APPLICATION OF MOBILE AND INTERNET TECHNOLOGIES FOR THE INVESTIGATION OF HUMAN RELATIONSHIPS WITH SOUNDSCAPES

Charlie MYDLARZ

Ph.D. Thesis **2013**

APPLICATION OF MOBILE AND INTERNET TECHNOLOGIES FOR THE INVESTIGATION OF HUMAN RELATIONSHIPS WITH SOUNDSCAPES

Charlie MYDLARZ

Acoustics Research Centre

School of Computing, Science and Engineering

College of Science and Technology

University of Salford, Salford, UK

Submitted in Partial Fulfilment of the Requirements of the

Degree of Doctor of Philosophy, July 2013

Contents

List of Figures

List of Tables

Acknowledgements

This thesis has been made possible by the support from a number of people, far too many to list in their entirety but I'll endeavour to mention the special cases.

Firstly, I am extremely grateful to my family who have been incredibly supportive throughout this whole process. My supervisor Dr Ian Drumm has been a fantastic source of support, inspiration and well timed humour. Thanks also goes out to my co-supervisor Prof Trevor Cox.

The acoustics researchers I have worked with over the years in G10/11 have more than anyone, made this experience so rewarding and enjoyable. In no particular order, a huge thanks goes out to: Rodolfo Venegas, Sarah Payne, Tobias Ackroyd, Jenna Condie, James Woodcock, Rob Oldfield, Andy Elliott and everyone else I've not listed.

A place also worth mentioning is The Crescent, where the debates over my work provided much inspiration.

I would also like to acknowledge the Engineering and Physical Sciences Research Council for the financial support provided for this work in its early stages.

List of abbreviations

3G

Third generation mobile communications technology providing data rates of at least 200 kbit/second

AGC

Automatic gain control - an adaptive system to provide variable input audio range control on mobile devices

App

Application running on mobile device

Bluetooth

Short range wireless data transmission generally between mobile devices

DSP

The study and implementation of signals in digital computing and their processing methods

GPRS

General packet radio service - A packet based mobile data service providing data rates of 56-114 kbit/second

GPS

Global Positioning System - a space-based satellite navigation system that provides location and time information to devices containing a GPS receiver

GUI

Graphical User Interface - a type of user interface that allows users to interact with electronic devices using images rather than text commands

HTTP

Hypertext Transfer Protocol - the foundation of data communication for the World Wide Web

IDE

Integrated Development Environment - a software application for software development

iOS

A mobile operating system developed and distributed by Apple Inc.

J2ME

A Java platform designed for embedded systems including mobile devices

MEMS

Micro Electro Mechanical Systems - very small electrical devices, namely PCB mounted microphones used in mobile devices

MIR

Music Information Retrieval - the interdisciplinary science of retrieving information from music

OS

Operating System - a collection of software that manages computer hardware resources and provides common services for computer programs

PDA

Personal Digital Assistant - a mobile device that functions as a personal information manager

SLM

Sound Level Meter - an instrument which measures sound pressure level

WM

Windows Mobile - a family of mobile operating system developed by Microsoft for smartphones and Pocket PCs

Abstract

This thesis presents a methodology for soundscape research, utilising consumer mobile and internet technologies. This has been used to gather objective environmental data, as well as subjective data from participants in-situ. A total of 323 untrained members of the public have submitted soundscape recordings from around the world. For the first time, participant choice has been factored into soundscape research, where members of the public decide which sound environments are investigated. Human relationships with their sound environments have been investigated, with a number of findings corresponding with those of other studies utilising entirely different methodologies. In addition, a number of new findings have been made to contribute to the field.

The two extracted subjective principal components of 'Appreciation' and 'Dynamics' has shown a solid validation of the project's methodology, due to their similarities with a number of other studies utilising different techniques of data retrieval. The distinctive groupings of the different soundscape types within this factor space defined by the extracted components reveals the perceptual differences between the soundscape categories: urban, rural, urban public space and urban park. The activity a person is involved in while making their submission has shown to be influential in soundscape appraisal, with relaxation and recreation situations resulting in increased soundscape appreciation. The reasons behind a soundscape submission have revealed significant differences in subjective response. The positive interpretation of the term soundscape has resulted in a majority of positive reasons for participation. Soundscapes that arise from a participant's daily routine are generally less appreciated than soundscapes containing a particular sound source focus. The highest levels of appreciation were observed in soundscapes whose focus is on a specific activity that the participant is involved in. The interest that a participant has on their soundscape is seen to result in raised levels of appreciation.

Chapter 1

Introduction

This thesis presents a novel methodology for the investigation of human relationships with soundscapes, utilising ubiquitous consumer mobile and internet technologies. A soundscape research tool has been designed, implemented and validated. This new methodology has been used to gather over 1300 submissions of objective environmental data, as well as quantitative and qualitative subjective data from participants in-situ. For the first time, participant choice has been factored into soundscape research, where members of the public decide which sound environments are investigated.

The research builds on existing work pertaining to soundscapes by utilising data collected as part of the study. The influential characteristics within these subjectively perceived sound environments were yet to be fully defined. In order to determine this, input from a large number of individuals was required to ensure any explorative analysis techniques are considered statistically robust. The pervasive nature of these mobile devices means that the potential geographical coverage attainable using these techniques is undoubtedly advantageous and improves on the reach of more traditional methods. For example, a number of past studies [1, 2, 3] focus on in-situ questionnaire/interview based research in specific locations. Whilst these can reveal a detailed insight into a particular location's soundscape, the reliance on the researcher being present at all locations limits their coverage and scope. By empowering the public with the means to control what is being monitored and allowing the individual to record specific impressions of these environments, a more representative and complete analysis of the acoustic environment can be achieved. Becker [4] claimed that participant observation of events, rather than interviewing about events provides a more complete record and understanding of the event in question.

For the purposes of this thesis the soundscape has been defined as: "the sounds around us, their effects and our perceptions of them". The sounds being the physical sound sources within a person's vicinity, or sound environment, the effects being the unconscious effects of these sources on a person, and the perceptions of them being the conscious acknowledgement and processing of these sources. In a purely objective sense; individual sound events are equal to sounds emanating from a source and interacting with the physical environment and the sound environment is equal to the summation of these sources.

This chapter continues by introducing the overarching themes of the thesis and places the motivation for the work into context. Thereafter, the rationale and goals defined for the investigation of the project are discussed, followed by a summary of the overall project. Finally, an overview of the thesis is given on a per-chapter basis.

1.1 Research aims

The main aim of the study is the implementation and validation of a novel methodology for soundscape research with enhanced geographical reach and the potential to recruit large numbers of participants. The main focus of the thesis is in the development, use and validation of this technique, by comparing its outcomes against more traditional research methodologies. Data gathered from the study forms the basis for the secondary aim, which is to investigate human relationships with their sound environments and explore the influential factors and trends that determine a person's appreciation of these environments. Specifically, the following research questions will be answered:

- Can a soundscape research tool based on internet and mobile technologies be implemented which is suitable for gathering meaningful data pertaining to human perceptions of their sound environments?
- Which factors affect a persons perception of their sound environment?

To date, no other study has made use of consumer mobile technologies to gather data of this kind for use in soundscape research.

1.2 Contributions of the thesis

This thesis has shown that new mobile and internet technologies can be used to create a novel and robust methodology for soundscape research. Crowd-sourced, large scale, participant driven soundscape data has been used to assess human relationships with their sound environment's for the first time. The findings have been validated against those from existing research methodologies utilising more traditional tools. Inferences have been made from the quantitative and qualitative responses of participants, as well as comparing these with objective metrics obtained from the soundscape recordings.

The research described in this thesis contributes to the literature in the following ways:

- **Pervasive mobile technologies** used in the study have shown their feasibility for use in soundscape studies and also provided a case study for future research
- **High resolution spatial data** from location aware devices provides new levels of spatial accuracy and data for a study of this scale and could be of significant use to soundscape policy makers
- **Participant discretion** for in-situ soundscape capture in a quantitative study is a new addition to the field
- **Public engagement** in soundscape research is a novel concept providing the means to gather large volumes of data

1.3 Project premise

The project was initially conceived as a purely research based project with no planned public involvement, titled: IMPRINtS (Internet and Mobile Technologies for a Public Role In Noise Surveying - EP/E06552X/1) and funded by the EPSRC (Engineering and Physical Sciences Research Council), with strong public engagement objectives. The project's public facing name was chosen to be "The

Sound Around You Project" and will be referred to as that for the remainder of this thesis. The project aimed to enable and encourage public participation in a large-scale mass participation soundscape survey. The initial objective was to equip members of the public with mobile phone applications and online tools that allowed them to acquire quantitative and qualitative environmental data directly from their surroundings. This data was stored in a format that could then be collated and analysed in a bid to uncover the influential factors and trends that determine a person's appreciation of their sound environment. The outcomes also aim to address the role that sound plays in the design process and appreciation of public spaces. The findings could then be compared with the findings of more traditional soundscape research methodologies [5, 6, 7], in an attempt to validate the techniques used.

This project has three main aspects in its methodology: the environment, mass participation, and an example of pervasive computing. Pervasive computing describes widespread computing devices that can be readily introduced and used within a human environment by its inhabitants [8]. The project's first objective was to enfranchise the public by letting them contribute to the science that might eventually inform legislation that will impact on the quality of their lives, an example of "research in the wild" [9], where participants are given the freedom to influence current and future research.

1.4 Mobile technologies

Advances in mobile computing offer the opportunity to allow many people to participate in soundscape surveys. Recent developments in mobile technology were utilised, including: mobile phones, mobile/PC connectivity and distributed application technologies from the studies website. The combined use of these technologies has contributed to the study in two respects:

- It enables environmental soundscape data from a large participant base to be automatically collated and analysed.
- It enables participants to include subjective responses to the soundscapes they inhabit, providing a more nuanced understanding of the context and reasons for human responses to environmental sounds.

The proliferation of these mobile devices has given rise to a number of projects [10, 11, 12, 13], all with similar intentions of tapping into a potentially huge, geographically and culturally diverse participant set. With many people taking their phones with them wherever they go, these mobile sensing units are then placed in a variety of environments, situations and specifically for this project; soundscapes. An example of the pervasive nature of mobile devices is seen in India, where more people access the internet from their phones than from a computer [14]; a trend that is expected to be mirrored globally in the future. At the end of 2011, The International Telecommunications Union estimated that there were 6 billion mobile phone subscriptions worldwide, equivalent to 87% of the world's population [15]. Figure 1.1 shows the increase from 2001 - 2011. Subscribers in the developed world have reached saturation point with at least one mobile device per person. In fact the mobile phone has been adopted faster than any other technology in human history [16].

This growth in mobile subscriptions is accompanied by falling costs of mobile data rates, with the majority of mobile operators providing customers with included monthly data allowances with their mobile subscriptions. Uptake and consumer

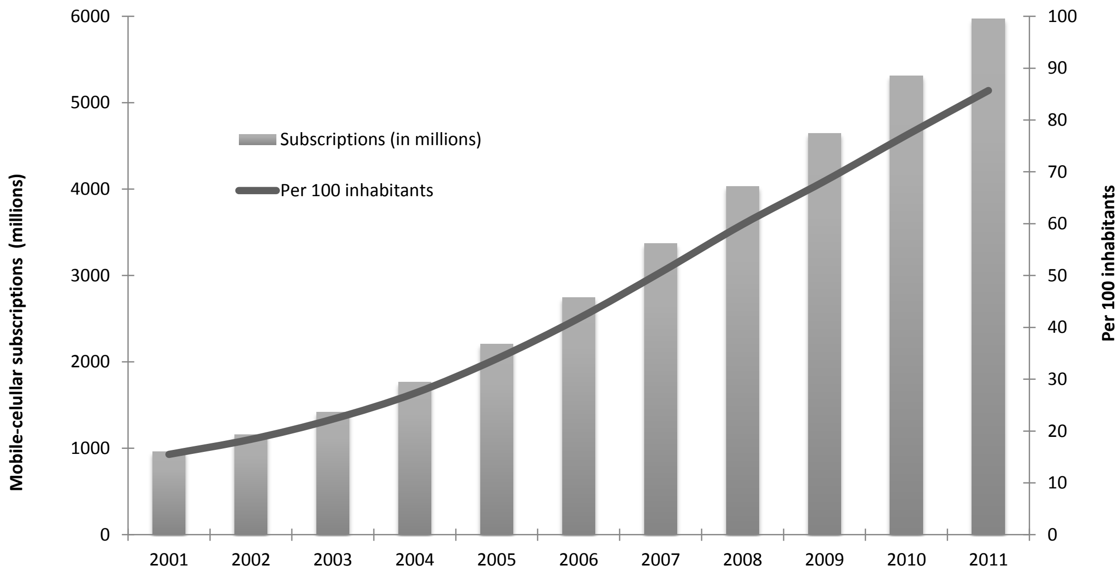

Figure 1.1: Global mobile-cellular subscriptions, total and per 100 inhabitants, 2001-2011 (Source: ITU World Telecommunication /ICT Indicators database)

adaptation of modern smart phones has also given rise to a surge in data consumption, due to the increased power and capabilities of these devices. In 2010, the monthly mobile data traffic per smart phone user stood at 149 megabytes (MB), which is predicted to rise to 3390 MB by the end of 2015 [17] with the rollout of high capacity 4G mobile data networks providing the equivalent to current home broadband high speed mobile connections. The adoption of mobile applications (apps) has also seen a vast surge since 2008 when Apple launched its App Store for iPhone and iPod Touch devices. 2009 saw 2.9 billion app downloads, with 30.1 billion in 2012 [17]. The mobile phone has very quickly become more than simply a communication device; it is now a consumption platform for a huge range of services. This rapid spread of mobile phone technology throughout the world offers a new and exciting means of data collection for a whole host of research applications. The enormous potential for public participation in soundscape surveying using mobile and internet technologies has yet to be utilised. With the extensive adoption of mobile technologies, the retrieval, collation and analysis of objective and subjective soundscape data from widespread locations is now possible. With the European Union Noise Directive [18] and Future Noise Policy [19] including aims to stimulate public discussion on noise issues, projects that enfranchise and engage the public might acquire much needed data for inclusion into policy considerations.

1.5 Challenges

The following issues were the central research challenges in the current project of mobile device based soundscape research:

- **Understanding of the term soundscape**
  How do members of the public interpret the term soundscape and will any provided definitions influence a participant's choice of submission?
- **Recruitment**
  How will people find out about the project and what will convince them to take part?
- **Technological suitability**
  Is the mediating technology suitable for the task and what are the limitations and compromises that will have to be made?
- **Data quality**
  Will the data provide a meaningful indication of soundscape appreciation

and can it be compared with datasets from existing research?

These challenges will be addressed in Chapter 3.

1.6 Thesis outline

The rest of this thesis is organised as follows: Chapter 2 reviews the literature relevant to the project. Chapter 3 describes the design of the project. Chapter 4 details the analyses of the dataset as well discussing any inferences made. Chapter 5 discusses the methodological approach. Finally, Chapter 6 closes the report, reviewing the work undertaken and draws conclusions about key parts of the work that was undertaken, ending with the discussion of future work, with a particular focus on the applications of the methodology and project dataset.

Chapter 2

A Review of relevant research

This chapter reviews the relevant literature in relation to the current project.

2.1 The soundscape

The relatively new field of soundscape research inherits a number of definitions of it's title from the varied disciplines that study the topic. The traditional, first cited and most well-known definition comes from Murray R Schafer, who in his 1977 book "The Tuning of The World" [20], considered the "soundscape" as an adaptation of the visual term "landscape", with the focus on the auditory environment. He expands on this by explaining that the soundscape includes all of the sound from a particular environment that reaches the human ear. Barry Truax briefly defines the soundscape in his 2001 book, "Acoustic Communication" [21] as: "the aggregate of all sound energy in any given context with an emphasis on how that environment is understood by those within it". A more holistic definition used frequently by Jian Kang is that the soundscape "is about the relationship between the ear, human beings, sound environments and society" [22]. In 2008, the International Organisation for Standardisation (ISO) put together a new expert working group (WG54) on soundscape studies named "Perceptual assessment of soundscape quality" (TC 43/SC 1/WG 54). This international group of academics and policy makers are tasked with understanding how people perceive soundscapes and to develop new research tools to fulfil this aim. A major outcome of this project will be an internationally recognised standard for measuring perceived soundscape quality, due to be completed by 2015.

Soundscapes describe differing acoustic environments, which each of us is subjected to throughout our lives. They play an important part in our lives, making us feel comfortable, productive and happy or uneasy and distracted [21]. Soundscapes provide the contextual references that contribute to our feelings of belonging and place [23]. The soundscapes we inhabit can be perceived as welcome hustle and bustle or as a noise nuisance [21]. Getting the balance right is a challenge for planning, development and construction, and a challenge for how we as individuals choose to spend each day. However, this luxury of choice is not always present, since many people have little or no control over their day to day sound environments. This lack of control highlights the need for soundscape research to be considered when designing and building spaces that are to serve the public [24]. Although there are strong guidelines in place to determine what is acceptable and unacceptable in terms of visual aesthetics [25], the sound environment is lacking in these considerations, and are therefore not widely incorporated into urban planning and assessment. A possible reason for this is the ease with which the visual environment can be captured, designed and replicated, compared to the difficulty of achieving this process in the acoustic environment.

2.2 Schaferian approach

The first serious discussions and analyses of soundscapes emerged during the 1960s with the work of the man who coined the term “Soundscape”, Murray R. Schafer. Schafer formed The World Soundscape Project in the late 1960s, and was created as an education and research group with its main aims shaped by Schafer’s initial attempts to highlight the plight of the sonic environment in the modern world. The ultimate aim of the project was to find solutions for an ecologically balanced soundscape where the relationship between the human community and its sonic environment is in harmony. The project produced two educational booklets, The New Soundscape [26] and The Book of Noise [27], as well as a number of noise by-laws. His issue with this approach was the negative connotations being fostered by the discussion of noise pollution issues. To develop a more positive outlook into the acoustic environment, Schafer published, “The Music of the Environment” [28] in which he attempted to detail both the positive and negative aspects of the sonic environment in a more balanced and considered way using predominantly questionnaire studies and a process of aural education and sound awareness through the recognition of sounds that should be preserved and encouraged.

He also introduced the notion of soundscape fidelity, where a low fidelity or “lo-fi” sound environment has a low signal to noise ratio, or in other words, discrete sounds cannot be heard clearly because of a high ambient noise level and vice versa in the “hi-fi” case. Schafer tended to highlight the negative aspects of the urban soundscape in his work, stating that both the voice and the ear are neglected physical entities in the city and stress is a common experience for the urban person [23]. Schafer’s work took a phenomenological approach to soundscape research, looking mainly at subjective impressions of the soundscape. His most notable contribution was the investigation into the social implications of the different soundscape elements. He proposed that the soundscape is the acoustic manifestation of “place”, where the sounds within it give its inhabitants a sense of place and the places acoustic quality is shaped by the inhabitants activities and behaviour [23]. He derived the perceptive descriptors: keynote, signal, soundmark, sound object, and sound symbol.

- **Keynote** refers to a ubiquitous and prevailing sound, usually in the background of the individual’s perception, to which all other sounds in the soundscape are related, e.g. running water, wind, traffic.
- **Signals** is a term borrowed from communication theory, and are defined as foreground sounds, listened to consciously, often encoding certain messages or information, e.g. train station announcement, warning alarms.
- **Soundmarks**, analogous to landmarks, are unique sound objects, specific to a certain place, e.g. the bells of St Paul’s Cathedral, London.
- **Sound objects** are defined as the smallest self-contained elements of a soundscape, e.g. chirping bird, passing car.
- **Sound symbols** are sounds which evoke personal responses based on collective and cultural levels of association, e.g. tribal chanting, chanting of sport fans.

The World Soundscape Project began the important work of connecting soundscape studies with the physical sciences, aesthetics, architecture, sociology, psychology and a range of other disciplines. Each of these fields has brought their own approach to soundscape research, providing a wide range of differing methodologies with a common aim of understanding and describing the effects soundscapes have on their inhabitants.

2.3 Acoustic communication

To understand how the sound environment affects the humans within it, an overview of how a soundscape is perceived by the listener is required. The outer ear receives the complex pressure wave that makes up the soundscape and focuses it onto the tympanic membrane or eardrum. The middle ear ossicles transmit these vibrations to the inner ear using a process of impedance matching to maximise the transmission of sound energy. In the inner ear the cochlea converts this sound energy into neural impulses. Inner hair cells transmit sensory information about sound, coupled with the outer hair cells which amplify the inner hair cell responses. The frequency information of the signal is coded by the tonotopic organisation of the basilar membrane and by the phase locking of auditory nerve fibres to the input sound signal. The auditory cortex then processes the encoded sound information in "streams" of neural analysis. These streams of auditory information are processed one at a time with their organisation determined by the individual's attentional sets. This is a process of selection that focuses cortical processing resources on the most relevant sensory information in order to maintain goal-directed behaviour in the presence of multiple, competing auditory distractions [29]. Albert Bregman states that the auditory information is prioritised based on expectation, prior knowledge and experience [30]. For example, the sound of a lions roar when heard in a shopping centre would invoke an attentional "interrupt" where it would be allocated a higher neural priority, because it is not an expected sound in that environment, you recall that it is the sound of a lion and your experience tells you that loosely: “lions mean danger”. Based on this rationale, it can be surmised that the focus and effect of different soundscape characteristics are not driven by the physical acoustic parameters alone. This also means that individuals within the same soundscape will have their own specific perception of it, based on their previous knowledge, experience, expectations and personal associations of the sources within it. The experience an individual may have of an acoustic environment will influence and shape their relationship to it. Figure 2.1, taken from [31] shows an adaptation of Barry Truax's model of the function of sound within an acoustic space. Truax uses this to describe the soundscapes capacity to convey information, acting as a mediator between the listener and the environment.

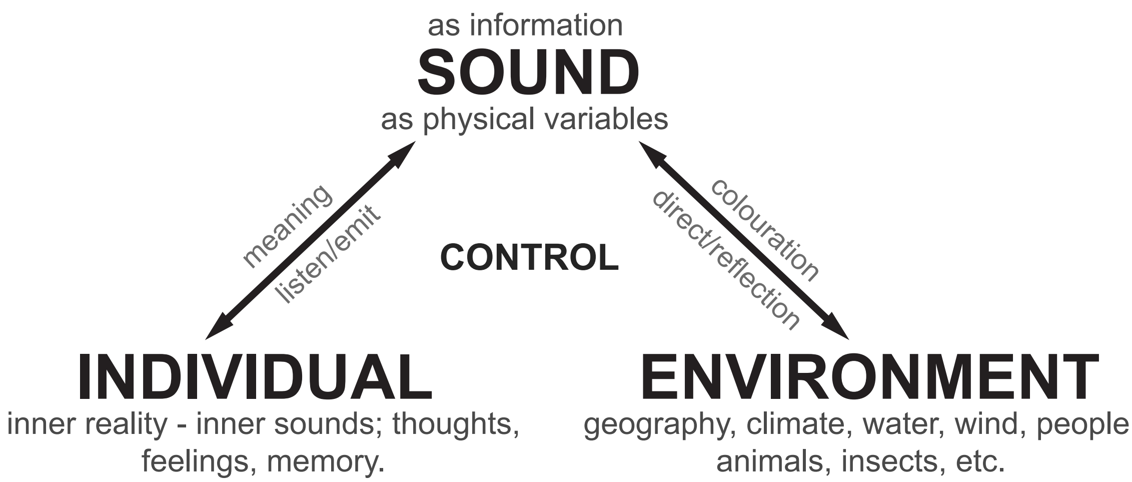

Figure 2.1: The mediating relationship of an individual to the environment through sound (modified from Truax 1984, 11)

A soundscape will simultaneously contain several sources, some of which attract the attention of the listener more than others, depending not only on the acoustic characteristics of the signal, but on its meaning and relevance to them. Invariant low-level background noise is, for example, less perceptually salient, even when it may be at a higher intensity than sounds such as birds or people talking [32]. The distinction lies in the pseudo random nature of these latter sources, which register as events within the soundscape, whereas sounds with a continual level produce the psychological reaction of habituation to that sound.

2.4 Listening states

Listening can be defined simply as the processing of sound information that is usable and potentially meaningful to the brain. It is the primary interface between the listener and their environment. Barry Truax defined his three listening states [21] (see below for definitions) in contention to the traditional assumption that listening always takes place with the full attention of the listener. While this may be true for immediate foreground sources, it does not take into account the more subtle elements of a soundscape that may be less perceptually salient. The complex and dynamic nature of environmental sounds mean these listening states are constantly switched as the individual moves through their environment.

- **Listening in search** is described as the most active form, where a conscious effort is made to search the environment for sound cues. Detail is key to this state and covers the ability to focus in on and perceptually isolate a specific sound amongst many others.

- **Listening in readiness** is an intermediate state in which the listener is in readiness to receive significant information but whose focus may be directed elsewhere. This predominantly occurs with sounds that are familiar to the listener, where strong associations have been built up over time and the identification of them can be performed using minimal auditory processing.

- **Background listening** is reserved for sounds which are a usual and expected occurrence within a soundscape with no immediate significance to the listener. Sounds in this state are habitually perceived, but are the sounds which reflect the fundamental characteristics of the environment. This state of listening can be seen as the absorption of a place's ambience.

A fair assumption would be that listening in search, with it being the most active form of listening, would bring about feelings of aural fatigue sooner than the other two states. However, the second form has the listener in a state of "dual processing" where the anticipation of a meaningful sound cue requires a higher cognitive load, with the addition of the load introduced by the individual's main auditory focus. An example of this is a person walking a busy street having a mobile phone conversation. Predominant auditory focus is given to the conversation, with the surrounding environment producing its own set of auditory cues that must be processed accordingly by the brain to determine if for example a car or bus is passing very close to the individual. An investigation into this phenomenon could provide a useful addition to soundscape research, affording an insight into how a location might be designed to reduce the potential of aural fatigue. The concept of listening states however, is rarely dealt with directly in soundscape research, as there is little evidence proving their existence. However, the concept has been utilised in large cross disciplinary studies such as the Positive Soundscapes Project [33] to determine suitable soundscapes for public spaces based on its intended use.

According to Truax, this kind of auditory processing takes place in the auditory cortex of the brain, an area whose study is covered by the field of auditory neuroscience. The Positive Soundscape Project [34] investigated the neural correlates associated with perceptual and affective responses to different sound sources within a soundscape by measuring on-going brain activity using functional magnetic resonance imaging [35]. In one study, sixteen participants were scanned while passively listening to a set of recordings made in urban spaces [36]. The recordings totalled 150, and contained multiple sound sources. They were all levelled at an average of 71dB(A), but differed in their ratings of pleasantness from a five point scale. The perceived pleasantness rating of each soundscape had an effect on the auditory systems response to that soundscape. One of the outcomes of the research showed that the loudness of the soundscape is not the deciding factor in peoples reactions to urban soundscapes, a point echoed by Truax in his book “Acoustic Communication” [21]. Another example of research into differing methods of cognitive perception of soundscapes is by Manon Raimbault. She introduced two methods of listening: “descriptive” and “holistic”, after performing psycholinguistic analysis of verbal comments made by users of urban locations in two French cities [37]. Descriptive listening is the identification and interpretation of individual sources within the environment by the listener, where their response is based on the meaning of each sound object. Holistic listening describes the interpretation of the soundscape as a whole, without the individual interpretation of sources within it.

2.5 Soundscape classification and analysis

The individual sounds that make up a soundscape are generally segregated into categories to define the type of source that they are [23]. The most common terms used in this reductionist approach are “natural”, “human” and “artificial” sounds, which have been used spontaneously by listeners in previous studies [38, 39] and identified by researchers [35], with examples shown in Figure 2.2. For these terms to be a representative definition of the source, contextual and visual information may be necessary, as certain sounds can be perceived differently but exhibit similar acoustic properties (e.g. motorway and waterfall).

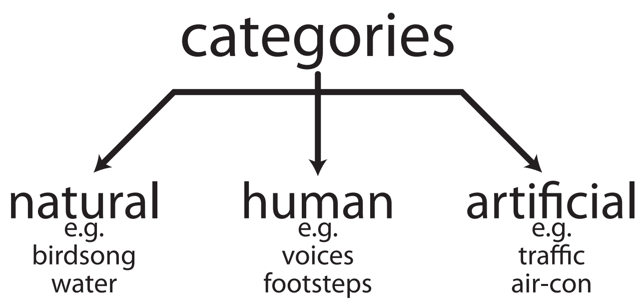

Figure 2.2: Sound source categorisation examples

Another approach to sound source classification utilised discourse analysis of participant’s responses to perceived low and high frequency sounds [40]. This determined two types of category. Recognisable sound sources were categorised as source events, while indistinguishable sources were categorised as background sound. One approach to automatic pattern recognition of audio signals represents signals as the long-term statistical distribution of feature vectors [41]. This method attempts to uncover the perceptive saliency of sound events based on their statistical typicality within a soundscape recording. The resultant output of the technique is a collection of “classes” containing soundscapes of a similar (human defined) type, e.g. avenue, park, urban street etc. This has been named the “bag-of-frames” (BOF) approach, an analogy to the “bag-of-words” (BOW) analysis of text data as a distribution of word occurrences without retaining their organization or context within phrases, traditionally used in textual data mining [42]. The signal to be analysed is cut into short overlapping frames (typically 50ms in length with a 50% overlap), and for each frame, a feature vector is computed. The features generally consist of a generic, all-purpose spectral representation such as the Mel Frequency Cepstrum Coefficients (MFCC) [43]. These feature vectors are fed to a classifier based on a Gaussian Mixture Model [44] which models the distributions of these features and assigns these distributions to varying classes. Distributions for each class can then be used to compute decision boundaries between classes.

When a number of these classes have been established by the model, a new soundscape recording is classified by computing its feature vectors, finding the most probable class for each of them, and taking the overall most represented class for the whole recording. This approach has proved effective for soundscape classification. Classification precisions of 91% have been reported on a dataset of 80 3-second sound extracts from 10 everyday soundscape classes (street, factory, football game, etc.) [45]. This technique is intended only to simulate (very crudely) the outcome of human cognitive modelling of a sound environment.

A study of eight different streets in a Japanese city used 11 semantic differential scales to rate the soundscape of these locations [46]. The results of the study were analysed using cluster analysis. This method revealed three different types of soundscape that comprised 1) large amounts of vehicle and human activity, 2) mixtures of human, transport and natural elements, and 3) mostly natural elements with few vehicle or human activity. The study showed that the relationships between source types can give rise to differing soundscape categories based on responses to quantitative, semantic scale based question sets.

A number of researchers have employed the technique of factor analysis to characterise soundscapes based on semantic differential scales. In one particular study [6], eighteen, seven point bipolar rating scales were used to evaluate how people perceived a selection of urban open soundscapes. Some of these scales were based on previous research relating to urban soundscape as well as product sound quality [47, 48, 49], and some were compiled for the study which applied to the soundscapes under investigation. Passing members of the public (N=491) were interviewed in two urban locations across all seasons and at different times of the day. Varimax rotated principal component analysis (PCA) was used to extract the orthogonal factors underlying these eighteen adjective scales. Four factors were extracted in the analysis. The first, accounting for 26% of the explained variance was mainly associated with relaxation, including the scales: comfort-discomfort, quiet-noisy, pleasant-unpleasant, natural-artificial, like-dislike and gentle-harsh. The second, accounting for 12% of explained variance was associated with communication, including: social-unsocial, meaningful-meaningless, calming-agitating and smooth-rough. The third, accounting for 8%, was mostly associated with spatiality, including: varied-simple, echoed-deadly and far-close. The fourth, accounting for 7% was principally related to soundscape dynamics, including: hard-soft and fast-slow.

Kang states that these four factors cover the main considerations of the acoustic design of urban public spaces: function (relaxation and communication), space and time. These four factors however, only cover 53% of total explained variance, which is notably lower than results from sound quality and environmental noise evaluations [49, 50, 51]. Kang suggests that this low value may be due to the complexity and diversity of soundscapes, with its inhabitants not fully understanding the presented terminology to describe it.

It is apparent from existing literature that soundscapes are considered to be complex entities made up of a large number of interrelated variables. Principal

Component Analysis allows us to perform a linear transformation that maps data from a high dimensional space to a lower dimensional space. The main drawback of using the PCA technique stems from its main assumption that the relationships among variables are assumed to be linear. The complex interactions between the numerous variables associated with soundscapes may require something more advanced than a linear model to explain. Nonlinear PCA (NPCA) is a viable alternative to the traditional approach as it is more robust to data based on ordinal scales, which may exhibit nonlinear interactions, as opposed to continuous data which is more suited to classical PCA [52, 53]. An investigation of this improved suitability for ordinal data was carried out in [54], showing a slight improvement over PCA when a larger sample size is used.

2.6 Laboratory research

Laboratory based research has also been utilised to further study an individual's reaction to simulated sound environments. The two main approaches either provide a simulation of high ecological validity, including contextual visual cues alongside the soundscape recreation [55, 56] , or they can be designed to exclude other factors to try and ascertain direct relationships between the acoustic variables, perception and evaluation [57, 58]. A simulation of high ecological validity is preferred in soundscape research, but experiments are not often designed with this consideration, with participants generally responding to stimuli with no contextual setup or described scenarios to set the scene. The conditions within the laboratory setting must provide ecological validity in their presentation of the environment, to transfer any findings to the real world. The recreation of these sound environments is an important consideration in laboratory experiments. Binaural recordings using a dummy head with microphones within the ear canals ensure a realistic representation of the environments acoustic properties if they were heard by a person within that environment [59]. Guastavino & Katz [60] state that the number and positioning of loudspeakers for soundscape simulation is important in faithful reproductions of acoustic stimuli. In experiments of this nature it is also worthwhile to ask the participant to rate their perception of the soundscape as being realistic [61].

2.7 Mobile phone sensing

Large scale field surveys are common in the field of health research. Paper based data collection has been the standard for decades but is prone to high error rates and is prohibited by the high cost of storage and data entry. Older style Pocket PCs have been used in the past for this type of data collection [62, 63] in a bid to overcome some of the limitations of the traditional paper technique. The major problems with these kind of devices is firstly their relatively high cost at the peak of their popularity and also their limited capabilities in terms of connectivity. Data from these kinds of devices would generally have to be downloaded to an internet connected computer, adding an extra level of complexity to the acquisition process.

Health research using more modern smartphones is gaining in popularity, with researchers leveraging the increased capabilities that they possess. In the case of a South African community health study by Tomlinson in 2009 [64], it was observed that low and middle income countries lacked the infrastructure to provide adequate fixed line internet access, whereas widespread wireless cellular networks allow access to communications and data services in regions where fixed access is limited. For a number of years in the health profession, smaller studies have been on-going into patient monitoring using mobile phones. In 2004, Anhoej, investigated the feasibility of using short messaging services (SMS) for patient asthma diary data collection [65]. The study showed an enthusiastic response to the technique, with response rates of 0.69 ($\approx$ 50% of participants provided two thirds of the requested diary data). The study highlighted that the pervasive nature of mobile phones was the driving force behind such a technique.

More recent attempts at utilising feature phones in research have highlighted the shortcomings of these older devices. One such study in Micro Blogging, [66] in 2008 made use of the Nokia N95 feature phone, providing internet connectivity and GPS localisation. The study investigated the feasibility of the platform for obtaining global user generated content for future human interaction analysis. User feedback was generally positive but the major issue came down to GUI design and usability. Participants complained of an unfriendly user interface which did not provide a smooth transition through the process of content submission, leading to a high percentage of submission drop-outs, something that could be attributed to the feature phones lack of large touch screen interface.

Studies focussed on audio sampling on mobile devices have also surfaced in recent years. The Apple iPhone was used in a 2009 study, looking into the possibility of acoustic sound event classification using the devices microphone and DSP capabilities [67]. The study showed correct classifications of around 80% for the categories: ambient sound, speech and music. An interesting application of this research was the building of audio diaries of a participant's day. With the potential for high classification rates on sound events, the app can build up an acoustic time history of sound events throughout the day. This is a promising endorsement of the capabilities of hardware DSP and microphone sampling on smartphones. The key problem with this methodology however, is the fact that the application has to be sampling continuously throughout the day, something that is not allowed by Apple on consumer iPhones (the use of iPhone sensors are only allowed while the application is running in the foreground). This limitation makes this system not practical for unsupervised large scale public use and also raises the issue of increased CPU and battery load. Whilst no mention of impaired battery life was made, increased CPU load and memory usage were observed but not at a level that would impact normal device usage.

The social sciences have also begun to exploit the widespread adoption of mobile phones. In one particular study [68], a participant's: location, Bluetooth activity and communication events were logged using their mobile phone. This observational data was used to infer friendship network structures within a set of volunteers and was also compared to self-reported survey data from the same group of participants about these networks. The study showed that whilst providing overlapping but distinct data sources from the participants, they could infer 95% of friendships based on the observational data alone.

Another research team have developed a technique for "Ambience Fingerprinting" [69]. This involves logging the output from all of a mobile phones various sensor devices (microphone, accelerometer, camera, etc.). They argue that ambient sound, light, colour and motion can be captured in a "photo-acoustic" signature, resulting in the accurate indoor localization of a person, which they have named their logical location. The fingerprint generated from a user's device is compared to an online database, which provides this location awareness. The team state that they could locate a person in the correct shop of a shopping centre 85% of the time. The experiment had participants constantly holding their mobile phones to capture camera images, something which rarely happens in real life, with mobile phones being mainly kept in pockets or bags. It does however show that researchers are recognising and utilising the power of mobile phones with their array of integrated sensors and internet connectivity.

The most prolific research project to make use of smartphones in a mass participation study is the Mappiness project from the London School of Economics [70]. This ambitious project recruits members of the public to install an iPhone app which periodically requests the participant's current location situation and mood. The project aims to better understand how peoples feelings are affected by features of their current environment such as: pollution and green spaces. The app also logs noise levels, but doesn't store or upload the raw audio samples. The projects website features a map displaying recent submissions and also average ratings of mood based on different UK locations. Thanks in part to a large amount of media coverage; the project has amassed over 50,000 participants totalling over 3 million individual mood submissions.

The key factors to the projects successes are: its use of the ubiquitous iPhone platform, the human interest element of its "Happiness" theme, its quick and easy submission process, the relevant feedback provided to participants on the app and website and crucially, the inclusion of daily "reminders" for participants to take part. The frequency of these reminders are user definable, but default to one a day, which provide an audible reminder and pop-up note on the device prompting the participant to rate their mood and submit to the study. This technique was seen as annoying to some participants, but its effectiveness is evident in the high submission counts seen.

2.7.1 Discussion

It is clear from this type of study that continued user involvement in mobile based research requires careful consideration. When asking the user to carry out voluntary tasks in order to acquire research data, some form of prompt is vital to ensure users do not simply forget to take part. The preamble of a public project such as this must also be well considered and written, so as to explain what is being done, how it benefits the user and what the outcomes are of the research.

All of these studies actually provide participants with mobile devices to use in the research. This means participants have previously agreed to take part in the study, so recruitment is dealt with before the data gathering stage begins. Projects that rely on publicly available mobile applications and voluntary participation have the added difficulty of convincing people to take part and subsequently convincing them to continue contributing to such a project. Assuming that a mobile phone based study is successful in its recruitment of voluntary participants; a whole host of challenges still remain. For example, what is the best way to validate these studies? How will the very large dataset be dealt with and analysed? What is the best way to collect ground truth data to assess the usefulness and accuracy of the data interpretation? How is the privacy of participants protected? How do researchers scale to the increasing demands of potentially thousands of participants?

2.8 Conclusion

The varied approaches to soundscape research are reflected by the multitude of disciplines studying it. The perception of these sound environments is evidently a complex process, inherently subjective and varied. It is clear from the literature that a mixed methods approach can reveal the salient aspects of these soundscapes. However, the traditional approach utilising the “stop and question” technique of data capture is limited in terms of its geographical reach and relatively high time and cost requirements. Laboratory based studies suffer from the lack of context and therefore a major aspect of the environment's perception is altered. Thus, a gap in the literature exists for a study incorporating large numbers of participants from a wide range of locations, that provides scalability that isn't cost dependent. There are also no studies to date which provide any element of participant choice into the sound environments investigated. It would seem that the application of crowd sourced consumer mobile based soundscape data may provide this more nuanced data in large amounts and from a wide range of locations.

Chapter 3

Project design

3.1 Introduction

In Chapter 2 the relevant literature in the field of soundscapes and the use of mobile technology in research were reviewed. This chapter describes the design and implementation of the studies methodology. Firstly, Section 3.2 describes the methodological approach taken for the study. Section 3.3 explains the data gathered. Section 3.4 describes the tools and methods that allowed this data to be gathered. Section 3.5 investigates the suitability of mobile devices for soundscape capture. Section 3.6 discusses the pilot studies carried out to refine the project's methodological processes. Section 3.7 explains the methods used to attract participants and the motivations behind participation. Section 3.8 details the objective acoustic feature extraction carried out on soundscape submissions. Section 3.9 then discusses the analysis undertaken on the subjective data.

