and skill acquisition), as well as uniform, collaborative exploration and strategy-building (deliberate manipulating the system to achieve something). This observation was once again made particularly in Phase I, in the user testing sessions of socio-ec(h)o. An example of this could be a problem-solving collaborative activity in a ubiquitous computing setting, where sound feedback would both recognize and respond to individual actions by individual participants, but also acknowledge and react with meaningful feedback towards group, collaborative actions.

Support local ecology: The acoustic ecology field proposes designing balanced acoustic soundscapes whenever possible, as clarity of sounds and wide acoustic horizon are crucial to the communication cycle of listener, sound and environment. This suggestion was further supported by our observations in Phase I where masking between different sound avatars was reported as a problem (whenever present) in the design workshops, while shrill soundscape content was reported as annoying and disruptive of speech in the socio-ec(h)o prototype testing. A local acoustic ecology is a situation where both the feedback and user verbalizations are clearly heard. In a larger, general sense, an ecology would also relate to the earlier discussed concept of acoustic community, where sound has an active role in fostering and preserving a sense of place, a magic circle, within which the ubiquitous computing environment resides. To enlarge the idea of ecology even more, the aural ecology would nest harmoniously within the
informational, visual, cultural and media ecologies in the environment. An example, extending further the computational aspects of ubiquitous computing, would be a space where the environment "listens" and dynamically adjusts the spectrum and/or amplitude of the auditory display to position it in a different sonic niche than speech so that both are clearly audible.

Support Ambiguity: Ambiguity was a fact present in all of the iterations of research and design workshops described here. Whether it was designed in (as in the case of socio-ec(h)o, where soundscapes increased in ambiguity to aid in game level difficulty as a core mechanic) or were a by-product of low-tech workshops, where human errors cause minor inconsistencies in the feedback. As well, many of the sounds and soundscapes that were used and tested in my investigations here were deliberately chosen to be complex, connotative, abstract and rich environmental compilations. Often to our surprise, participants dealt well with, and welcomed the challenges of ambiguity in the sonic feedback. At times they were frustrated, but often cleverly found ways to compensate for and return to reinterpret the sound changes, as observed in the socio-ec(h)o user tests. In trying to deal with ambiguity, we also observed many emergent approaches to interaction with the system, with the space, and each other, that resulted in better understanding of the environment. For these reasons, then, an effective auditory display for ubiquitous computing context would allow users to experience and deal with ambiguity of the system’s rules and ambient response patterns in a manner that fosters experiential learning, cognition-in-action and interactivity. An example of this guideline would follow the idea of socio-ec(h)o where sound feedback in a responsive environment conveys information in an increasingly (or decreasingly – depending upon activity requirements) ambiguous manner so as to allow and support users to learn how to interpret it themselves. This kind of approach would take advantage of ambiguity and foster analytical everyday listening rather than frustrate with uniform-level, abstract response and result in distracted, detached listening attentions.
Figure 7.3: A representation of the relationships between main design areas identified here as central to the framework for auditory display design for ubiquitous computing.

Framework Comparison

Before I finally return to and address the research questions articulated in the first chapter, one final and very important step is to ask where does the new framework fit within and compare to the two paradigms of conceptualizing sound and sound design – acoustic communication and auditory display research? Based on what I have learned thus far in my inquiries, what distinctions and re-articulations am I able to make? Let’s start with the acoustic communication framework. Its categorization of sonic elements in the environment into keynotes, soundmarks and signals appear to fit well with my propositions regarding sound displays in ubiquitous computing, and furthermore, fits well with the here-presented model of intensity-based gradient approach to sound. As I have mentioned earlier, a keynote carries meaning, rather than information and it situates someone in a place, in an acoustic field. Listeners extract information from the familiar patterns in the sound’s structure and characteristics, acquired over the years (Truax, 2001, p.30). The intensity gradient sound feedback model strives particularly to actively convey information in a continuous, ambient way, even when the context or soundscape is not readily familiar to the listener. The notions of listening positions in the acoustic communication framework also fit extremely well within the new auditory display model, as the ambient, persistent nature of ubiquitous computing predicates shifting mental and
perceptual attentions of the users. Perhaps one way of combining the idea of ambient yet informative sound with the concept of fluid listening positions, and relating it back to the acoustic communication framework is by invoking the notion of ‘programming flow’ that Truax uses to describe the deliberate sonic design on the part of media producers to manage and direct listener attention to different program sections over long periods of time (2001). However, while programming sonic flows is a commercial endeavour, the auditory display model proposed here is to be used to help, guide and inform active, embodied participation in an interactive exchange, rather than influence passive viewers. To take further another concept from the acoustic communication framework – that of the mediating role of sound – it is again clearly emergent in contexts of ubiquitous computing. What I have tried to do beyond simply recognizing this important role is detail and describe some of the mediational strategies that people have used in interpreting, experiencing and using auditory feedback in contexts of embodied interaction.

Now to take a look at the field of auditory display design, I have to note first that because of the vast and disparate areas that this field covers, I have for the most part been referring to psychoacoustic and perception research, developments in auditory icon displays, design frameworks for auditory displays, and most of all, sonification. Regarding the first phase of research, considerations such as spatialization of sound, perception of amplitude, timbre and pitch, reverberation, as well as masking and composition were included in the sonic display. While the workshops were exploratory and open-ended we did use loose guidelines from psychoacoustic research and they appeared to hold well and support embodied interaction in the space, mediate the relationship between participants and system. The design approach to sound in socio-ec(h)o incorporated even more guidelines from auditory display, specifically sonification, trying to take advantage of the framework proposed by Walker & Kramer (1996) proposing scaling, polarity and data-to-parameter mapping as some of the core elements of sonification. This framework fit extremely well for our purposes in socio-ec(h)o, and was effective in conveying information. Further, we build upon these sonification guidelines by using complex, representational sounds as a base, instead of simpler synthetic tones, and incorporated deliberate ambiguity in the approach to sonification, making sound display part of the game mechanics of the activity. Again, this type of approach seemed to support and compliment well embodied, collaborative interactions.
There are three sides to the contributions that this new framework affords (see Figure 7.3). One is that it attempts to put together two existing models that have not necessarily found a practical realization in a prototype system, or have been theoretically combined together. The other contribution is in advancing and building on guidelines, concepts, contexts, settings and application domains from both existing models of sound for ubiquitous computing environments.

The last, but not least contribution of the new framework is in the hybridization of methodological approaches for investigating auditory displays for ubiquitous computing contexts. While the acoustic communication framework falls easily within the confines of design research with its focus on the human experience and holistic interplay of sound and listener through the mediation of the environment, the general analytical approach of auditory display research, and specifically that of data sonification, provide valuable constraints that when built into participatory workshops afford the possibility of isolating and examining specific auditory display design issues in the form of research questions, as exemplified in the second phase of research.

### Table 7.2: A breakdown of the comparison between main features of the two existing frameworks and the proposed new framework for auditory display design.