3.2 Methodological approach

The approach taken in this study is one of mixed methods, leaning towards a post positivist epistemological viewpoint [71]. The complexity of the inhabitants of soundscapes and of the soundscapes themselves requires an approach that accepts that this complexity will not result in straight forward dichotomous answers. A purely positivist paradigm simply does not adequately cover all of the required facets of a field such as soundscape research with its deep rooting within the social sciences. The main data collected from participants will be quantitative, however there will be a stage of qualitative analysis on participant responses. The qualitative aspects will serve as a complementary approach to the quantitative data analysis [72, 73, 74], as they will be quantified for exploratory analysis with the quantitative data. Throughout the development of the studies' methodologies and especially the ever changing nature of its main research tool of mobile technologies, the research approach has also changed. Information gathered from participants using these mobile technologies has shaped the approach taken. Initially, with earlier devices providing limited means to enter qualitative information, it seemed that the only data retrievable was quantitative in nature. However, with the inclusion of smartphones in the methodology providing a more fluid method of textual response entry, the option for qualitative data retrieval and analysis has become available. This adaptation to the changing potential of the research tools available has been termed Bricolage by Kincheloe [75]. The concept is derived from the French term bricoleur, a crafts-person who utilises the tools at hand to get a particular job done. The mixed methods approach used in this study, facilitated by the pervasive and multifunctional mobile research platform is well described by the concept of a researcher-as-bricoleur. In a sense, the approach could also be considered to apply elements of grounded theory, whereby the studies initial pilots and trials have dictated how the tools employed have been utilised and what kind of hypotheses can be addressed through their use.

The post positivist tendencies have come about mainly because of the choice of data acquisition tool and its intended user. Volunteer citizen scientists with mobile phones can only be expected to carry out short survey type tasks to take part in a study of this kind. This limits the data available to a small number of quantitative scale based responses, single word and short open ended qualitative questions. The participant is also unsupervised, so must have no issues in interpreting the presented question set. The context and meaning behind these soundscapes is buried within the responses to these scales. Whilst it provides a holistic indication of a persons perception of aspects of their environment, the underlying context and meaning may serve to add "noise" to the dataset, blurring any potential relationships between the objective and subjective, if these even exist at all. This meaning and context is however fundamental to the perception of any environment, so its removal would reduce the usefulness of the scales to asses perception. The use of a number of sources and types of subjective data should serve to reduce the inherent variability within each of these response types.

The human centred nature of the study and its epistemological approach brings with it the ontological perspective that each participant will experience their sound environment with their own point of view and thus their own perception of reality. However, the dataset will be dealt with in terms of a combination of a realist and relativist paradigm, where the reality that participants inhabit is "real" in terms of its objective features, but is perceived in a probabilistic way that deserves a combination of subjective quantitative approaches to attempt to understand it.

3.3 What was gathered

3.3.1 Objective data

The majority of the project's objective data was collected automatically, without any user intervention or input either through the mobile application or the web interface, both of which are described in Section 3.4. Table 3.1 lists the objective data collected from each version of the project's applications and web interface.

Objective data	Description	Web	J2ME	WM	iOS
Time and date	Time and date of capture	✔	✔	✔	✔
Soundscape recording	Audio file with analysis detailed in Section 3.8	✔	✔	✔	✔
Location	Latitude & longitude of soundscape	✔	✔	(✔)	✔
Weather	Condition, wind speed/direction, temperature, humidity				✔
Area house price	Average house price of soundscape area	✔	✔	✔	✔
Demographics	Participant age & gender	✔	(✔)	(✔)	✔
Model	Device model		✔	✔	✔

Table 3.1: Gathered objective data & support; ✔= supported (✔) = gathered at web upload stage

A tick in brackets denotes an item of data which is entered at the website upload stage and not in-situ on the device. For example, some devices did not feature GPS capabilities which meant the user would be required to manually locate the soundscape when they came to upload it via the project website.

A number of soundscape studies have recorded weather data [76, 77], but none found have commented on any effects it may have on soundscape perception. As this project's study benefits from the potential to gather large amounts of geolocated data, alongside free online services to retrieve weather data, any effects that may exist can be investigated with relative ease. As the project dataset contains time, date and accurate location data for each submission, online APIs were utilised which provide historic weather data automatically from a simple web HTTP (Hypertext Transfer Protocol) call to the Wunderground.com Weather API [78]. All weather data is collected from a weather station within 3 miles of the soundscape location and within 30 minutes of the soundscapes record time. Soundscapes outside of this 3 mile distance were not used in any weather focussed analysis. The weather in which the soundscape recording was made may have an influence on a participant's appreciation of the soundscape in general. The wind speed value also has the potential to be used in the identification of recordings with high levels of wind noise. Although this was not performed in the present study, future projects involving user contributed audio recordings could utilise this technique for data screening.

House price data was gathered to investigate the relationship between the average price of an area and the perception of soundscapes from within that area. Again, free online services make gathering this data an easy and free process. Past research investigating residential quality of life has looked for indications that excess noise may have an affect on perceived reductions in house price [79]. The location and time data for each submission was used to gather average house price data from the area at which the soundscape was recorded, using the Nestoria API [80]. Data was available from these countries: Australia, France, Germany, Italy, Spain and the United Kingdom. For each submission the geometric mean price is calculated for each house in the area, all properties beyond two median absolute deviations from the geometric mean are then removed, which leads to the calculation of a new geometric mean unaffected by any very high priced houses skewing the estimate. These average house prices were then averaged again over 12 months from October 2011 to October 2012 to give the final 12 month mean house price. However, this value does not factor in fluctuations in house price across counties or countries.

Demographic data was collected to identify any potential age or gender bias in soundscape response. This information also provides a basic profile of the type of participant that is attracted to take part in a study of this nature to better inform future projects of this kind.

The device model information provides firstly an indication of how the participant has submitted to the project (web/mobile device) and secondly, which device was used to submit. This will provide details on preferred submission routes and give a breakdown of the specific device types used.

3.3.2 Subjective data

The subjective characteristics of a soundscape can be evaluated using semantic differential scales, first proposed by Osgood in [81]. The use of these scales has been successfully applied to judge the perception of complex sounds and is often used to characterise urban sound environments [6]. Participants are presented with a set of bipolar scales (see Table 3.2) that they use to judge a particular sound environment. The scales use discrete points for each increment and are not continuous. The semantic differential terms are placed at each end of the scale as shown and discussed in Section 3.4. A larger number of rating scales than the expected number of actual psychological dimensions on which the entities can be judged is essential if any further dimensionality reduction techniques are used to determine the underlying dimensions on which the stimuli are evaluated, such as principal component analysis [82].

Table 3.2 describes the subjective measures asked of the participant for each soundscape captured using the project's mobile software and web based interface. The semantic differential scales used in the project were based on those used in previous research [6, 83, 84, 85, 3], these had determined the most influential measures of soundscape quality to be the four semantic differentials shown in Table 3.2. These scales ask participants to rate a soundscapes: pleasantness, tranquillity, eventfulness and how exciting it is. Previous studies [34, 86] have also identified these measures to constitute the principal components in the assessment of soundscapes, as they produced the highest explained variance.

Subjective measure	**Description** (scale from 1 – 9)
Participant activity	What is the participant doing there (multiple choice: Passing through, Working, Relaxing, Recreation, Waiting)
Overall location quality	Rating of the location in general (bad ↔ good)
Soundscape quality	Rating of the soundscape in general (bad ↔ good)
Soundscape pleasantness	Rating of how pleasant (unpleasant ↔ pleasant)
How exciting	Rating of how exciting (boring ↔ exciting)
How eventful	Rating of how eventful (uneventful ↔ eventful)
How tranquil	Rating of how tranquil (chaotic ↔ tranquil)
Positive sound	Text entry of one of the sounds that contributes positively to the soundscape
Positive sound prominence	Rating of how much the sound stands out within the soundscape (small ↔ large)
Negative sound	Text entry of one of the sounds that contributes negatively to the soundscape
Negative sound prominence	Rating of how much the sound stands out within the soundscape (small ↔ large)
Capture reason	Open response: "Why did you choose to record this soundscape?"

Table 3.2: Descriptions of subjective question set including semantic differential scales

The predominant use of semantic differential scales is mainly due to the restrictions on data entry that the devices introduce. The semantic differential technique also provides the most transparent form of questioning, allowing the participant to answer the questions unsupervised. A help button can be pressed by the mobile device user when answering any question which provides a short text expansion on it, in a bid to clear up any confusion they may have. A more holistic assessment of the soundscape can also be achieved using forced range semantic differential scales as the phrasing of the question is such that the participant is asked to consider the soundscape as a whole rather than its individual elements.

3.4 Gathering the data

All of the objective data was gathered using the project's mobile application or audio recording device in the form of a soundscape recording with its associated time and location data. The subjective data was gathered from the device application or the project website.

3.4.1 Mobile devices

Up until the proliferation of modern mobile phones, mobile device based research required specialised devices to be fabricated to fit the research aims [87, 88]. The software applications running on these devices had to be manually installed and setup for each device. Whilst innovative, this resulted in small scale studies due to the high cost and complexity of bespoke hardware and software. More recently, the advancement of this platform allows the current project to exploit the following technological advances:

- **Sensor array & user input**
  Initially included to enhance user experience (e.g. accelerometer used to change display orientation), the logging of acoustic, motion, altitude, location, time and user experience data is now possible. Keypads and touch

screens allow for rich subjective data entry, easily storable on the device's built-in memory.

- **Open development**

  Through modern phone Operating System's (OS's) Application Programming Interface's (API's), most aspects of the device can be accessed and utilised. Programming complexity has been greatly reduced in order to promote the development of engaging applications for these devices.

- **Application deployment**

  Older style feature phones supporting the Java Mobile platform were not supported by a dedicated online portal for users to access applications for their device. Newer devices feature access to an application store for their platform, allowing developers to deliver apps to large numbers of users worldwide. This has transformed the deployment of mobile apps and allows for the collection and analysis of data beyond the scale possible using previous technology.

- **Device connectivity**

  Older GPRS networks allowed devices to use data services with relatively high costs and limited speed. Newer devices utilising 3G and WiFi allow for in-situ data upload with their increased network speeds and data packages.

The integrated sensors within these devices have seen large increases in numbers over the years. Figure 3.1 shows this increase from feature phones through to more recent smart phones such as the iPhone 4.

Figure 3.1: Sensor capabilities of feature phone compared to iPhone 4

Device types

Since the project began in late 2007, there have been a number of advances in the technological capabilities of these mobile devices. With an average product life cycle of around one year, and high consumer demand for the latest mobile communication technology, the project has had to adapt to these changes. In the lifetime of this project, three main mobile device types have emerged and established themselves into the consumer lifestyle:

- **Feature phone**

  A lower end mobile device with limited internet capabilities and a smaller screen with no touch control interface. These generally run on a proprietary operating system (OS), but support third party software via Java. They also boast a cheaper retail price and are considered easier to use. In 2011, twice as many feature phones were shipped over smartphones, however, the annual growth of feature phones is down 2.9%, while smartphone growth is up 61.3% in 2011 [89]. Access to the phones hardware can be made via third party applications; however this access is limited and requires any application to seek user permission.

- **Pocket PCs (PPCs)**
  Introduced as a palm top computer with mobile phone capabilities, these devices generally run a scaled down version of the Windows OS (Windows Mobile). They feature touch screen control and internet connectivity, as well as Wi-Fi capability. PPCs are now largely considered obsolete due to the widespread adoption of smartphones. Third party applications can be readily introduced on this platform, which provides full access to the device's hardware.

- **Smartphone**
  These provide advanced mobile capabilities beyond a typical feature phone. Screen size is generally larger than the previous two device types and offers an advanced touch screen interface (multi-touch). They combine the functionality of mobile phones, digital video/stills cameras, GPS navigation units, and full web browsers. All smartphones provide Wi-Fi and cellular data network connectivity, high resolution touchscreens and GPS receivers. One of the most significant differences between smartphones and feature phones is that the advanced application programming interfaces (API) on smartphones have a better integration with the phones OS and hardware. The most common OSs found on these device types are: Windows Phone [90], Blackberry OS [91], iOS [92] and Android [93].

Software considerations

In the project's early stages, participants were encouraged to download a small application for use on their mobile device. A number of versions of the mobile software were available to cater for the differing capabilities of current and future devices and to ensure compatibility and maximum participation potential. There were two major versions of the mobile software during phase one of the study. These were a Java 2.0 Mobile Edition based (J2ME) version with the other based on the Windows Mobile 5.0+ platform (WM). The Java mobile software exploits the audio capture functionality of the mobile phone through the Java JSR135 (Multi Media API) [94]. This Application Programming Interface (API) allows the capture of audio through the device's microphone. The possible length of recording is limited by the phone's hardware capabilities. The J2ME version could potentially be run on a diverse range of mobile phone platforms with varying capabilities. The high level nature of the Multimedia API meant that recordings could not be streamed directly to a file. The recording limit was therefore defined by the amount of random access memory (RAM) available to Java. This could be anything from a few megabytes down to a few kilobytes depending on the current state of the phone. To account for this, a ten second limit was set for all mobile recordings on J2ME devices. Whilst this ten second limit on older devices may seem short for soundscape characterisation, the majority of the objective metrics used in Section 3.8 do not require longer signal lengths. The length of the recording may however mean that the positive/negative sources identified by the participant are not necessarily captured in the recording. Due to the large sample size, each recording cannot be scrutinised, but any errors that this will cause will be limited by the large sample sizes attained. However, this is an assumption that could affect the results.

iOS devices allowed for a longer record time, capped at three minutes to ensure upload times over mobile networks were not too long.

As well as the capturing of audio, all device configurations allow for the logging of subjective response data from the participant. The software prompts the user to enter short worded responses and select values from semantic differential scales. The data gathered from the participant while immersed in the soundscape provides a more accurate impression of the impact of the soundscape on the individual, as they are responding in situ and not relying on memories, which can create inaccuracies.

The amount of data requested from the participant had to be kept to a minimum to ensure the user does not get bored or frustrated. The cross device support required on a project of this kind brought a number of problems in terms of the successful acquisition and integrity of the acoustic data that these devices could capture. The Java ME platform only provided high level access to the microphone and file system functions in the mobile devices. It was necessary to purchase a Thawte Code Signing License [95] so the app would sit in the “Trusted” domain, removing the need to constantly ask the user for permission to perform tasks. Early versions of the mobile app were prompting users, on average, 20 times per 10 second soundscape recording, which obviously frustrated and confused participants. These prompts also rose suspicions as it made it seem like the app was attempting to perform malicious operations on the device.

The differing audio recording capabilities of mobile phones were another factor. Newer models allow uncompressed wave file capture up to sampling rates of 44.1kHz. Older models however could only record in a compressed AMR-WB format (Adaptive Multi Rate Wideband Compression) optimised for speech encoding. This format provided a maximum bit rate of around 16kHz at 13 bit producing a resultant signal with a filtered frequency range of 50-7000 Hz. Recordings made using this codec were quite rare (and were removed from the data analysis stage in Section 4).

Submission paths

The J2ME and WM editions of the project application did not feature a direct upload function due to connection and data cost factors, so uploading of audio and response data was done via the project's website shown in Figure 3.2. Both of these versions embedded the subjective responses and objective meta data into the header of the audio file which was extracted and uploaded at the project website upload stage. Participants using their own digital recording equipment were prompted to enter their subjective ratings of the soundscape retrospectively at the upload stage.

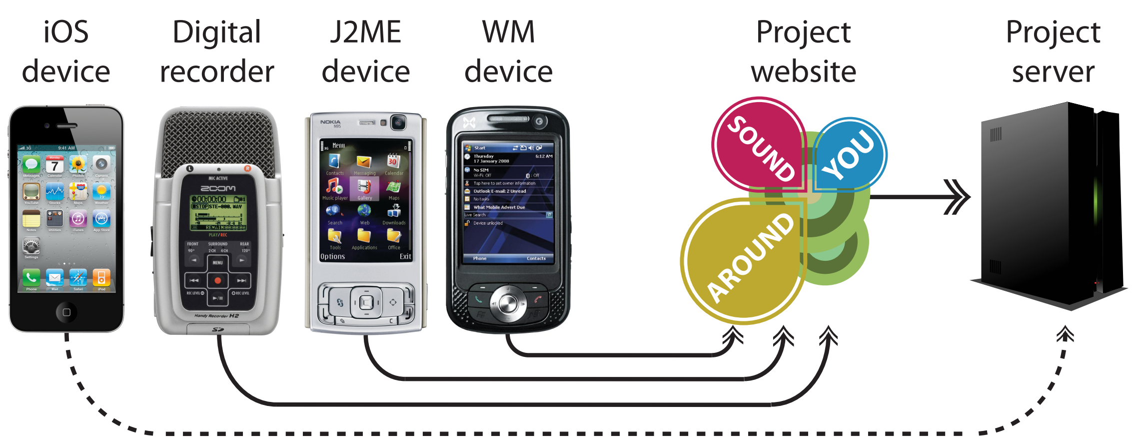

Figure 3.2: Differing submission paths from supported devices

The iOS app, which was named i-SAY, bypasses the website upload stage allowing participants to submit their soundscape recording and responses directly from the device to the project server. The dashed line depicting the iOS apps submission path in Figure 3.2 signifies a wireless transfer of data, whereas all other methods require some form of wired or bluetooth connection to a computer for soundscape transferral and upload via the project website.

Response process

Figure 3.3 gives examples of the older J2ME and WM versions of the project application. The WM version provided navigation via a touch screen interface controlled with a small stylus with text input via a pop-up on screen qwerty keyboard. The J2ME version was controlled using a directional keypad with text entered using the number keypad.

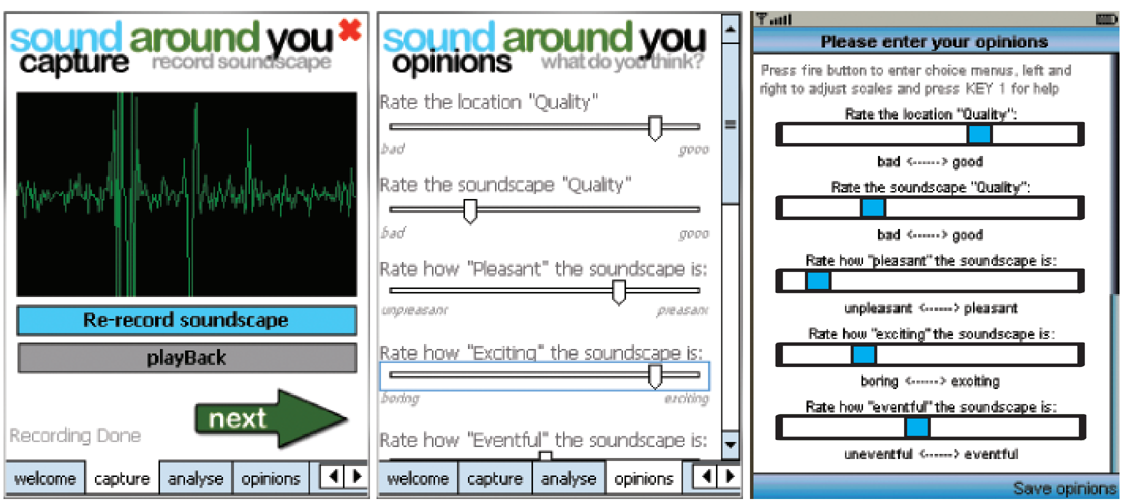

Figure 3.3: Screenshots of WM (left two) & J2ME project application showing capture and opinion screens

Figure 3.4 shows the functionality and major processes of the J2ME and WM applications. This older version was essentially a sound recorder that also logged associated subjective responses from the user.

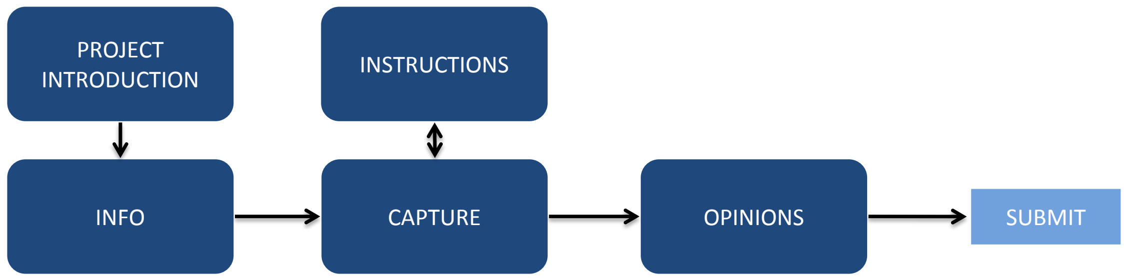

Figure 3.4: Participant flow through the J2ME & WM application

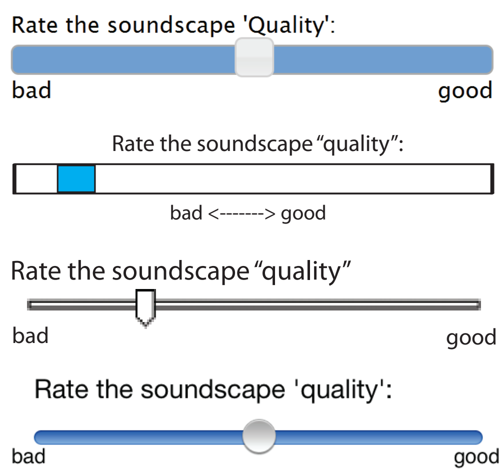

Figure 3.5: Web, J2ME, WM & iOS opinion sliders (top to bottom), default position at center

Figure 3.5 shows the different slider designs across platforms. Although they are not all visually identical, their functionality is the same with all using unlabelled discrete scale points ($1 \leftrightarrow 9$). The web slider is dragged using a mouse, the J2ME slider is incremented/decremented by pressing the directional pad on the device, the WM slider is dragged using the device's stylus and the iOS slider is dragged by the participant's finger.

Text input for the positive/negative source identification questions is done using the device keypad on the J2ME application using Multi-tap (number pad represents collections of letters per key, with multiple taps on each key allowing for desired letter selection) or T9 Predictive text entry [96]. WM and iOS devices provide a software based touch screen QWERTY keyboard for text entry, including automatic and manual spelling correction features.

Initially a PC application was also included to act as an intermediary between the mobile application and the project server, easing the upload and comment process. It became clear during the pilot studies (see Section 3.6), that the effort required to participate needed reducing. The extra application installation and use was seen as a barrier to participation for early adopters, and so the PC stage was removed. The revised process allowed members to upload directly to the project's servers from their mobile device connected via USB to their home computers. Throughout the final pilots, users responded favourably to the new cut down methodology and commented that it made the whole process easier and more appealing.

All versions of the application provided instructions to the user so as to reduce any kind of user error, resulting in less than adequate soundscape recordings. The iOS version with its larger touch screen interface allowed for a clear instruction page shown in Figure 3.6. On the app's first load, this information page had to be viewed before a soundscape recording could be made. This was to ensure that users were aware of any issues that could affect recordings, such as minimising wind noise and making sure that no microphone peripherals are plugged in.

Figure 3.7 shows the three main stages of the iOS version of the project app. In all versions of the project application the screen orientation rotated 180° before the soundscape recording commenced. This was to ensure that the device's mi-

Figure 3.6: Instruction page for iOS users shown before recording commences

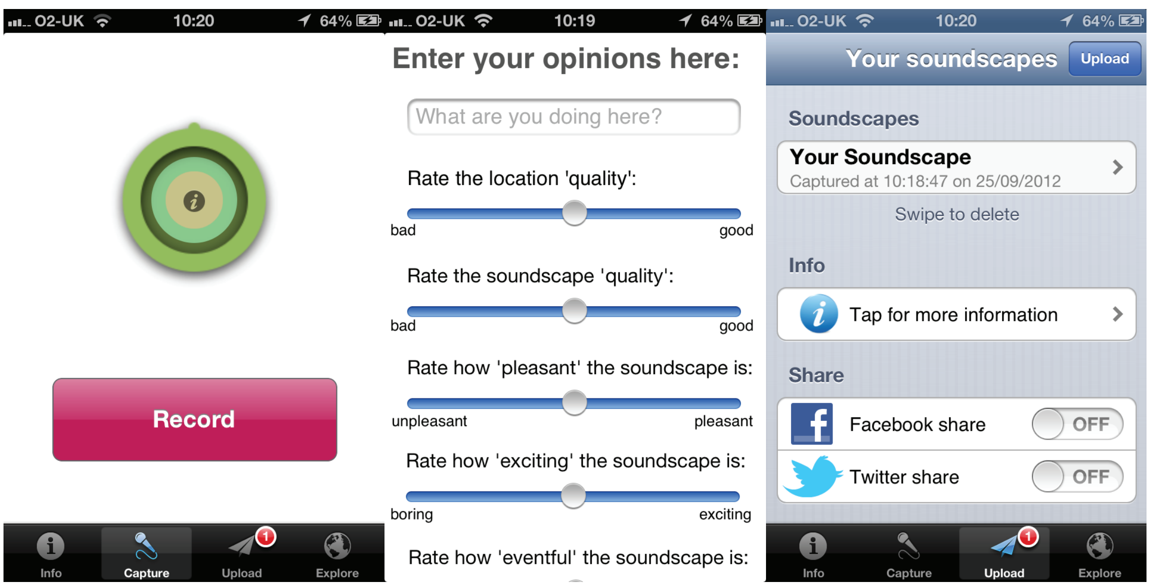

Figure 3.7: Screenshots of iOS version showing capture, opinions and upload stages

crophone, which is generally located at the bottom, is kept free and not blocked by the users hand, something which was identified in the pilot study discussed in Section 3.6. The far right screenshot in Figure 3.7 also includes the Share options available to users submitting soundscape recordings, which are explained in Section 3.7.

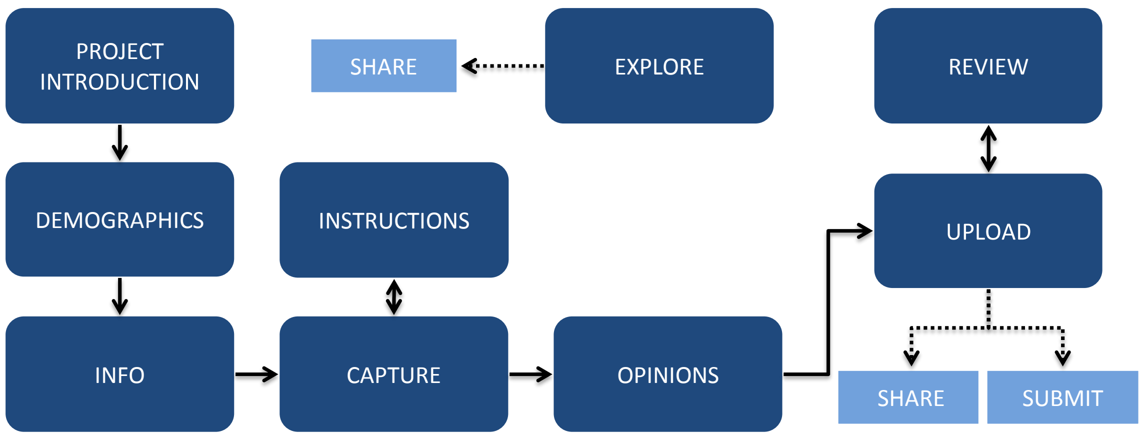

Figure 3.8: Participant flow through the iOS app

The advanced functionality and features of the iOS version are shown in Figure 3.8. As all data capture and submission is carried out on the device, the extra demographics and upload stage mark the major differences from the previous version. The Explore map allows the user to browse the project's submissions to date on an interactive world map, where soundscapes submitted by them appear as red markers rather than green. Another new addition is the Share feature, allowing participants to post on Facebook or Twitter when they have uploaded a new soundscape or when they have found one which they find particularly interesting on the Explore map. This concept of soundscape sharing is an important aspect in keeping the project alive and attracting and maintaining interest from new and existing participants (see Section 3.7). The review page provides a place for participants to view and playback their yet to be uploaded soundscapes. This allows them to delete and rerecord soundscapes or keep them purely for personal interest.

3.4.2 Calibration

The desire to obtain accurate sound level measurements brought about lengthy discussions on how to carry out calibration with participants whom may not have any prior acoustics knowledge. A repeatable sound source with a known sound level was needed that could be generated in a domestic situation. Calibration was also considered to be a potentially educational and enjoyable experience, providing participants an insight into an important process in acoustic measurements. A number of interesting ideas were produced to resolve the calibration issues, including asking participants to drop a fixed weight object from a fixed height onto a known object, measuring the impulse produced and so providing a known level that could be used as a calibration signal. However, tests found these types of domestic calibration signals varied greatly in sound level. With the project already asking much of its participants, it was concluded that this extra stage of calibration was likely to be seen as an annoyance to potential contributors and so was removed from the process. The drawback of this is that all of the project's submissions are un-calibrated in level. However, the objective metrics calculated in Section 3.8 mainly consist of spectral or relative level features, which are not entirely reliant on absolute level.

The potential for level discrepancies between iOS devices is minimal, due to the similarities in hardware across the Apple range. The largest contributor to level inaccuracies will be user handling differences across submissions. A participant holding the device closer to the body will result in a reduced overall signal level compared to holding the device out from and above the body. The location of the phone cannot be determined when recording, therefore the effect cannot be compensated for when the submission has been received. The large sample size should mitigate this effect to an extent, however, significant discrepancies may be observed because of this.

3.4.3 How we got the applications to the people

Developing a methodology that requires large numbers of participants is common in many fields of research. Getting a critical mass of participants was a major concern in the project. It was clear that one of the major barriers to people taking part was how they would access the project's mobile application.

- **Feature phones and PPCs**
  For a potential participant to install the Java ME edition of the project application on their device, two options were available. First of all, two files are downloaded onto a home computer, which are then copied to a mobile device. The participant must then locate these files and launch them to install using the device's file browser. Another option available is the setup of an online website which provides a link that can be accessed on a mobile device providing a faster route through the applications installation, called Over The Air Provisioning (OTA Provisioning).

- **Smart phone** Smartphone OSs provide a dedicated application distribution channel, allowing for application installation via a very quick and easy process. The Apple App Store charges developers a yearly fee of $99 to allow apps to be submitted.

3.4.4 Web contributions

Another vital tool for data acquisition is the project websites Contribute feature. The nature of the project attracted participants with an interest in sound and environmental recording, who preferred to use their own digital recording equipment to capture soundscapes. Other participants wishing to take part using mobile devices not able to run the project's software could use any recorder on their device to record soundscapes. To cater for these participants and to gather more data, the project's website was enhanced to include a soundscape upload feature and survey to collect the users opinions of them. This was implemented using HTML, JavaScript, Flash, MySQL and PHP, including extensive use of the Google Maps API [97]. Figure 3.9 shows the demographics form users must complete before providing them with the upload and comment features. These are only asked once as when completed the details are stored locally on the users machine in the form of a browser cookie [98].

Once the user has completed the form shown in Figure 3.9, they are presented with the Contribute form, depicted in Figure 3.10. Firstly instructions are given to the user on how to proceed, beginning with the selection of a soundscape audio file to upload. Size and format checks are performed after selection to ensure the file is of the correct type and below 20MB in size. After file upload the user locates the soundscape using the Google Maps draggable marker seen in Figure 3.10. The final stage is the completion of the survey question set using the mouse to

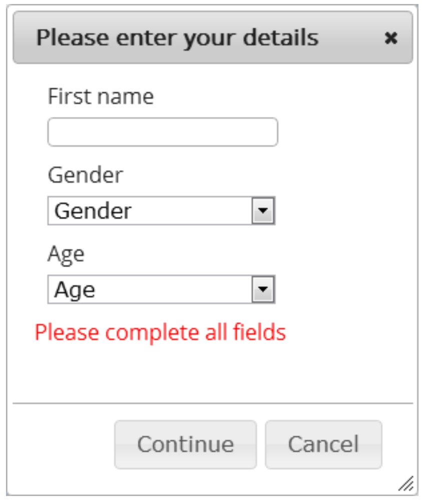

Figure 3.9: Web based demographics form showing data completion check

drag the semantic differential sliders and keyboard for text entry questions.

Once all elements of the form are completed, the Submit button is enabled which inserts all user responses into the MySQL database. Submissions will then be viewable on the soundscape map shown in Figure 3.11, which allows visitors to listen back to submitted soundscapes, view subjective responses, share soundscapes and if a Google Streetview [97] street is nearby, visualize the location.

3.4.5 Data storage

The website was hosted on a consumer server package including a MySQL database. All project data was stored securely on a single database table, making the process of data collation much faster and more efficient. Data analysis applications such as Matlab and SPSS also include relevant database support, allowing the most up to date project data to be downloaded on demand. Common to the majority of hosting companies is the addition of automatic database backups ensuring

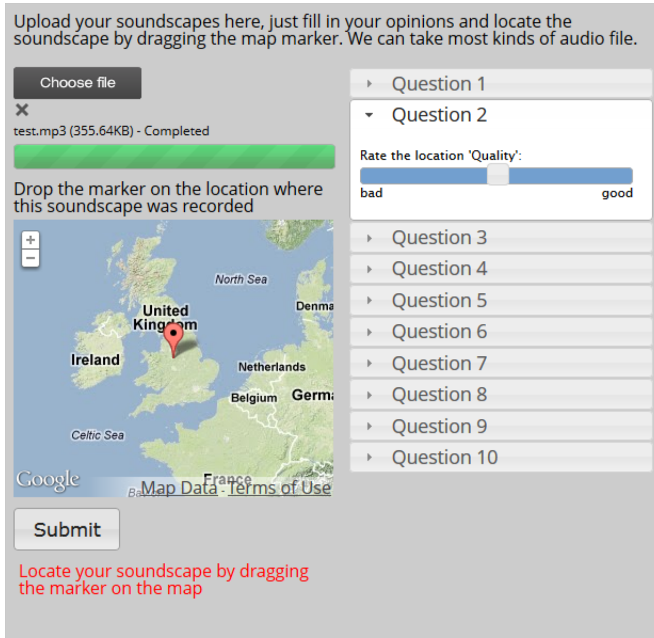

Figure 3.10: Web based upload, locate and opinions form

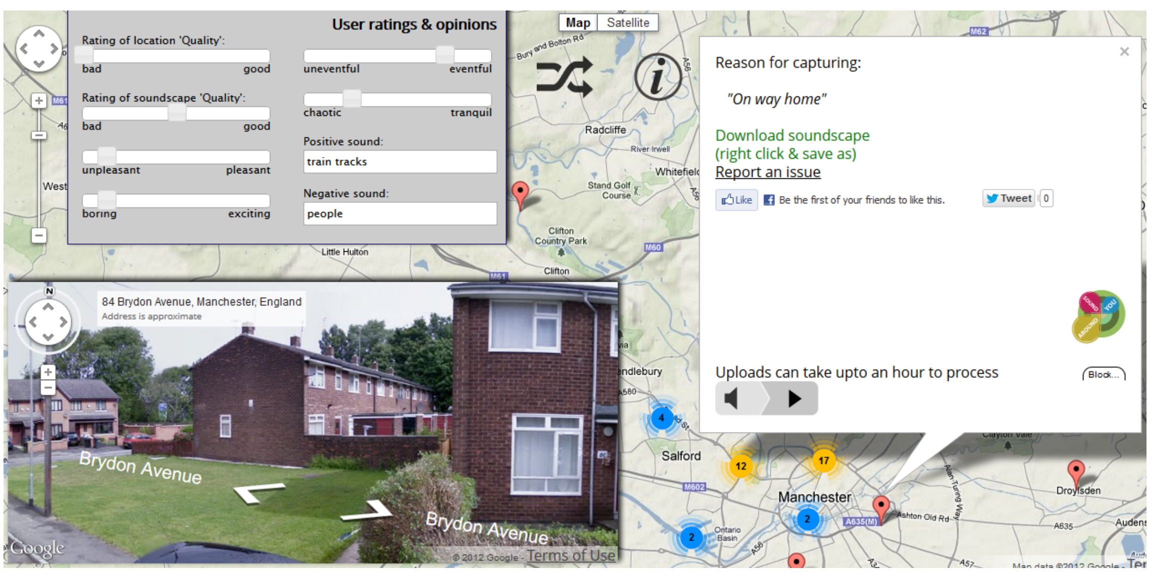

Figure 3.11: Soundscape map showing submissions with associated responses and Streetview

that all research data is securely and safely stored.

3.4.6 Web and mobile response considerations

An important methodological difference lies between the web and mobile paths to submission. Whilst the questions asked of participants utilising the project's web and mobile phone submission paths are identical, the time at which these responses are made will be different. Participants using the project's mobile applications provide their subjective responses as soon as the soundscape recording has been made, meaning they are given in-situ whilst the participant is still immersed in the sound environment. Web submissions are presumably responded to at the participant's home computer in a reflective manner, where the responses will be based on their memories of the locations soundscape. The participant may have also listened back to their recording before making their web submission, however, this eventuality was not logged.

It is important to acknowledge this difference between the web and mobile phone submission routes and any differences between these responses in terms of their subjective ratings will be investigated and quantified in Section 4.2.1.

3.5 How good is a phone at recording audio?

In order to gauge the effectiveness of a mobile device for the recording of soundscapes, a set of test signals were designed which subjected the device microphone to a logarithmically swept amplitude signal repeated at a number of increasing octave frequency bands. Very few studies have carried out testing of this nature on mobile devices, where the closest identified research compared iOS devices to Type 1 sound level meters (SLM) in terms of their suitability for sound pressure level measurements [99]. A custom J2ME application was created that captured extended length audio at the device's highest possible sample rate, which in the case of the J2ME devices tested was 44.1kHz, 16 bit, uncompressed wav files. The iOS devices ran another custom application capturing uncompressed wav files of 44.1kHz, 16 bit.

All response plots begin at 40Hz in order to exclude content that is too low in frequency for the devices to transduce and end at 16kHz as this is the upper frequency limit at which recordings are made on the release version of the project app due to device memory and processing restrictions.

The following test signals were generated to ascertain the devices suitability for this application.

1. **Logarithmically swept amplitude signal at octave centre frequencies:** Measure dynamic range and identify any compression effects including effects on differing frequencies

2. **Continuous level 500Hz tone with increasing amplitude 0.5s pulses 2.5s apart:** Identify auto gain control with response to differing pulse amplitudes

3. **Filtered white noise (boosted/attenuated to produce a more uniform frequency response on playback):** Measure frequency response of device

Comparisons were made with a reference logging type 1 sound level meter (Svantek 959), under anechoic conditions. Each device was mounted at the centre point between the calibrated anechoic hearing defender speaker array [100] as shown in Figure 3.12. The device was mounted with the microphone port facing upwards, mounted in the same position for all devices. The hearing defender rig consists of 4 speakers surrounding the device all at a distance of 1 metre.

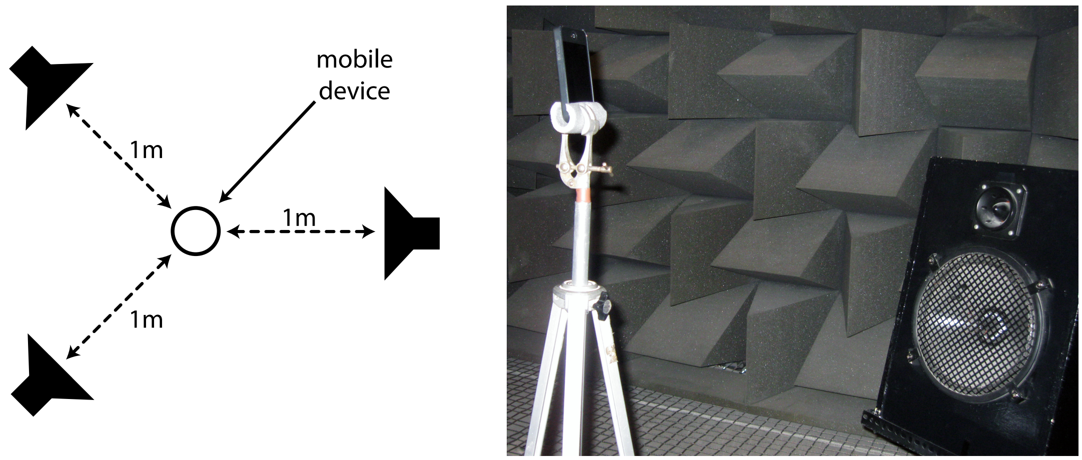

Figure 3.12: Test setup in anechoic chamber: top down view and device mount

The frequency response curves of the device microphones were produced using the substitution technique, alongside the point by point method in the frequency domain, as specified in British Standard 60268-4 [101]. The calibrated measurement microphone was used to record 60 seconds of filtered white noise to be used as a reference measurement. The white noise signal was created to increase measurement accuracy at low frequencies due to the loudspeakers observed roll-off in reproduction level at around 100Hz. The measurement microphones frequency response is assumed to be flat from 20-20,000 Hz. The device was then placed with its microphone port in the same position and orientation as the measurement microphone and then subjected to the same white noise signal. These two signals were then converted to the frequency domain using a moving average 44100 point Short Time Fourier Transform (STFT) using a hanning window of 1 second in length with 50% overlap across consecutive windows.

All measurements were taken 3 times, ensuring that the microphone port was in the same position each time. Device response was finally produced by dividing the mean of the device spectrum by the reference spectrum, with all responses normalised at 500 Hz as this was identified as the frequency at which all devices displayed the most consistent response. Apple devices running the iOS operating system are displayed separately from the older J2ME and Windows Mobile devices to ease plot interpretation. Frequency response plots were also smoothed to aid in visual comparisons using third octave averaging.