<table>
<thead>
<tr>
<th>Framework Comparison</th>
<th>Acoustic Communication</th>
<th>Auditory Display Design</th>
<th>New Framework for Sound in Ubiquitous Computing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain of Theory</td>
<td>The natural sonic environment, media soundscapes</td>
<td>Alarms, alerts, earcons, auditory icons, sonification, virtual audio</td>
<td>Ambient or non-ambient soundscapes and sound feedback systems for ubiquitous computing</td>
</tr>
<tr>
<td>Domain of Knowledge</td>
<td>Psychoacoustics, socio-cultural aspects of listening, socio-cultural and physical aspects of acoustic and electroacoustic sound</td>
<td>Auditory perception, psychophysics, psychology, cognitive science, human-computer interaction</td>
<td>Situated auditory perception, embodiment theory, socio-cultural aspects of listening and cognition, interaction design</td>
</tr>
<tr>
<td>Methodology</td>
<td>Grounded theory, media analysis, field work, interviews and observation</td>
<td>Controlled experiments, usability testing, qualitative studies</td>
<td>Situated inquiry, naturalistic experiments informed by design methods, participatory design workshops</td>
</tr>
<tr>
<td>Framework Comparison</td>
<td>Acoustic Communication</td>
<td>Auditory Display Design</td>
<td>New Framework for Sound in Ubiquitous Computing</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------------------</td>
<td>-------------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Components of Interest</td>
<td>Present in framework Y/N</td>
<td>Present in framework Y/N</td>
<td>Present in framework Y/N</td>
</tr>
<tr>
<td>Acoustic Ecology (balance of sonic elements in space)</td>
<td>Y</td>
<td>Not explicitly as ecology, but from a perceptual perspective care is taken to make sure auditory displays are heard</td>
<td>Y – Here the concern is both for functionality and for ecology</td>
</tr>
<tr>
<td>Role of Space (as an architectural, as well as symbolic entity – including immersion and idea of “place”)</td>
<td>Y</td>
<td>N – While virtual audio puts emphasis on spatialization and spatial hearing, other auditory displays rarely consider architectural or cultural space</td>
<td>Y – While the role of cultural and physical space and “place” is important, it has not been systematically explored in this thesis</td>
</tr>
<tr>
<td>Role of Context (type of activity at hand, type of engagement required, social aspects)</td>
<td>Y</td>
<td>N – Worth noting however, sonification and earcon/auditory icon research that strive to consider type of activity/task as part of design protocol</td>
<td>Y – This consideration is mostly illustrated with anecdotal evidence and observation here, however, it is nevertheless crucial to the new framework</td>
</tr>
<tr>
<td>Role of Embodiment (the level of physical, full-body engagement with system)</td>
<td>Y – as an aspect of listening, physically being in an acoustic space</td>
<td>N</td>
<td>Y – as a design factor, influencing perception and interaction. Explored and detailed in Phase II of this thesis research</td>
</tr>
<tr>
<td>Role of Mediation (transformative epistemic practices)</td>
<td>Y – as a relationship between listener, soundscape and environment</td>
<td>Not explicitly</td>
<td>Y – as user-driven actions to derive meaning from sound</td>
</tr>
<tr>
<td>Employing traditional auditory perception principles</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Sonification design framework (scaling, polarity, data-parameter mapping)</td>
<td>N</td>
<td>Y</td>
<td>Y – modified principles to fit different activities and social settings</td>
</tr>
<tr>
<td>Framework Comparison</td>
<td>Acoustic Communication</td>
<td>Auditory Display Design</td>
<td>New Framework for Sound in Ubiquitous Computing</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------------------</td>
<td>-------------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Components of Interest</td>
<td>Present in framework Y/N</td>
<td>Present in framework Y/N</td>
<td>Present in framework Y/N</td>
</tr>
<tr>
<td>Cognitive aspects of interpreting auditory information</td>
<td>N</td>
<td>Y - in certain areas of auditory display design, such as human-computer-interaction</td>
<td>Y - interpretation and meaning-creation is established as highly important, however, not specifically explored in this work</td>
</tr>
<tr>
<td>Auditory Perception of complex everyday and environmental sound</td>
<td>Y</td>
<td>N - only a small section of the field related to auditory icons research looks into recognition of everyday sounds</td>
<td>Y - because of the rich potential of complex, representational sound, this framework is highly interested in its auditory perception</td>
</tr>
<tr>
<td>Auditory Perception of complex changing sound (Intensity gradient)</td>
<td>Y - not the way it is used in this thesis as a basis for intensity gradient feedback, however deals with perception and interpretation of subtle changes in everyday sounds and sound qualities, temporal structure and attributes</td>
<td>N</td>
<td>Y</td>
</tr>
</tbody>
</table>
CHAPTER 8: LIMITATIONS AND FUTURE WORK

"Meaning inheres the world as we find it. The central element of our existence is to interpret that meaning through the ways in which we encounter the world."  
(Dourish, 2001, p. 108)

Design Methods: Limitations and Future Work

Design-based research is a way of investigating issues through making things, where design is central to the research inquiry, yet knowledge creation is the ultimate goal. The pivotal challenge of design as research lies in separating design goals from research goals, and design outcomes from research outcomes. Hybrid, interdisciplinary fields such as interaction design have embraced design methods and design-based research as a way of tapping into ways of investigation that are participatory and contextual; focus on artifact creation; and position people in close-to-real technological and social settings as a site of conducting research. The workshops presented in this thesis are essentially design workshops focused on sound – they employ artifacts (tools or entire systems) in order to better understand the role of sound in relation to embodied interaction, as well as to derive patterns, concepts and guidelines adding to the base of knowledge on auditory displays for ubiquitous computing environments. The reason why I have chosen to explore design workshops on their own or as a basis for experimental studies is because of their potential for generating a multitude of rich, contextual units of data that are not possible with other qualitative methods such as fieldwork or ethnography, or with quantitative methods such as controlled experiments. At the same time, the iterative nature of this process allows me to relate all of the workshops together - take a step back and consider all research units as data points on a continuum that both produce individual results, but also work together to create a fuller picture towards my research inquiry – a framework for auditory display design for ubiquitous computing contexts.