3.5.1 Java mobile edition and Windows mobile devices

Frequency response

Figure 3.13 shows the Java Mobile Edition (J2ME) & Windows Mobile (WM) devices free field frequency responses relative to the measurement microphone. The Nokia N95 was the J2ME device and the MWG Atom was the WM device.

The first aspect of note is the unusually high, low frequency response of the MWG. This is due to the very high noise floor of this device. A clearly audible rhythmic buzz is heard on all recordings, producing a recording only able to adequately capture higher level audio events due to quieter events being masked by the very high noise floor of the device. Large troughs at 1.4, 3.5 and 9.4kHz show attenuations of upto 21dB, which could be attributed in part to the high frequency components of the noise floor. These response fluctuations could also be due in part to the device's casing disrupting the passage of sound to the microphone, which is located within a small duct at the device's base to ensure it is protected. The opening to this duct is around 1mm in diameter and around 5mm deep, which would explain the device's highly varied response at higher frequencies, potentially

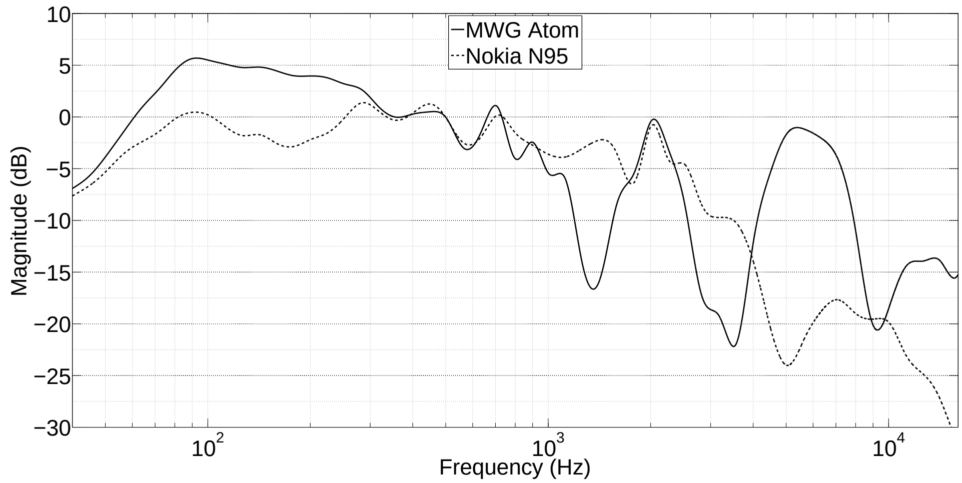

Figure 3.13: J2ME & WM device frequency responses (1/3 octave smoothed) relative to measurement microphone (normalised at 500Hz)

being caused by this duct and the device casing, increasing the directivity of the microphone setup. A low frequency roll-off of 8dB/octave begins at around 90Hz, presumably in place to reduce the effects of wind and operator handling noise. The Nokia N95 shows reduced low frequency sensitivity with a slightly slower low frequency roll-off than the MWG. The steep high frequency roll-off beginning at 2kHz drops as low as 24dB at 5kHz, which has the potential to effect the outcome of any frequency content based analysis of soundscapes from this J2ME device type.

The frequency response of the MWG was also measured using white noise signals of varying amplitude to highlight any change in frequency response at differing levels, where 100% represents the maximum sound pressure level the device can record at a distance of 1 metre from the loudspeaker array.

As can be seen in Figure 3.14, at frequencies upto 1kHz there is only very

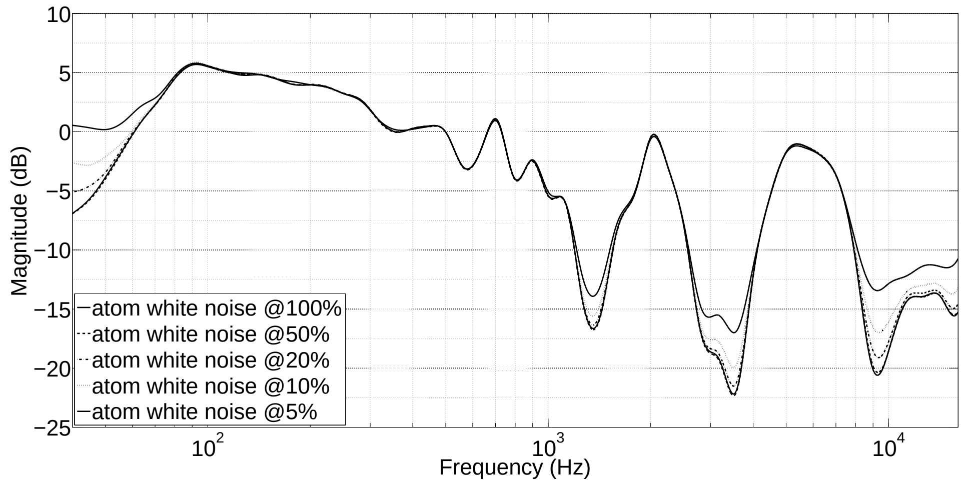

Figure 3.14: Frequency response of MWG Atom at 100% , 50% , 20% , 10% and 5% source level (normalised at 500Hz)

small differences in response. The troughs mentioned previously do show some variation between source levels, which after investigation, proved to be caused by the noise floor of the device containing components at these frequencies.

Another outcome of these tests was the identification of an initial impulse, introduced at the commencing of recording, shown in Figure 3.15 for the MWG Atom device. These were seen in all J2ME and Windows Mobile based devices and seem to be an artefact of the call to the record start function of the J2ME version and the beginning of the audio buffer copying on the Windows Mobile version. This sharp impulse would affect any further objective analysis of the signal so it had to be removed.

To remove this initial impulse from the final signal, a linear amplitude gradient/slope was applied to the first 100ms in order to gradually fade in the audio signal. This was performed before the signal was saved to file on the device.

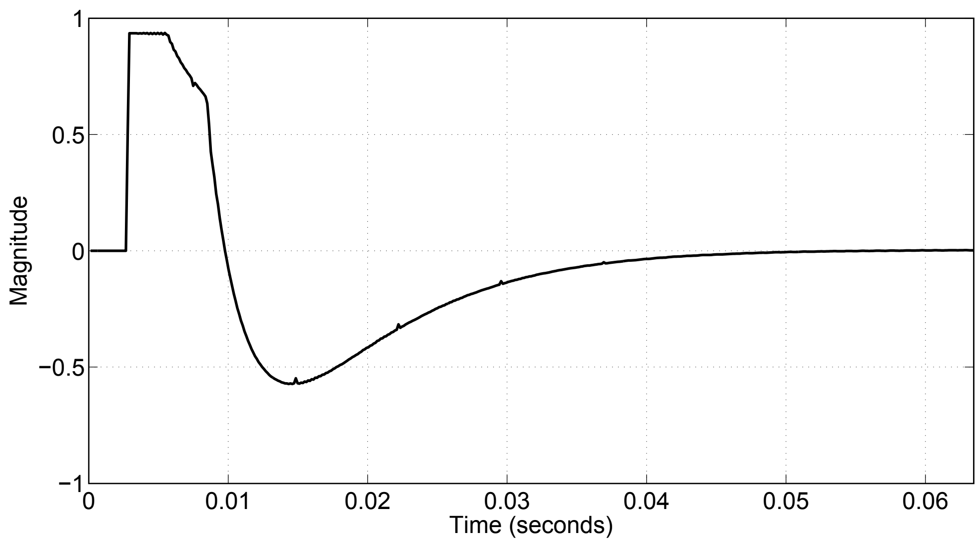

Figure 3.15: Initial settling of recorded signal (50ms pulse) MWG Atom

Level response

To investigate how these devices react to varying sound levels, octave band pure tones were logarithmically swept up in level in an anechoic chamber beginning at 20dB (ref. 2×10^{-5} Pa.). Measurements were made of these sweeps with all devices and a Type 1 Sound Level Meter (SLM) was used as a reference, to be compared with the mobile devices. Figure 3.16 shows the MWG Atom level response to a 1kHz pure tone, sweeping up from 20dB to 108dB at 1 metre, smoothed using an envelope detection averaging process to ease plot interpretation.

The high noise floor is evident at around 65dB, which effectively means that this device cannot adequately capture sound levels below this floor, as the signal will be masked by the device's internal noise. The highest level captured at this frequency is around 88dB, providing a rather small practical dynamic range of 23dB. The device only really matches the reference response between 77 and

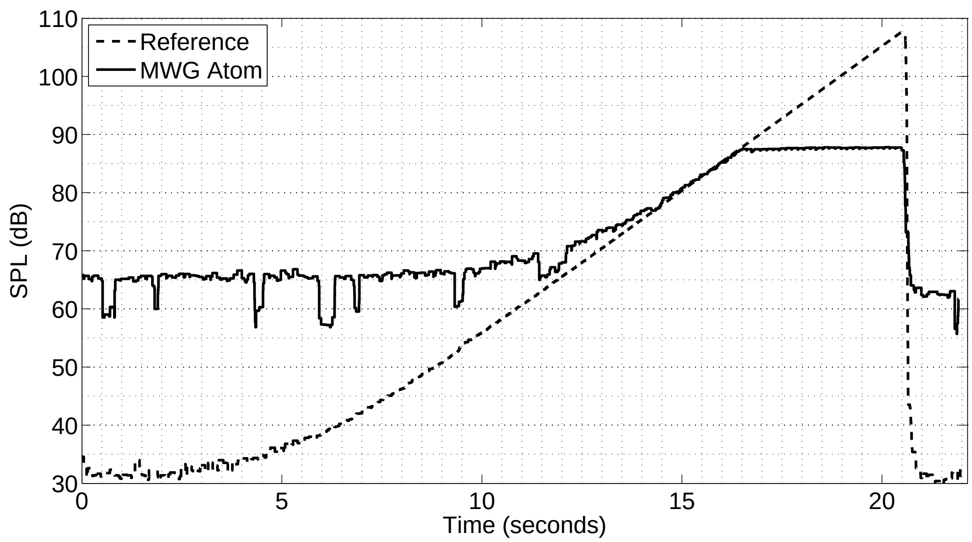

Figure 3.16: MWG Atom 1kHz tone level response compared to Type 1 SLM (reference)

87dB, a major disadvantage when attempting to capture the high dynamic range of some soundscapes. The Nokia N95, a J2ME device with very similar level response to the Windows Mobile MWG Atom shows another interesting trait, as the recorded signal does not distort when it hits its maximum level; a form of compression is applied by automatically adjusting the sensitivity of the microphone level. This can be observed in Figure 3.17, where the peak level is reached and then attenuated by 10%, which then rises over 150ms, where the process repeats.

This is commonly known as Automatic Gain Control (AGC) and is in place to handle the varying distances a mobile phone user may hold their phone microphone away from their mouth to maximise signal to noise ratio and reduce distortion. Unfortunately, this mechanism appears to be hard coded into the device's functionality so there was no way to bypass it via the J2ME Multimedia API

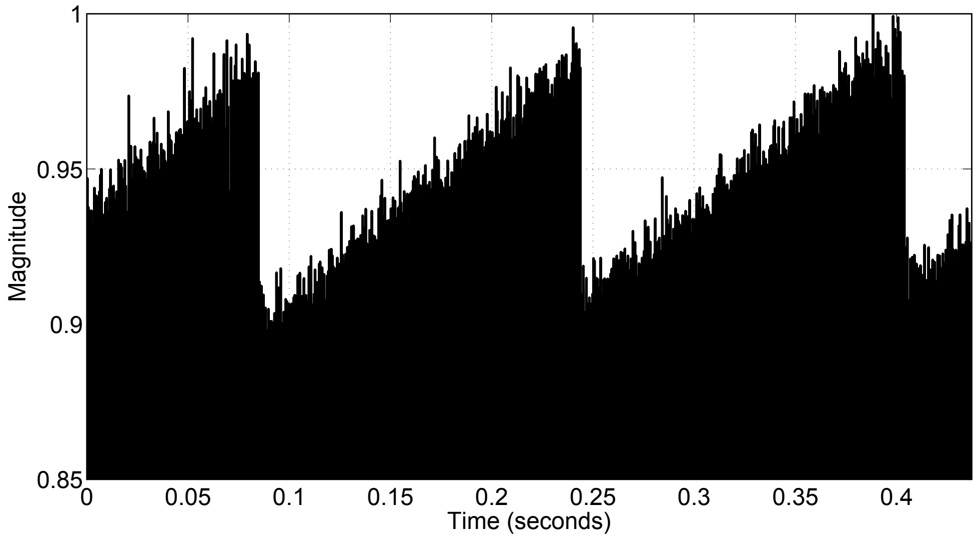

Figure 3.17: Close-up of N95 level response at 1kHz saturation point showing AGC

or Windows Mobile API. While recordings made using these devices may be limited in terms of their level information, one benefit of AGC is that it reduces the side effects of distortion, such as the creation of harmonic content, colouring the recorded soundscape in the frequency domain, which may have an effect on any frequency content focussed objective metrics used in the analysis Chapter 4.

The MWG response to the pulsed 500Hz signal in Figure 3.18, shows the AGC attempting to compensate for a signal that could potentially distort. The final two pulses depict the action of the AGC system, drastically attenuating the signal when it is subjected to an instantaneous rise in level. The AGC reacts differently to higher level pulses as can be seen in the central peaks. The system is triggered when the level exceeds a certain threshold and attempts to reduce the gain until it reaches the systems' optimum input level before clipping and distortion occurs.

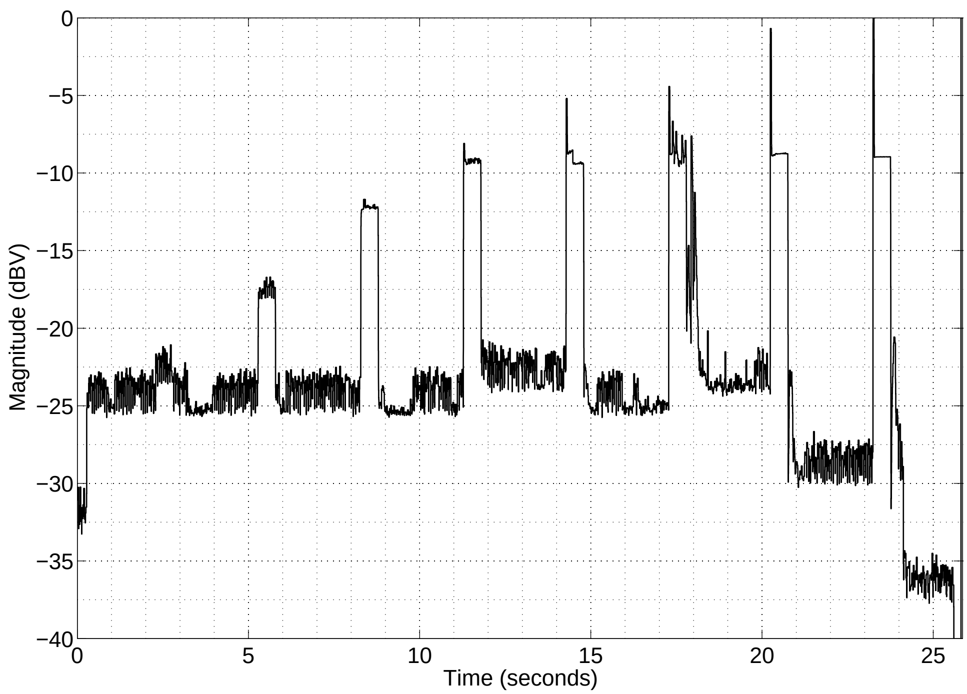

Figure 3.18: AGC response of MWG Atom to increasing amplitude pulses at 500Hz

The final two pulses also reveal the systems reduced input gain after the pulses have ended, illustrated by the reduced amplitude portions directly after the pulses.

3.5.2 iOS devices

Because of the number of iOS devices tested, the devices running the previous version of iOS (iOS 4) will be presented separately from the more recent iOS 5 devices. The iOS 4 devices were the older iPhone 3G and the next model iteration, the iPhone 3GS. The iOS 5 devices were the iPhone 4 and the larger tablet device, the iPad 2.

Frequency response

Figure 3.19 shows the iOS 4 device free field frequency responses relative to the measurement microphone.

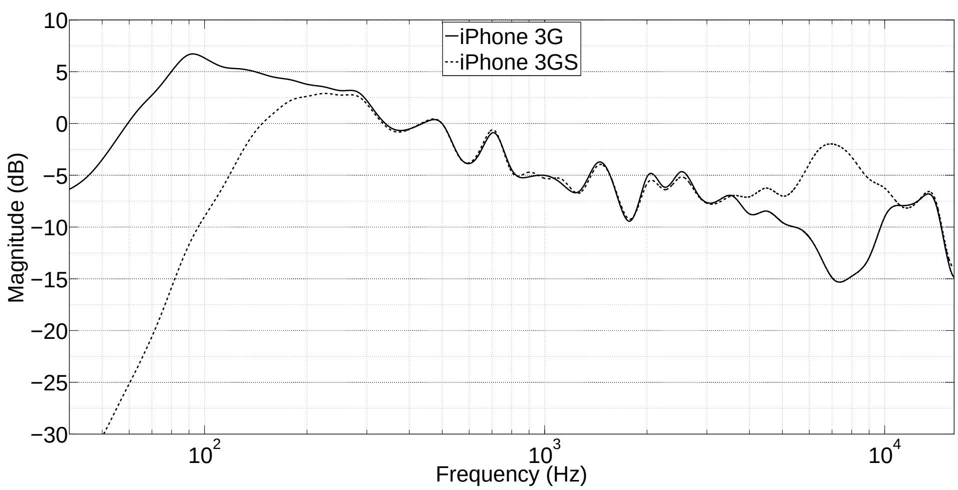

Figure 3.19: iOS 4 device frequency responses (1/3 octave smoothed) relative to measurement microphone (normalised at 500Hz)

It seems apparent that the microphones in these two devices are very similar, exhibiting roughly the same response from 300Hz to 3kHz. However, the low frequency roll-off seen on the iPhone 3G begins much lower in frequency at 90Hz and also occurs at around 12dB/octave. The 3GS exhibits a much earlier and steeper roll-off of around 24dB/octave, beginning at 200Hz. This enhanced high pass filter on the 3GS is clearly an attempt at further reducing wind and handling noise and is unfortunately hard wired into the input stage of the microphone system. Any objective metrics concerned with low frequency energy of signals may be skewed by this feature, but any attempt to boost the signal in this range will only serve to amplify the noise floor and colour the overall soundscape. The differences seen at high frequencies could be due to small variations in measurement position and the age difference of the devices, with the 3G being in much poorer physical condition, with more chance of dust blockages over the microphone port. These differences will be more prevalent in the field due to the differing conditions of participant devices and device orientations used when making soundscape recordings. These measurements do however provide a useful comparison between the type of devices that may be used by participants around the world.

Level response

The level responses to white noise shown in Figure 3.20 show the iPhone 3GS's consistency in frequency response under different level conditions. To determine its noise floor and dynamic range, the phone was subjected to a 1kHz tone increasing in level in anechoic conditions. Figure 3.21 shows an improved noise floor level of 48dB over the MWG Atom discussed in Section 3.5.1.

The dynamic range of the device at this frequency also shows a large improve-

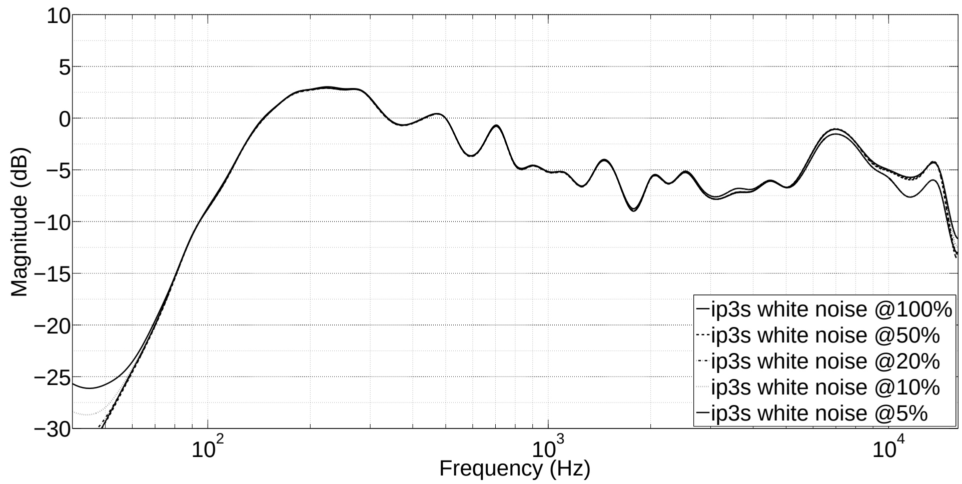

Figure 3.20: Frequency response of iPhone 3GS at 100% , 50% , 20% , 10% and 5% source level (normalised at 500Hz)

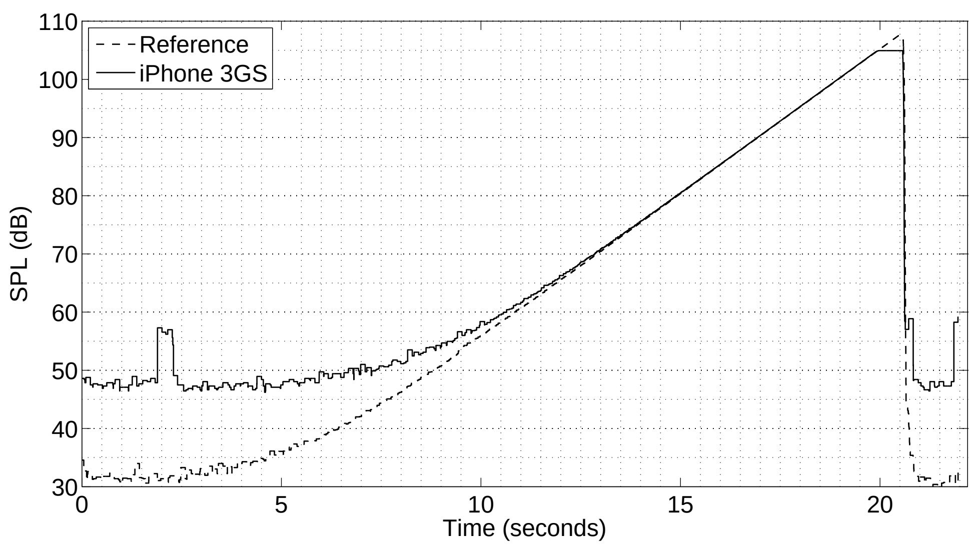

Figure 3.21: iPhone 3GS 1kHz tone level response compared to Type 1 SLM (reference)

ment over non iOS devices, with a maximum level of 105dB and practical dynamic range of 57dB. The 3GS closely matches the reference response from around 60dB upto its 105dB limit, providing an adequate range for soundscape recording.

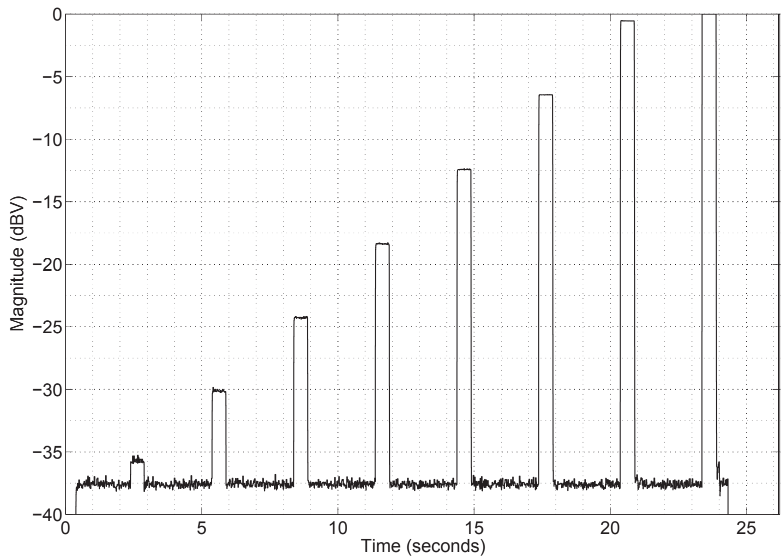

Figure 3.22: Response of iPhone 3 to increasing amplitude pulses at 500Hz

The iPhone 3G shows a clear improvement over non iOS devices as it does not seem to employ any form of AGC to the input signal. Figure 3.22 shows no attenuation to the pulses themselves or the period after the pulse. This was observed for all devices tested running iOS 4. Figure 3.23 shows the iOS 5 device free field frequency responses relative to the measurement microphone. The high pass filter can be seen at the low end as was observed in the iPhone 3GS.

At low frequencies the response between devices is similar, however at higher frequencies the response of the iPad 2 shows large variations around 2 and 7kHz. This will be caused by the larger size of the iPad and the different placement of its microphone within the body. The iPhone 4 microphone is mounted on the bottom of the handset, whereas the iPad has the microphone placed slightly back from

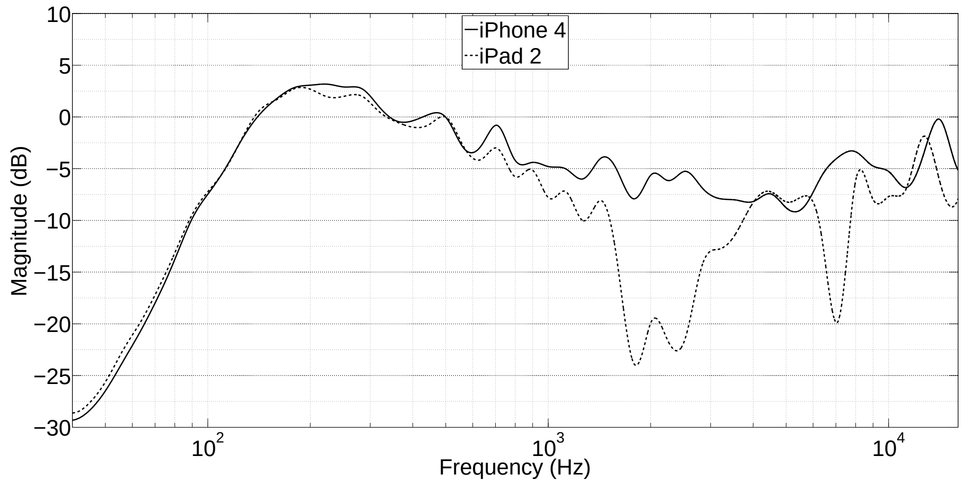

Figure 3.23: iOS 5 device frequency responses (1/3 octave smoothed) relative to measurement microphone (normalised at 500Hz)

the top of the device. Diffraction of high frequency signals around the casing is likely to produce frequency troughs such as those observed in Figure 3.23.

The iPhone 4's response to differing levels of white noise does not seem to affect its frequency response, seen in Figure 3.24. Loss of high frequencies can be seen when the level is at 5%, which could be attributed to the frequency content of the device's noise floor masking some of these frequencies.

The iPhone 4's response to an increasing amplitude 1kHz tone seen in Figure 3.25 shows that its noise floor sits at around 45dB, with a maximum level of 100dB, giving a practical dynamic range of 55dB, slightly less than the iPhone 3GS, but still suitable for soundscape recording. This response was also closely mirrored by the iPad 2. An interesting effect is seen when these iOS 5 devices are subjected to the pulsed signal. Figure 3.26 shows a large dip in amplitude after the final three pulses.

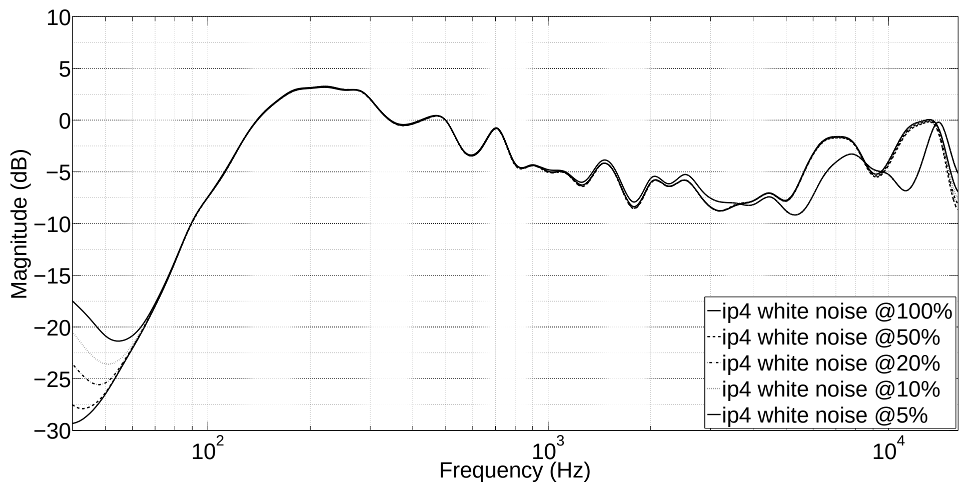

Figure 3.24: Frequency response of iPhone 4 at 100% , 50% , 20% , 10% and 5% source level (normalised at 500Hz)

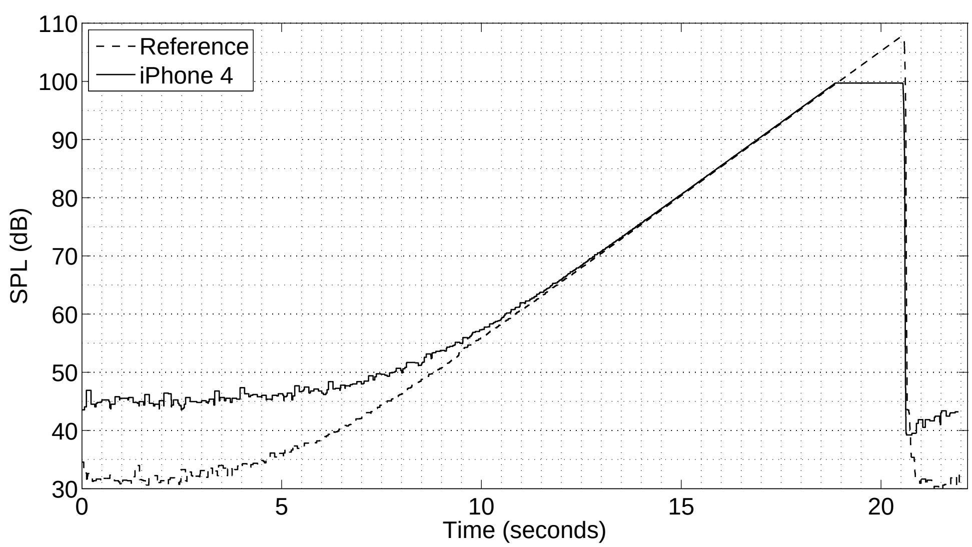

Figure 3.25: iPhone 4 1kHz tone level response compared to Type 1 SLM (reference)

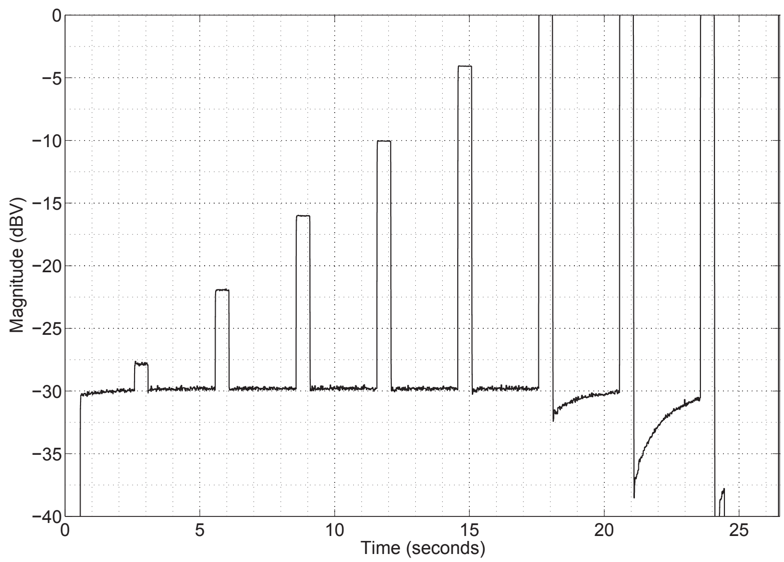

Figure 3.26: AGC response of iPad 2 to increasing amplitude pulses at 500Hz

This is a clear indication of some form of AGC, attempting to reduce amplitude in response to a high level signal. However this differs from the response of the MWG Atom in Figure 3.18 as a high level pulse is held at the maximum level before clipping and not attenuated by such a large degree. After the pulse has ended the AGC then gradually increases the gain until the gain returns to its original level. As previously mentioned, this is in place to handle the varying distances a mobile phone user may hold their phone microphone away from their mouth to maximise signal to noise ratio and reduce distortion. Again, this mechanism appears to be hard coded into the devices functionality.

Whilst the major drawback of this AGC may be a reduced dynamic range of soundscape recordings, the inevitable distortion due to high levels signals that would occur without it would be more detrimental to the soundscape recordings. An issue that may arise from this is that a loud impulsive sound experienced by these devices will cause the AGC to reduce the input gain, which will mean that any quieter sounds that are captured immediately after this louder source will be attenuated.

In order to gather a consistent and comparable set of soundscape recordings, all submissions made via the web method were high pass filtered using the same roll-off and attenuation characteristics as observed on the iOS 5 type devices. This means that each soundscape's low frequency content below around 200Hz is equally represented prior to feature extraction.

3.6 Pilot studies

3.6.1 Schools pilot

A key issue was ensuring that the mediating technologies and survey semantics were fully stable and comprehensible for members of the public who naturally have a wide range of technical abilities. The pilot studies with 14-16 year old secondary school students were invaluable in the identification of any initial problems as the students were perfect beta testers of the mobile apps, freely vocalising any issues they had. This active feedback allowed a dialogue to be set up which resulted in an efficient method of development iteration for the project's mobile software and question set. The feedback was acquired using a set of activities detailed in Appendix B and a follow-up questionnaire which can be found in Appendix C. In total 112 students took part in the pilot study, using the project's J2ME application on provided devices and on their own, if supported.

Technology

The variation in participant technical ability and the fluidity of the technology was a key factor in determining participant uptake of the project. A number of mobile devices were brought into schools with the project's software preloaded, as well as a laptop with bluetooth, to transfer the application out to the students' own phones. Students shared a device if there were shortages of supported handsets.

The initial issue found was with the transferral of the software onto devices. J2ME applications require that a certain security certificate be installed on the device to provide the application access to the phone's functions without constant security confirmations. Some devices missing this certificate were effectively rendered useless when the application was run due to the constant dialog boxes asking the user to confirm an action. After learning this a Thawte security certificate [102] was purchased to place the application in the "Trusted" domain, which allowed for device function use without incessant pop-ups.

The students were also confused by the graphical displays of signal frequency and time domain amplitude, which was dropped in future versions of the application. The students found the process of capturing and commenting on soundscapes rather simple, with 88% finding soundscape capture easy and 82% understanding the process of commenting on the soundscapes they recorded. The only major issue was with the students who struggled using the software, with 46% commenting that there was not enough feedback and information provided by the application. The addition of instruction pages and prompts helped with this situation in later versions of the application.

Survey question set

In a series of pilot studies with the software in local Manchester schools, pilot participants ran through the project's methodology and were then asked questions on the semantics used in the project's survey question set. This process has been based on previous cognitive pre-testing techniques used for questionnaires as shown by Collins [103]. The semantic differential scales used had been utilised and pretested in previous soundscape research, as discussed in Section 3.3.2.

Data collected from the student feedback questionnaire in Appendix C revealed that the students responded positively to the survey question set. The students required little to no help when understanding the terminology used for the scales, with 94% understanding what the questions were asking without requiring any help or an explanation. Of the students questioned, 86% found that the language used was understandable before an introduction to the topic of soundscapes, which increased to 96% after these introductions had been given.

Comments were mainly around the term "soundscape", which confused some students, but after a short explanation, they found the concept easy to grasp. This resulted in a number of short soundscape definitions on all device applications and the project website.

3.6.2 Live pilot

The project's live pilot was run with a small group of beta testers ($\approx$ 20) before the main national launch. This was the stage to iron out any issues and ensure that the processes involved in taking part were not too complicated. It became immediately apparent that simplicity and ease of use were far more important than functionality and technical impression. It seemed that potential members were not won over by technical wizardry, but rather the notion of taking part in something important and meaningful to them, which would also take a minimum amount of effort to achieve.

There were a number of issues that arose while live piloting the technologies. This led to methodological changes to ensure that the most reliable data was being gathered. The unsupervised recordings of the project were prone to user error so it was important to provide concise instructions to participants or preferably, implement ways to avoid these issues without using written instructions, which are generally ignored. An example of this was when participants were observed holding mobile phones whilst making soundscape recordings but blocking the microphone port with their hands. This was solved by rotating the display of the mobile app by 180° which meant the user had to rotate the mobile itself whilst they were taking the soundscape recording, leaving the microphone free from any obstructions. However, this solution brought up the issue of wind noise on the microphone saturating the recording. After much deliberation, the solution was simply to provide instructions advising the participant to stand out of direct wind or position themselves so that their body blocked the wind from the phone's microphone.

The content and functionality of the technologies involved had seen a number of manifestations since the project's beginnings with the most notable being the simplification of the mobile interface. Original plans to provide real-time visualisations of the soundscape on the participant's device were removed after the school pilots. It was found that maximum engagement and willingness to participate was achieved through a majorly scaled down interface whose only functionality was to record and comment on the sound environment. It was important to consider the participant's experience of the mobile applications as the success of the project relied heavily on its public uptake and essentially their collection of the research data.

Initial plans were to specifically target mobile phone users. After beta testing and initial piloting, the methodology was opened up to allow audio submissions from any digital recording device. This made the project accessible to more participants but meant the logging of subjective responses was omitted at the recording location. To provide for this, a printable response log was provided for members to document their perceptions in situ. With only around 50 downloads of this log, it was apparent that this was not regarded as a worthwhile effort by the project's participants.

3.7 Participant engagement and attrition

Subjective experiences such as soundscapes are difficult to quantify and explain using a traditional positivist approach involving purely quantitative data collection. This warrants a more flexible approach that places aspects of the experimental method at the discretion of the participant. Crowdsourcing [104] or Citizen Science [105] style soundscape experiments provide this flexibility, allowing participants to choose what type of soundscape is investigated, as opposed to presenting them with an environment to which they must respond. Placing this element of choice on the participant may provide a more nuanced perspective of soundscapes in general, but may introduce a bias that stems from the types of soundscapes people will be inclined to choose. The problem of experimenter bias [106] should be raised at this point, as the participant is making the research location choices and the times when they make them. This contributes to a research yield with a greater person–focussed approach. An interview led experiment that is testing a hypothesis can sometimes inadvertently lead to participants behaving in a way that confirms the hypothesis [107]. The use of the project's methodology described in Section 3 is a step closer to the ideal of an absolutely objective, dispassionate interview technique. The mobile technology being used will not treat the participant differently or introduce any underlying emotion in terms of the research hypothesis. The participant does however know what is being asked of them, which will introduce a heightened awareness of their surrounding soundscape. This participant bias is known as the Hawthorne Effect and describes how a participant's behaviour is affected by the knowledge that one is part of an experiment and is therefore important to the studies success [106]. This relates to the discussion of listening states in Section 2.4, where a participant might take on a more active listening state when making a soundscape submission for the project.

3.7.1 Communication

Engagement with potential and active participants was carried out using a number of channels. A visual identity was created that was intended to be visually appealing and also consistent across all our communication outlets. A distinctive logo was created with a colour palette and font used across the project's applications and content. This is shown in Figure 3.27.

Through the collaboration with an internationally recognised PR company, the media coverage generated for the project launch was substantial, with worldwide coverage. The PR company has a successful track record with science related public engagement PR campaigns with past clients including The London Science Museum and the Institute of Physics. The main benefit of this collaboration was

Figure 3.27: The Sound Around You project logo

the ability to create direct links with journalists and editors working for the relevant sections of the different media outlets, which could not be easily achieved without the contacts that these PR companies posses. The company also handled the scheduling and setup of meetings and interviews for print media and radio.

Website

A project website was created early on into the development process at: www.soundaroundyou.com, with screenshots shown in Appendix A. This provided: information on the project, the world soundscape map, a path to the social media pages, and access to the web soundscape contribution features. The website carries the logo as well as the University of Salford logo to emphasise that the project is a non-commercial, academic research study. An information button opens up a page describing the project's aims and potential outcomes as well as: providing an example soundscape recording, a YouTube video of a soundscape being captured and a recent academic conference paper for anyone with a greater interest in the project. A shuffle button also allows people to randomly fly around the soundscape map listening to soundscapes. All of these features were implemented to engage visitors with the concepts and ideas behind the project and hopefully entice them to take part. The website will also detect if the user is accessing the site from an iOS device and display a pop-up box prompting them to click a link which opens up the App Store on their device on the i-SAY, iOS app download page.

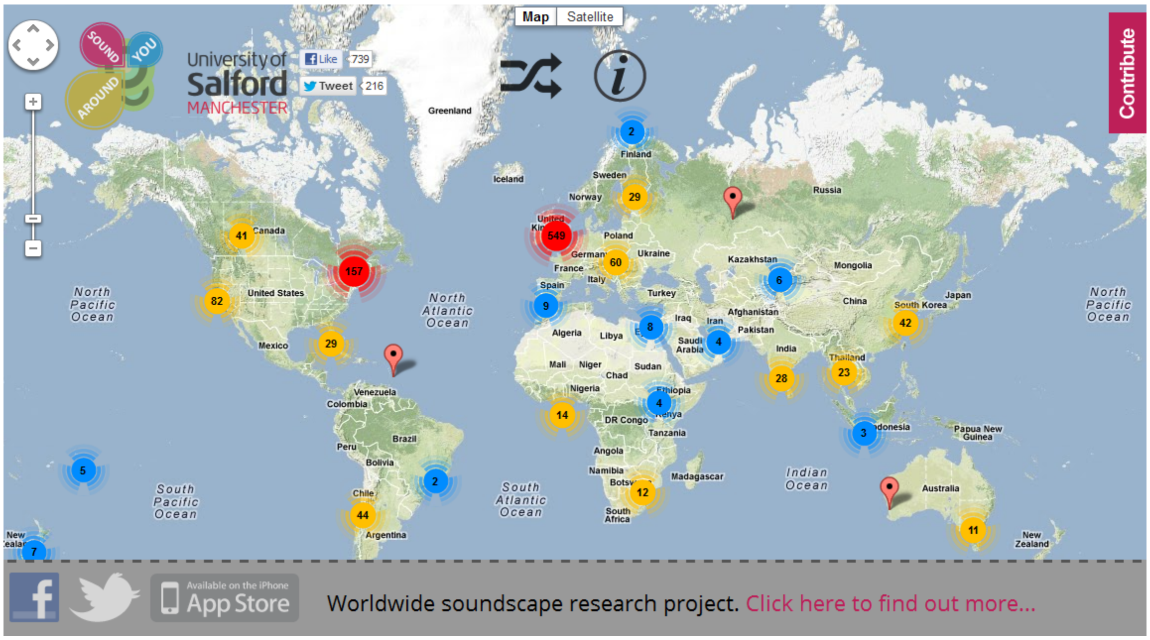

Figure 3.28: The Sound Around You project soundscape map

The websites most striking feature is the world soundscape map, seen in Figure 3.28, allowing interested parties or existing participants to browse all of the soundscape submissions to date. A Google Maps interface provides a familiar navigation process where visitors can playback, download and share soundscapes of interest from around the world. Participant responses are shown alongside the soundscape information, providing an insight into their opinions and reactions to the place and its sound environment. If a Google Streetview scene is within 10 meters of the soundscape location a rotating 3D image of the place is also shown, making for a more visually engaging experience.

Traditional media

At the launch of the project's first feature phone stage a press release was produced and sent out to all the large media organisations and relevant editors and journalists. The full text of the release is included in Appendix D. Press resources were also made available via links on the project website, including a set of images depicting project participation and a set of some of the more interesting soundscape recordings submitted for use on TV & radio shows or web news sites. The major media coverage included:

- Feature story on the BBC's flagship technology programme, BBC Click - broadcast on BBC 2 & worldwide on BBC News World Service (web screenshot shown in Figure 3.29)
- TV interview on Canada's major news channel CBC
- Feature story across two pages in the Times 2 supplement, Telegraph, Independent, Guardian, Metro, Observer and Manchester Evening News (scan shown in Figure 3.30)
- Radio interviews on: BBC World Service, Good Morning Scotland, BBC Radio Scotland, Good Morning Wales, Radio 5 Live Breakfast, BBC Manchester Breakfast Show, Smooth Radio, Rock Radio, LBC Radio
- Online stories on: BBC Online Technology, TimesOnline, Telegraph.co.uk

Social media, video sharing and blogs

The project included pages on: Facebook (soundaroundyou), Twitter (@say_project) and YouTube (cmydlarzSAY). The social media elements of the project were indis-

Figure 3.29: Feature story on BBC Click website

Looking for sound of the overground

PHD Student Charlie Mydlaz from Salford University in an accoustic sound deadening chamber

STREET sounds across Britain are being collected by researchers at The University of Salford to create a 'sound map of Britain'.

Members of the public are being urged to record noises from everyday locations with their mobiles and upload the clips onto a virtual map online.

Researchers, who launched the study as part of the Manchester Science Festival, hope to create an audio equivalent of Google street view, where web users can experience the sounds of different locations.

PhD student Charlie Mydlarz, who is leading the study, said: "We're asking people to capture any environment they choose, and that includes both public and private spaces, so recordings could capture anything from a family car journey to a busy shopping centre."

An existing noise map of Britain - provided by the Department for Environment, Food and Rural Affairs - displays noise levels using colour coding.

Areas on the map shaded purple are loud and those shaded green are quiet.

The results are based on traffic flow, ground shape and the extent to which buildings absorb acoustics, but are not created using recordings.

The map enables users to assess areas for volume but does not allow people to hear the sounds.

The Salford university map will be a first and does not focus on loud noise as a negative but portrays all sounds as part of a place's atmosphere.

Mr Mydlarz said: "By using everyday technology to get people involved, this has the potential to be the largest study of its kind.

"The findings could have many uses from psychological research to town planning."

To upload your audio clips and listen to the map of Britain go to: www.soundaroundyou.com

Figure 3.30: Manchester Evening News print story

pensable in driving people to the project site and also for maintaining continued participation from existing members. An example of this is discussed in Section 3.7.3. The Facebook pages main purpose was a discussion platform for interested parties to discuss soundscape issues. Twitter was mainly used as a means to post new developments from the project, such as new uploads, features, issues and questions. The YouTube page hosted videos showing the project's tools being used in interesting places around the world, informing potential participants how the project works and raising interest in soundscape issues. All of these social media outlets provided a showcase for the project and helped to increase interest and ultimately promote people to take part. It is difficult to quantify how many people were driven to take part by the project's social media presence, however, statistics provided by the project's web host, show that on average around 50% of new users to the site over a month are referred from these social media sites.

3.7.2 Participant motivation

Participant recruitment was purely opportunistic. A study into response rates and participant motivations in mail surveys, revealed three main factors which effect participation rates [108]. The present studies opportunistic approach to participant recruitment shares many similarities with mail based surveys, therefore these identified motivations will still apply in this context.

- **intrinsic motivation**, based on a pre-existing interest in or enjoyment of the tasks associated with participation in the project
- **altruistic motivation**, through assisting research aimed at producing knowledge that may be of social benefit

- **novelty motivation**, created by the relatively unusual survey platform and the use of new technologies

During the live pilot of the project a number of soundscape submissions were received in a very high quality audio format. These had come in from professional sound recordists, who would presumably have been motivated to take part through a pre-existing interest in the field. The issue of participant bias should be mentioned at this point as a participant with prior experience in the field of environmental sound recording may provide skewed subjective responses, but the quality of his objective submissions will be very high. Altruistic motivations may also promote the submission of soundscapes which people either want to preserve or on the contrary want to highlight as being unsuitable and in need of a change. Finally, the novelty motivator could lead to non-serious submissions as a participant experiments with the project applications. The effects that these bring are almost impossible to identify, quantify or compensate for, however, the potential for the acquisition of a large dataset should serve to reduce the effects that these factors may introduce.

The voluntary nature of participation, does however lead to a potentially very large and diverse set of submissions. With no stipulation in any of the marketing material about exactly what to capture and submit, the range of locations that these devices move through is extremely diverse. To date no other soundscape study has had the potential to reach this range of locations and gather subjective and objective soundscape data.

In the early stages of the project the idea of providing incentives to participants to increase submission numbers was discussed; such as prizes for especially interesting submissions or awards for every 1000th submission. These were ruled out due to financial restrictions and more importantly the fact that they may introduce bias to the project's submissions. Placing any kind of leading request to participants might produce a skew in the types of responses received, where people interpret an "interesting" soundscape as one that is inherently positive. An award for every 1000th submission could also result in participants uploading a large number of meaningless or blank submissions in a bid to win the prize.

3.7.3 Retention and attrition rates

Attrition is a concern in projects of this kind as attrition rates are generally high when project researchers are out of sight and possibly out of mind for the duration of the survey period [109]. Three techniques were used to aid participant retention:

- Facebook and Twitter provide an API that allow anyone to place a "share" button on their site and provide a custom URL linking people back to whatever it is they would like to promote. This was used on the project website to allow visitors to share soundscapes of interest on their own social media time-line or feed, illustrated in Appendix E, Figure E.1.
- The iOS app provides Facebook and Twitter Share switches shown in Appendix E, Figure E.2, which, when turned on, automatically post to a participant's time-line or feed that they have uploaded a soundscape using the project software, including a link to the soundscape itself. This was switched off by default so as not to annoy participants by posting items that they did not want posted. The Explore tab of the iOS app also allowed people to post any interesting soundscapes to Twitter.

- The soundscape of the week notification produces a pop-up alert on a participant's iOS device (shown in Figure 3.31 every Saturday at 2pm (local time), which when tapped, opens up the project app with the weeks most interesting soundscape open. The benefit of this is that a participant can be alerted even when the app is not running. The main purpose of this is to remind the participant about the project and to hopefully promote them to make another submission. The day and time were chosen to catch people in a variety of situations so if any submissions were made after seeing the notification they would be potentially from a wide variety of soundscape locations. It was assumed that the majority of participants would be at work during the week, which may lead to an influx of workplace soundscapes if a weekday was chosen, potentially skewing the dataset. A weekly, rather than daily alert was chosen so as not to annoy participants, leading them to delete the app.

All of these techniques were predominately employed to drive traffic to the project website and also remind existing participants to continue contributing. They are all voluntary and were designed to be as non-intrusive as possible. The real danger is in annoying people and either pushing them to leave the project or putting them off taking part in the first place.

As the current study is investigating the samples response as a whole to their sound environments, the minimum input required from an individual participant is a single soundscape submission. Whilst repeat submissions will benefit the study in terms of numbers of responses, it is not crucial that members submit more than once. However, participants sharing their submissions on social media outlets raises the profile of the study and should serve to drive more new participants

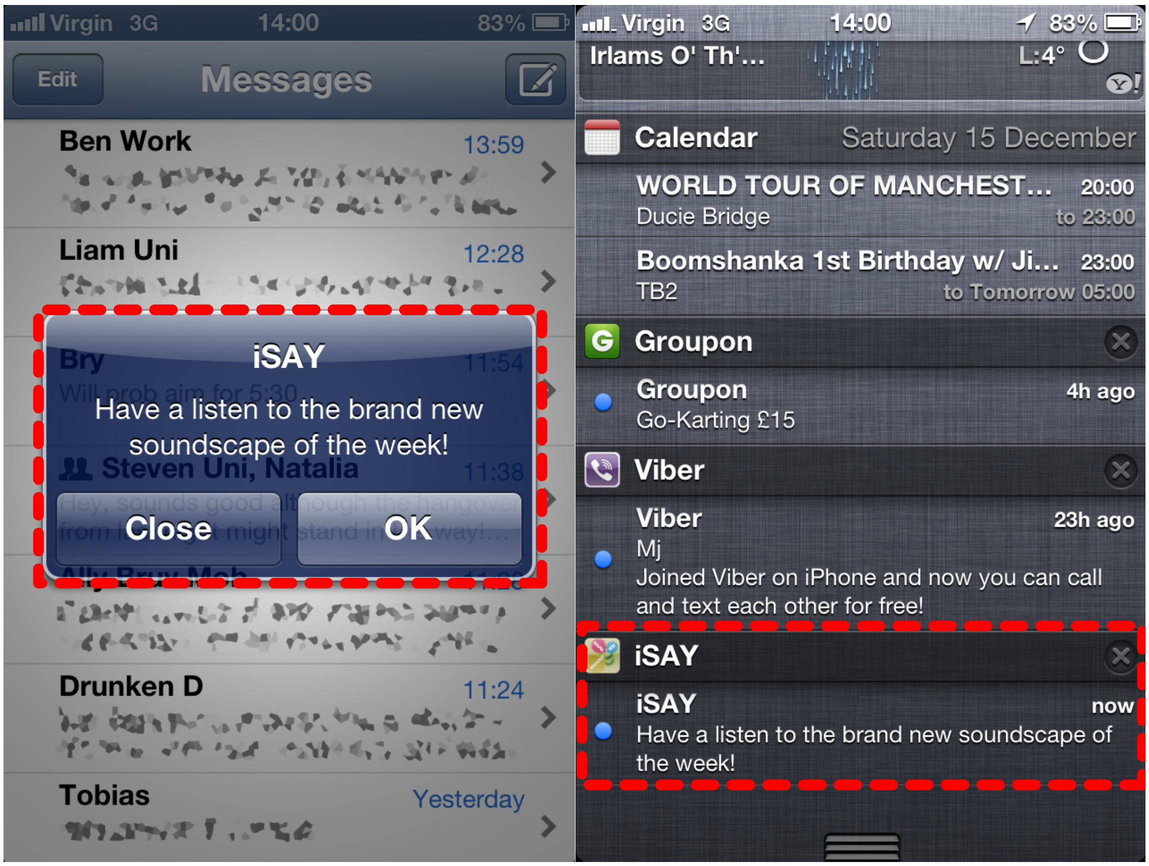

Figure 3.31: Soundscape of the week notifications

to take part. An issue to consider is the participant who makes the soundscape recording, but then skips past the opinions stage. Although a participant can only proceed past the opinion stage when any one of the sliders has been interacted with, there is the potential to proceed by randomly moving a slider to activate the Done button to proceed to the upload stage. These submissions are hard to identify, but one solution could be to check for subjective data which contains no sound identification entries and the majority of slider values are still set to the default value. This issue is discussed in Section 3.9.1.

3.7.4 Participant privacy and submission legality

The demographics collected directly from participants was intentionally kept to the bare minimum. Only demographic data considered important for further analysis was asked of people. The use of only a first name (for user soundscape identification), age range and gender request was intended to put participants at ease at the surveys start, as none of this information could be used to identify them. An assumption may be that if a person believes they can be identified by their demographics, they will change their answering behaviour to fit with social norms [110]. This however has not been shown to influence participant responses to an online study testing for changes in answering behaviour and drop out rates [111].

Although the demographic data gathered from participants cannot be used to identify them, it is transmitted and stored securely on the project's server behind a firewall. Submissions are posted on the project's website including: the soundscape recording, its location and the subjective responses to it. Participants taking part in the study are required to agree to the project disclaimer shown in Appendix F. This disclaims liability for the user generated content, including the submission of copyrighted material inadvertently recorded. Each soundscape point on the websites map also has a “Report an issue” link, which presents a pop-up box with “OK” and “Cancel” buttons, stating: “If you think this soundscape entry is inappropriate in any way, click OK and it’ll be removed for review”. When an entry is flagged as inappropriate it is immediately removed for manual review.

3.8 Objective acoustic analysis

Each soundscape recording submitted is stored on the project’s server, ready for conversion and analysis. Each filename is a unique 10 digit code that links the soundscape to its subjective data, stored on the project’s MySQL database. As the project began accepting files from any kind of audio recorder or mobile device, there are a number of different file formats stored online. These range from the low quality Adaptive Multi-Rate (AMR) recordings of low end J2ME based mobile devices, storing compressed speech band optimised files sampled at 8kHz up to professionally recorded uncompressed wav files sampled at 48kHz. To maintain the quality and consistency of the project’s objective dataset, it was decided to exclude recordings with a sampling frequency of less than 22kHz and that were compressed (lossy compression). Recordings of the AMR type were of too low a quality and the compression artefacts that the AMR format introduced, could potentially affect the outcomes of any future objective acoustic analysis on the dataset. Out of the total number of submissions stored, 11% were excluded from the objective analysis.

3.8.1 Pre-processing

Each raw soundscape file underwent a process of pre-processing before its analysis, using the Matlab technical computing language [112]. The first stage was the download and conversion of the varied soundscape formats to the uncompressed wav format, extracting only the left channel if the original file was recorded in stereo. These files were then resampled to 44kHz, 32 bit. Any DC offset present was removed using the signals arithmetic mean. The soundscapes were then filtered using a FIR Butterworth low pass filter with a cut-off frequency of 15kHz. This was done to remove any high frequency aliasing, seen on some iPhone recordings. Finally the soundscapes were trimmed of zeroes and exponentially tapered at each end to slowly fade them in and out over 0.2 seconds. An example of this is shown in Figure 3.32. This was done to ensure that each end of the recording was zero crossed to remove any impulsive artefacts that may affect the objective analysis.

3.8.2 Feature extraction

While human listeners intuitively judge the similarity or dissimilarity between audio signals, the task of comparing signals with a computed algorithm is far more complex. As the temporal and spectral evolutions of different signals cannot be compared directly, any implementation of similarity rating has to rely on a number of quantised signal characteristics called “audio features”. In order to take into account the dynamic evolution of these signals, the analysis is carried out on short term segments in which the signal is assumed to be stationary. For these reduced segments, analysis is performed that extracts a number of audio descriptors or low-level features. These can be combined to produce a high-level feature

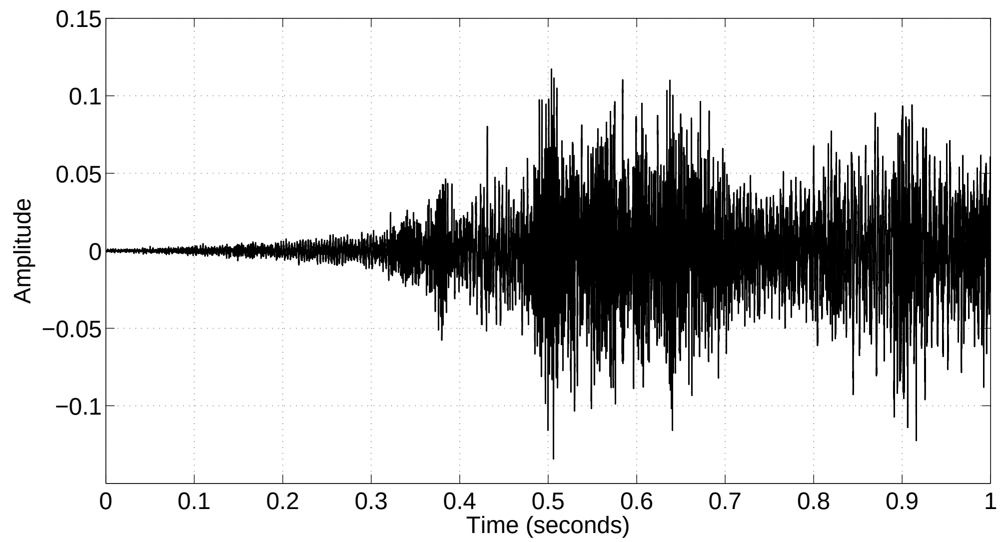

Figure 3.32: Exponential soundscape tapering over the first 200ms

representation of the signal. A selection of these features describe time varying characteristics, therefore averages and values of distribution will be used in further statistical analysis in these cases. Since a number of the features are also highly correlated, their use in further soundscape analysis must be carefully considered. A selection of these features are explained, beginning with level based features, then features that reflect the dynamics of the soundscape and features associated with the timbre of a signal are summarised.

After the pre-processing stage, a number of spectral and temporal features were extracted from each soundscape for analysis and comparison. The MIRtoolbox musical feature extraction package for Matlab [113] was used to extract most features with the remaining calculated using custom Matlab scripts with sources referenced. The objective analysis of soundscape recordings using a musical information retrieval (MIR) approach has been previously attempted and proved successful in soundscape classification [41, 114]. The combination of features extracted using the MIR techniques seem to provide a good set of metrics to describe soundscapes objectively and for later use in objective/subjective comparisons.

The following objective metrics were run through a principal component analysis in a bid to reduce the high dimensionality of the data (see Section 4). The aim of this process was to find a structure in the acoustic setup of these varying soundscapes and identify and exclude any redundancy in the extracted features. The components extracted were then used in a stage of correlational analysis investigating any relationships between the objective and subjective features of a soundscape. Table 3.3 briefly describes the extracted objective features, followed by a more detailed description for each.

Feature	Description	Soundscape use
RMS	Arithmetic mean of amplitude squared	✔
L_{AeqT}	Continuous A weighted sound level	✔
L_n	Statistical sound pressure level	✔
N_n	Zwicker loudness including percentiles	✔
L_{CeqT} - L_{AeqT}	Difference between C and A weighted sound pressure level	✔
CoG	Spectral centre of gravity	
Spectral spread	Spectral spread relative to CoG	✔
Spectral flatness	Ratio between geometric & arithmetic power spectrum mean	✔
Spectral roll-off	Descibes of signal spectrum	
Zero cross rate	Signal sign changes per second	
Brightness	Measure of high frequency energy	
MFCC	Mel-frequency cepstral coefficients	✔
Roughness	Measure of dissonance	
Irregularity	The amount of variation between partials	
Entropy	Describes a signals frequency distribution	✔
L_{A10} - L_{A90}	Difference between sound pressure level exceeded 10% and 90% of the time	✔
Low energy rate	Temporal distribution of signal energy	
Crest factor	Ratio between peak and RMS of signal	
Spectral flux	Amount of spectral variation across time	
Event density	Estimation of average frequency of events per second	

Table 3.3: Objective feature descriptions and use in soundscape research

Relative levels

As all submissions are uncalibrated in level, these features are all considered relative to each other. There may be some skew associated with the varied dynamic range from submissions from different devices, however, the size of the dataset should reduce this influence. The issue of participant handling differences will also produce variations in relative level as discussed in Section 3.4.2. The majority of these values are expressed in full scale decibels (dBFS), meaning they have a ceiling of 0dB, representing the maximum signal level handled by the system.

- **RMS Root Mean Squared**
  A measure of the average amplitude of a signal over time. It represents the arithmetic mean of the amplitude squared of each discrete point of a waveform, given in Equation 3.1. Used in this case as an overall value representing soundscape level.

$$x_{rms} = \sqrt{\frac{1}{n}(x_1^2 + x_2^2 + \cdots + x_n^2)} \tag{3.1}$$

- **L_{AeqT} Equivalent continuous sound pressure level**
  The measure L_{Aeq} is the equivalent continuous A-weighted sound level, which characterises fluctuating instantaneous sound levels during a time interval T as an equivalent steady state level. This has traditionally been used as a primary objective indicator in soundscape research because of its ease of measurement and calculation and it correlates reasonably well with perceived loudness and specific annoyance [115]. The time interval T will be 10 seconds for all soundscape recordings, calculated from the first 10 seconds of a soundscape signal for consistency across submissions of differing

length.

- **L_n Statistical sound pressure level**
  This measure quantifies the sound pressure level exceeded for n percent of the duration of the measurement period. If N measured instantaneous sound pressure levels are obtained with a fixed interval over a given time period T and listed in ascending order, then L_n is the $(\frac{100n}{N})^{th}$ SPL in that list [116]. They make up approximate measures of peak (L_1), intrusive (L_{10}), median (L_{50}) and background (L_{90}) sound levels.

- **N Zwicker loudness (sone)**
  This "classic" measure of loudness for steady state sounds was proposed by Zwicker in 1960 [117]. The original method of calculation involved: (a) A fixed filter representing the effect of the outer and middle ear; (b) Calculation of excitation patterns of output from auditory filter banks; (c) Transformation of the excitation pattern to a specific loudness using a power law relationship; (d) The summation of the specific loudness values then provides the overall loudness N [118, 119]. Previous studies on environmental noise have shown that annoyance can be highly correlated with loudness [120]. For steady state sounds, perceived loudness is well predicted by the Zwicker model, but the problem lies in the temporal state of a soundscape, which is anything but steady. To overcome this and attempt to judge the overall loudness of time-varying sounds, Fastl and Zwicker adapted their original technique, incorporating the methods of the statistical sound pressure level mentioned above. The statistical or percentile loudness level N_5 (the loudness level reached or exceeded for 5% of the measurement duration) was used successfully to show correlation with the judgement of the loudness of

transport noise [121].

Spectral

- **L_{CeqT}-L_{AeqT} Difference between C and A-weighted sound pressure level**
  The difference between C and A-weighted equivalent continuous sound pressure level, has been used as a measure of the relative proportion of low-frequency sound within a soundscape [5].

- **CoG Spectrum centre of gravity (1/3-octave spectrum)**
  The acoustic centre of gravity of a spectrum can be seen as the “mean” frequency of a signal, or the distribution of a signals power spectrum. It has strong perceptual links to the “brightness” and “sharpness” of a sound [122].

$$f_{centroid} = \frac{\sum_{n=0}^{N-1} f(n)x(n)}{\sum_{n=0}^{N-1} x(n)} \tag{3.2}$$

  The “centroid” frequency can be found using unweighted third octave band measurements using Equation 3.2 (where x(n) = band n magnitude and f(n) = band centre frequency) or derived from the wide band output of a signals Fourier Transform.

- **Spectral Spread**
  The spectral spread describes the spread of the signal spectrum in relation to the spectral centroid. It can be interpreted as a measure of tonality, where noisy signals that display a broad-band spectrum will have a higher spread

than tonal sounds that are confined to narrow-band peaks. This measure has been used in studies involved in the categorisation of environmental sound sources [123].

- **Spectral Flatness**

  Spectral flatness is defined as the ratio between the geometric and the arithmetic mean of the signal's power spectrum, with its calculation defined in Equation 3.3:

$$flatness = \frac{\sqrt[N]{\prod_{N-1}^{n=0} x(n)}}{\frac{\sum_{N-1}^{n=0} x(n)}{N}} \tag{3.3}$$

  where $x(n)$= the magnitude of the STFT bin n. This feature is also known as the tonality coefficient and corresponds to the "noisiness" of a soundscape. This metric has been previously used to improve soundscape classification algorithms in [114].

- **Spectral Roll-Off**

  The spectral roll-off loosely describes the shape of the signal spectrum and has been used in musical feature extraction. It is a measure of the frequency range where the main part of the signal energy is found. The roll-off frequency value is found by summing the signal energy across the frequency spectrum and finding the frequency where a certain percentage of the signal energy lies (generally defined as 85%) [124].

- **Zero cross rate**

  The zero-crossing rate describes the number of times the signal amplitude crosses zero per time unit, i.e. the number of sign changes in the time domain. This can be interpreted as a measure of noisiness or tonality and is also correlated with the signal pitch, since a noisy signal will tend to change signs more often. This measure has been utilised in music information retrieval [125].

- **Brightness**

  The spectral brightness is a measure of high-frequency energy content similar to the spectral roll-off measure. The approach differs from the roll-off calculation in that a threshold frequency is fixed at 1.5kHz and the percentage of energy above that cut-off frequency is computed [126].

- **MFCC Mel-Frequency Cepstral Coefficients**

  Mel-Frequency Cepstral Coefficients (MFCCs) represent the spectral characteristics of a signal in a condensed way. They are used in speech processing and coding algorithms because the cepstral domain, defined as the inverse transform of the logarithmic signal spectrum, can be useful in extracting and manipulating the spectral envelope of a signal [127]. Studies have shown that MFCCs are seen to work well as audio features for musical and soundscape analysis [128, 41], with the ability to classify soundscapes as discussed in Section 2.5. The MFCCs are calculated as follows: firstly, the signals magnitude spectrum is determined, by computing the FFT. The magnitude spectrum is then filtered by a Mel filter bank, a filter bank of triangular filters that group together frequency components according to the Mel scale, a frequency scale based on the human perception of pitch. The resulting groups are summed and the logarithm is calculated, in an attempt to roughly mimic the behaviour of the human cochlea, where neuronal impulses are evaluated in groups, representing frequency bands. In the final stage of the calculation, the values obtained from the filter bank are transformed into the cepstral domain using the Discrete Cosine Transform (DCT) [129].

- **Roughness**

  Roughness is a measure of the dissonance that is produced by two sinusoidal signals with regard to the frequency ratio between them and is derived from Plomp and Levelt's concept of tonal consonance [130]. Roughness is estimated by computing the dissonance between all maxima in the signals spectrum. The final roughness value is gained by calculating the average over all roughness values.

- **Irregularity**

  The irregularity measure describes the amount of variation of the distance between successive partial tones in a harmonic sound [131]. This can be interpreted as an indicator for the tonality of a signal, where the irregularity will be lower when the partials are harmonically related.

- **Entropy**

  The spectral entropy of a signal describes the curve of the frequency distribution of a signal. A single value of relative Shannon entropy [132] is calculated for each soundscape, where high values of entropy equate to spectral distributions that tend towards random white noise and low values tending towards zero refer to signals made up of a single pure tone. This feature has been used in the past for: soundscape identification [133], sound source identification [134] and biodiversity appraisal [135].

Dynamics

- **L_{A10}-L_{A90} Difference between A-weighted sound pressure level exceeded 10% and 90% of the time**

  The difference between A-weighted SPL exceeded 10% and 90%, has been used as an indicator of soundscape variability [5].

- **Low energy rate**

  The low energy distribution provides an assessment of the temporal distribution of energy within each soundscape segment. The rate is defined as the percentage of segments showing less than average energy compared to the remaining signal [124].

- **Crest factor**

  The crest factor is the ratio between the peak and RMS of a signal, shown in Equation 3.4. A related measure is the peak-to-average power ratio (PAPR), which is the peak amplitude squared divided by the RMS squared. This can be more usefully expressed in dB, in Equation 3.5.

$$C = \frac{|x|_{peak}}{x_{rms}} \tag{3.4}$$

$$PAPR_{dB} = 10 \times \log \frac{|x|_{peak}^2}{x_{rms}^2} \tag{3.5}$$

  Crest factor is used to determine sound level impulsiveness and is also reported to bear a significant impact on perceived annoyance [136].

- **Spectral flux**

  Spectral flux defines the amount of frame to frame frequency variation in time. It is computed by calculating the energy difference between consecutive short term fourier transformed (STFT) frames, shown in Equation 3.6.

  $$SF = | X^d_f - X^d_{f-1} | \quad (3.6)$$

  X denotes the magnitude components and superscript f and $f-1$, the current and previous frame. It has been used to identify speech in music signals [137].

- **Event density**

  Makes an estimation of the average frequency of events per second, using a peak identification process defined in [138].

3.9 Subjective response analysis

3.9.1 Data cleaning

To ensure that the subjective data is of the best quality possible, a process of data cleaning was carried out before the analysis stage. The default rating for each semantic differential scale is 5, in other words, if the scale is not changed it will remain on 5. A non-response to the subjective scales is characterised by a rating of 5 for every scale. This means that a participant skipped through the opinion stage without responding to the soundscape. Entries of this kind will be excluded from any subjective analysis. Out of all responses, 15% were identified as a non-response and excluded.

The textual source identification questions were trimmed of any leading and trailing white-space and converted to lower case for use in further analysis. A number of submissions used words such as “no”, “nothing”, “nowt”, or entered nothing to signify that no positive or negative sound was identified. All occurrences of these words were replaced with the word “none” to keep consistency across all submissions. This situation could arise if a person has not entered because of a lack of interest or because they could not identify a positive or negative sound. It is difficult to identify these different cases, except where the participant has performed a non-response, mentioned previously, in which case the submission is excluded from analysis.

3.9.2 Semantic differential analysis

The purpose of this stage of analysis is to identify factors that characterise the soundscapes submitted to the study, based on their associated subjective responses from the semantic differential ordinal scales, participant activity, positive & negative source identifications, participant age and participant gender.

Variations in soundscape response will be investigated between the different participant activities of: Passing through, Working, Relaxing, Recreation and Waiting. This will serve to identify soundscape appreciation with respect to the type of activity a person is involved in when the responses are made. The assumption that a person involved in relaxation would have chosen to be in a place whose soundscape has high ratings of soundscape quality, pleasantness and tranquillity. Whereas a person passing through a location might not have factored in the soundscape in their choice of route through the place, resulting in lower ratings of quality, tranquillity and pleasantness.

Correlational analysis will be performed between each subjective differential rating with particular focus on the soundscape quality scale, to uncover the subjective variables which contribute or otherwise to a persons appreciation of a soundscape. Alongside this, a principal component analysis will be performed to identify any overlap between these subjective ratings. Each component can then be checked to see if any correlations exist with soundscape quality. Identified components can also be compared with those of previous research as a validation of the present study's methodology.

Initially a full correlational analysis will be made across all combinations of subjective and objective metrics to uncover any linear relationships. The principal components extracted from the subjective metrics will then be used to further investigate any potential correlations with the objective metrics. This process will then be repeated, using the principal components extracted from the objective variables to determine if there are any association between the objective and subjective parameters of a soundscape.

Individual soundscapes will be plotted within a two dimensional space defined by the extracted subjective principal components in a bid to discover any trends or associations with the objective components extracted.

3.9.3 Small group study

To provide further validation of the projects methodology, a small participation study was carried out using a small group of subjects. A total of 14 subjects were taken to 2 locations, representing:

- An **urban soundscape (Peel Park, University of Salford)** located close to a large road with a green area sparsely covered with trees and surrounded on three sides by various styles of large buildings

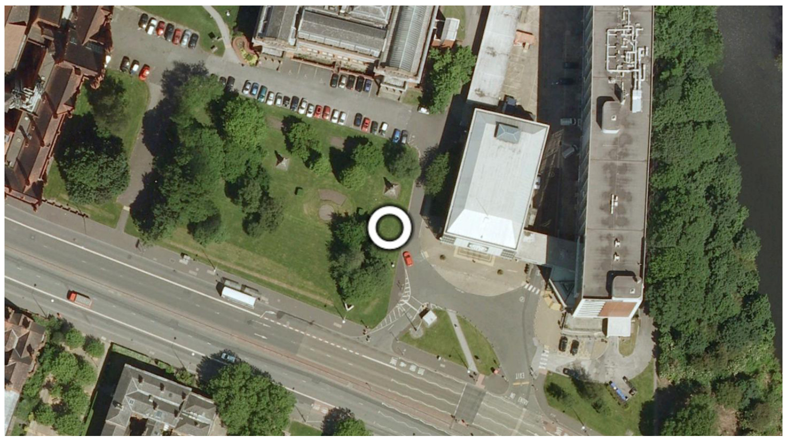

Figure 3.33: Satellite view of urban soundscape used in small group study with subject location marked

- An **urban park soundscape (next to main University of Salford building and main road)** surrounded by trees and a number of large buildings

Whilst at the 2 locations, subjects were asked to listen to the soundscape in silence for around 30 seconds, then respond to it using the project question set on paper. They were also asked not to confer between themselves while responding. Subjects were Engineering Mathematics undergraduate students, made up of 10 males and 4 females, with 12 aged 22-27 and 2 aged 34-39. Subjects were given the same information provided on the project website and app before the sites were visited. The first site visited was the roadside urban location followed by the urban park. For comparison, a sample of 10 UK soundscapes, similar in nature

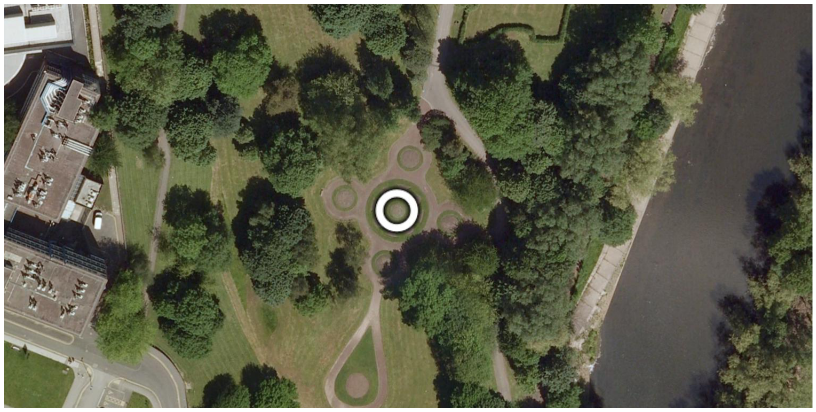

Figure 3.34: Satellite view of urban park soundscape used in small group study with subject location marked

to the small participation locations were chosen from the project's main dataset. Only 10 were chosen because of the lack of suitable submissions matching the required criteria for a fair comparison. The similarity judgement was made based on satellite imagery of the location and the actual audio recording of the site itself. The age and gender distribution of these submissions also matched that of the small study group.

3.9.4 Textual responses

A potentially troublesome set of variables to categorise and quantify are the open response questions asking the participant to identify a positive and negative sound source within the environment. The issue of synonymy is a major hurdle in this analysis, as it describes the phenomenon where different words describe the same thing [139]. Another issue is Polysemy, where a singular word may have multiple but related meanings [140].

Past research has utilised a technique of semantic analysis of free verbal responses from 77 participants [38], whereby psycholinguistic analysis of spontaneous verbal descriptions was conducted to identify semantic categories of environmental sounds and relevant sound quality criteria for urban soundscapes. This technique however, is not viable for this project's projected large scale data-set as it is performed manually and would be too larger job for a single person.

An alternative method of analysis for these responses is to employ Latent Semantic Analysis (LSA) [141]. This technique is an established method for automatically inferring the contextual similarity of words from a large collection of text descriptors. It is primarily utilised when dealing with large bodies of text to extract their semantic structure. The technique can be used alongside dynamic clustering to group descriptors based on their conceptual similarity using a Singular Value Decomposition Model [142]. Before this method can be implemented, the identification of similar terms for source descriptors must be performed. This process spots multiple uses of source identifiers, taking into account pluralisation and the harder to identify misspellings of these words. Pluralisations can be compensated for computationally using simple logic string comparisons. The detection of valid but misspelt sources requires a more in-depth process of extraction. Due to the method of entry on mobile devices, there are two main erroneous entries of sources that need to be compensated for: predictive text entry faults whereby the wrong word is entered due to the user allowing the phone's dictionary to wrongly assume the desired word and straight spelling mistakes where the user simply enters the desired word but spelt wrong. Misspellings of words will be harder to catch and will require comparing them with a dictionary database. Each of these techniques will also require a list of potential source choices to compare to, made up of the existing set from the non-erroneous entries gathered from the project.

Firstly the textual responses are pre-processed by: trimming any whitespace around single words and removing any special characters. A dictionary is then built up from the words and entries such as: “nothing”, “nowt” and blank entries, signifying that the participant has not identified a positive/negative source are replaced with “none” to keep consistency. The responses are then converted to a Term Matrix using the Text to Matrix Generator [143], using Matlab. This generates a matrix of weightings for each of the dictionary entries, based on their prevalence within the response set. This algorithm also disregards common terms such as “the” and “a”.

The term matrix is then fed into a clustering process based on a principal direction divisive partitioning (PDDP) clustering algorithm using k-means [143]. This algorithm determines the distances between clusters by computing the Euclidean distance from every entry within a particular cluster to every other point in every other cluster. The number of clusters is progressively reduced as each entry is assigned a “scatter” value that denotes its distance from the clusters centroid. In this instance, this process served to categorise around 90% of all entries into a number of clusters. These clusters were then manually identified as either: miscellaneous sounds, human sounds, natural sounds or artificial sounds. There were also a small number of clusters containing terms which the algorithm could not classify as they were too obscure, in a foreign language or spelt wrong. These had to be manually classified.

Once each sound source is classified, investigations can be made into soundscape response with respect to the type of positive/negative sounds identified. This will serve to determine the influence these source types have on appreciation and a persons perception of the soundscape. The prominence rating associated with each sound source will also help to uncover the strength of this influence as a function of their perceived prominence within the soundscape. This broad categorisation may serve to over generalise certain sound types, but will allow the project's findings to be compared to those from past research.

Using a more standard implementation of the LSA technique, the open responses to the question: "Why did you record this soundscape?", submitted by participants will also be analysed, providing an insight into the reasons why the user chose a particular soundscape to record. This information (captured at the online upload stage & on iOS devices) is limited to 8000 characters with the question asked of the participant: "Why did you choose to record this soundscape".

3.10 Summary

This chapter has detailed the methodology design behind the project. The methodology has provided the means to capture a large amount of localised data about a particular soundscape, the majority of which involve no user intervention or input requests. The ability to capture the precise record time and location of the soundscape with ease allows for the acquisition of a wide range of information about the location at that particular time. The actions required of the participant have been kept to a minimum, with the only requests (after the initial project discovery) requiring an active response being: demographics, subjective responses, the recording of the soundscape itself and the uploading of these. The automatically collected data, such as weather conditions and area house price is sourced after submission using the location, time and date data.

The main tool used for data collection is the project's custom mobile phone applications and web interface. The iOS route to submission is the easier and faster option as it consolidates all of the processes required to submit in one package, whereas the web interface and older Java mobile application require additional stages to complete a submission. The method of response entry across all submission paths was kept consistent with the use of discrete point sliders to rate each semantic differential scale. The terminology used in the project's question set has also been piloted and shown to be easily understood and interpretable by novice participants in the schools and live pilots.

It was identified that public engagement and communication was vital to the project's success due to the voluntary nature of participation. A number of methods were employed to reach people and engage with them to take part. Through a combination of easily accessible technology and marketing, the motivations behind participation can be exploited to ensure the maximum possible uptake of the project. Methods to aid in the retention of participants was also discussed, utilising social media and iOS weekly reminders.