As shown in the previous chapter, the workshops have yielded an impressive collection of data such as user-generated artifacts, patterns of embodied and social interaction; user engagement with technological systems and contexts of activity, performative actions, gameplay actions, response patterns, interviews and discussions,
and video/audio capture. In fact, quite often all of these data points are present in a
given study. While it becomes difficult to then untangle and analyze them in isolation, I
suggest that it may not be necessary or even desirable to do so. As discussed in earlier
chapters, there is a strong need to find more holistic ways of analyzing embodied
interaction and human activity, the "creation, manipulation and sharing of meaning
through engaged interaction with artifacts" (Dourish, 2001, p. 126) in contexts of
technology and multi-modal displays.

All in all, what I have is a series of individual workshops that both fit into an
overarching database of sound design concepts for ubiquitous computing environments,
and at the same time serve as independent explorations of a number of different sets of
auditory display problems. So, searching for validity in this research undertaking is
two-fold, as it may refer to both the individual objectives of each workshop (specifically
the ones structured as experimental studies), as well as the objective of answering the
larger research endeavour of building theory around a new framework for sound design
in ubiquitous computing. Clearly, ecological validity is the key contestant here, as all of
the workshops are otherwise qualitative in a non-controlled fashion and thus not
statistically generalizable. However, taken together, the experimental studies can begin
to yield some analytical generalizations as presented in the previous chapter – through
iterations of varying embodiment and technological involvement, played out in a range of
ecological correspondence to the intended context. Many more would need to be
conducted and the experimental design will have to improve and evolve in order to truly
address issues of situated auditory perception in a way that completes the picture.

Considered together, all of the iterations in this thesis begin to tell a story about
sound's role in an environment and its relationship to embodied interaction, ecology,
social engagement, listening attention, perception and cognition. The validation of this
story is further complicated by what Fallman calls "the recognized problems of the
structured design methods...not being able to show evidence of the same kind of
control, structure, predictability and rigourousness" as other standard research methods
(Fallman, 2006). However, as Fallman points out, the practices of iteration and HCI
sketching allow the designer/researcher to move fluidly "between the stages of analysis,
synthesis and evaluation...[since] one does not really know the problem until one starts
working on its solution." (2006). HCI sketching, analogous to sketching in architecture, is
essentially low, mid and high-tech prototyping and having people interact with that
prototype. Along with iteration, prototyping and evaluation form the “design dialogue” as an unfolding story building knowledge as it grows. Now, if one were to use design methods as a basis for more constrained, experimental-style research, as I have attempted in Phase II, there is a lot more work to be done on polishing the methods and approaches taken to such an inquiry. What I hope I have achieved with my contribution is showing that it may be possible to do so, and offering one way of structuring a research investigation using design methods as a basis.

Auditory Display Framework: Limitations and Future Work

The auditory display framework for contexts of ubiquitous computing proposed here is in its early stages of development. While a lot of knowledge has been gained from the experience of conducting the research and design inquiries presented in this thesis, the specific context in which they were conducted only scratches the surface of the vast and diverse domain of ubiquitous computing, varying both in application focus, as well as in technological and physical settings. Furthermore, the way I have organized and analyzed the knowledge gained from the research experience is only one way of structuring and building a theory about the requirements of auditory display design for ubiquitous computing environments. Lastly, the specific approach to auditory display design synthesized after Phase I – the intensity-based gradient approach to sound feedback, is also only one way of conceptualizing and designing system sound display. Other applications of ubiquitous computing may have different requirements and characteristics, resulting in different approaches to sound design. However, it is the qualities of the approach presented in this thesis and the core concepts, that I suggest are well suited and fit within the paradigm of ubiquitous computing. Its characteristics of being ambient, informative and ecological I believe respond to and complement well the features of ubiquitous computing that call for social and physical engagement, collaboration and embodied interaction (see Table 7.3).

Future work will have to include not only more exploratory participatory workshops, delving into significant features of sound feedback, possibly ones that I have missed or overlooked here; but also more and better designed naturalistic experimental studies that utilize design methods with constraints and focus the inquiry specifically on sound. As well, I believe, other kinds of inquiry may benefit this area as well, such as conducting fieldwork into existing audio-enhanced ubiquitous computing environments to
try and classify, categorize and compare the approaches used in designing sound but also the way it is experienced by listeners. This type of study, unlike design-based approaches, will allow potentially looking into a much larger array of applications and setting domains of ubiquitous computing, and thus add another layer of knowledge to the base presented here.