The mobile devices themselves have been extensively tested for use in soundscape research, with the iOS devices showing themselves to be better suited to the recording of soundscapes. Issues around the inbuilt automatic gain control on all device types have been identified, but the conclusion is that this should not have a serious effect on the objective metrics extracted. The predominant use of feature averages means that the effect these do have will be reduced by this averaging across the large dataset. The frequency responses of the iOS devices was also investigated, revealing the presence of a sharp high pass filter, rolling off at around 100Hz. Urban soundscapes containing a high amount of low frequency energy caused by an abundance of road traffic will not be accurately captured using these devices, something which, unfortunately, cannot be remedied.

The objective and subjective data "cleaning" was then discussed. This included the conditions in which a soundscape recording or set of subjective responses were excluded as either a non-response or of too low quality in terms of the audio signal. The analysis processes were then detailed, ending with the methodological approach taken to the research carried out.

A final issue that has been considered is the difference in responses that may exist between responses made in situ using the projects iOS application and responses made retrospectively using the older Java application and web interface. These differences will be investigated in the analysis Chapter 4.

Chapter 4

Analysis of project dataset

4.1 Overview

This chapter details the analysis of the projects dataset. The analysis has been split into sections to focus on each type of data gathered, with inferences and validations running in-line with the presentation of the results. Section 4.2 describes the frequencies of submissions in relation to: submission method, location and app statistics. Section 4.3 details the demographics of participants including their submission habits. Section 4.4 details the analysis of the project's subjective data, with Section 4.5 covering the objective analysis. As of writing, the project has received over 1300 submissions, however, a large number of these arrived after the dataset had been analysed and processed so were not included in the following analyses.

The large majority of data shows non-normal distributions, therefore non-parametric methods have been used, such as Mann-Whitney U and Kruskal Wallis tests for group average comparisons [144]. In some cases, objective variable scaling such as Z transforming was employed due to the differing measurement scales of these metrics.

4.2 Submission statistics

As of January 2013 the project's number of individual submitted soundscapes stood at 826. Out of this total, 120 or 14.5% subjective responses were excluded as being non-responses, with the objective data having 81 or 9.8% excluded due to the file having a low sample frequency ($\leq$ 8000Hz). Therefore, the projects dataset included 706 subjective responses to soundscapes and 745 recordings which were used in the following analyses. The total number of submissions as of writing is much larger than this, however, these arrived after the analysis cut-off date.

4.2.1 Web and iOS comparison

Subjective responses made in-situ (mobile phone submissions) and reflectively (web submissions) were subjected to a Mann-Whitney U test to compare their distributions. The null hypothesis of this test is that the distributions are derived from the same population. A non-significant result would mean that the null hypothesis is rejected, and the two groups are considered to be homogeneous and have the same distribution.

	Quality	Loc. quality	Pleasant	Exciting	Eventful	Tranquil	Pos. prom.	Neg. prom.
M-W U	43509	44524	42832	45043	46525	46189	62388	59416
Z	-.224	-.936	-.821	-.694	-.012	-.167	-1.125	-.749
Asymp. Sig.	.743	.349	.416	.488	.990	.868	.261	.512

Table 4.1: Mann-Whitney U test for web and mobile phone groups

Based on the non-significant differences between the iOS and web submission responses seen in Table 4.1, the null hypothesis that the distributions of responses in each group are similar is accepted. As a result of this, any further analysis will use the combined data from both groups as one large dataset to increase the statistical power of these analyses.

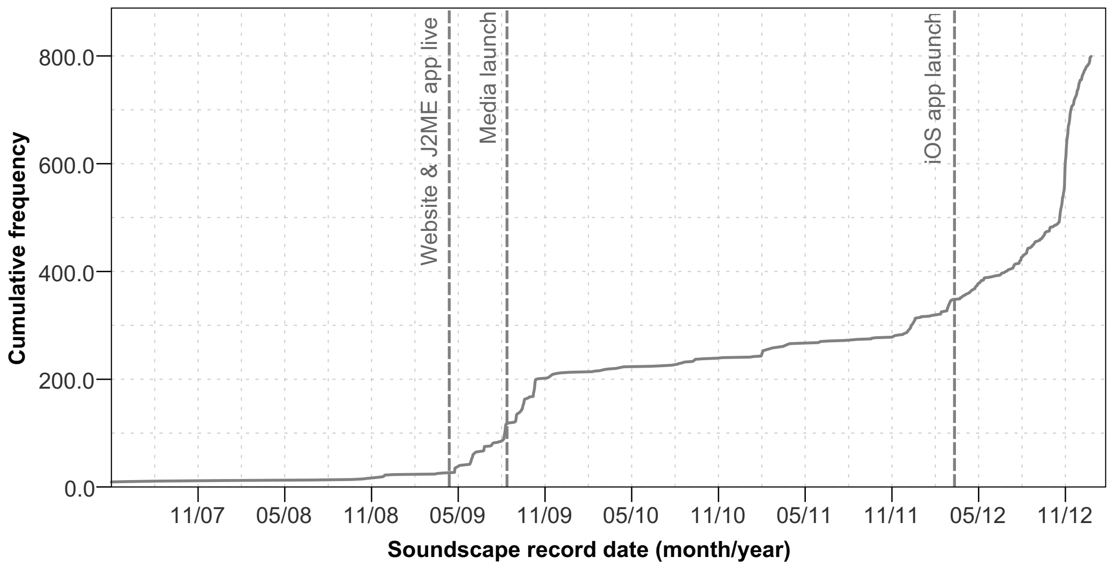

Figure 4.1: Cumulative soundscape submissions over the project lifespan (total N=826)

The major events in the projects lifespan are marked on Figure 4.1, which define the upload rate of soundscapes. A surge of submissions is seen with the launch of a new participation technique or especially when the project gains press attention as seen after the media launch point.

The iOS method is seen to be the biggest contributor to submission numbers since the project start date up until January 2013, illustrated in Figure 4.2. However, this may be misleading as the iOS app only went live in April 2012, whereas

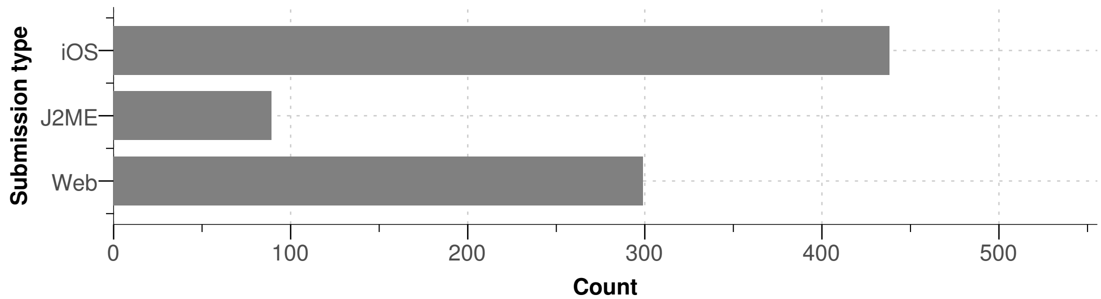

Figure 4.2: Submission device groups of total submissions (N=826)

the web interface has been live since April 2008. With an average monthly upload count of 48 against 7, the iOS submission rate shows a clear improvement over that of the web interface. The J2ME average monthly upload rate from its live launch in April 2008 to January 2013 stands at 2, highlighting the complications involved with submitting using this method.

The iOS application is clearly far more successful than the web or J2ME routes. The monthly submission averages reveal the ease of use of the app when compared to the other methods. Whilst a retention rate of 18% may seem low, app use research by Flurry [145] across 230,000 applications show that the average retention rate over 90 days for all app types is 35%. This means that after 90 days, an average of 65% of apps have only been used once. The project application has seen 35.6% of participants submitting more than one soundscape, which can be seen as an indication of participant retention as being in line with this studies average. In future studies, a response to the project server from the app after the install process indicating a new user would provide more information on retention rates.

4.2.2 App statistics

Since the iOS app went live on April 2012, a total of 1070 downloads has been logged as of January 2013. As there is no method of tracking app use with the current app version, it is difficult to identify the number of users who have downloaded the app but not submitted.

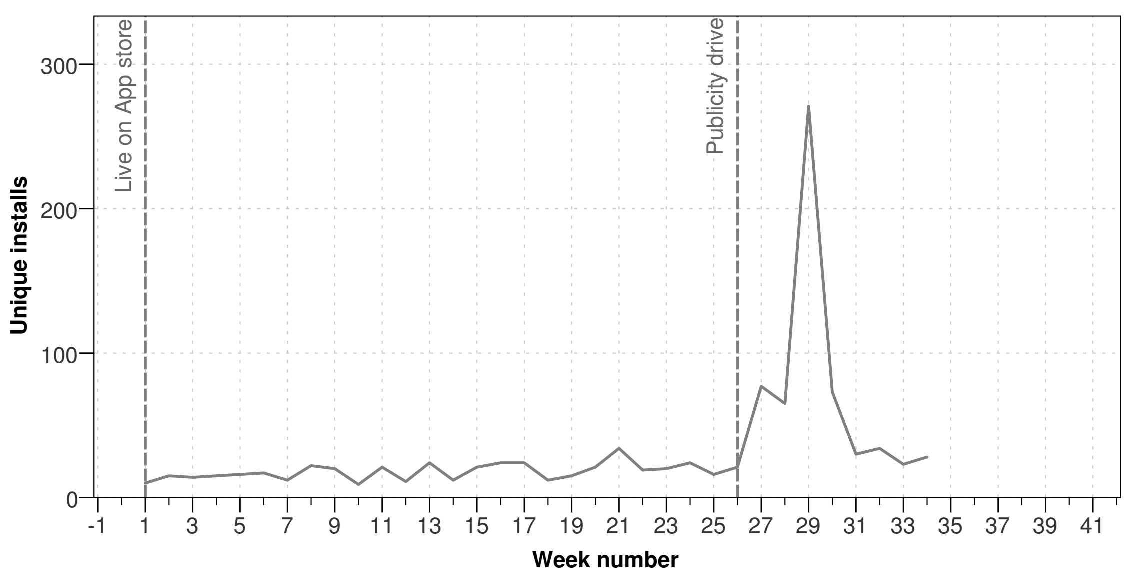

Figure 4.3: App downloads per week (total N=1070)

App downloads since May 2012 are shown in Figure 4.3. With individual iOS participants at 196 and total iOS downloads at 1070, this translates to a retention rate of 18% over a period of 9 months. Where the retention rate means the percentage of people who use the app to submit a soundscape after they have installed it.

Figure 4.4 shows the number of app downloads per country. The Others category is made up of downloads from countries with a frequency of one.

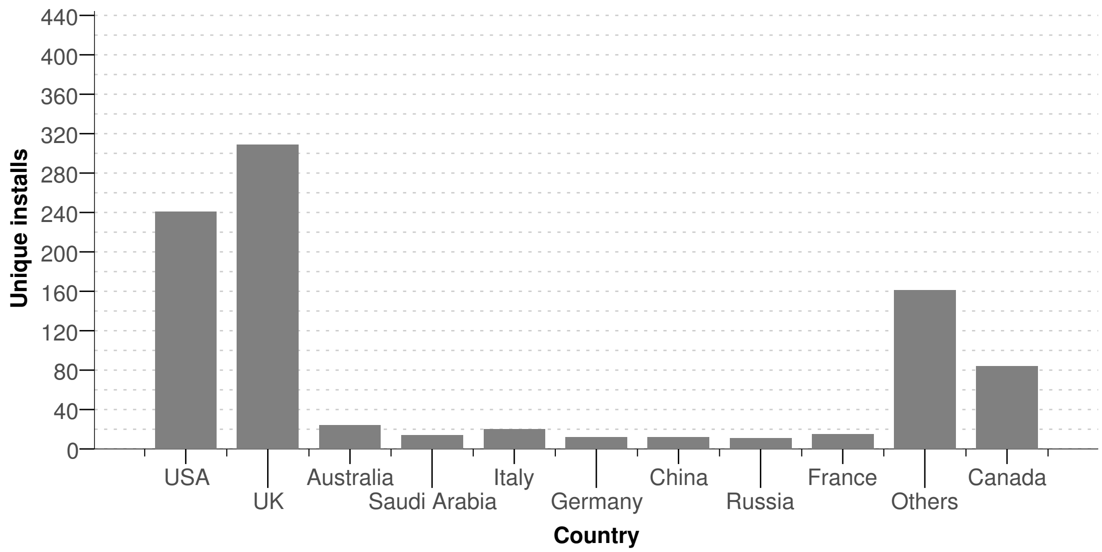

Figure 4.4: App downloads by country (total N=1070)

4.2.3 Submission country

The majority of submissions are made up of Great British soundscapes, followed by those from the United States. Figure 4.5 shows the list of countries from which soundscapes have been submitted from. The Other category is made up of a number of countries where under 2% of total submissions have come from. However, soundscape appreciation was not found to differ significantly between the two largest contributing countries, the US and GB.

The top three countries downloading the project app are English speaking countries, which is understandable, as the app is advertised on the App Store as being only available in an English language version. No differences in soundscape appreciation were observed between those recorded in the United States or Great Britain. However, this data cannot reveal the complexities of cultural dif-

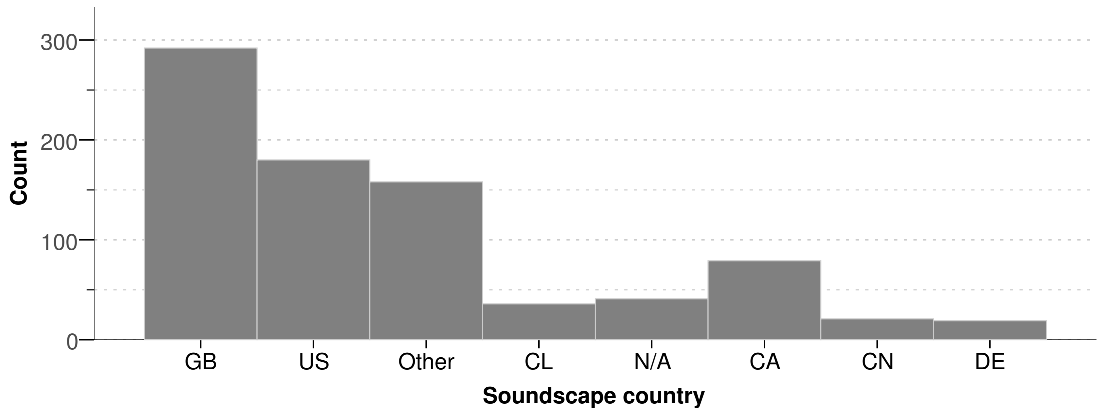

Figure 4.5: Submissions by country frequencies

ferences in terms of soundscape appreciation, due to the fact that the participants providing these worldwide submissions may not actually be a native of the country they were submitted from. More Participants downloading the application from non native English speaking countries are worthy of a mention as the assumption would be that English is their second language, which may mean that their interpretation of the questions asked will differ from that of a native English speaker. Worldwide studies of this kind would benefit from an extra demographic question asking about a persons native language and even their level of understanding of the English language. Whilst translation into all possible language variants may not be feasible, a limited number of versions of the app could be created to improve coverage of non native English speaking countries. This however will raise the issue of translation and whether the semantic differential terms hold the same meaning and are interpreted in a similar way across different localised versions of the app.

4.2.4 Soundscape capture time

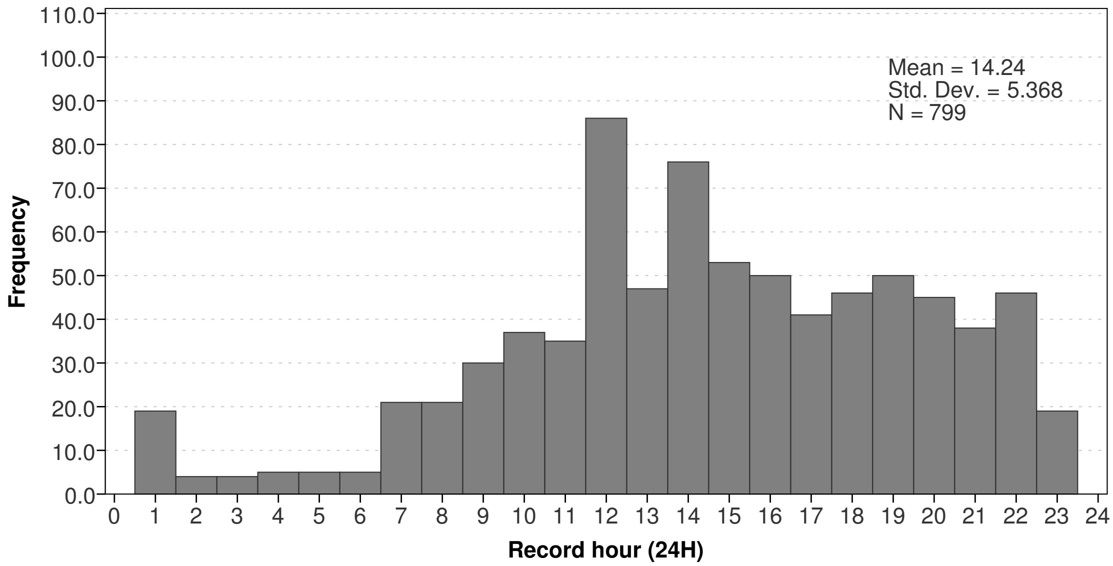

Figure 4.6: Soundscape record time distribution

The time of day that a soundscape was recorded at is shown in Figure 4.6. For web submissions this was extracted from the audio file information, which is absent in some cases, resulting in a user submitted record time which defaulted at 12:00, explaining its dominant spike on the plot. These were excluded from this stage of analysis. iOS and J2ME record times are logged at the time of recording. Sunrise and sunset times for the day and location that each soundscape was recorded at was retrieved using the Earthtools - Sunrise & sunset API [146]. Soundscapes captured after sunrise and before sunset were given a “Day” label, with all others labelled “Night”, in a bid to uncover any significant differences in soundscape response between the two. The only subjective variable which showed a significant difference was tranquillity (U=42341, Z=-2.293

p=.022) which showed that the median daytime rating was 5, with a night-time rating of 4. The differences observed between day and night time soundscapes, suggests that participants chose slightly more chaotic environments when making night-time soundscape recordings.

4.2.5 House prices

The distribution of average house prices are shown in Figure 4.7, exhibiting a strong positive skew.

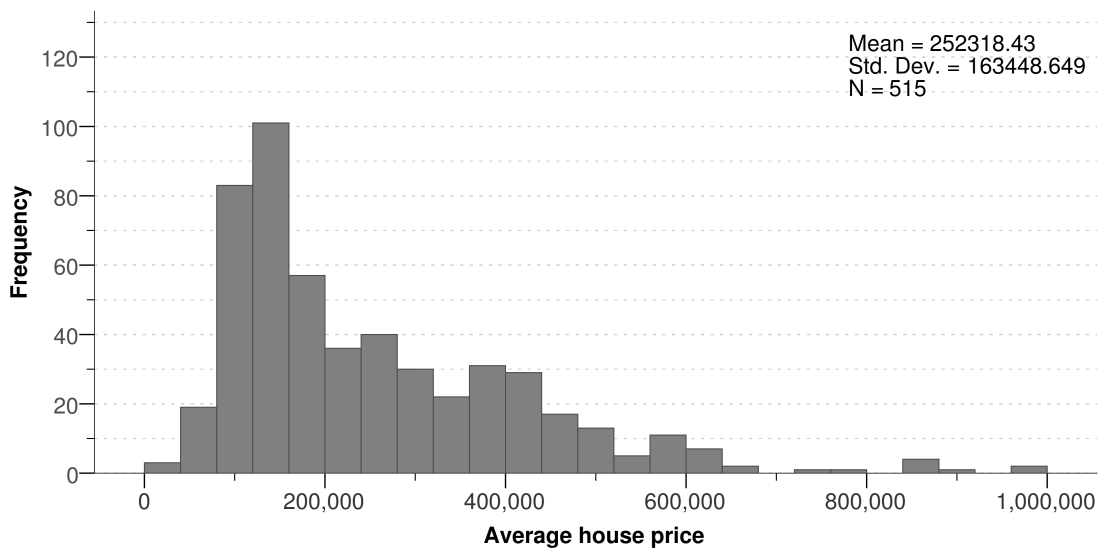

Figure 4.7: Soundscape area average house price distribution

After non-parametric Kruskal-Wallis ANOVA between subjective ratings and average house price no significant differences were observed, therefore the affluence of the area around which a soundscape was recorded in has no significant affect on the sound environment appreciation.

4.2.6 Weather

The weather API provides a wide range of weather descriptions, which were grouped into the more general categories shown in Figure 4.8. There were no significant differences observed in subjective ratings between the weather groups, suggesting that weather conditions do not have an affect on soundscape appreciation. The weather variables of temperature, humidity and wind speed also showed no significant relationships with subjective soundscape response.

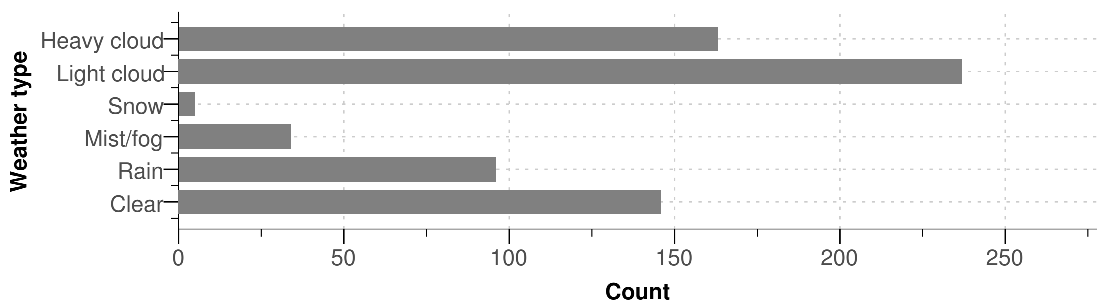

Figure 4.8: Weather group frequencies (N=681)

The current weather and local average house price of the area in which a soundscape was recorded in did not seem to have any effect on soundscape response. The lack of any weather effects is surprising as a previous study by Nyunt [76] has shown that good weather does influence a locations perceived “calmness” and “pleasantness”, which could be expected to translate to its soundscape appreciation. This however may be due to cultural differences as this particular study was carried out in New Zealand.

4.3 Demographics

As of January 2013 the project's number of individual participants submitting soundscapes stood at 323. Participants are relatively young, with 53.6% aged under 34, and 90.7% under 52. The gender distribution of participants shows a clear male bias of 72.1%.

	# of participants	% of participants
Gender		
Female	90	27.9
Male	233	72.1
Age group		
Under 10	25	7.7
10-13	18	5.6
14-17	6	1.9
18-21	19	5.9
22-27	43	13.3
28-33	62	19.2
34-39	37	11.5
40-45	47	14.6
46-51	36	11.1
52-57	17	5.3
58-64	7	2.2
Over 65	6	1.9

Table 4.2: Demographic characteristics

Figure 4.9 shows the majority of Web submissions against those from the J2ME app highlights the difficulties involved in distributing applications to these older devices, with only 8 participants using this application for soundscape recording. The dominance of iOS can be clearly seen as the submission method of choice.

The distribution of age groups submitting to the project is shown in Figure 4.10, with high percentages coming from the 22-27 and 28-33 age range. The gender bias within age groups is moderately significant (*U(2)*=8238, Z=-1.949, *p*=.051),

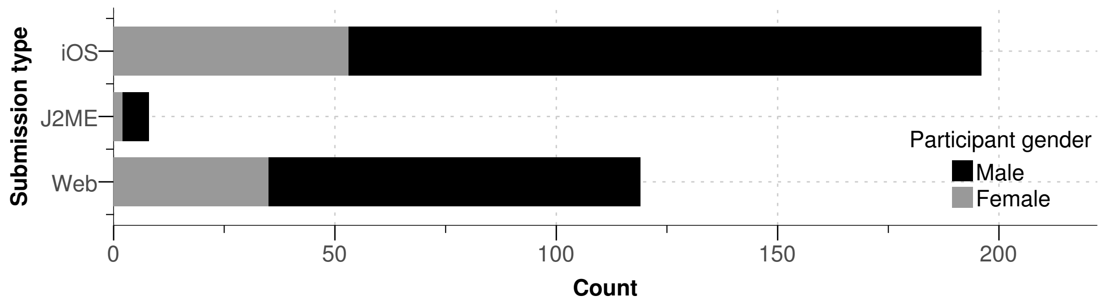

Figure 4.9: Submission device groups of project participants stacked by gender (N=323)

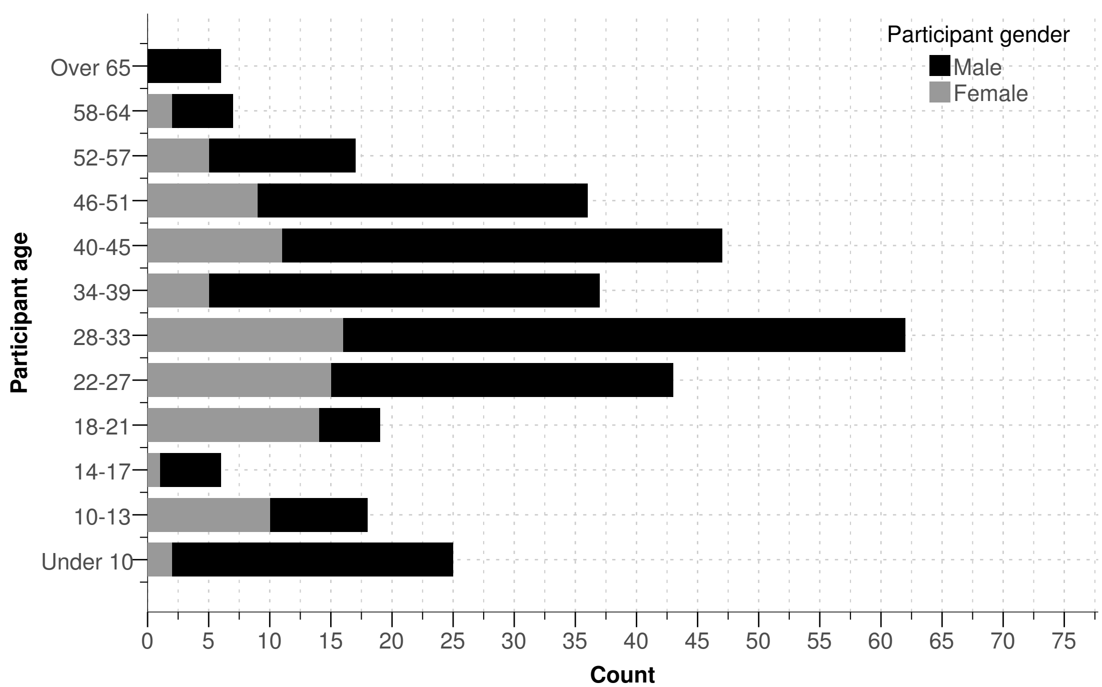

Figure 4.10: Age groups of project participants stacked by gender (N=323)

with a more pronounced male bias within the 34-57 age groups.

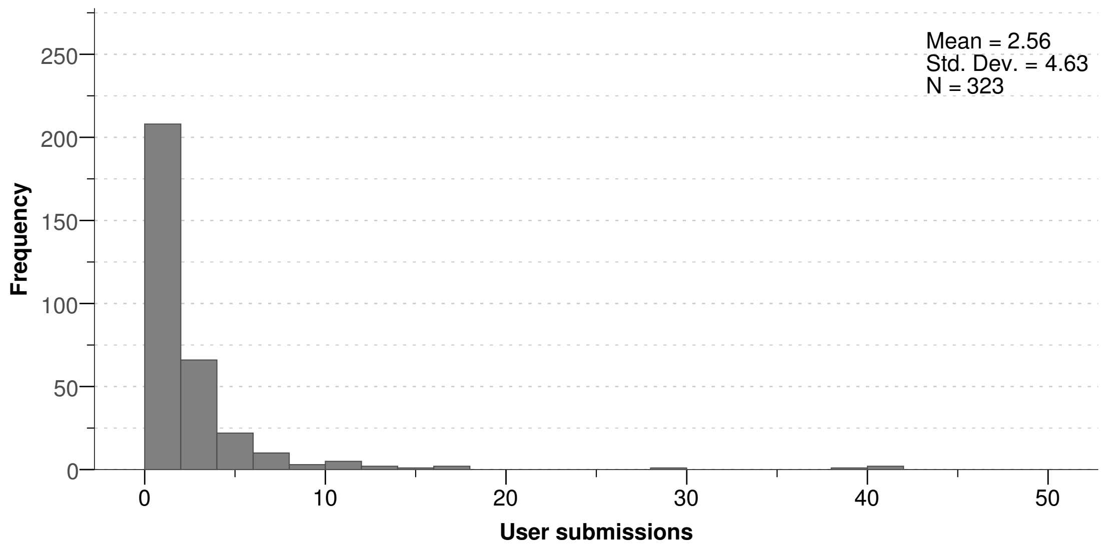

Figure 4.11: Submission frequencies of individual participants

The number of responses per participant ranges from 1 to 41 (mean=2.56, σ=4.63) seen in Figure 4.11. The distribution of participant submission numbers reveals a small number of especially proactive contributors that skew the distribution heavily to the right. A more accurate measure of the central tendency of participant submissions is the median, which stands at 1. It is also worth noting that there were no significant differences observed in the number of submissions made between genders or age groups. The spread of submission frequencies resembles an exponential distribution, which is also observed in studies investigating group dynamics, specifically, those looking into discussion participation in group situations [147].

The large reliance on participants with iPhones clearly restricts the sample's demographic profile. Statistics on iPhone user demographics in the US from 2011

[148] shows the dominance of 25-34 year olds in iPhone ownership, but only a 52% gender bias. The current studies participant demographics show a similar age trend to these statistics, as well as those of more traditional soundscape studies [24]. It is also apparent that participants were hindered in taking part by the involved process of installing the project's J2ME mobile application. The lack of an on device upload feature on these devices also means the participant must undertake the extra stage of copying the soundscape recording from their device to their computer for upload using the web interface. The combination of these two factors has resulted in a poor take up rate of participants using this method. The clear dominance of iOS as the participatory method of choice highlights its ease of use in comparison to the J2ME and Web routes.

4.4 Subjective analysis

Whilst all of the semantic differential scales may not be required to adequately characterise a soundscape, the validation of this project's findings require comparisons with a number of previous studies, which make use of a number of different scales. A principal component analysis may also create redundancy in the numbers of these subjective ratings. For correlational analysis with objective metrics, a larger number of subjective dimensions is also preferable.

4.4.1 Participant activity

Figure 4.12 shows the distribution of activities that the participant was involved in when the soundscape was recorded and commented on. It can be seen that the majority were captured whilst they were being passed through on the way to somewhere else. To investigate the difference in mean scores between participant activity types for the subjective responses a Kruskal-Wallis one-way ANOVA was performed on the different activity groups, with results shown in Table 4.3.

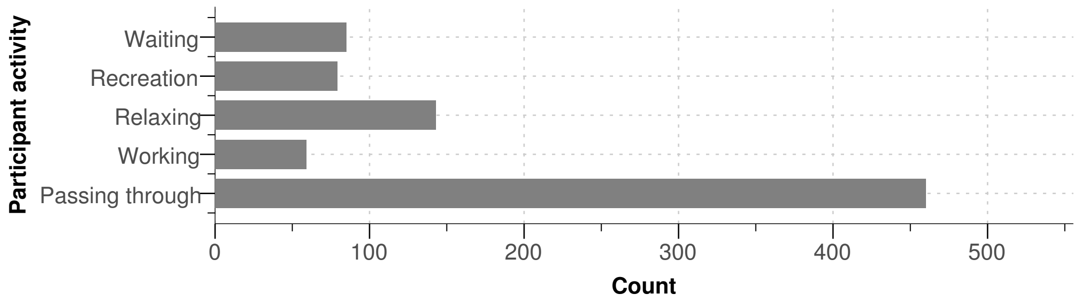

Figure 4.12: Participant activity frequencies (N=826)

	Quality	Loc. quality	Pleasant	Exciting	Eventful	Tranquil
Chi-Square	40.621	33.577	52.460	11.362	19.565	30.007
df	4	4	4	4	4	4
Asymp. Sig.	$<$.000	$<$.000	$<$.000	.023	.001	$<$.000

Table 4.3: Kruskal-Wallis one-way ANOVA of subjective responses between activity groups

Significant differences are seen across all subjective response ratings. To investigate further into this, mean ratings for each subjective descriptor are plotted for each activity grouping showing on average how participants responded when involved in different activities.

Figure 4.13 illustrates the differences in subjective ratings between activity. With significant differences in subjective ratings observed between activity groups, it can be assumed that the activity a person is involved in will have an effect on their perceptions of the soundscape they are immersed in. On average, participants perceived soundscape quality as being higher when their activity was logged as

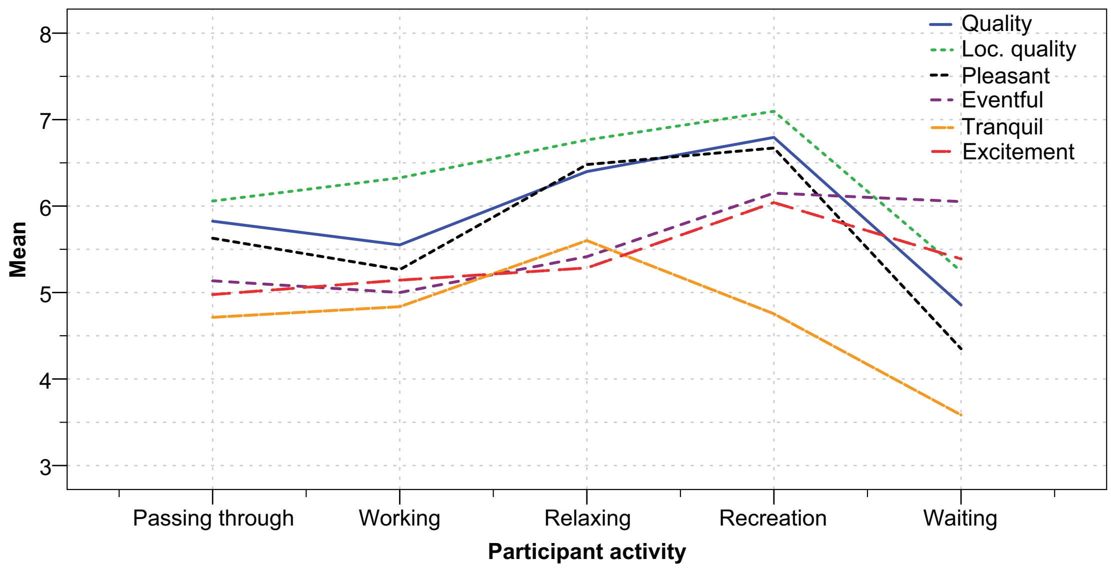

Figure 4.13: Mean subjective ratings grouped by participant activity

"relaxing" and "recreation". This implies that a person chooses to relax in a place that they regard as being of good quality in terms of the location itself and its soundscape, an expected and coherent finding. The perceptions of pleasantness are also generally higher when participants are relaxing or taking part in recreation. Pleasantness and quality are in close agreement throughout each activity, implying that there is a close relationship between these two factors. The eventful mean is more consistent across the activity groups suggesting that this factor is not considered so much when choosing a location to spend time in. The raised levels of perceived eventfulness and ratings of excitement in recreation situations, alongside the high ratings of location and soundscape quality suggest that the action of choosing a recreational activity means that even if it is high in activity it's soundscape is still rated highly. The action of waiting somewhere suggests less control over the participants location, therefore ratings of soundscape quality, location quality, pleasantness and tranquillity are significantly lower than when more choice in location can be assumed, such as when people are relaxing and taking part in recreation. As expected, the tranquillity average is at its highest when respondents were relaxing. Research using questionnaire based methodologies have also identified the influential nature of participant activity and control over their location on soundscape appreciation [1, 149]. It is worth mentioning however that the participant was actually "actively" listening to the soundscape when these responses were made. This means that their perception of the soundscape would have been generally heightened due to the process of taking part in this study.

4.4.2 Semantic differential analysis

As the subjective responses were mainly measured using these ordinal scales, the data is not normally distributed, therefore the following analyses use non-parametric approaches. This lack of normality was confirmed statistically using a one sample Kolmogorov-Smirnov test [144]. Firstly the inter-variable correlations and distributions are presented, followed by a non-parametric analysis of variance (ANOVA).

	Quality	Loc. quality	Pleasant	Exciting	Eventful	Tranquil	Pos. prom.
Loc. quality	.690**						
Pleasant	.739**	.659**					
Exciting	.225**	.153**	.151**				
Eventful	.145**	.122**	.063	.661**			
Tranquil	.385**	.389**	.572**	−.294**	−.360**		
Pos. prom.	.337**	.255**	.355**	.293**	.231**	.098*	
Neg. prom.	−.229**	−.149**	−.284**	.156**	.204**	−.251**	−.201**

Note: *p<.01, **p<.001

Table 4.4: Spearman's Rho correlation coefficients for subjective descriptors

Table 4.4 shows the Spearman's rank correlation coefficients for each combination of subjective rating. As expected there is strong correlation between soundscape quality and location quality, suggesting that soundscape appreciation does have an influence on the appraisal of a locations overall quality. Pleasantness and tranquillity are also significantly correlated with soundscape quality making them strong contenders for influential factors. Ratings of excitement and eventfulness, whilst significantly correlated, do not explain a large enough amount of variance (when compared to the other variables correlations) to be considered influential factors of soundscape quality. However, the positive interpretation of perception of soundscape excitement is evident with it showing a stronger relationship with perceived soundscape quality over ratings of eventfulness. The prominence of the positive and negative sound sources are positively and negatively correlated to quality respectively, indicating the influence of these sources in quality appraisal. A soundscapes rating of eventfulness and how exciting it is are highly correlated indicating that they stand as factors together when describing a soundscape. Eventfulness and excitement are also, as expected, negatively correlated to tranquillity, showing that a tranquil soundscape is generally regarded as being low in activity.

Since the dataset was of sufficient size, the appropriate measure of central tendency was the mean, where a non-parametric Wilcoxon test was used to investigate the differences between the mean values of the subjective descriptors. Figure 4.15 illustrates this difference graphically. The Wilcoxon test results in Table 4.5 show that participants rated the locations overall quality slightly higher than the soundscape quality. The largest difference seen is that between soundscape quality and tranquillity, as depicted by the higher Z value in Table 4.5. Z values based on negative ranks mean that this particular variable is showing a higher average value than the average value of soundscape quality and vice-versa for

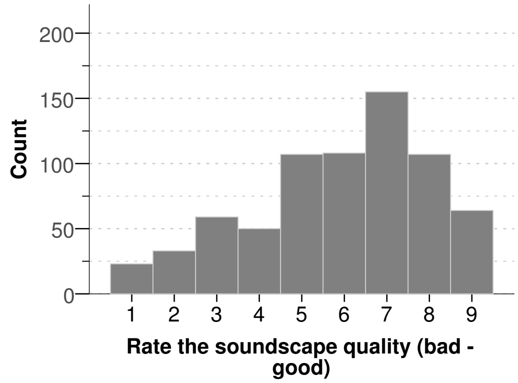

Figure 4.14.a: Soundscape quality

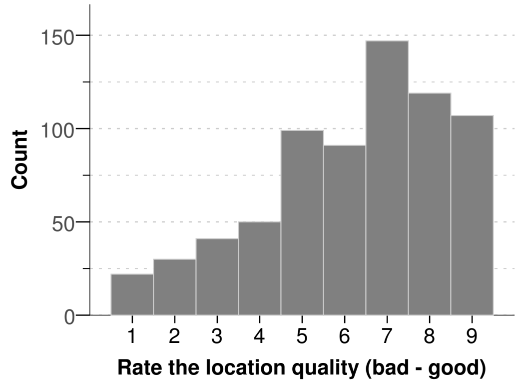

Figure 4.14.b: Location quality

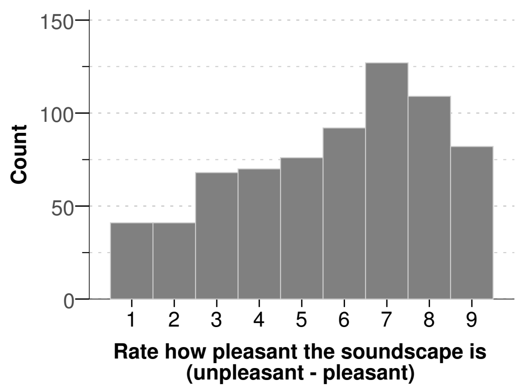

Figure 4.14.c: Soundscape pleasantness

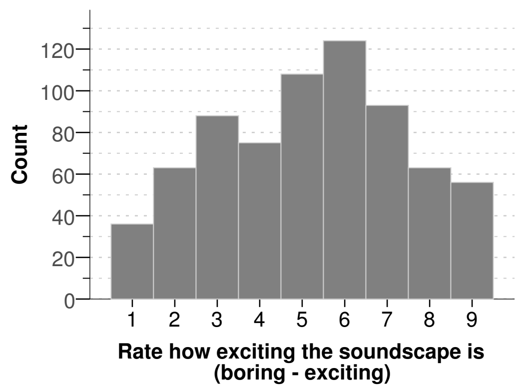

Figure 4.14.d: Soundscape excitement

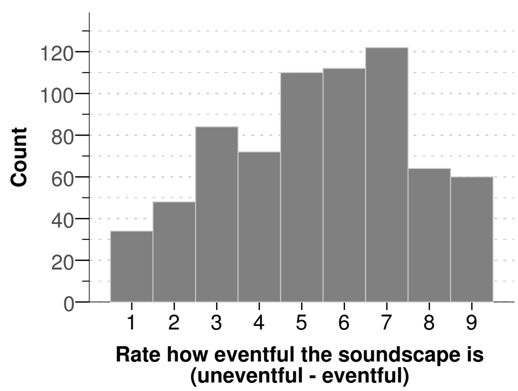

Figure 4.14.e: Soundscape eventfulness

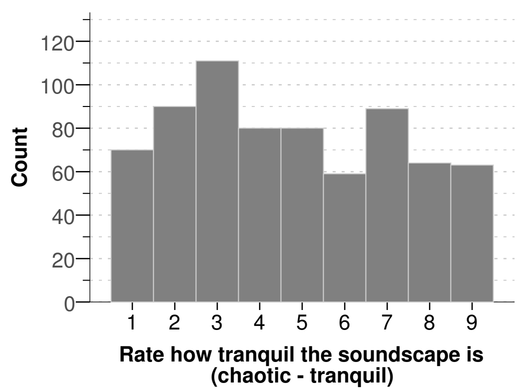

Figure 4.14.f: Soundscape tranquillity

Figure 4.14: Subjective response distributions (N=706)

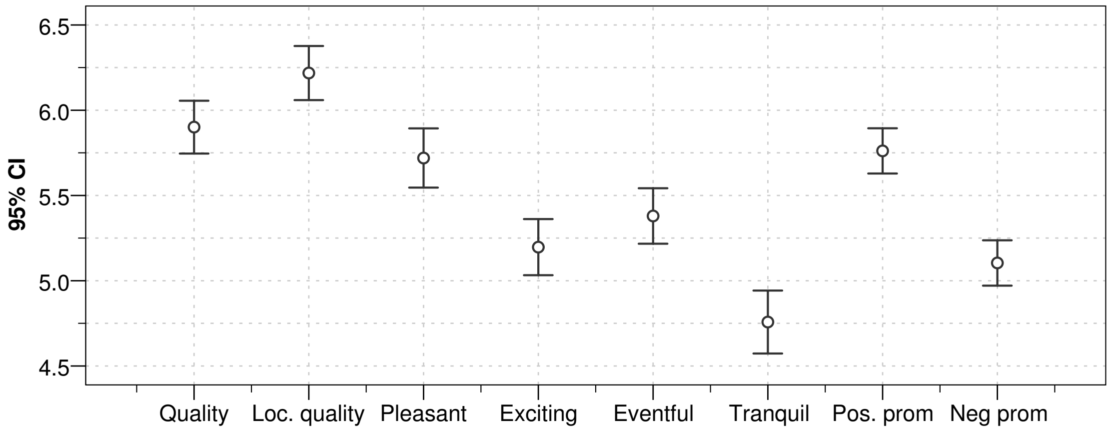

Figure 4.15: Subjective rating means with 95% confidence intervals

positive ranks.

	Loc. quality	Pleasant	Exciting	Eventful	Tranquil	Pos. prom.	Neg. prom
Z	-4.584[a]	-2.577[b]	-6.996[b]	-5.125[b]	-10.531[b]	-.282[b]	-6.538[b]
Asymp. Sig.	<.000	.010	<.000	<.000	<.000	.778	<.000

a. Based on negative ranks
b. Based on positive ranks

Table 4.5: Wilcoxon signed ranks test for comparison of subjective rating mean ranks against mean rank of soundscape quality

The strong correlations observed between soundscape quality and location quality seem to suggest that soundscape appreciation factors in the appraisal of a locations overall quality. It may be that the location quality itself can act to offset ratings of soundscape quality, as it will be incorporating more of the visual within the appraisal. Pleasantness and tranquillities close relationship to soundscape quality also mean that these measures are influential in defining soundscape appreciation. Ratings of how exciting and eventful seem to stand apart from the measures of soundscape appreciation, suggesting that they form their own factor when it comes to overall soundscape appraisal. With the large differences in mean values observed between tranquillity and quality, it is also apparent that tranquillity may not be as closely linked to soundscape appreciation as the measures of pleasantness and location quality.

The large numbers of submissions to the project's dataset means the statistical power of these analyses is relatively high. As a result of this, very low *P* values have been observed in most analysis stages. The intention for this project was the acquisition of as large a dataset as possible, as the inferences being made were not all fully defined and anticipated. In particular, the statistical robustness of the later performed principal component analysis in Section 4.4.5 relies on large sample sizes, reflected by a high value of Kaiser-Meyer-Olkin Measure of Sampling Adequacy. If the project had more of a focussed approach to the particular environments under scrutiny, then a minimum sample size could be established which would provide the required level of confidence (generally 95%).

4.4.3 Source identification

To visualise the distribution of terms used to identify a sound within a soundscape, positive and negative word clouds were created using the Wordle online tool [150, 151]. The size of the font is determined by the frequency at which the word occurs in the project's dataset.

With a lot more variation in more frequently used terms in Figure 4.16, the prominence of human sounds are immediately obvious. Natural sounds including: birds, water & wind seem to occur with high frequency, suggesting that these will prove to be another popular choice as a positive sound source. It is interesting that the majority of all mentions of human sounds refer to something a person may be doing, for example: talking, conversation or chatter. Negative sources seen

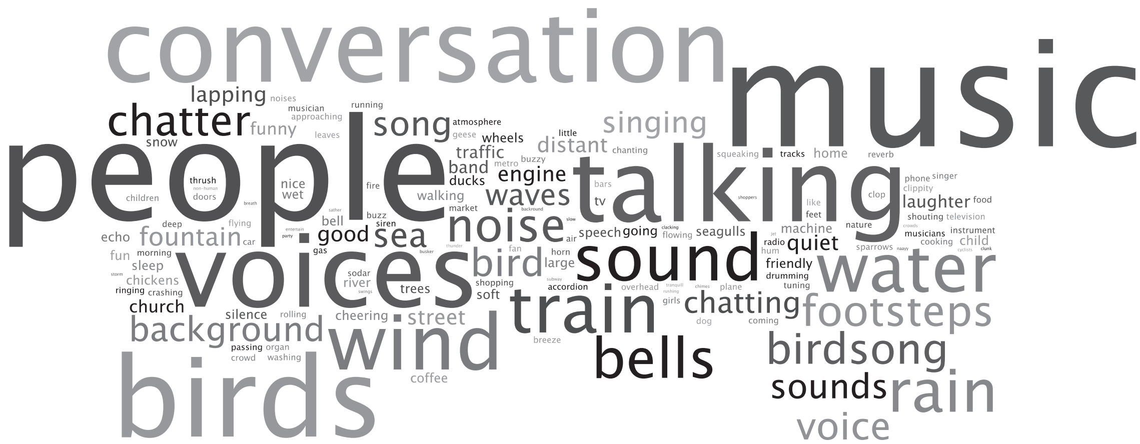

Figure 4.16: Positive source identifications size weighted by frequency

in Figure 4.17 tend to refer to the sound sources themselves, such as: traffic, cars or engine. This observation mainly occurs when subjects refer to human and artificial sounds. Human sounds are commonly lumped into a human group activity whereas artificial sounds generally have a specific target from which the sound emanates from.

Negative sources are clearly dominated by the term traffic as well as the more vague but negative term noise. The dominance of artificial sound sources is also apparent and to be expected due to the large majority of urban soundscapes submitted. These identifications of positive/negative sounds have the advantage that they are not led by a researcher, who might suggest a sound to inspire the participant to respond, as in more traditional methodologies. The responses also follow the same pattern as those provided in a study carried out by Guastavino [38]. She analysed free responses of participants about the “Ideal urban soundscape”, where the ideal soundscape was predominately described in terms of the object source producing the sound (76% of occurrences). Human and natural sounds

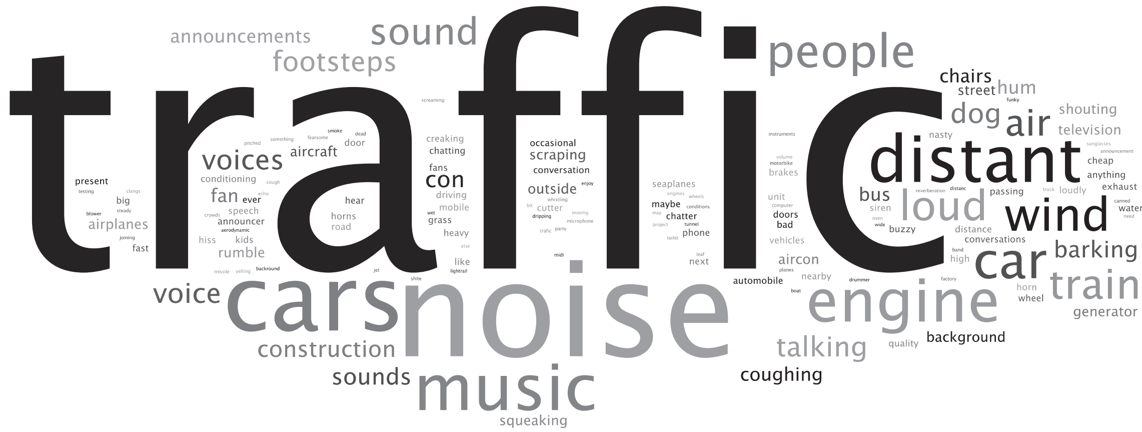

Figure 4.17: Negative source identifications size weighted by frequency

also were more prominently described using positive terms with artificial sounds being generally described using negative terms, as seen in the present study. The negative identification of “traffic”, “cars” and “engine” sounds also suggests that this is an appraisal of the source itself. For example, all of these commonly identified negative sources mentioned produce pollution which is riddled with negative connotations. The negative identifications therefore stray further from their simple auditory perception much more so than their positive counterparts. The acceptance of human sounds in urban environments is presumably due to the fact that cities are human constructs where the absence of humans would result in a very alien place. Negative transport noise in urban situations is potentially avoidable, seen in the large numbers of pedestrian areas in cities. These responses may be indicative of the element of choice within urban environments, where it is very hard to avoid being around people so their sound is accepted and appreciated but transport noise can be avoided thanks to these pedestrian zones and parks, resulting in a reduced appreciation of it.

Positive and negative sound responses were then manually categorised after LSA clustering. After the positive/negative source identifications had been categorised, it was possible to carry out explorative and statistical analysis on these now categorical variables. Source identifications left blank or identified with terms such as “none”, “nothing” or a response which bore no relation to individual sources made up 36.7% of positive submissions and 54.5% of negative submissions. These were excluded from further analysis. Whilst these may seem salient, it is impossible to distinguish between a true response of no positive/negative sounds present and a non-response, meaning that the subject has simply skipped the question.

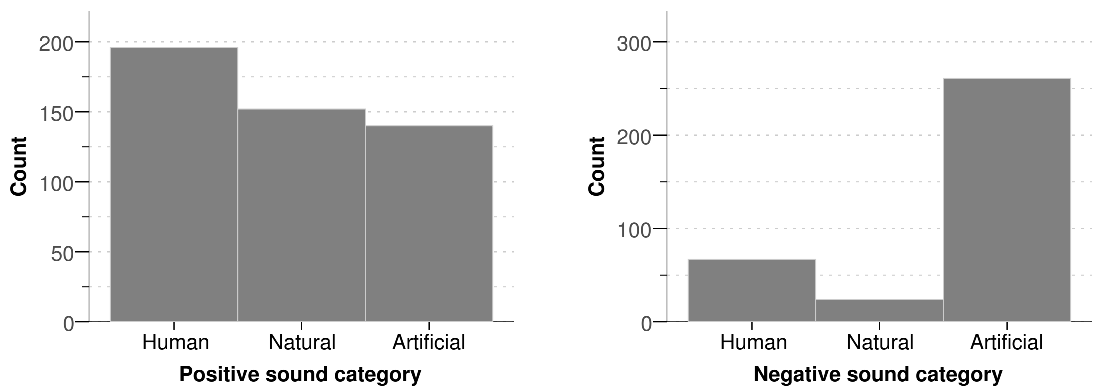

Figure 4.18.a: Positive source

Figure 4.18.b: Negative source

Figure 4.18: Positive and negative category distributions

The predominance of human sounds is clear in the positive identifications, whereas artificial sources make up a clear majority of the negative sources. Figure 4.18 shows this distribution graphically. Table 4.6 describes the central tendencies of the subjective ratings highlighting any significant differences between source type groupings. With significant rating differences seen between categories for soundscape quality and pleasantness, the type of identified positive source does seem to have an effect on these responses, although a rather small one as seen by the low Chi-Square values for each and the minimal change in means between groups. How exciting, eventful and tranquil, however show a much stronger significance and larger Chi-Square values. A soundscape identified with a human type positive sound is seen to be more exciting than one with a natural positive sound. A more pronounced difference is observed in ratings of eventfulness between human and natural sound groups.

Mean	Quality	Loc. quality	Pleasant	Exciting	Eventful	Tranquil
Human	5.93	6.24	5.88	5.67	5.93	4.30
Natural	6.53	6.74	6.56	4.72	4.86	6.20
Artificial	5.79	6.34	5.44	5.38	5.53	4.20
K-W test						
χ^2	10.770	6.816	18.123	18.712	22.367	61.792
Asymp. Sig.	.005	.033	<.000	<.000	<.000	<.000

Table 4.6: Mean ratings by positive source category with Kruskal Wallis ANOVA test

There were no significant differences found for the case of the identified negative sound categories.

Participants also rated the "prominence" of these identified sound sources. This rating was dichotomised, with ratings from 1-5 being classed as "Not prominent" and ratings of 6-9 as "Prominent". Table 4.7 compares the means of each subjective rating between these new categories.

A strongly significant difference is observed between all subjective ratings with the exception of tranquillity. This suggests that the prominence of an identified positive sound source does not have an influence on perceived soundscape tranquillity. This finding is mirrored in the lack of any correlation between tranquillity and positive sound prominence seen in Table 4.4.

When focussing on the positive sounds identified as prominent within a sound-

Mean	Quality	Loc. quality	Pleasant	Exciting	Eventful	Tranquil
Not prominent	5.27	5.65	4.95	4.65	4.89	4.57
Prominent	6.40	6.64	6.30	5.64	5.79	4.84
M-W U test						
U	38537	40982	37171	40692	41882	52161
Z	-6.765	-5.766	-7.299	-5.867	-5.386	-1.203
Asymp. Sig.	<.000	<.000	<.000	<.000	<.000	.229

Table 4.7: Mean ratings by positive source prominence category with Mann-Whitney U test

Mean	Quality	Loc. quality	Pleasant	Exciting	Eventful	Tranquil
Human	6.27	6.43	6.20	6.02	6.34	4.13
Natural	7.06	7.11	7.17	4.87	4.82	6.56
Artificial	5.86	6.50	5.61	5.62	5.66	4.22
K-W test						
χ^2	17.427	8.826	25.208	16.441	26.922	55.873
Asymp. Sig.	<.000	.012	<.000	<.000	<.000	<.000

Table 4.8: Mean ratings by prominent positive source category with Kruskal Wallis ANOVA test

scape, significant differences are seen across all subjective ratings, shown in Table 4.8.

Mean	Quality	Loc. quality	Pleasant	Exciting	Eventful	Tranquil
Not prominent	6.22	6.38	6.13	5.06	5.16	5.14
Prominent	5.27	5.87	4.85	5.49	5.85	3.88
M-W U test						
U	38141	43988	35221	44134	40137	34916
Z	-5.018	-2.509	-6.250	-2.439	-4.150	-6.372
Asymp. Sig.	<.000	.012	<.000	.015	<.000	<.000

Table 4.9: Mean ratings by negative source prominence category with Mann-Whitney U test

After the categorisation of identified negative sources into prominent/not prominent groups, significant differences are seen across all subjective ratings shown in Table 4.9. Of note are soundscape quality, pleasantness and tranquillity, where a decrease in rating value is seen when the negative source is prominent within the soundscape. When compared, the prominent negative source types of human, natural and artificial only showed significant differences in ratings of quality (χ^2=8.975 p=.011) and pleasantness (χ^2=6.895 p=.032). Prominent artificial sources are seen to reduce perceived soundscape quality and pleasantness.

Positive source identifications reveal the majority of human sourced sounds, largely focussed on human activities. It would seem that being within hearing distance of other people and the activities they are involved in is generally considered to be favourable. The tendency for people to identify the actual source of negative sounds might indicate that the participant has also made a visual identification of the object emitting the sound source.

The Latent Semantic Analysis (LSA) clustering of sound sources did prove to be relatively successful when categorising the identified positive and negative sounds, however, the amount of human intervention required to actually identify the groupings and place sounds within these make this technique only a slight improvement over a purely manual approach.

The clear dominance of human and natural sound source types within the positive sound group, has been previously identified in a number of studies, including those of Hall and Kang [152, 77]. Ratings of soundscape appreciation between these groupings were highly significant but small, suggesting that the type of positive source within a soundscape only has a small affect on appreciation. It would seem that the perceived “action” of a soundscape is significantly affected by the presence of certain types of sources, especially in the case of an identified positive human source, which tends to increase eventfulness and ratings of excitement, a similar outcome was seen by Axelsson in [5]. Based on this, soundscapes containing a positive natural sound source rather than a human or artificial one are perceived as less eventful. Tranquillity ratings also see a marked increase when a natural source is identified.

The similarities seen between these findings and those of previous studies are promising, as all previous studies mentioned utilised in-situ and lab based questionnaire techniques to gather soundscape responses. This agreement in subjective response across data acquisition types is the first indication of a validation of this studys methodology to gather meaningful soundscape response data.

The prominence of the identified positive sources is also shown to have an affect on soundscape responses. A prominent positive source unsurprisingly is associated with high ratings of soundscape quality, location quality and pleasantness. Increases in ratings of how exciting and eventful are also observed, which could mean that this dominant positive source is right at the foreground of the soundscape and is adding to the action and dynamics of environment. The lack of any effect on tranquillity suggests that this subjective rating is not determined by individual sources but by the soundscape as a whole.

Prominent natural sounds show a marked perceived increase in soundscape quality, pleasantness and tranquillity over soundscapes containing prominent artificial sources. Soundscapes containing prominent human sources are perceived as more exciting and eventful than those with prominent natural sources. Once again these findings have been observed in previous studies utilising more traditional data gathering techniques [5].

As expected, prominent negative sound sources have a negative effect on the subjective ratings of: quality, location quality, pleasantness and tranquillity. Specifically, prominent artificial sources seem to have the most effect in reducing perceived soundscape quality and pleasantness. These findings, again confirm the ability of the methodology to capture these subjective reactions to human sound environments. Whilst it may be obvious that a highly prominent negative sound source will negatively affect perceptions of soundscape quality, the fact that the data gathered is confidently showing this, is further validation of the methodologies ability to gather meaningful subjective data.

4.4.4 Manual soundscape categorisation

In order to focus on a smaller number of distinct soundscape types, all submissions were manually categorised into three groups. This is something which could be asked of participants in future studies.

- **Urban:** inner city exterior location generally surrounded by large buildings and roads
- **Interior:** a submission made from inside any type of building
- **Rural:** exterior location a substantial distance from a densely populated area

The size of the dataset necessitated these rather broad groupings, due to the time required to categorise each submission. The factors influencing category choice were:

- **Satellite imagery:** where urban and rural locations could be identified from their surrounding landscape
- **Reason for recording:** the location is generally described by the subject in their submission
- **Populated area proximity:** the distance to the nearest densely populated area
- **Audio:** the soundscape itself can be auditioned

If a submission could not be classified using at least three of these factors, it was placed in a miscellaneous category and excluded from further analysis. Using a combination of these factors, 97.3% of soundscapes were categorised.

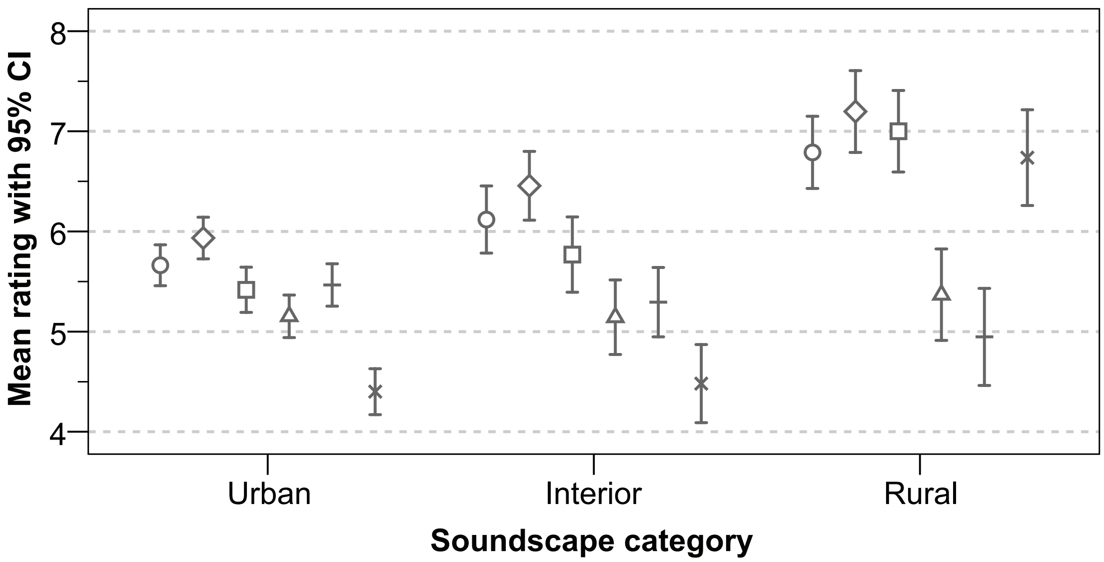

Figure 4.19: Subjective rating means by category with 95% CI (◯ = Quality, ◇ = Loc. quality, □ = Pleasant, △ = Exciting, – = Eventful, X = Tranquil)

Mean	Quality	Loc. quality	Pleasant	Exciting	Eventful	Tranquil
Urban (N=414)	5.67	5.94	5.43	5.15	5.47	4.41
Interior (N=161)	6.13	6.47	5.79	5.15	5.30	4.51
Rural (N=76)	6.79	7.20	7.00	5.37	4.95	6.74
K-W test						
χ^2	22.525	28.574	31.861	.521	4.472	54.835
Asymp. Sig.	<.000	<.000	<.000	.771	.107	<.000

Table 4.10: Mean ratings by soundscape category with Kruskal Wallis ANOVA test

Figure 4.19 shows the mean subjective ratings between soundscape categories with the significant differences shown in Table 4.10. Soundscape and location quality ratings are on average lower in urban situations, with progressively higher mean ratings for interior and rural locations. Pleasantness rating differences are more pronounced between locations, with a significantly higher rating in rural situations. Tranquillity ratings are similar in urban and interior settings, but show a large increase in rural situations. Surprisingly, ratings of how exciting and eventful these soundscape categories were did not show any significant differences, as a fair assumption would be that the urban soundscape would be considered more dynamic in nature.

The differences in ratings between rural and urban soundscapes reveal that urban sound environments are rated significantly lower for all aspects of appreciation. The surprising non significant differences in how exciting and eventful these locations are maybe due to participants being accustomed to the nature of these environments and therefore do not consider them to be more exciting. The positive connotations of the rating of excitement are seen in the similar quality ratings in urban settings, but show large differences in rural settings, presumably because of the more sedate nature of rural soundscapes.

Figure 4.20 shows the percentages within each soundscape category of identified positive/negative sounds. Urban situations show the participants preference for human sounds closely followed by natural then artificial. Interior soundscapes show a clear dominance of positive human sources which makes sense as interior spaces are generally reserved for human occupation. Interior soundscapes containing positive natural sources are presumably made up of situations in which external sounds such as birdsong can be heard inside a building. Unsurprisingly, in rural situations, the place you would expect to find more natural sources does show a large majority of identified positive natural sources. Identified negative sources are once again dominated by artificial sounds across all soundscape categories.

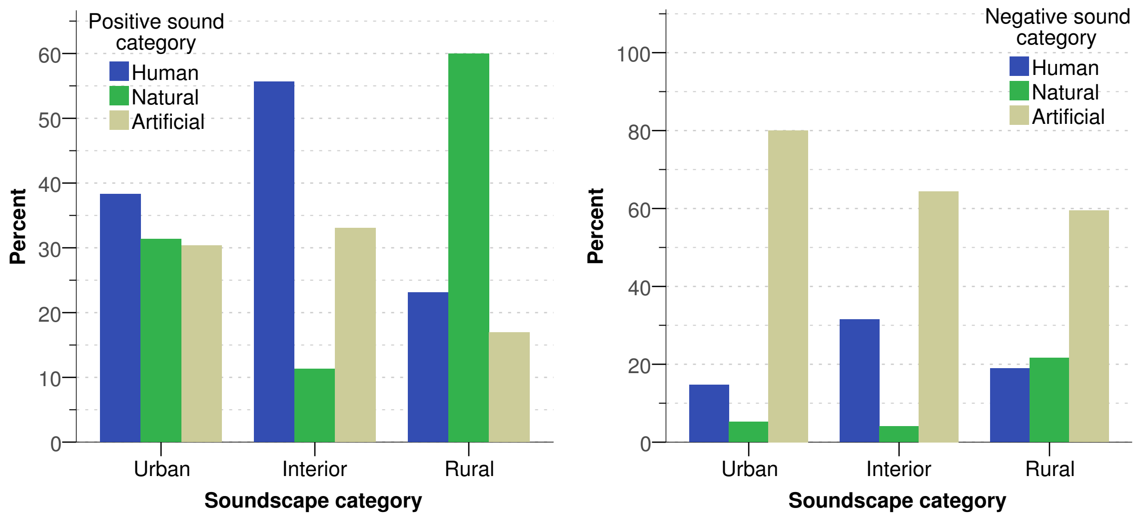

Figure 4.20: Positive & negative category distributions between soundscape categories

The preference for human sounds within urban environments could be attributed to the expectation that people have of humans being in urban environments. An urban place with no human sound sources would possibly feel lonely and abnormal. Interior soundscapes are not surprisingly dominated by positive human sources and the rural submissions were again expectedly filled with positive natural sources.

4.4.5 Principal component analysis

To investigate the influential and distinct subjective characteristics of these soundscapes, a principal component analysis (PCA) was performed on the responses to the semantic differential scales using a variance maximizing rotation of the original variable space (varimax). The regressed scores from each factor will then be used in further analysis to help explain and summarise the complex objective interactions involved in soundscape perception. Table 4.11 shows the two components extracted with a criterion factor of eigenvalue >1, which account for 77.4% of variance (Component 1 45.8%, Component 2 31.6%, Kaiser-Meyer-Olkin Measure of Sampling Adequacy = .710).

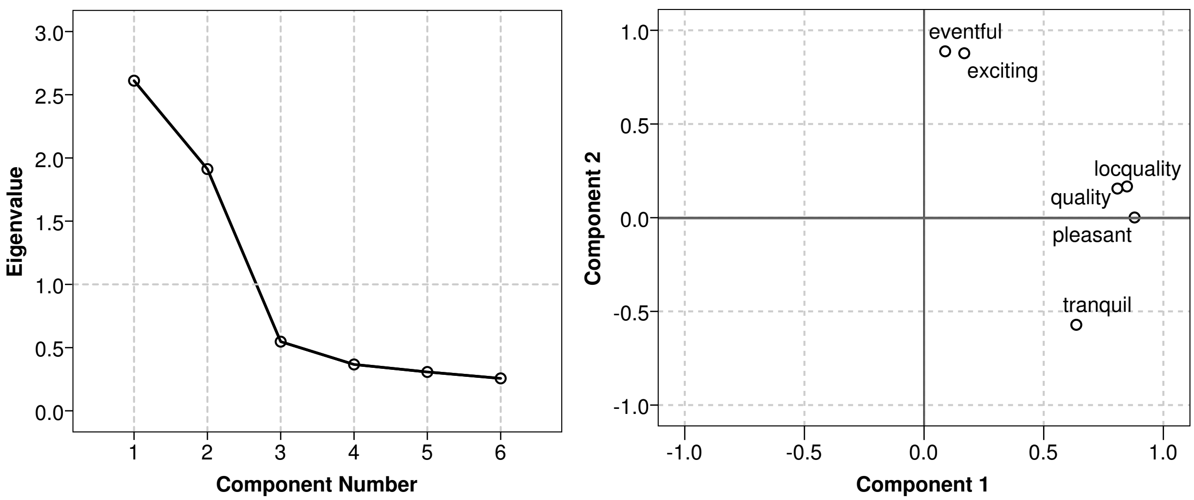

Figure 4.21.a: Scree plot

Figure 4.21.b: Rotated factor plot loadings

Figure 4.21: Scree plot and rotated factor plot for subjective rating variables

Figure 4.21 shows the scree and loading plot for each variable within the extracted components. Component 1 consists of descriptions of “appreciation”, containing the variables: soundscape pleasantness, tranquillity, location quality and soundscape quality. Component 2 describes the “dynamics” of a soundscape, made up of ratings of excitement and eventfulness. Whilst all the other variables load very strongly along each component axis, tranquillity shows a strong positive loading on the relaxation component, as well as a comparable negative loading on the dynamics component.

The two extracted subjective components of “Appreciation” and “Dynamics”

	Component 1	Component 2
How pleasant	.881	
Quality	.851	
Loc. quality	.810	
How tranquil	.656	−.546
How eventful		.892
How exciting		.881

Table 4.11: Varimax rotated component matrix for subjective descriptors (loadings< |0.2| removed)

exhibit distinct regions within the factor space. The exception is the variable tranquillity, which has a relatively strong negative loading on the dynamics factor. This suggests that tranquil soundscapes are perceived as more relaxing and chaotic soundscapes are seen as being more dynamic in nature. Whilst this may seem obvious, it confirms participants correct interpretation of the studies semantic differential terms and the robustness of the data gathered. The distinctive loadings on the dynamics component of Exciting and Eventful suggest that these soundscape components do not play a direct role in the subjective rating of quality.

Research	Factor	Expl. variance	Description
Kang et al. [6]	1	26.0%	Relaxation (comfort–discomfort, quiet–noisy, pleasant-unpleasant, natural-artificial, like-dislike, gentle-harsh)
	2	12.0%	Communication (social-unsocial, meaningful-meaningless, calming-agitating, smooth-rough)
Axelsson et al. [5]	1	49.0%	Pleasantness (pleasant, appealing)
	2	19.0%	Eventfulness (eventful, lively)
Raimbault et al. [153]	1	67.0%	Assessment (pleasant–unpleasant) linked to strength (quiet–loud)
	2	15.0%	Sound dynamics: temporal balance (steady–unsteady), spatial arrangement (organised–disorganised)
Viollon et al. [154]	1	46.6%	Affective impressions, preferences (pleasant, comfortable, rural, friendly, silent)
	2	18.0%	Activity due to sound presence of human beings (bustling, marked by living creatures)
Kawai et al. [155]	1	25.0%	Preference (irritating–relieving, unpleasant–pleasant, artificial–natural)
	2	16.8%	Activity (lively–deserted, joyful–empty, exciting–gloomy)

Table 4.12: First two factors emerging in PCA of soundscapes based on semantic differentials

The PCA outcome shows similarity with findings made by Kang, Axelsson and others detailed in Table 4.12, where the main principal components of relaxation/pleasantness were extracted (Kang - Relaxation, Axelsson - Pleasantness) from semantic differential perceptions of soundscapes. Axelsson described his second component as Eventfulness, closely matching the outcome from the present study, whereas Kang describes his as Communication. It is worth mentioning however that both of these studies focussed on urban soundscapes and utilised a large number of individual subjective variables. The limitations imposed by the use of mobile devices and unsupervised voluntary participation meant that requesting responses from subjects had to be kept to a minimum. This voluntary choice available to participants also meant that there was no control over where people may capture these sound environments, meaning that urban and rural soundscapes are analysed as one. A supplementary question asking participants to label the soundscape as “rural” or “urban” would benefit future studies of this kind.

Another study with comparable results is that of Kawai et al [155]. They used a different methodology using a subjective card sort technique to evaluate the structure that lies at the basis of peoples psychological evaluation of environmental sounds. The components extracted were: Preference and Activity, which correspond to the components extracted in the present study. Numerous other studies have also been found that have extracted comparable subjective principal components [156, 7].

To utilise these extracted components, 40 submissions were then manually selected to represent 4 subgroups of UK soundscapes to compare the subjective factor scores of these differing acoustic environments. This number was selected because of the time taken to manually ascertain if they could be grouped into the following:

- **Urban:** inner city locations generally surrounding by large buildings and roads
- **Rural:** locations surrounded by countryside such as farmland
- **Urban public space:** inner city spaces, such as squares, plazas and markets
- **Urban park:** inner city green space generally surrounded by large buildings and roads

These soundscapes factor scores are plotted in Figure 4.22, with the x-axis defined by the component of Appreciation and the y-axis representing Dynamics.

The urban soundscapes form a cluster, predominately located on the left of the factor space. With low scores of Appreciation, it is clear that this soundscape group is considered to have an inferior sound environment. The Dynamics scores of the urban type show spread across the x-axis, signifying a lot of variation in the perceived dynamics of these urban soundscapes. Rural soundscapes sit in the opposite half of the factor space to urban group, showing slightly lower dynamics scores, but a marked increase in appreciation. The urban public space group have high scores on dynamics and have an even spread of appreciation scores, suggesting that these places are met with mixed emotions, but are generally considered to be exciting and eventful spaces. Surprisingly, urban parks exhibit generally lower perceived scores for dynamics than rural soundscapes. The energy of an urban environment should serve to increase perception of eventfulness and excitement, but this finding is contrary to that assumption. The generally large dynamic range of rural soundscapes may be perceived as more more exciting and

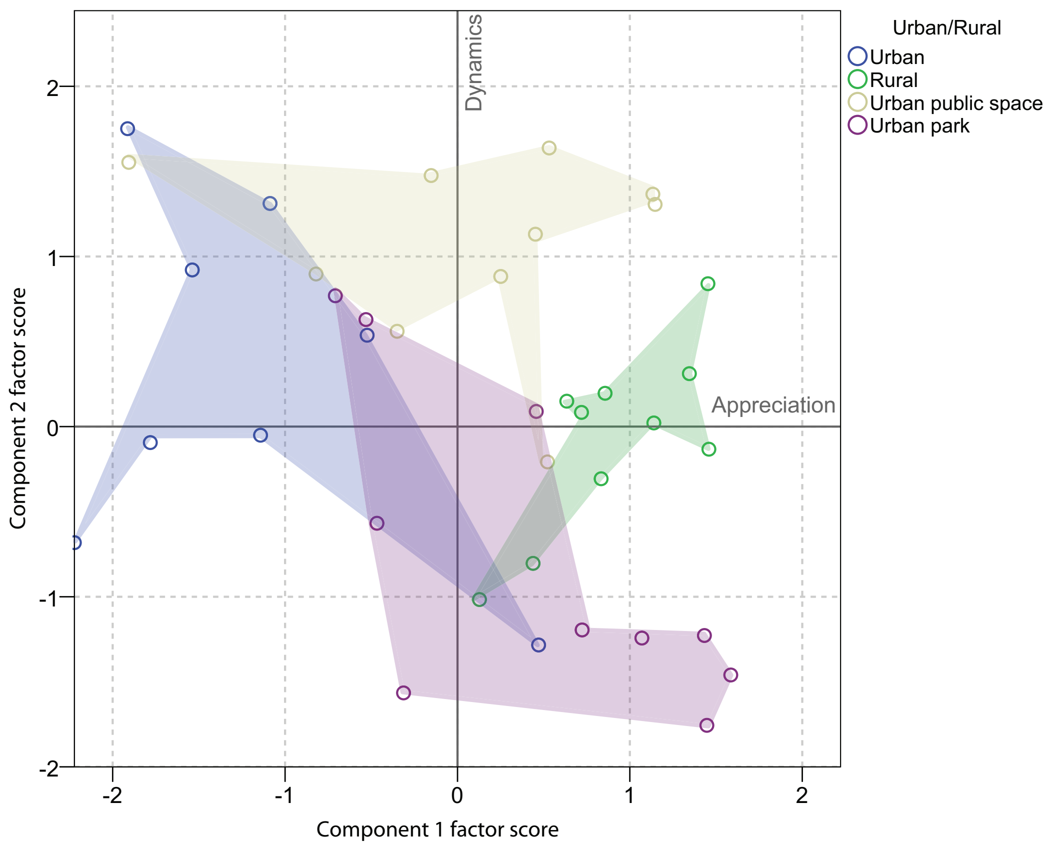

Figure 4.22: Subjective component scores showing: urban, urban public space and rural soundscapes

eventful as the individual sound sources could be more defined against the low background levels.

The manually extracted soundscapes of type: urban, rural, urban public space and urban park, revealed varying and distinct subjective opinions of each. Urban soundscapes were considered to have an inferior sound environment when compared to the rural group, with lower appreciation scores in general. The perceived dynamics of urban soundscapes showed a wide range of dynamics ratings, with rural soundscapes being perceived as less dynamic. The urban public spaces scores suggest that these places are met with mixed emotions, but are generally considered to be exciting and eventful spaces. The low scores for urban park dynamics when compared to rural soundscapes may be due to subjects rating urban parks as relative to the high dynamic nature of an urban soundscape, which you can assume they have just passed through to enter the urban park.

4.4.6 Small group study

The results from the small group study were compared with those from the main study, with Figure 4.23 showing the difference in ratings between the two locations and Table 4.13 detailing these mean values and significance of these differences.

Mean	Quality	Loc. quality	Pleasant	Exciting	Eventful	Tranquil
Urban roadside	3.43	4.14	2.93	4.64	6.50	2.21
Urban park	6.21	7.29	7.00	3.79	3.79	7.29
M-W U test						
U	20.5	6.5	2.0	69.0	17.5	2.0
Z	-3.606	-4.258	-4.468	-1.369	-3.778	-4.480
Asymp. Sig.	<.000	<.000	<.000	.171	<.000	<.000

Table 4.13: Mean ratings by location with Mann-Whitney U test

To compare subjective responses made between the two groups, the mean

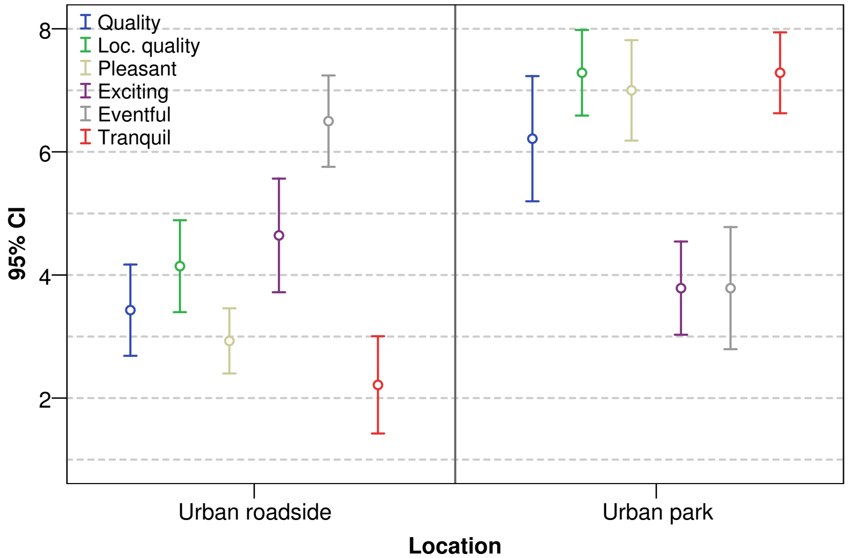

Figure 4.23: Subjective mean scores with 95% CI between locations

ratings and Mann-Whitney U tests are shown in Table 4.14. Mean ratings are provided to give a clearer indication of any differences present between groups.

Mean	Quality	Loc. quality	Pleasant	Exciting	Eventful	Tranquil
Small group	3.43	4.14	2.93	4.64	6.50	2.21
Random sample	3.60	3.50	3.60	4.50	5.90	3.70
M-W U test						
U	65.0	51.5	61.5	60.5	67.5	35.5
Z	-.301	-1.102	-.515	-.562	-.152	-2.091
Exact Sig.	.796	.285	.625	.585	.886	.042

Table 4.14: Urban roadside soundscape comparison

With the null hypothesis being that the two groups median values are equal, the clear non-significance of all tests but the rating of tranquillity show that there is significant evidence to accept the null hypothesis. The difference in tranquillity rating could be due to a number of factors, but the low number of subjects involved in the test may have produced this close to non-significant test result. This trend is also seen in the urban park comparison in Table 4.15, where there is significant evidence to assume equal median values for the two groups. In the case of the urban park location, all subjective variables show this significant average equality.

Mean	Quality	Loc. quality	Pleasant	Exciting	Eventful	Tranquil
Small group	6.21	7.29	7.00	3.79	3.79	7.29
Random sample	6.60	7.60	6.20	3.50	4.00	6.10
M-W U test						
U	57.5	53.0	59.5	54.5	64.0	52.0
Z	-.743	-1.040	-.631	-.931	-.363	-1.082
Exact Sig.	.472	.341	.546	.371	.752	.312

Table 4.15: Urban park soundscape comparison

The small group study served to further validate the studies large scale methodology. The supervised small group has shown significant agreement with the findings made from the random sample taken from the project dataset. Whilst the observed similarities does not fully validate the projects novel methodology, it does support it in terms of its potential to gather meaningful soundscape research data in light of its different methodology. A more robust validation would involve much larger numbers of subjects in the small group to improve the statistical robustness of the comparison.