**Future Applications**

As reiterated many times throughout this thesis, embodied interaction and embodied learning employs perception, cognition and actions that are qualitatively different from their counterparts in traditional print or digital applications. One obvious area of future application for informative sound feedback in ubiquitous computing environment is to benefit visually impaired users, as well as people with motor disabilities in rehabilitation. The utility of the here-presented intensity-based gradient approach to sound feedback is in that if interpreted correctly, learners would know not only if they are on a right track, but also, how close they are to completing a task or realizing a goal. Any situation of hand-free and eyes-free interaction where a system still needs to convey meaning and information to users could benefit form a well-designed auditory display.

The new framework could also benefit educational environments where learning through doing is supported by a system of multi-modal displays in a ubiquitous computing space. Finally, worth mentioning is the domain of health care for the elderly, where interactive technologies that support and enable ageing people to continue living independently, are increasingly harnessing the affordances that sound, light, haptic and other displays have to offer.
CHAPTER 9: CONCLUSION

In this thesis I have presented a rationalization and investigation of a novel auditory display framework for ubiquitous computing environments. First, I argue that sound carries a great deal of importance to ubiquitous computing environments and that there is a gap in that particular field of interaction design when it comes to systematically harnessing, exploring and building knowledge about auditory feedback. I then introduce two existing paradigms of conceptualizing sound as feedback, the auditory display analytical approach to design, and the acoustic communication framework; and suggest creating a hybrid framework for sound feedback for ubiquitous computing, by harvesting valuable knowledge from both existing ones. In order to flesh out this new framework I have presented a research investigation in two phases. In the first phase, I describe the design process and development of an auditory display system for a ubiquitous computing prototype. The aim is to elucidate what can be learned from the experience of designing and testing such a system of auditory display in a practical fashion. Following this experience, I outline and redefine propositions and questions relating to the model of auditory display feedback applied in socio-ec(h)o - the intensity-based gradient approach to sound. I then present several experimental studies informed by design methods in order to investigate the new research questions specifically dealing with intensity-based sound and its interplay with embodiment. These exploratory experimental studies preserve the situated, contextual nature of design methods and introduce relevant constraints that allow for empirical evidence gathering, ecological coherency, cross-comparison, and analytical generalization. Both phases of research are then analyzed by extracting important design concepts concerning sound in ubiquitous computing environments, resulting in a synthesis of guiding principles, meant to open a design dialogue for knowledge creation and future investigation of auditory displays in interaction design.

I hope this thesis offers a twofold contribution. One is in attempting to formulate and solidify a framework for auditory display design for ubiquitous computing environments. Such a framework, as I have demonstrated in my present research, has to recognize, examine and actively design for the role of context, space, level of
embodied interaction and types of mediation within an environment. It also has to allow for group and social engagement, leave room for shifting attention and listening positions, and support users making sense of ambiguity and feedback. Finally, in order to actually convey information, such a framework has to make use of not only traditional auditory perception research, but also socio-cultural aspects of listening, knowing and being; as well as further build and develop ecologically valid knowledge about people’s auditory perception of complex, everyday sound. The second contribution of the thesis is in demonstrating how modified design methods could be used to query different research questions relating specifically to sound feedback, in a contextual, embodied and ecological fashion.

I’d like to finish by recognizing that the proposed framework is developing and further explorations will help add to and improve it, as well as possibly refute some of its propositions. In addition, I want to acknowledge that while I am focusing specifically on the sonic aspects of ubiquitous computing environments, I actually believe such complex technologies and spaces that involve multiple modalities and forms of cognition and perception, have to ultimately be explored in ways that consider all of their aspects, feedback mechanisms and modalities together, because it is in the holistic form that such environments are experienced and used. Of course, existing methodologies and forms of analysis stretch to make sense out of this level of complexity, especially when there are gaps in knowledge about their individual components. Until then, explorations such as the present one will have to take the place of holistic inquiries, by pushing the limits of situated, contextual inquiry and design-oriented research.
Appendix 1: Full Table Breakdown of Passive Listening Task