4.4.7 Qualitative analysis: motivations

The question asked of participants: “Why did you choose to record this soundscape?” provided responses that were used in this analysis stage, with the aim of attempting to answer the main question: “What motivated a participant to record a particular soundscape?”, whilst also investigating:

- What people find interesting within a soundscape

- The positive/negative distribution
- If people focus on particular objects or activities within a soundscape
- If the social media revolution has impacted on these motivations

To begin with, each response was allocated to one or a number of categories:

- **Positive:** the response takes a generally positive angle
- **Negative:** the response takes a generally negative angle
- **Routine:** the soundscape forms part of a daily routine of sound environments
- **Activity:** the participant is involved in a particular activity
- **Focus:** a sound within the soundscape is the main focus
- **Misc:** no discernible theme in the response (excluded from analysis)

These categories are not mutually exclusive and a participant response can be grouped into any number of them.

Examples of responses from each category are shown below, where each response has only been assigned to one category:

- **Positive:** "*It complimented the warm afternoon sunshine perfectly*"
  The response is generally positive with no references to any other category
- **Negative:** "*It was annoying me!!*"
  This is generally negative

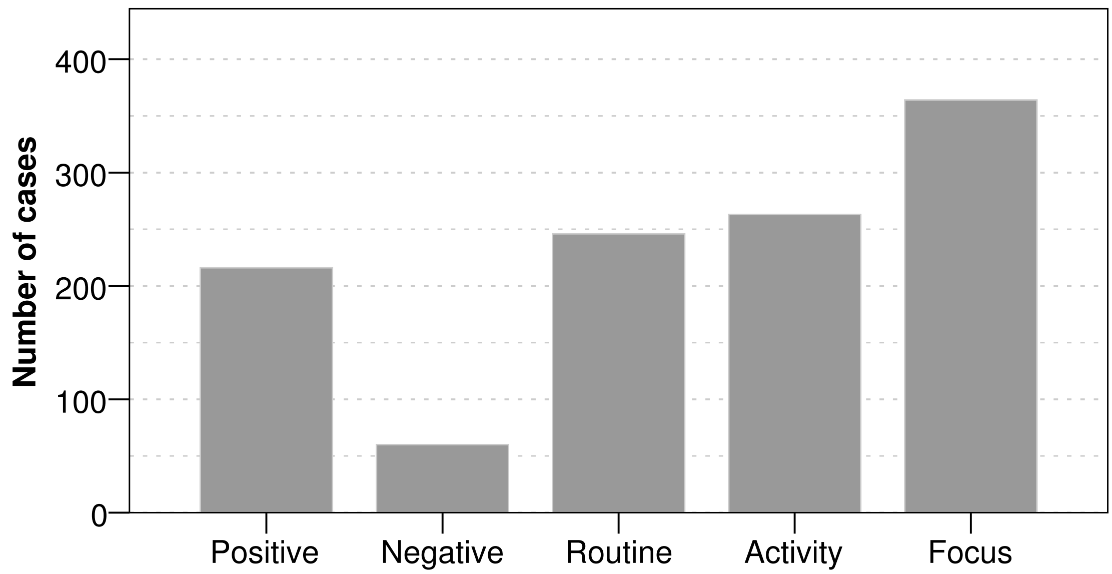

Figure 4.24: Response category counts

- **Routine:** "*The Moscow commuter's bread and butter soundscape*"
  This only refers to an everyday soundscape of routine

- **Activity:** "*To share the experience of sailing*"
  This explicitly refers to an activity the participant is involved in

- **Focus:** "*Wedding bells on a Saturday afternoon*"
  The focus is on a particular sound source

Based on Figure 4.24 it would seem that more soundscapes were captured because of positive rather than negative reasons. This suggests that participants were more inclined to share a positive aspect of their acoustic lives. The terms used in this solely positive response group are shown in Figure 4.25. They generally consist of positive adjectives used to describe the soundscape itself.

Figure 4.25: Solely positive response group terms with frequency weighting font size

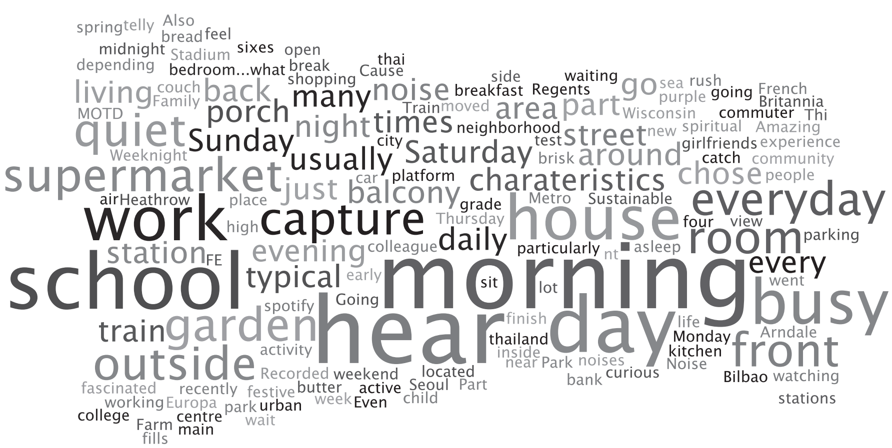

Figure 4.26: Solely routine response group terms with frequency weighting font size

The routine response terms in Figure 4.26 are mainly made up of words that describe a specific time of day and location that signify a part of a person's every-day soundscape.

Figure 4.27: Solely activity response group terms with frequency weighting font size

Figure 4.27 seems to mention specific places in reference to the activity that is being carried out.

Figure 4.28: Solely focus response group terms with frequency weighting font size

The focus group whose terms are shown in Figure 4.28 are mainly made up of references to individual sources such as: people, birds and traffic.

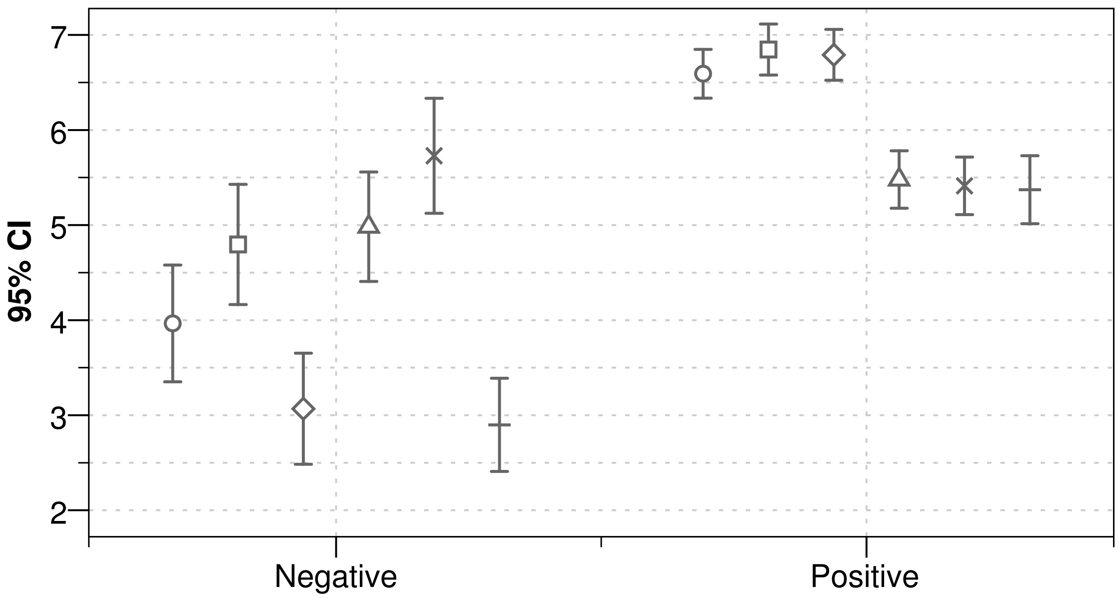

Figure 4.29: Subjective rating means with 95% CI between positive & negative responses (◯ = Quality, □ = Loc. quality, ◇ = Pleasant, △ = Exciting, X = Eventful, – = Tranquil)

Figure 4.29 shows the mean subjective rating scores with 95% CI bars for the groups containing postive/negative reasons for capturing. Significant differences are seen between all scale means apart from those of the Exciting and Eventful (see Table 4.16).

Table 4.17 shows the significantly different mean values for each opinion group. These groups are those which have only been assigned one of the categories: routine, activity or focus. Therefore the participant has given a specific reason involving only one of the categories as to why the soundscape was recorded.

Mean	Quality	Loc. quality	Pleasant	Exciting	Eventful	Tranquil
Negative (N=60)	3.97	4.80	3.07	4.98	5.73	2.90
Positive (N=215)	6.59	6.85	6.79	5.48	5.41	5.37
M-W U test						
U	2306	2983	1448	5062	5111	2586
Z	-7.083	-5.704	-8.817	-1.461	-1.362	-6.482
Exact Sig.	<.001	<.001	<.001	.144	.173	<.001

Table 4.16: Postive/negative response mean comparison with M-W U tests

Mean	Quality	Loc. quality	Pleasant	Exciting	Eventful	Tranquil	Pos. prom.	Neg. prom.
Routine (N=76)	5.48	5.56	5.48	4.44	4.59	4.68	5.32	5.12
Activity (N=72)	6.38	6.80	6.38	5.56	6.36	4.84	5.99	4.82
Focus (N=160)	5.79	6.07	5.82	5.43	5.39	5.13	6.20	5.09
K-W test								
χ^2	7.234	11.702	4.774	11.384	22.943	1.785	10.577	0.877
Asymp. Sig.	.027	.003	.092	.003	<.001	.410	.005	.645

Table 4.17: Routine, activity, focus only response mean comparison with K-W tests

Figure 4.30 shows these significant subjective response differences graphically.

The question asked of participants: “Why did you choose to record this soundscape?” provides an insight into the reasons why a person chose to take the time to record and comment on their sound environment. The reasons why a person chooses to actively engage with their sound environment (in the form of submitting a soundscape recording to the project) could provide information on what kind of sound environments evoke stronger responses from people and also gives an indication of the context in which the soundscape was captured in. The dominance of positive reasons behind submissions can be attributed to the interpretation of the project’s marketing approach. The use of the term soundscape, which may be new to some subjects, has been interpreted with a positive bias. If the project had been publicised as a noise survey for example, it could be assumed that submissions would be inherently biased towards the negative aspects of the sound

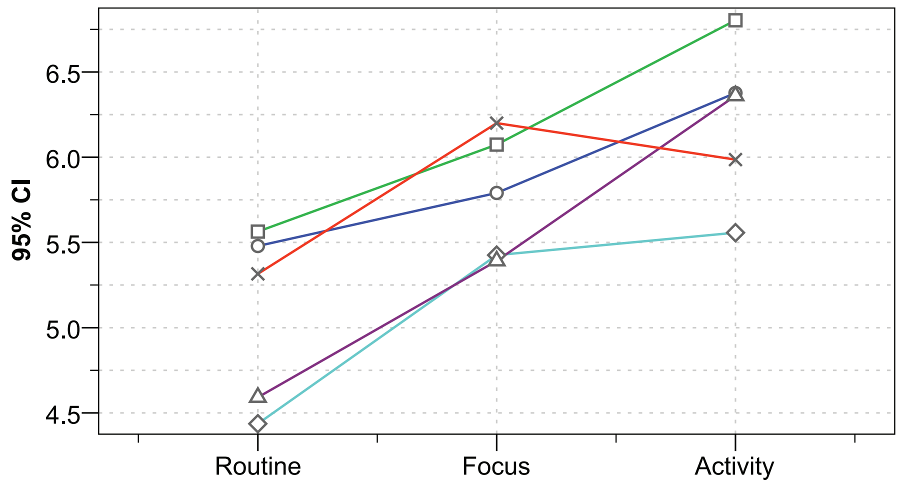

Figure 4.30: Significantly different subjective rating means between response groups (◯ = Quality, □ = Loc. quality, ◇ = Exciting, △ = Eventful, X = Pos. prom)

environment, as the use of the term noise infers negative connotations. Studies that make use of the term noise, are generally motivated by complaints, sleep disturbance and learning impediments, which may produce a negative bias in subjective response [157, 158].

The differences observed in the subjective ratings between positive and negative response groups mirrors those found in Section 4.4.2, where it was determined that ratings of how exciting and eventful a soundscape is were not largely influential in overall appreciation. Pleasantness ratings exhibit the largest mean difference between groups, suggesting that this plays an integral part in the overall appreciation of a soundscape. In fact, the differences observed between these means correspond to the factor loadings seen in Section 4.4.5, where quality, overall quality, pleasantness and tranquillity load heavily onto the Appreciation factor.

Quality ratings are seen to be at their lowest when a person is involved in some kind of a routine activity, such as travelling to work. The highest ratings of quality are observed when a subject is involved in some kind of focussed activity. The overall location quality shows a more marked increase over the soundscape quality with a high rating in the activity group, presumably due to the persons choice in undertaking this activity in a location that they also rate highly. Ratings of how exciting a soundscape is show a lower value when the subject has indicated that they have submitted whilst taking part in some kind of daily routine. This lower rating could be attributed to the fact that subjects will have been subjected to these soundscapes regularly and it therefore does not induce feelings of excitement as perhaps it once did. Eventfulness ratings show higher ratings within the activity group, and also the most variation between groups. It would seem that submissions involving a specific activity are considered more eventful in terms of their soundscape. The positive source prominence within a soundscape shows a higher rating when there is a focus on a particular sound source mentioned in the reasons for recording. This suggests that the identified sound source is predominantly considered a positive sound with a high prominence within the soundscape.

The linear increase in subjective ratings seen in Figure 4.30, as a persons motivations for capturing change could be considered as an increase in “interest”. The control over the soundscape in a routine situation could be considered less than a situation with a specific sound focus and when a participant is involved in an activity. Location quality, soundscape quality and eventfulness ratings show a significant, roughly linear increase as this “interest” scale is traversed. There is undoubtedly a certain element of interest present in each submission, as the participant has chosen to capture the soundscape and submit it, but a measure of this level of interest may provide an insight into which aspects of a soundscape invoke these feelings. Asking the question: “How interesting do you find this soundscape?”, alongside the existing question set could provide the basis for an ordinal scale of interest, which could be assessed, based on the responses made to the project's existing set of semantic differential scales.

4.5 Objective analysis

From the acoustic analysis of the project's soundscapes submissions, 30 metrics were computed. These explained the acoustic features of the soundscape recordings, in terms of their temporal, spectral and statistical features. Uncalibrated statistical measures of Zwicker loudness, RMS level and dBFS were also calculated for use in this analysis. These uncalibrated measures are considered relative to other recordings within the dataset, where the calibration issues concerning these devices has been discussed in Section 3.4.2.

4.5.1 Soundscape length

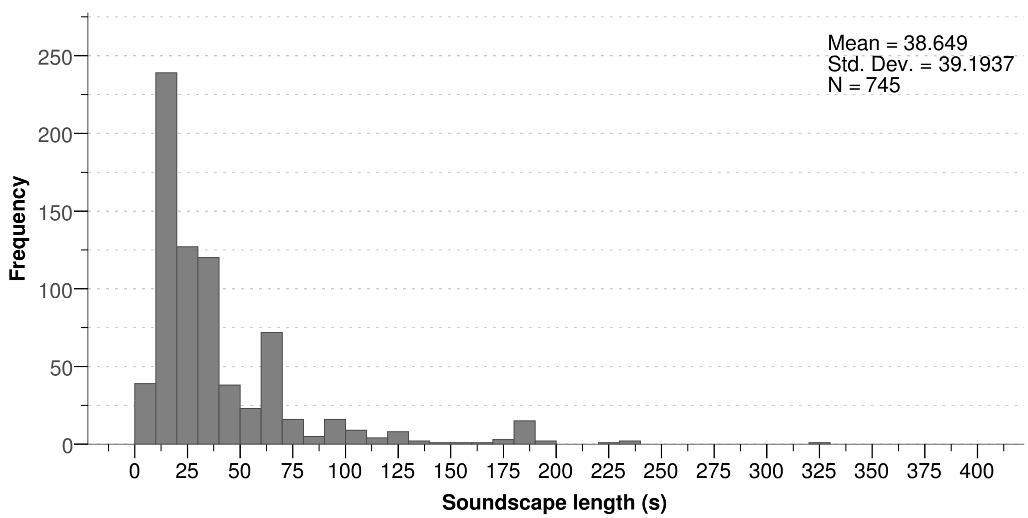

Figure 4.31: Soundscape length distribution

The mean length of soundscape submission seen in Figure 4.31 is 38.6 seconds. A small number of longer length submissions skew this mean with the longest submitted recording at 322 seconds. The iOS app limits recordings to 180 seconds, however, the web interface only limits submissions that exceed 20MB in size, allowing for longer soundscape uploads. The length of submission should have no affect on the objective analysis due to the predominant use of feature averages.

4.5.2 Descriptives

The objective variable descriptives are listed in Table 4.18.

4.5.3 Relative soundscape level

The unsupervised nature of these soundscape submissions leads to the assumption of high variations in relative level between submissions due to participant's holding their device in different ways. The rotation of the app interface and provided instructions mean that the microphone should at least be unobstructed by the hand, however, the distance from the participant's body and the direction in which the phone is pointing will have an affect on record level. The large numbers of submissions should serve to reduce the affect that this inherent error has on further inferences.

Variations in record level will also be present between devices. With no log of device type from web submissions, it is impossible to quantify this variation. iOS submissions will be more consistent due to their similarity in overall level response between devices as seen in Section 3.5.2. The level response limitations of the

Level	N_{max}	N_1	N_5	N_{10}	N_{50}	N_{90}	N_{95}	L_1	L_{10}	L_{50}	L_{90}	RMS	L_{AeqT}
Mean	−17.6	−20.4	−22.9	−24.2	−29.1	−34.5	−42.5	−25.2	−32.4	−43.0	−59.9	−31.8	−41.5
SD	11.4	12.2	13.0	13.3	14.4	16.7	19.9	13.4	13.8	13.8	13.9	13.1	13.2
Min	−58.1	−70.8	−76.0	−77.5	−82.2	−97.0	−97.0	−69.0	−74.2	−82.4	−100.9	−71.0	−83.3
Max	0.9	−0.6	−2.1	−2.5	−4.3	−6.0	−6.6	2.2	−2.6	−13.5	−29.3	−4.8	−13.9

Spectral	Centroid (Hz)	Brightness	Spread (Hz)	Roll off 95% (Hz)	Roll off 85% (Hz)	Entropy	Flatness	Irregularity
Mean	2669	0.44	106	8959	5437	0.78	0.10	0.34
SD	1286	0.17	100	2830	2731	0.07	0.04	0.18
Min	722	0.07	0	2187	490	0.46	0.03	0.00
Max	10 744	0.98	823	14 542	13 680	0.92	0.20	1.42

Spectral	MFCC	Zero cross (N/s)	L_{CeqT}-L_{AeqT} (dB)	Roughness (asper)
Mean	0.84	1486	3.33	0.06
SD	0.74	1215	3.03	0.14
Min	−2.56	50	−1.88	0.00
Max	3.33	10 530	18.10	1.81

Dynamics	L_{A10}-L_{A90} (dB)	Low energy (×100%)	Crest factor (dB)	Spectral flux	Event density (N/s)
Mean	27.5	0.58	21.6	660	2.32
SD	5.1	0.08	6.8	151	1.89
Min	21.2	0.22	5.0	6	0.10
Max	69.2	0.87	48.9	989	10.58

Table 4.18: Objective variable descriptives (Loudness N metrics in sones, level L metrics in dBFS, remaining labelled or dimensionless)

iOS devices tested in Section 3.5.2 will mean that the majority of soundscapes will have a limited dynamic range due to their noise floor and upper recording limit. The automatic gain control could act to reduce overall level measures when large numbers of impulsive events are captured at levels which exceed the device's upper limits.

	Quality	Loc. quality	Pleasant	Exciting	Eventful	Tranquil	Appreciation	Dynamics
N_{max}	−.11**	−.07	−.09*	.17**	.15**	−.17**	−.13**	.20**
N_1	−.11**	−.06	−.08*	.19**	.17**	−.17**	−.12**	.21**
N_5	−.11**	−.06	−.08	.19**	.19**	−.17**	−.12**	.22**
N_{10}	−.10*	−.05	−.08	.20**	.19**	−.17**	−.11**	.23**
N_{50}	−.08*	−.03	−.06	.20**	.19**	−.14**	−.09**	.22**
N_{90}	−.09*	−.04	−.07	.19**	.20**	−.13**	−.09*	.21**
N_{95}	−.05	−.02	−.07	.10*	.09*	−.15**	−.08*	.13**
L_1	−.10*	−.04	−.07	.20**	.18**	−.16**	−.10*	.22**
L_{10}	−.08*	−.03	−.05	.21**	.19**	−.15**	−.09*	.23**
L_{50}	−.08*	−.02	−.05	.21**	.19**	−.14**	−.08*	.23**
L_{90}	−.08	−.02	−.05	.21**	.20**	−.15**	−.08	.24**
RMS	−.07	−.03	−.06	.21**	.20**	−.13**	−.08	.23**
L_{AeqT}	−.09*	−.04	−.06	.20**	.19**	−.16**	−.10*	.23**

Note: *p<.05, **p<.01

Table 4.19: Spearman's Rho correlation coefficients between subjective responses and level metrics (N_x = Zwicker's relative percentile loudness and L_x = relative percentile level)

The Spearman's rho correlation coefficients are shown in Table 4.19, where values without an accompanying asterisk are non-significant at the 5% level. Ratings of quality show only weak correlations with any of the level features of the submitted soundscapes. The more significant negative correlations are observed between the higher level Zwicker loudness levels exceeded <5% of the time suggesting that an excess of louder level sound events within a soundscape has some effect in reducing overall soundscape quality judgements. Ratings of excitement and eventfulness show stronger positive correlations with loudness measures. This perception of the dynamics of the soundscape as discussed in Section 4.4.5, seems to have a weak but significant relationship with soundscape metrics.

The relatively weak but highly significant correlations between the subjective and objective variables reveal that an excess of louder level sound events within a soundscape has some effect in reducing overall soundscape quality judgements. Based on this, an abundance of more impulsive sound sources can lead to reduced perceptions of soundscape quality and tranquillity. In the extreme case, the example of being in close proximity to a pneumatic drill would lead to high N_1 and N_5 values and therefore low values of perceived soundscape quality and tranquillity. The increase in perceived soundscape dynamics with increases in overall level reveal the relationship between relative level and soundscape activity and energy. Again, impulsive elements within soundscapes are showing increased activity perceptions. Whilst these coefficients of correlation are generally weak, they do show some agreement with past research [159]. This study observed negative correlations between soundscape quality and measures of statistical sound pressure level (L_{A50} -.55, L_{A95} -.53) and Zwicker loudness (Overall loudness N - .53). This large difference in explained variance could be attributed to the fact that this study was performed in 16 urban park type locations with structured stop and question style interviews using ten minute calibrated acoustic measurements. The present studies varied mix of participant chosen soundscape types and the identified limitations of the recording platform could be a reason for the reduction in explained variances including the possibility that certain relationships between the objective and subjective are being completely obscured by these factors.

4.5.4 Spectral features

The spectral features extracted will be affected by the iOS devices high pass filter. The main influence this will have is the introduction of a high frequency bias on the majority of these metrics. Soundscape containing a large amount of low frequency energy below 200Hz will not be represented accurately. With the high pass filtering of all submissions prior to analysis, discussed in Section 3.5.2 the recordings themselves, (discounting issues around device high frequency response and user handling error) have been captured on a relatively similar platform in terms of its reaction to spectral content.

The non parametric correlation coefficients seen in Table 4.20, show the metrics describing the spectral features of the objective soundscape dataset and their correlations with the gathered subjective ratings. No significant correlations of note are observed between soundscape quality and any spectral features. Location quality shows highly significant negative correlations with Roll off 85% (high frequency energy) and Entropy (frequency distribution feature). Stronger correlations are observed with the dynamic subjective rating of the soundscape. A significant negative correlation is observed between perceived dynamics and Roll off X% suggesting that as the high frequency content of the soundscape increases, a decrease in perceived soundscape dynamics is recorded. Spectral flatness and entropy also show negative correlations with perceived dynamics.

The lack of any relationship between the spectral features and ratings of soundscape quality, pleasantness and tranquillity suggest that these features have no effect on appreciation. Ratings of location quality however, show correlations between roll-off 85% and spectral entropy. This finding seems to suggest that a reduction in high frequency content can result in improvements in overall loca-

	Quality	Loc. quality	Pleasant	Exciting	Eventful	Tranquil	Appreciation	Dynamics
Centroid	−.06	−.10*	−.07	−.13**	−.17**	−.04	−.09*	−.15**
Brightness	−.06	−.09*	−.05	−.09*	−.12**	−.06	−.08	−.10*
Spread	−.06	−.06	−.02	.11*	.14**	−.03	−.05	.12**
Roll off 95%	−.03	−.09*	−.06	−.15**	−.18**	.01	−.06	−.17**
Roll off 85%	−.06	−.12**	−.07	−.16**	−.19**	−.04	−.09*	−.17**
Entropy	−.08	−.11**	−.08*	−.13**	−.15**	−.07	−.11**	−.14**
Flatness	−.03	−.11*	−.07	−.16**	−.18**	−.02	−.08	−.17**
Irregularity	.01	.03	.02	.15**	.14**	−.04	.01	.17**
MFCC	−.00	−.09*	−.02	−.09*	−.04	.01	−.03	−.07
Zero cross	−.06	−.07	−.04	−.09*	−.14**	−.06	−.08	−.11**
L_{CeqT}-L_{AeqT}	.10*	.06	.06	.02	.08	.11**	.10*	.03
Roughness	−.04	.00	−.01	.13**	.12**	−.05	−.02	.13**

Note: *p<.05, **p<.01

Table 4.20: Spearman's Rho correlation coefficients between subjective responses and spectral metrics

tion appraisal. It also appears that a spectrally rich soundscape may result in the location itself being rated lower in quality.

Increases in measures of a soundscapes high frequency content have revealed reductions in their perceived dynamics. Increased flatness and entropy values indicate a flat frequency response tending towards white noise. Based on this, a soundscape that is more spectrally random in nature will be prone to be perceived as less dynamic in nature. Soundscapes exhibiting spectra with defined peaks will tend to be perceived as being more eventful and exciting. The positive correlation between spectral spread and the perceived dynamic metric suggest that soundscapes that are spectrally "richer" can be perceived as being more dynamic. Negative correlations between perceived dynamics and the measures of high frequency content such as centroid, brightness, zero cross and roll off, indicate that a soundscape which has more low frequency energy is perceived to be more dynamic. This could be attributed to low frequency traffic noise which was mentioned as a source within a large number of the submitted urban soundscapes.

4.5.5 Soundscape dynamics

The objective dynamics of a soundscape show limited correlations with subjective variables, seen in Table 4.21. Subjective appreciation seems to have no links to these dynamic measures, however a weak but significant negative correlation is observed between ratings of subjective dynamics and crest factor.

	Quality	Loc. quality	Pleasant	Exciting	Eventful	Tranquil	Appreciation	Dynamics
L_{A10}-L_{A90}	−.05	−.04	−.01	.02	.01	−.02	−.04	.01
Low energy	−.01	−.02	.02	−.09*	−.08	.02	.01	−.09*
Crest factor	−.01	.00	.00	−.16**	−.14**	.06	.01	−.17**
Spectral flux	.00	−.00	−.05	−.06	−.04	−.03	−.04	−.04
Event density	.02	.00	.04	.06	−.00	−.03	.02	.05

Note: *p<.05, **p<.01

Table 4.21: Spearman's Rho correlation coefficients between subjective responses and dynamics metrics

The indication of a negative correlation between crest factor and perceived dynamics is contrary to the assumption that a signal with a high crest factor value will be made up of impulsive sources. One study however, showed a positive correlation between crest factor and annoyance from road and railway sources [160]. This finding does correlate to the findings of the current study, as ratings of excitement have positive connotations, whereas annoyance is a negative response.

4.5.6 Manual soundscape categorisation

Using the categories defined in Section 4.4.4, objective metric averages between urban and rural soundscapes were compared. Mean values and significant Mann–Whitney U statistics are given in Table 4.22.

The rather small differences in RMS and entropy, whilst significant, do not reveal any salient differences between these two soundscape types. The spectral

Mean	RMS	Centroid	Brightness	Entropy	Zero cross	L_{A10}-L_{A90}	L_{CeqT}-L_{AeqT}
Urban	0.842	2579	0.424	0.778	1383	27.4	3.49
Rural	0.770	2989	0.494	0.793	1851	29.1	2.82
Δ	0.072	410	0.070	0.015	468	1.7	0.67
M-W U test							
U	13658	13765	13729	14460	14017	14676	14055
Z	-2.825	-2.735	-2.765	-2.150	-2.523	-1.968	-2.491
Exact Sig.	.005	.006	.006	.032	.012	.049	.013

Table 4.22: Soundscape category means and Mann-Whitney U statistics for selected significant objective metrics

features however do reveal a distinct difference between them. Firstly the urban group seems to contain more low frequency energy as indicated by the lower values of spectral centroid, brightness, zero-cross rate and L_{CeqT}-L_{AeqT}.

The objective differences in high frequency content seen between urban and rural soundscapes does make intuitive sense, as the majority of urban submissions contained mentions of traffic or transport noise, explaining these higher levels of low frequency energy. Also worth noting is the difference in L_{A10}-L_{A90}, used to describe soundscape variability. The perceived dynamic nature of urban and rural soundscapes as discussed in Section 4.4.5 follows the same trend as this outcome, where rural soundscapes are perceived as less dynamic than urban soundscapes.

4.5.7 Principal component analysis

All objective metrics were initially run through a Varimax rotated principal component analysis in a bid to reduce the high dimensionality of the data. The aim of this process was to find a structure in the acoustic setup of these varying soundscapes. Each variable was normalised prior to analysis by taking its Z transform, to eliminate the influence of the differing variable scales [161]. Variables with loadings <0.2 on only one component were removed and the analysis was repeated in a bid to aid component identification and increase the explained variance of these factors.

Two principal components were extracted with an eigenvalue greater than 1 with the inflection point shown in Figure 4.32.a, consisting of 22 objective variables. A value for factor loadings of .40 was used as the cut-off to identify items which loaded onto each factor [162]. These factors collectively explained 86% of variance between variables. Component 1 explains 65% of variance, component 2 explains 21%.

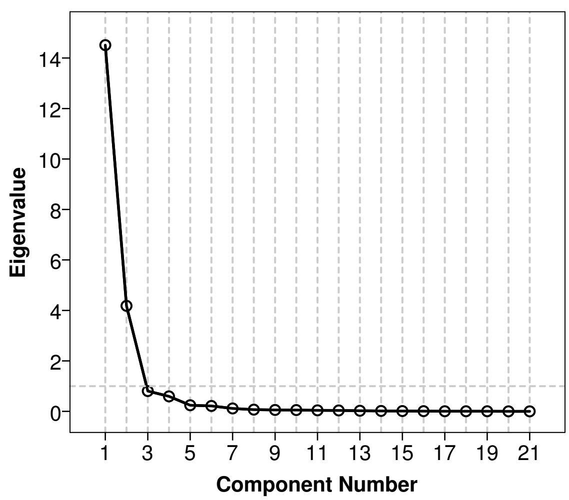

Figure 4.32.a: Scree plot

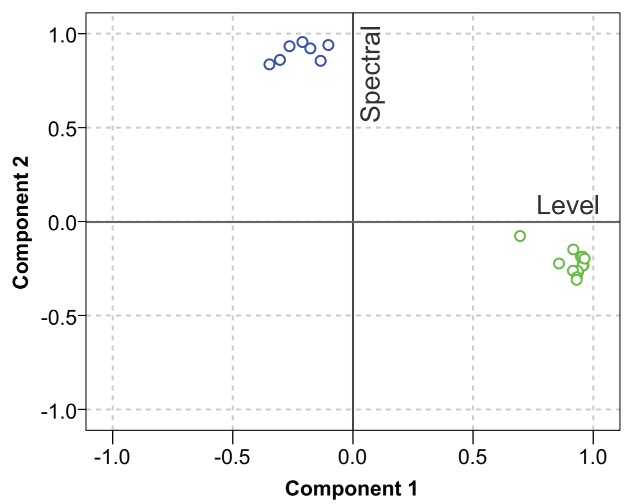

Figure 4.32.b: Rotated factor plot loadings

Figure 4.32: Scree plot and rotated factor plot for objective metric variables (right cluster = level, top cluster = spectral)

Component 1 consists of statistical and overall average level measures. Component 2 is made up of a soundscape's spectral features, where the majority of included variables describe a signals high frequency content. The plot in Figure 4.32.b illustrates how the variables load onto each factor, where the middle right cluster are made up of statistical and overall level measures and the middle top consists of the spectral features, with the actual rotated component loading values shown in Table 4.23.

Variable	Component 1	Variable	Component 2
N_{max}	.918	Centroid	.958
N_1	.950	Brightness	.940
N_5	.959	Roll off 95%	.861
N_{10}	.961	Roll off 85%	.934
N_{50}	.937	Entropy	.922
N_{90}	.854	Flatness	.835
N_{95}	.686	Zero cross	.855
L_1	.956		
L_{10}	.959		
L_{50}	.936		
L_{90}	.915		
RMS	.932		
L_{Aeq}	.965		
L_{Ceq}	.931		

Table 4.23: Varimax rotated component matrix for objective metrics (loadings $< |0.4|$ removed)

An increase in the value of component 2 factor score translates to an increase in high frequency content, based on the loadings of: centroid, brightness, roll off & zero cross. This component also includes descriptions of a signals spectral distribution, where high values of entropy and flatness equate to spectral distributions that tend towards random white noise. Therefore an increase to this components factor score can also be considered as an increase in the signals complexity in terms of its frequency content. Distinctive sources within the soundscape such as speech with its clear formant structure would serve to reduce the value of this factor score due to its effect on entropy and flatness.

Firstly, these extracted objective component scores were compared to the previously extracted subjective component scores from Section 4.4.5. Spearman's Rho correlation coefficients were calculated between each set of objective and subjective component scores, resulting in these findings:

- A significant but weak correlation was observed between soundscape appreciation and soundscape level (Spearman's Rho = -.115 p=.005). A finding which suggests that the overall level of a soundscape does have a negative influence on soundscape appreciation.

- Perceived soundscape dynamics were correlated with soundscape level (Spearman's Rho = .217 p=<.001), indicating that a persons appraisal of a soundscapes dynamic nature is positively related to its overall level.

- A soundscape's spectral features are correlated to perceived dynamics (Spearman's Rho = -.126 p=.002). This shows a weak but significant relationship between a soundscapes perceived dynamics and its high frequency content and spectral complexity.

The two extracted objective principal components describe the overall level and spectral complexity of a soundscape recording. The investigation into any links between these objective and subjective components revealed that the overall level of a soundscape does have a negative influence on soundscape appreciation. A persons appraisal of a soundscapes dynamic nature was also seen to be positively related to soundscape level. Also, it would seem that an increase in a soundscapes high frequency content and spectral complexity can result in a decrease in its perceived dynamics. Therefore, a loud soundscape with low variation and levels of high frequency content will generally be perceived as lower in quality, but high in its dynamic nature.

In a bid to investigate further into these subjective/objective interactions, all of the project's subjective scale responses were dichotomised in the following way:

- **Low group:** ratings including values of 1 through 5
- **High group:** ratings including values of 6 through 9

Taking this reductionist approach to ordinal data does have it drawbacks, many of which are detailed in [163]. However, the large variations in responses observed with this kind of subjective environmental data warrants a more generalised approach to the data analysis. The delineation point was chosen because of the odd number of scale items and the general negative skew in the variables distributions meaning larger N values for higher ratings. Each variable was grouped in the same way to aid in the interpretation of the results. Table 4.24 shows the mean scores by subjective variable group for the objective level component. The higher level score in the low soundscape quality group reaffirms the influence that level has on soundscape appreciation seen in Section 4.5.3. Perceived soundscape dynamic measures of excitement and eventfulness show higher objective level values when in the high rating category. Objective level scores show the largest difference between tranquillity rating groups, suggesting that this component has the most influence over the perceived tranquillity of a soundscape.

Table 4.25 uses the same technique as in Table 4.24, but uses the objective spectral component scores. In this case, the only significant differences are observed between the excitement and eventfulness groups a trend seen previously in Section 4.5.4. This outcome suggests that a soundscape with higher levels of high frequency energy can lead to reduced ratings of a soundscape's dynamics.

By dichotomising the subjective rating variables, the effect of level on soundscape appreciation is once again seen, with higher ratings of quality in lower level soundscapes. High ratings of excitement and eventfulness are also present in soundscapes with higher overall levels. Tranquillity also shows a marked domi-

	Group	Mean score	Δ	U	Z	Sig.
Quality	Low	0.14	−0.21	38071	-2.404	.016
	High	−0.07				
Loc. quality	Low	0.10	−0.14	37767	-1.495	.135
	High	−0.04				
Pleasantness	Low	0.04	−0.06	43082	-.668	.504
	High	−0.02				
Excitement	Low	−0.11	0.24	38250	-3.397	.001
	High	0.13				
Eventfulness	Low	−0.12	0.25	37790	-3.638	<.001
	High	0.13				
Tranquillity	Low	0.14	−0.34	33987	-4.393	<.001
	High	−0.20				

Table 4.24: Level component mean scores by subjective response group with M-W U tests

	Group	Mean score	Δ	U	Z	Sig.
Quality	Low	0.02	−0.10	39669	-1.637	.102
	High	−0.08				
Loc. quality	Low	0.06	−0.16	36977	-1.885	.059
	High	−0.09				
Pleasantness	Low	0.02	−0.10	41474	-1.427	.153
	High	−0.08				
Excitement	Low	0.08	−0.24	38631	-3.220	.001
	High	−0.17				
Eventfulness	Low	0.08	−0.23	38999	-3.074	.002
	High	−0.15				
Tranquillity	Low	−0.04	−0.01	42494	-.315	.753
	High	−0.05				

Table 4.25: Spectral component mean scores by subjective response group with M-W U tests

nance of low ratings when rated in soundscapes with a higher overall level, once again confirming findings made in Section 4.5.3. The perceived dynamics variables of eventfulness and excitement show lower ratings in conditions with increased levels of high frequency content. Again, this outcome have been previously identified in Section 4.5.5.

4.5.8 Prominent source extraction

Considering the comparatively weak associations between objective and subjective variables when considering the soundscape as a whole, a more focussed study into the positive sources within a soundscape was carried out. 20 soundscapes were selected from the main dataset which contained an identified positive sound source with a prominence rating of >7. These sources were then isolated in the frequency and time domain as shown in Figure 4.33 and extracted to create a sound source file, with a mean length of 3.2 seconds across all 20 extractions.

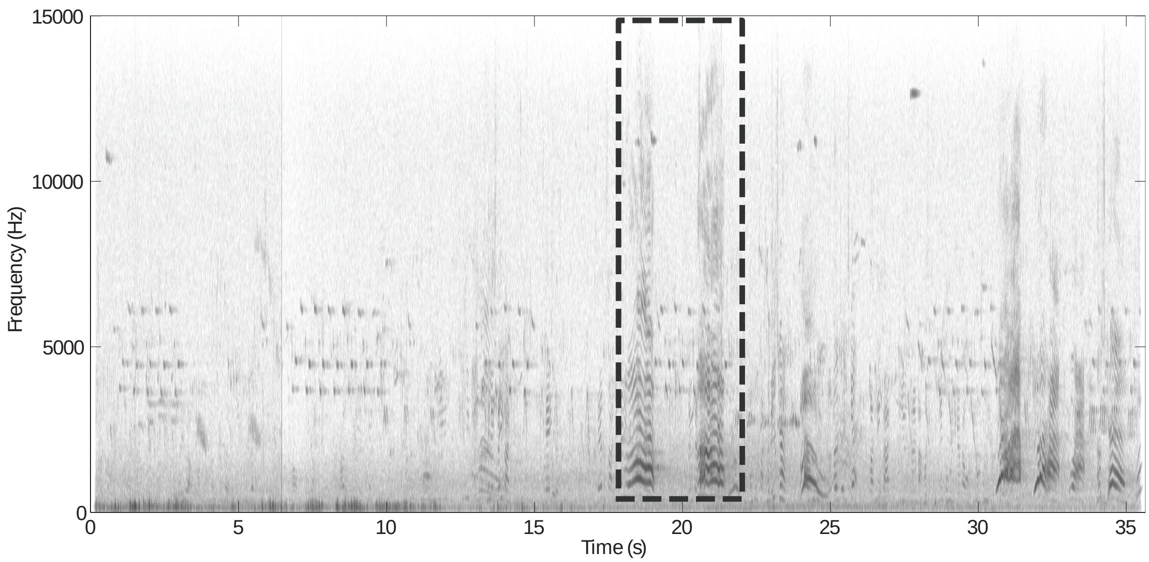

Figure 4.33: Sound source extraction

The extraction and filtering process was achieved using Adobe Audition 3.0 audio software. The resultant sound source files maintained the same sample rate of 44.1kHz. The source types extracted were: animals, bells, voices and live instruments. Sources were chosen which inhabit a distinct region in the time and frequency domain to make extraction possible and ensure that the predominant source within the extracted segment is the one which is under scrutiny. A commonly identified positive source such as flowing water inhabits a broad range in frequency and time so was therefore not easily extracted and not suitable in this stage of analysis.