<table>
<thead>
<tr>
<th>Type and Degree of Change (intended)</th>
<th>Graphic Waveform View (actual)</th>
<th>Participant-Generated Graphs (interpretive)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sound 1: Sine tone at 100Hz rising steadily to 440Hz</td>
<td><img src="image1.png" alt="Waveform View" /></td>
<td><img src="image2.png" alt="Participant Graphs" /></td>
</tr>
<tr>
<td>Sound 2: Sine tone starting at 550Hz and slowly falling to 300Hz</td>
<td><img src="image3.png" alt="Waveform View" /></td>
<td><img src="image4.png" alt="Participant Graphs" /></td>
</tr>
<tr>
<td>Sound 3: Pulsating sine tone of 440Hz decreasing in tempo from 87% to 37% (pre-existing scaling system)</td>
<td><img src="image5.png" alt="Waveform View" /></td>
<td><img src="image6.png" alt="Participant Graphs" /></td>
</tr>
<tr>
<td>Type and Degree of Change (intended)</td>
<td>Graphic Waveform View (actual)</td>
<td>Participant- Generated Graphs (interpretive)</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>-------------------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Pulsating sine tone of 220Hz increasing in tempo from 12% to 87%</td>
<td><img src="image1" alt="Waveform View" /></td>
<td><img src="image2" alt="Graphs" /></td>
</tr>
<tr>
<td>Sound 4</td>
<td><img src="image3" alt="Sound Waveform" /></td>
<td><img src="image4" alt="Sound Graphs" /></td>
</tr>
<tr>
<td>Abstract musical [vocal] sound through a low-pass filter - starting from crisp and high to muffled and high, to crisp, then drops to a lower pitch.</td>
<td><img src="image5" alt="Waveform View" /></td>
<td><img src="image6" alt="Graphs" /></td>
</tr>
<tr>
<td>Sound 5</td>
<td><img src="image7" alt="Sound Waveform" /></td>
<td><img src="image8" alt="Sound Graphs" /></td>
</tr>
<tr>
<td>Gurgling water sound starting at normal pitch and decreasing from 100% to 0% of normal pitch/speed.</td>
<td><img src="image9" alt="Waveform View" /></td>
<td><img src="image10" alt="Graphs" /></td>
</tr>
<tr>
<td>Sound 6</td>
<td><img src="image11" alt="Sound Waveform" /></td>
<td><img src="image12" alt="Sound Graphs" /></td>
</tr>
<tr>
<td>Fire sound through a low-pass filter. Starts muffled (most frequencies are filtered out) and quickly is unfiltered (0% to 100%)</td>
<td><img src="image13" alt="Waveform View" /></td>
<td><img src="image14" alt="Graphs" /></td>
</tr>
<tr>
<td>Sound 7</td>
<td><img src="image15" alt="Sound Waveform" /></td>
<td><img src="image16" alt="Sound Graphs" /></td>
</tr>
<tr>
<td>Type and Degree of Change (intended)</td>
<td>Graphic Waveform View (actual)</td>
<td>Participant- Generated Graphs (interpretive)</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>---------------------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Crisp sound of sticks clanking increasing from normal pitch 0% to 100% of normal pitch</td>
<td><img src="image" alt="Waveform 8" /></td>
<td><img src="image" alt="Graphs 8" /> sound eight</td>
</tr>
<tr>
<td>Sound of crackling rocks and pebbles increased from 37% to 100% using existing scaling system - start at normal pitch, then increase in both pitch and tempo, and in the end, a new similar sound is faded in</td>
<td><img src="image" alt="Waveform 9" /></td>
<td><img src="image" alt="Graphs 9" /> sound nine</td>
</tr>
<tr>
<td>Abstract musical vocal sound (same as above) changing from 0% to 50% using existing scaling system - sound starts fast and high-pitched, and steadily drops in both pitch and tempo</td>
<td><img src="image" alt="Waveform 10" /></td>
<td><img src="image" alt="Graphs 10" /> sound ten</td>
</tr>
<tr>
<td>Pulsating water stream sound increase in tempo and pitch from 25% to 100% using existing scaling system</td>
<td><img src="image" alt="Waveform 11" /></td>
<td><img src="image" alt="Graphs 11" /> sound eleven</td>
</tr>
<tr>
<td>Type and Degree of Change (intended)</td>
<td>Graphic Waveform View (actual)</td>
<td>Participant-Generated Graphs (interpretive)</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>--------------------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Crackling fire sound starts loud and booming and subsides to quieter, softer crackles at 50% to 0%</td>
<td><img src="image1.png" alt="Waveform 1" /></td>
<td><img src="image2.png" alt="Graph 1" /></td>
</tr>
<tr>
<td>Abstract musical vocal sound (same as above) starts quiet and increases in loudness 50% to 100%</td>
<td><img src="image3.png" alt="Waveform 2" /></td>
<td><img src="image4.png" alt="Graph 2" /></td>
</tr>
<tr>
<td>Ambient forest sound becomes richer and louder with bird songs</td>
<td><img src="image5.png" alt="Waveform 3" /></td>
<td><img src="image6.png" alt="Graph 3" /></td>
</tr>
<tr>
<td>Ticking clock sound through a low-pass filter – starts crisp, then becomes muffled then crisp again.</td>
<td><img src="image7.png" alt="Waveform 4" /></td>
<td><img src="image8.png" alt="Graph 4" /></td>
</tr>
<tr>
<td>Type and Degree of Change (intended)</td>
<td>Graphic Waveform View (actual)</td>
<td>Participant-Generated Graphs (interpretive)</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>--------------------------------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>Same as #5. This is the only sound excerpt with more than two types of distinct change in it. I use it as a control to see if it would be represented or described any differently.</td>
<td><img src="image" alt="Waveform View" /></td>
<td><img src="image" alt="Participant-Generated Graphs" /></td>
</tr>
</tbody>
</table>

Sound 16
### Appendix 2: Table View of Passive Listening Task Written Captions

<table>
<thead>
<tr>
<th>Sound #</th>
<th>Participant 1</th>
<th>Participant 2</th>
<th>Participant 3</th>
<th>Participant 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sound 1</td>
<td>Higher pitch, increased volume, increased frequency</td>
<td>Higher frequency</td>
<td>Fairly steady rise in intensity</td>
<td>Deep penetrating intense “sucking down under” feeling</td>
</tr>
<tr>
<td>Sound 2</td>
<td>Tone chips – 60Hz? – then drops lower</td>
<td>Lower in frequency</td>
<td>Starts high and drops slowly</td>
<td>Falling down intense tone</td>
</tr>
<tr>
<td>Sound 3</td>
<td>Frequency of sound increased</td>
<td>Same frequency throughout</td>
<td>Higher note that increases in frequency</td>
<td>Probing, irritating, submarine-like</td>
</tr>
<tr>
<td>Sound 4</td>
<td>Low sound, repeats often, then less often</td>
<td>Sound woofer slowed down</td>
<td>Low sound that starts fast and slows down over time</td>
<td>Vibration, ends in repetition, beckoning</td>
</tr>
<tr>
<td>Sound 5</td>
<td>Sounded fuller, harmonic, then scratchy</td>
<td>Sound peaked out</td>
<td>High, then dips tone then starts to slow</td>
<td>Irritating, choir-like, yet mechanical</td>
</tr>
<tr>
<td>Sound 6</td>
<td>Sounds like surface of water then below and slower</td>
<td>Sound retreated</td>
<td>Sounds like being above water, then below water, sound becomes more muted</td>
<td>Soothing, familiar, sense of being under</td>
</tr>
<tr>
<td>Sound 7</td>
<td>Sounds like a rocket about to take off. Burns faster over time</td>
<td>Soundwave frequency increased</td>
<td>Sounds different then moves increasingly closer like its right beside you</td>
<td>Gravel road surface, ends intensely</td>
</tr>
<tr>
<td>Sound 8</td>
<td>Sounds like object dropping in a plastic bottle, repeat frequency, increases over time</td>
<td>Sound frequency increased and became higher in pitch</td>
<td>Frequency increases from slow to way fast</td>
<td>This instrument – rain stick – familiar. Soothing at lower volume</td>
</tr>
<tr>
<td>Sound 9</td>
<td>Sounds like shaking less repetition over time</td>
<td>Noise definition changed (altered)</td>
<td>Sounds like its moving around, coming closer and then farther</td>
<td>Circular motion of sound – steady throughout</td>
</tr>
<tr>
<td>Sound 10</td>
<td>Not much change except for different sound</td>
<td>Vibrato decreased by number of five per second</td>
<td>Pitch is going up and down, frequency is increasing</td>
<td>Neutral, wobbly sound, at the end – noticeable shift</td>
</tr>
<tr>
<td>Sound 11</td>
<td>Sound repeats more over time</td>
<td>Loudness, frequency and definition increased</td>
<td>Speeds up – starts slow and goes very fast</td>
<td>Roller coaster – intense ride, uneasy feeling</td>
</tr>
<tr>
<td>Sound 12</td>
<td>Deeper sound then higher sound – fire?</td>
<td>Background noise diminished (only foreground sound continues to be heard)</td>
<td>Starts loud and gets quieter</td>
<td>Water sounds soothe me, so “drop down”</td>
</tr>
<tr>
<td>Sound</td>
<td>Volume increase</td>
<td>Loudness increases</td>
<td>Starts quiet, then quickly gets louder</td>
<td>Irritating – edgy near end</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------------------------</td>
<td>--------------------</td>
<td>----------------------------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>13</td>
<td>Traffic, then birds – constant changes</td>
<td>Ambient noise increases</td>
<td>Gets increasingly louder, plus it seems to go from more focused to more diffused</td>
<td>Two sounds mixing together – one playful and the other steady soothing</td>
</tr>
<tr>
<td>14</td>
<td>Sound slows then starts again</td>
<td>Sounds become muffled, then clear again</td>
<td>Speed doesn’t change, but pitch dips down, then comes back up, it’s like sound gets fuzzier</td>
<td>Up, down, up, down, down – wheels of a train changing steadily along</td>
</tr>
<tr>
<td>15</td>
<td>Sounds starts high, then low, then slows down too</td>
<td>Sounds become muffled, then clear again, then louder, then frequency of soundwaves decreases</td>
<td>Loud, then becomes muffled, then loud, then slows down</td>
<td>Circular motion in space, dropping distinctly at the end</td>
</tr>
</tbody>
</table>
Appendix 3: Scenario-Based Report of Geography Map Game (second blindfold turn)

<table>
<thead>
<tr>
<th>Timeline</th>
<th>Player 1 – Giver</th>
<th>Player 2 - Receiver</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time: 0 to 15 sec</td>
<td>No feedback – then slow, rhythmic, quiet clicks, getting quieter, then a bit louder</td>
<td>Starts in the middle and begins moving the cursor left, slowly, and stops, reverses her direction roughly then starts slowly moving back.</td>
</tr>
<tr>
<td>Time 15 sec to 30 sec</td>
<td>Quieter, stop, get even more rare for a while, then a click, then silence</td>
<td>She stops again for a moment and makes a curve down form her previous position, making a shorter trajectory, then stops again, just for a moment before she continues up, circling back around, again slowly</td>
</tr>
<tr>
<td>Time: 30 sec to 45 sec</td>
<td>Silence, one click, two clicks,</td>
<td>Moves up even slower, stops slowly, moves down, and veers off twice in both directions, comes back to main trajectory</td>
</tr>
<tr>
<td>Time: 45 sec to 1 min</td>
<td>One click, silence, one more click,</td>
<td>Moves in shorter trajectories, in a spiral pattern, then veers off to the left, to the right, down, moving in an expended spiral, circle shape</td>
</tr>
<tr>
<td>Time: 1 min to 1:15 min</td>
<td>Silence</td>
<td>Keeps moving sideways, down, in a spiral, in short strides</td>
</tr>
<tr>
<td>Time: 1:15 min to 1:39 min</td>
<td>Clicks start, get louder, then a little faster and louder, then faster, quieter, then louder and fast.</td>
<td>Moves to the left, straightens trajectory, and keeps moving in a line, farther and farther, then moves down, slowly and finally reaches the goal</td>
</tr>
</tbody>
</table>

**Scenario:**

When Lisa* started right away going in the wrong direction, Rebecca* questioned whether we had decided on silence as a form of feedback or not...likely, because she is thinking, “how do I tell Lisa that she is on the wrong track already?” As Lisa slowly went further in the wrong direction, Rebecca decided to start giving her some feedback, so that Lisa could get oriented by its change on where to go. Lisa did actually start going towards the right place, but feedback was not significantly more intense, it was fairly sparse and low-key. Rebecca starts getting more specific, so unless Lisa’s cursor is pointing towards the right way, she gives no feedback. Around halfway through, Lisa is back in an area that she was in the beginning, on the other side of the world, from where
the secret location is. Yet this time, Rebecca is giving her some feedback any time she takes a course towards the goal. In any other occasion, Rebecca gives no feedback. Lisa is thus stuck with experimenting with directions and eliminating them one by one. This may explain the fact that when she does finally begin on the correct trajectory, and hears feedback, she continues, and as soon as the feedback intensifies, she continues more, faster, with more confidence. The clave sticks become even faster and louder, so from there Lisa quickly finds the right place. Essentially, it takes her all of 15 seconds to find it once she has gotten on the right trajectory.

*Names of participants are fictional to protect their privacy
REFERENCE LIST