These isolated sources were then subjected to the same objective acoustic analysis as was performed on the entire soundscapes in Section 4.5. The most prominent relationship observed was between tranquillity ratings and spectral flatness, the variable which describes the sources frequency distribution by calculating the ratio between the geometric and arithmetic means. A significant negative correlation was observed between the two variables (Spearman's Rho = -.51 p=.02). It would seem that soundscapes containing a prominent sound source with a more varied frequency distribution are considered less tranquil. Another pronounced relationship is seen between ratings of excitement and spectral entropy (Spearman's Rho=.48 p=.03). A signal made up of a single pure tone will produce an entropy value tending towards 0, whereas random noise produces a value that tends towards 1. This can be interpreted as the spectral complexity of the sound source, which in this case is showing a relationship with how exciting a soundscape is perceived to be.

The extraction of prominent sources within a selection of submitted soundscapes was a technique to investigate the soundscape in an object orientated way. This means making the assumption that a soundscape is simply a collection of individual sources, with each having an effect on a persons perception of the entire soundscape. Extracting the positive prominent source identified by a participant is an attempt to gauge if it really is these individual sources that determine a persons response to their sound environment. From this analysis, it would seem that ratings of tranquillity are reduced when a positive source that is more

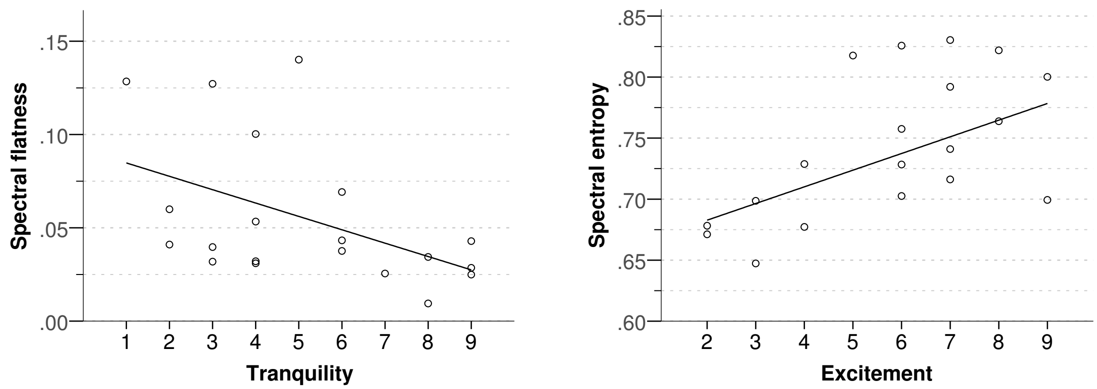

Figure 4.34: Extracted positive source objective metrics vs. subjective soundscape ratings (N=20)

spectrally rich is prominent within the sound environment. This outcome suggests that sources exhibiting distinct spectral peaks can have a detrimental effect on perceived tranquillity, e.g. an air conditioning unit with a number of spectral harmonics due to fan rotation will result in increased spectral flatness readings and therefore a decrease in perceived tranquillity. Excitement sees a positive relationship with spectral entropy, therefore a sound source with white noise-like qualities such as running water, will generally be perceived as more exciting.

4.6 Summary

The analysis and inferences made from the project's dataset have been presented. Validation of the various outcomes has been found from a number of studies utilising differing methodologies. A number of highly significant subjective and objective correlations have been observed, revealing the influences of a soundscape's level and spectral content on a person's perception of that soundscape. Also, the activity a participant is involved in, the type of sound sources within their soundscapes and the reasons behind them submitting all have an influence on the appraisal of their sound environment. Chapter 5 continues to discuss the methodological approach in more detail.

Chapter 5

Methodology discussion

5.1 Overview

The participation routes provided to participants have shown varied popularity. The global success of the iOS application shows this routes accessibility and ease of use. The public engagement and publicity aspects of the project have been instrumental in gaining a large dataset as seen by the submission increases as a result of media coverage. The retention values observed for the project application also compare to the averages observed from large scale international studies. Gender and age biases have also been observed in the project's dataset, however, no differences in response between gender and age groups was evident.

The objective and subjective complexity of the submitted sound environments make adequate feature extraction a difficult process. With a large combination of sound sources contained in every soundscape, each with varying levels of perceptual saliency, the difficulty lies in determining which sound sources are influential in a persons response to a soundscape. The subjects identification of a positive sound and its associated prominence has helped to uncover one of these influential foreground sounds, but the influence this has on overall soundscape perception will still be skewed or even masked completely by the remaining sounds contributions. The expected effect this will have is to reduce the explained variance possible using objective metrics when relating to subjective responses. To fully validate these theories, a process of subjective testing would be required with individual sources, but this lies outside of the scope of this study.

5.2 Summary of methods

The mixed methods approach taken has enabled the retrieval and analysis of quantitative and qualitative data from a large number of participants. The use of semantic differential ordinal scales has produced a rich dataset of quantifiable responses of humans to their sound environment. The qualitative data has afforded a more contextual insight into the reasons a person has chosen to contribute a soundscape to the project. The soundscape recordings themselves have provided a timely snapshot of the acoustic environment at the time when the participant has also responded to it.

The involvement of members of the public to gather soundscape research data has revealed a number of advantages over existing studies. Firstly, the number of submissions obtained is relatively high, something which a project of this type requires due to the amount of statistical analysis involved. The complex interactions humans have with their sound environments has been difficult to confidently gauge using a citizen science methodology. The limitations imposed by the voluntary nature of participation has meant that the questions asked of participants have to be kept to a minimum. The percentage of variance explained in the correlational analysis was significantly lower than more traditional soundscape studies, however, the statistical significance was generally very high. It would seem that the major trade-off that this methodology has is in the blurring of these discovered relationships, when compared to more traditional research methodologies. A number of similar findings have been found, but with a notable drop in explanatory power. The main cause for this is assumed to be the wide range of submitted soundscapes. Where existing studies investigate a single location or location type, such as “urban parks”, this study has investigated soundscapes ranging from suburban gardens to deserts. To reduce the influence of this factor, categorisation of soundscapes and separate sets of analysis on each category would be preferable, to enhance the explanatory power of the statistical methods used. The focus on a particular pre chosen location could be implemented using the technique of geo-fencing, discussed in Section 6.4.1.

The citizen science style survey has proven its suitability for soundscape research. With humans being the central factor within soundscapes, it makes sense to involve the user in more depth than simply questioning them opportunistically. The fact that people had the choice of what to capture has revealed the soundscapes that represent their lives or have affected them in a positive or negative way. Throughout the process of participation, the subject is aware that their submissions will be immediately accessible on the project website once uploaded. This sonic snapshot of their daily lives can be compared to a submission someone might make to a social network, such as the upload of an image with an associated caption. The social networking share features of the project app and website actually make this a reality, where participants can choose to automatically post a link to their submission on Twitter or Facebook as it is uploaded. The large numbers of people submitting soundscapes that represent aspects of their daily lives has shown that people are using this as a means to share their sonic lives. A large number of tweets and emails have been received from participants saying that they have used the project's technologies to allow them to add a sonic commentary to their social media presence.

5.2.1 Recruitment

The importance of public attention for a project of this nature cannot be over emphasised. Without a process of proactive public engagement, the project will be rarely discovered online. The relatively high level of participation required of participants mean that to ensure a large number of submissions, a very large number of potential participants must be aware of the project, assuming that a percentage of people aware of the project are deterred from taking part because of the effort involved.

The recruitment methods used proved successful in driving people to take part in the study. The inclusion of a single web portal to access the project's resources was important when it came to recruitment. A single URL could be provided to media outlets and other websites which allowed potential participants to access information about the project, its submissions to date and the ways to get involved. It became clear very early on that the accessibility of the web portal was paramount in ensuring people took the steps necessary to participate. The information provided to people on the website was also deliberately vague it defining what sort of soundscapes the project required as submissions. This provided the wide range of soundscape types eventually contributed.

5.2.2 Participation paths

The success of the iOS application is clearly down to its fast installation process, ease of use and upload from device functionality. The J2ME application hindered the process of participation through the additional stages required. Voluntary participation can immediately cease to happen if a person is frustrated by the actions required of them. The web interface has also proved very popular for participants who do not have an Apple device and professional sound recordists. The retrospective responses made by people using the web method has shown no differences to those made in-situ using the project app, which warrants the combined analysis of these two submission types.

5.2.3 Participant choice

Allowing the participant to choose what soundscape to capture raises a number of issues in the analysis of the projects dataset. Studies that focus on specific locations such as urban parks, gather data from subjects that is specific to that location and time. The response to a single location by a large number of people provides an average indication of how that location is perceived by its inhabitants. The uncertainty of this indication is produced by the variation in human perceptions of the place. The present study incorporates singular submissions from a wide range of locations and times, meaning that not only does it contain a level of uncertainty inherent in human perception, it also has the uncertainty stemming from the perception of a large number of different places with different forms and functions. Whilst it is difficult to quantify this variation and account for it directly, the grouping of submitted soundscapes into categories served to reduce the uncertainty caused by the different environments and reveal the differences in perception between these locations.

The voluntary nature of participation, has led to a very large and diverse set of submissions. The lack of any leading suggestions in any of the marketing material about exactly what to capture and submit, coupled with the diverse range of locations that these devices move through has resulted in a very broad dataset of soundscape types. To date no other soundscape study has had the potential to reach this range of locations and gather subjective and objective soundscape data. However, this diversity has had the detrimental effect of generating a dataset that could be considered too broad in the types of soundscapes submitted for a general analysis to be performed across the entire set of submissions. A more focussed approach to the analysis would provide the benefits of being able to understand the soundscape appreciation of a particular class of location such as a set of pre-selected urban parks. Through a more targeted marketing strategy, participants could be advised to capture the soundscape of a particular local park or urban space. The data that this would generate would be more focussed and could potentially provide a clearer insight into the soundscape appreciation of a particular place, something which the current methodology does not yet have the ability to do.

5.2.4 Mediating technology

Whilst it is clear that the technology utilised in the present study is not designed for use in soundscape research, it could be argued that the advantages of using these pervasive mobile recording devices can actually provide a more representative insight into a persons sound environment and their perceptions of it. In the case of the iOS and J2ME app, the subject is responding to the soundscape immediately after the audio recording is made. It could be argued that this has advantages over research techniques that take longer (but admittedly higher quality and generally calibrated) recordings in different locations and stop and question subjects within these locations, such as [164]. The recordings made and analysed using these more traditional methodologies may not feature the acoustic events that were commented on, so attempting to compare this objective and subjective data may not serve to reveal the true relationships between the acoustic environment and its perception.

The reach attainable using the current studies methodology is only limited by the places a participant can physically go, opening up all kinds of locations to the researcher, such as domestic or workplace environments. Currently, the only other soundscape research approach which could gather this range of sound environments is small participation sound diary studies [165].

The uncalibrated nature of the soundscape recordings, does raise some issues when it comes to the extraction of level metrics. The relative levels utilised in the study did reveal a number of significant relationships with subjective metrics, however, a calibrated level comparison would provide more of a quantifiable result which could be used by policy makers and urban design practitioners. A reliable measure of sound pressure level could be logged alongside a participants subjective responses in future versions of the app. The condition would be that only Apple devices were used due to their small product range.

5.2.5 Data quality

The issue of data quality in studies such as this is an important consideration. Past research has assessed the quality of data based on a number of dimensions that came about from stakeholder interviews [166]. These dimensions have been condensed and applied to data gathered as part of citizen science style experiments by Hunter et. al. [167]. These have been adapted and listed to assess the quality of the current project's dataset.

1. **Accessibility - The extent to which the data is available, or easily and quickly retrievable**

   The data is easily accessible due to its online storage. A number of applications have the capability to access online databases making the retrieval of it a trivial process.

2. **Appropriate amount of data - The extent to which the volume of data is appropriate for the task at hand**

   The limitations of the research platform mean there was an inherent limit on the amount of subjective data that could be gathered. For the project's aim of assessing human relationships with their sound environments, it would seem that the complexity of this relationship is not easily described using a small number of semantic differential scales. The large number of submissions made to the project, does however mean that the statistical power of the analysis on the dataset was of large enough order to gain meaningful and reliable results.

3. **Believability - The extent to which the data is regarded as true and credible**

   The subjective data credibility is difficult to assess as any kind of reaction to a place's sound environment should be assumed to be believable and correct. Cases where a non-response was detected were excluded from the analysis stage.

4. **Completeness - The extent to which the data is not missing and is of sufficient breadth and depth for the task at hand**
   Non-responses can be considered as missing and were excluded from analysis, as well as audio submissions that did not meet the minimum sample rate criteria. To accurately describe a persons appraisal of their soundscape, the data gathered could be lacking in depth. The explanatory power of the subjective variables used could be enhanced by asking more from participants to expand on their responses to each scale. This however may be off putting to participants due to the time required to answer these additional questions.

5. **Concise representation - The extent to which the data is compactly represented**
   The predominately quantitative nature of the project's dataset mean that it is inherently compact and concise, as well as being easy to interpret by a lay audience.

6. **Consistent representation - The extent to which the data is presented in the same format**
   The confined method of data entry online and on devices means that the data is consistent across all submissions. A participant's reason for submission and source identifications could be considered the least consistent as they prompt for an open response. However, this open response was manually categorised in the analysis stage to maintain data consistency in the final analysis.

7. **Ease of manipulation - The extent to which the data is easy to manipulate and apply to different tasks**
   The open response data did require a stage of manual classification before analysis, which was time consuming and including its own subjective interpretation of the different terms used by participants.

8. **Free-of-error - The extent to which information is correct and reliable**
   The reliability of the data was a major consideration in this study. The subjective responses were assumed to be correct for each valid submission, provided that each person interpreted the questions in a similar way. The soundscape recordings contain known errors in terms of the limitations of the recording device and less quantifiable errors such as blocking the phone microphone when capturing.

9. **Interpretability - The extent to which data is in appropriate languages, symbols, units and the definitions are clear**
   Semantic differential scale data provides a clear interpretation of what it is measuring and the extremities of this measurement. The categorisation of the source identification responses and reasons for capturing questions are open to a broader interpretation due to the researcher having to manually make their own choice when determining categories.

10. **Objectivity - The extent to which the data is unbiased, unprejudiced and impartial**
    A number of biases occur in the projects dataset. Participant gender has a strong male bias, the participant base is generally young, the focus on a particular smart phone brand again biases the type of person who will

participate. The project also really only caters for English speakers due to the lack of any translations of the project question set. The soundscape recordings have a high frequency bias as the majority of devices exhibit a low frequency roll-off which cannot be bypassed.

11. **Relevancy - The extent to which the data is applicable and helpful for the task at hand**
    The data is directly applicable to the task of soundscape analysis as that is the only focus of the research.

12. **Reputation - The extent to which the data is highly regarded in terms of source or content**
    The reputation of the data is difficult to quantify as this is the first study of this kind and will potentially be used to gauge the reputation of data gathered in future studies.

13. **Security - The extent to which access to data is restricted appropriately to maintain its security**
    The majority of the project's dataset is freely available online. With no data linking submissions to participants, its low sensitivity warrants its public presentation. In terms of its storage, the data itself has numerous backups and access checks so cannot be lost or stolen.

14. **Timeliness - The extent to which the data is sufficiently up-to-date for the task at hand**
    The data itself is very current with the potential to analyse submissions as soon as they arrive at the server.

15. **Understandability - The extent to which the data is easily comprehended**
The subjective data can be easily understood, however, the objective acoustic metrics extracted from each submission may be a little harder to relate to the signal itself as heard by an individual.

16. **Value-added - The extent to which the data is beneficial and provides advantages from its use**
The large amounts of data from a variety of locations is unobtainable using traditional soundscape research techniques due to time and cost factors. Therefore, the data that this project has produced can be used as a case study for similar future projects.

5.2.6 Time and cost

The development of the project's mobile applications and refining of the interfaces did cover a large part of the project's timeline. The voluntary nature of participation meant that the interface and functionality of the applications had to be as close to faultless as possible to ensure that people weren't immediately put off taking part due to a complex method and constant crashes. Once the development and piloting had concluded, the commencement of the actual live study required a constant level of commitment to handle technical support for the app and website. With the dataset being stored in an online database, the process of data retrieval for analysis was quick and easy. There was no time intensive process of data entry involved, as all data was already collated for analysis. The ongoing participation from members of the public also mean that the project has extended longevity. Updates to the project application can be made which allow for the retrieval of additional data from the large participant base.

The costs involved to develop and run a project such as this are relatively low compared to traditional studies of this scale. Researcher time taken to interview participants, collate data and transcribe responses can add up, along with the associated costs. Carrying out additional studies generally requires repeating the whole process with a new set of participants. The present study already has a large subject base running the project's application, which can be updated with additional questions, or a related but different study with a new focus.

5.3 Comparison with traditional methodologies

5.3.1 Advantages

- Can amass large amounts of research data
- Wide geographical reach
- Data from traditionally inaccessible locations
- Human choice behind submissions uncovers new insights
- Relatively low cost
- Subjective responses captured in-situ alongside audio data
- Raises awareness of soundscape issues
- Rewarding to participant
- Reduced interviewer bias
- Scalable and flexible platform for soundscape research

5.3.2 Disadvantages

- Limited amount of questions can be asked of participants
- Uncalibrated soundscape recordings
- High pass filtering of recorded soundscapes on iOS devices
- Relatively high numbers of non-responses

5.4 Summary

The strong reliance on voluntary participants has been identified and measures have been taken to mitigate for this. One of the major approaches to ensuring adequate take-up was ensuring ease of participation, which in turn was dependent on the ease of use of the mediating technologies. The addition of the iOS application served to drop the existing technological barriers to participation.

The methodology developed has proven successful in gathering large amounts of meaningful soundscape research data. The data itself has proven to corroborate with other studies involving varying methodological approaches. The novel inclusion of participant choice has also given rise to new findings around soundscape perception.

The methodology also exhibits a number of advantages over more traditional methodologies. There are also disadvantages, some of which will surely be removed with advances in technology and API access.

Chapter 6

Conclusions

6.1 Introduction

This chapter summarises the outcomes and achievements of this thesis. A novel methodology for soundscape research has been presented, utilising ubiquitous consumer mobile and internet technologies. A soundscape research tool has been designed and implemented that can be used to gather objective environmental data, as well as quantitative and qualitative subjective data from participants in-situ. Large numbers of untrained members of the public have submitted soundscape recordings from around the world. Human relationships with their sound environment's has been investigated, with a number of findings corresponding with those of other soundscape studies utilising entirely different methodologies. In addition, a number of new findings have been made to contribute to the field.

6.2 Findings

This section summarises the corroborative and novel findings of the present study.

6.2.1 Corroborative

Identified positive and negative sound sources within a soundscape have shown their influence on soundscape perception. Human sounds are generally appreciated more than natural, and certainly artificial sounds. Prominent human sounds also serve to increase perceptions of soundscape activity much more so than prominent natural sources. Tranquillity ratings are seen to increase when a natural sound source is identified.

Urban soundscapes have exhibited significantly lower appreciation ratings when compared to rural soundscapes. The number of urban inhabitants taking part cannot be quantified, but this preference for the rural soundscape may be caused by people appreciating what they do not usually have. An urban dweller spending time in the countryside may look more favourably on their surroundings as they have taken time out of their usual city environment to enjoy the countryside. Artificial sound sources are consistently viewed as negative influences within soundscape perceptions across all environments.

The two extracted subjective principal components of 'Appreciation' and 'Dynamics' has shown a solid validation of the project's methodology, due to their similarities with a number of other studies utilising different techniques of data retrieval. The distinctive groupings of the different soundscape types within this factor space defined by the extracted components reveals the perceptual differences between the soundscape categories: urban, rural, urban public space and urban park.

High levels of impulsive sources within captured soundscapes seem to reduce overall appreciation and increase perceptions of dynamics. High frequency content within sound environments resulted in reduced ratings of location quality and perceived soundscape dynamics. A soundscape containing a more varied and distinct set of spectral features generally results in higher ratings of perceived dynamics.

The small participation study revealed comparable findings to the main dataset, serving as further validation to the project's methodology.

6.2.2 Novel

The activity a person is involved in while making their submission has shown to be influential in soundscape appraisal, with relaxation and recreation situations resulting in increased soundscape appreciation. Recreational activities also exhibit increased perceptions of soundscape activity.

The reasons behind a soundscape submission have revealed significant differences in subjective response. The positive interpretation of the term soundscape has resulted in a majority of positive reasons for taking part. Soundscapes that arise from a participant's daily routine are generally less appreciated than soundscapes containing a particular sound source focus. The highest levels of appreciation were observed in soundscapes whose focus is on a specific activity that the participant is involved in. The interest that a participant has on their soundscape is seen to result in raised levels of appreciation.

The extracted objective components of "Relative level" and "Frequency content" have shown significant correlations with the extracted subjective components of ’Appreciation’ and ’Dynamics’. A soundscape’s level correlated with its perceived dynamics and appreciation. Its spectral content also correlated with its dynamics.

The extraction of prominent sources from the project’s dataset revealed a relationship between perceived tranquillity and the frequency distribution of these source’s. Sound sources with a spectral distribution tending towards white noise have shown increased perceptions of the general rating of soundscape excitement.

6.3 Contributions

This thesis has shown that new mobile and internet technologies can be used to create a novel and robust methodology for soundscape research. Crowd-sourced, large scale, participant driven soundscape data has been used to assess human relationships with their sound environment’s for the first time. The findings have been validated against those from existing research methodologies utilising more traditional tools. Inferences have been made from the quantitative and qualitative responses of participants, as well as comparing these with objective metrics obtained from the soundscape recordings.

The context of participant activity and the reasons behind soundscape submissions has revealed that the interest a participant has in their environment positively influences their appreciation of its soundscape. The positive interpretation of the term soundscape has revealed a persons inclination to focus on positive sound environments when taking part in the study. The activities of relaxation and recreation exhibit increased levels of soundscape appreciation.

6.4 Further work

6.4.1 Mobile technology

Universal soundscape research platform

The current project's application and web based infrastructure could be utilised by other researchers within the field of soundscapes. A particular research project could specify the questions that need answering by participants and in what format they need asking in. The introductory text explaining the research and what is required of participants would also be submitted, alongside the research institutes branding material. These could then be reproduced on an initially unbranded mobile application, which is then skinned in the research groups brand style. With the capturing and transmission of this audio and associated data already a function of the application. This then has the potential to become a flexible and practical platform to gather worldwide soundscape data for wide range of academic groups.

Geo–fencing

The concept of geo–fencing has come about as a result of the proliferation of always–on GPS enabled mobile devices. A geo–fence is a virtual perimeter around a real world geographic area [168]. A mobile device can have an action triggered when it enters this virtual area, such as an app prompt or alert. The use of this technique allows for the implementation of focussed soundscape research studies on particular areas such as urban parks, squares or entire cities. The existing user base of the project could be alerted when they are entering a chosen location and asked to record and comment on its soundscape. This location focussed approach could serve to produce a dataset with increased explanatory power to describe the soundscape of a small location. This could also be complimented with the addition of data from stop and question style research using researchers present in the park armed with tablets for data input and recording.

Gamification

Gamification of soundscape research involves adding a competitive task based element to the collection of in-situ research data. Participants would be tasked with gathering soundscape recordings and responding to them in response to a challenge from the researcher, maybe in competition with another participant. The use of this technique would need to be carefully considered in terms of the bias that may be introduced from competition based submissions. Small numbers of these techniques have even been discussed within the realms of noise pollution research using mobile phones [169]. One such example application that is not necessarily research related is the geocaching practice [170], which involves people using their GPS enabled smartphones to discover locations where a previous participant has left a small item to find. This idea could be converted for use in soundscape research, by posting geocaches of “sonic interest” which people must find and comment on. For example, town planners and local councils could gauge peoples perceptions of the acoustics of a particular building or outdoor public space using a technique such as this.

Tablet use in location surveying

The more traditional techniques of soundscape surveying could also benefit from the use of internet connected wireless devices. In particular, tablet computers could be used as data entry devices where researchers could gather subjective responses and high quality recordings using external hardware attached to these devices. With wireless internet connectivity on the majority of tablets, the gathered data could be tagged, collated and uploaded to a server, ready for analysis when the researcher returns to their office. This would reduce the time taken to convert collected field data into a database format ready for analysis. The recording capabilities of these devices also allow for the capturing of spoken interviews with participants, resulting in the consolidation of a number of research tools into one compact device.

Sound diaries

The ubiquity of mobile phones also means that they are carried around with people at all times. Studies that rely on providing participants with recording devices to capture and document their sonic lives would benefit from a custom application which incorporated all of the required functionality to carry out sound diary style soundscape research. The cost and time savings that this application could provide would be very advantageous to a studies of this kind. Recording equipment purchase and loaning, coupled with the time taken to collate the qualitative responses make this form of research very time consuming and limited in the number of participants that can be handled. A mobile device platform could greatly reduce the costs and logistics issues, as well as opening up the participant base to far larger numbers.

6.4.2 Project dataset

The project's ever increasing dataset can be utilised in future research with the aid of a clear and thorough set of meta data associated with each submission. The implementation of a web based API to access the audio recordings and associated data would mean researchers could access the data by keyword, specifying a particular set of keywords to gain access to specific soundscape types. An API call such as:

```
soundaroundyou.com/data.php?loc=GB&type=urban-park&fs=44100
```

This call would only return soundscape recordings and subjective responses from urban parks in Great Britain, with an audio sample rate of 44100Hz. However, the issue of adequate tagging of each soundscape is paramount in accessing data in this manner.

Meta tagging

An important aspect involving the legacy of the project's dataset and use of it is in the correct tagging of each submission. The audio file itself has an associated set of data already included with it, such as its location, time, date and subjective responses, however there would need to be a more fine grained set of terms used to describe the recording itself. Keywords would need to be added manually to accurately identify the sound recording in terms of its content and descriptive location. The objective metrics extracted from each recording could also be used as a means to select soundscapes that, for example have a large amount of low frequency energy or have a high number of impulsive events within them.

Objective soundscape metrics

The objective metrics extracted from each soundscape recording characterised the soundscape using a number of musical feature descriptors. A future stage of extraction would benefit from using objective metrics devised explicitly for soundscapes. The use of these in further analysis may serve to uncover more correlates between the objective and subjective. The following metrics have been used in the past in soundscape research:

- **The Slope Indicator**

  The Slope indicator is related to the time history of a signals SPL and is a measure of how often events appear within it and how they emerge from the background. It is the only current metric devised specifically for soundscape analysis. A peak in the signals low pass filtered spectrum evidences a repetitive event during the signals duration. The value of Slope is the measure of correlation between events appearing in this time history [2]. The slope indicator has been able to detect greater variation in manipulated soundscapes than some other acoustic indicators such as loudness, sharpness and fluctuation strength [171].

- **1/f**

  A 1/f spectrum is when the amplitude of a signal is inversely proportional to frequency, especially at very low frequencies such as 1 Hertz. When the magnitude spectrum of a signal is plotted on a log-linear scale, it produces a 1/f pattern. The 1/f pattern is found in numerous scenarios, including speech, music, radio static and even abstract paintings. This measure has also been used to examine the temporal structure of soundscapes, by observing the amplitude and pitch fluctuations [172]. In general, rural soundscapes tend to portray more 1/f characteristics than urban soundscapes [173].

- **NCN Number of noise events (Sounds exceeding L_{95} by 10dB)**
  The number of noise events is defined as the number of times the sound pressure level exceeds a threshold level, usually set by the statistical background sound level of L_{95}. Using this technique, a noise event can be considered as any sound that exceeds the L_{95} background level by at least 10 dB(A). The choice of threshold values has been shown to correlate well with the number of perceived vehicle passes in a study by De Coensel & Botteldooren [86]. The total duration of the exceeding of this threshold can also be measured, giving Tcn. It has been suggested that the best way to retrieve this value is by human observation of the time history of the recording. An automatic method could also be implemented which simply worked on the threshold levels defined and made a log of all events which exceeded these limits. This method would however disregard any contextual information which may be significant, for example close proximity speech would be considered as a number of noise events, rather than as a single event.

Appendices

Appendix A

Website screenshots

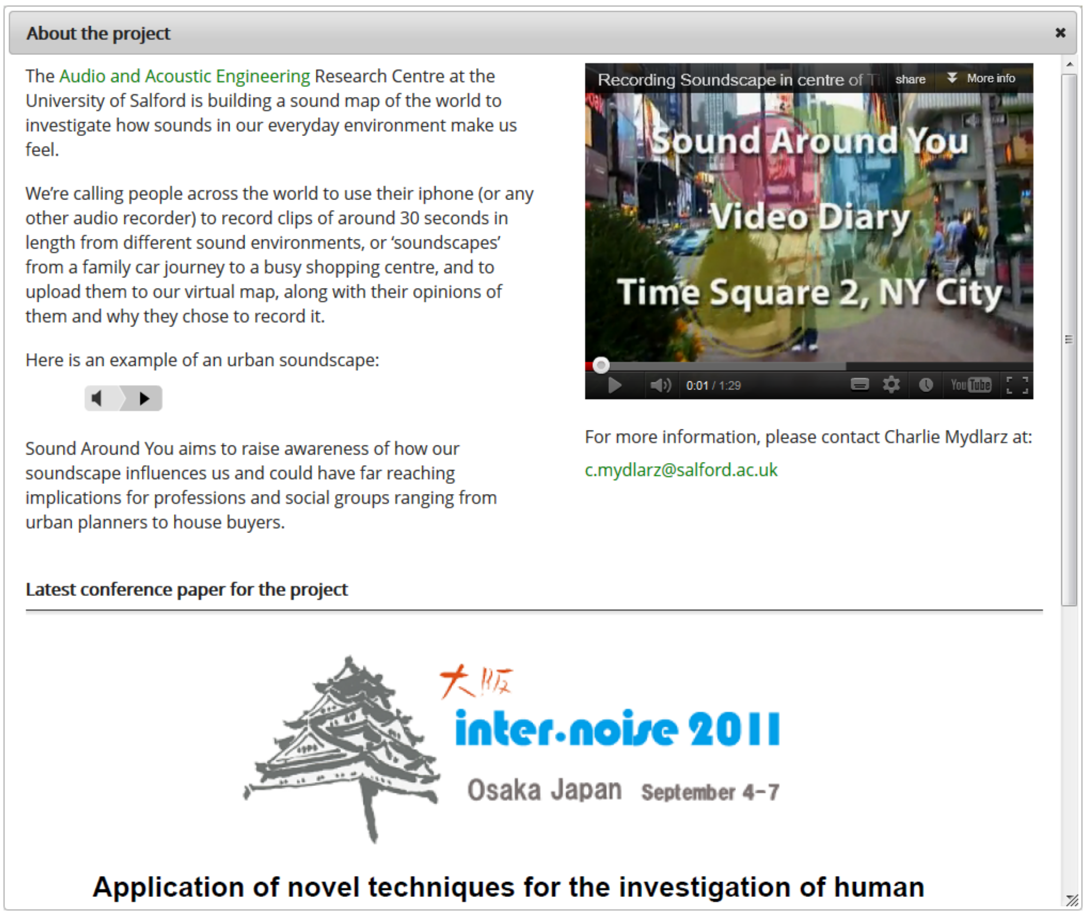

Figure A.1: Project website about view

Appendix B

Pilot study feedback activities

Aim

This is a group activity using SoundAroundYou.com resources to help participants assess, measure and appreciate their everyday soundscapes. Any data uploaded to SoundAroundYou will contribute to a major soundscape research project that has the potential to better inform environmental planning and law.

Objectives Summary

- Introduce the basic concept of soundscapes
- Take participants on a soundwalk of indoor or outdoor locations
- Record 10 second audio clips and opinion data of sounds in several locations
- Upload the data to a worldwide soundscape map
- Reflect on the nature and context of soundscapes measured

.

The Group Activities

Introducing Soundscapes to group To start the group activity you may wish to initialise a discussion to introduce the concept of soundscapes and why measuring and assessing them might be useful. You could do this with some questions to the group members about their everyday soundscapes; such as:

- What's a good and bad soundscape for relaxing?
- What's a good and bad soundscape for working?
- Can you identify some positive and negative sounds (written on post-its or paper scraps)?
- Is a loud soundscape always bad or a quiet one always good?

Sound walks A sound walk is where a group, sub groups or individual members in their own time go for a walk and in doing so choose locations to stop, listen, record and assess. The locations can be indoor or outdoor. Typically youd want to record and assess five to ten soundscapes at varying locations that members might choose as good, bad or especially interesting.

Group Reflective Practice Listening back to soundscapes on the PC and discussing the corresponding in-situ opinions can be enlightening and often surprising. For example the soundscapes members rated highly in-situ might sound unpleasant when listening back (and vice versa). Soundscape quality judgements depend on many factors beyond the sound itself including what our others senses experienced, our mood, expectations, the context, what were used to and our desired activity at the location. Further more the meaning of sounds, our control over them and who or what's making them all influence our judgements.

Upload soundscapes This can be done at the base PC by the group leader or by registered group members in their own time. Please ensure members correctly locate the soundscapes via our Google maps interface, give their opinion data and give some representative demographic data (age, etc) when prompted.

Impressions of session How did the session go for its participants? Short questionnaire given out to all members asking how they found using the mobile software, PC application, question set, website etc. How could any step be improved? Which steps worked?

Appendix C

Pilot study feedback questionnaire

IMPRINTS

How did it all work out?

Answer the questions by ticking the box of your choice and add any comments in the grey box to help us make the project better

Soundscape mobile phone software comments

Were you able to capture soundscapes easily? yes ☐ no ☐

Comments:

Was there enough feedback and information whilst recording? yes ☐ no ☐

Comments:

Did you know what to do after recording? yes ☐ no ☐

Comments:

Did the opinions questions make sense? yes ☐ no ☐

Comments:

Did you understand what was going on while using the software? yes ☐ no ☐

Comments:

Soundscape questionnaire comments

Did you understand what the questions meant? yes ☐ no ☐

Comments:

Would you need the questions explained to understand them? yes ☐ no ☐

Comments:

Was the language used understandable? yes ☐ no ☐

Comments:

Could you answer the questions on your own? yes ☐ no ☐

Comments:

Figure C.1: Schools pilot feedback questionnaire

Appendix D

Press release

23 October 2009

Acoustics experts to create sound map of Britain

Acoustic engineers at the University of Salford will, this week, launch a unique project to create a 'sound map of Britain.' From Monday, the public will be invited to use their mobile phones to record around five 10 second audio clips from different environments - such as a local park or a street - and upload them to a virtual map at www.soundaroundyou.com with their opinion on how it makes them feel and why they recorded it.

The study, which is being launched at the Manchester Science Festival (24 Oct 1 Nov), aims to get a better understanding of what gives a place 'character' and how opinions and attitudes to 'noise' vary.

PhD student Charlie Mydlarz, who is leading the study, explains: “We're asking people to capture any environment they choose, and that includes both public and private spaces, so recordings could capture anything from a family car journey to a busy shopping centre”

“And by using everyday technology to get people involved, this has the potential to be the largest study of its kind. We’ll be producing the first ever sound map purely for research purposes - the findings of which could have far reaching uses from psychological research to town planning.”

While existing studies tend to focus on volume, with loud assumed noisy/ undesirable and quiet as desirable, this study will investigate the idea that there is no such thing as ’noise’, simply sound that is out-of-place or context. For example, the sound of a busy street and shouting voices may be unpleasant and out-of-place in a quiet residential area but is an essential part of the ’atmosphere’ and personality of a market.

Sound, especially in urban areas, is an increasingly important issue, as recent debates about potential noise levels from a third runway at Heathrow demonstrate. While there are many bodies concerned with how our environments look (e.g. English Heritage or CABE) how they sound has, until recently, often been overlooked.

With more people living in cities, urban areas being redesigned and new technologies such as electric or hybrid cars offering the potential for quieter streets, there is a need to reassess our understanding of ’noise’.

This sound map of Britain could be useful in a variety of ways, for example, for urban planners or people checking out an area ahead of buying a house;

If used along side Google street view, end users could not only see a place they could hear it too, thereby getting a far better understanding of what a place is actually like.

In raising awareness of how our sound environment influences us, researchers hope that participants will embrace a new ’language of sound’. For example, rather than ’landmarks’ and ’landscapes’ we might describe distinctive features of our sound environment or 'soundscapes', as 'soundmarks' and value them as highly as an attractive country vista or dramatic urban skyline.

The research is being run by a team including PhD student Charlie Mydlarz, Prof. Trevor Cox and Dr Ian Drumm at the Department of Acoustic Engineering, University of Salford.

ENDS

For press enquiries or to interview project leader Charlie Mydlarz please contact Susie Hartley at mission 21 on Susie.hartley@mission-21.com

Appendix E

Share options

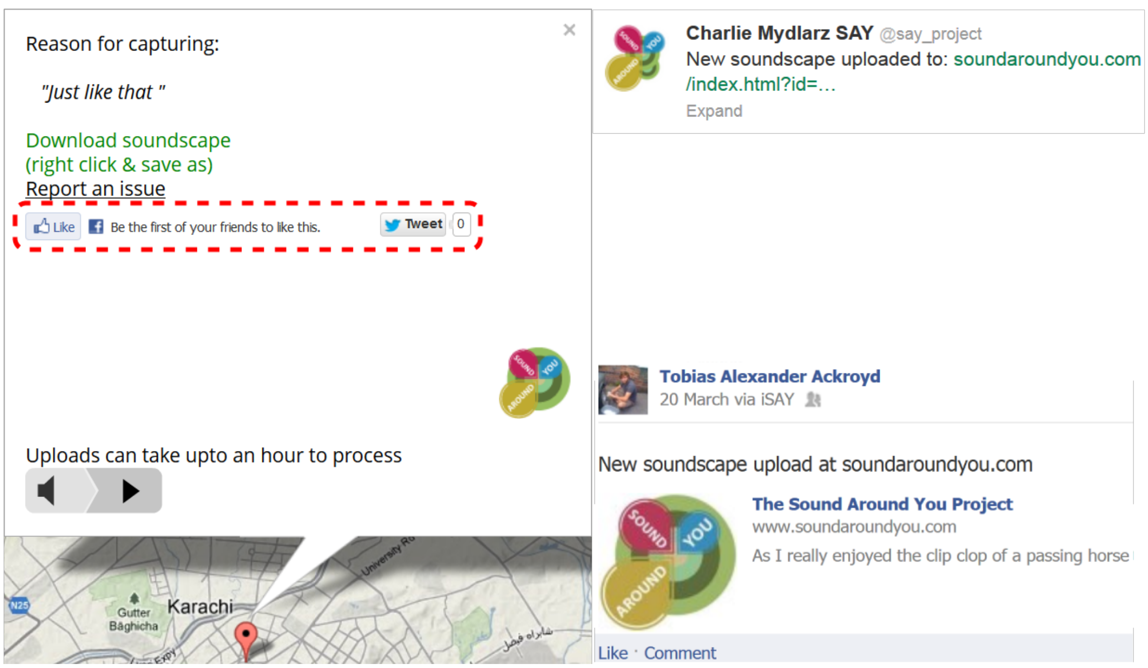

Figure E.1: Share buttons on project website (far left, looped in dashed red) & shared soundscapes posted to Twitter (top right) & Facebook (bottom right)

Figure E.2: Share buttons on iPhone app Upload & Explore tab (looped in dashed red), including tweet page from Explore page showing @say_ project & soundscape hashtag

Appendix F

Project disclaimer

Use of soundaroundyou.com

- You agree to use soundaroundyou.com only for lawful purposes, and in a way that does not infringe the rights of, restrict or inhibit anyone else's use and enjoyment of soundaroundyou.com. Prohibited behaviour includes harassing or causing distress or inconvenience to any person, transmitting obscene or offensive content or disrupting the normal flow of dialogue within soundaroundyou.com.

Contributions to the Sound Around You Project

- By sharing any contribution (including any text, photographs, graphics, video or audio) with the project you agree to grant to soundaroundyou.com, free of charge, permission to use the material in any way it wants (including modifying and adapting it for operational and editorial reasons).
- Copyright in your contribution will remain with you and this permission is not exclusive, so you can continue to use the material in any way including

allowing others to use it.

- In order that soundaroundyou.com can use your contribution, you confirm that your contribution is your own original work, is not defamatory and does not infringe any UK laws, that you have the right to give soundaroundyou.com permission to use it for the purposes specified above, and that you have the consent of anyone who is identifiable in your contribution or the consent of their parent / guardian if they are under 16.
- Please do not endanger yourself or others, take any unnecessary risks or break any laws when creating content you may share with soundaroundyou.com.
- If you do not want to grant soundaroundyou.com the permission set out above on these terms, please do not submit or share your contribution to or with soundaroundyou.com.

About your uploads

- Contributions must be civil and tasteful.
- No disruptive, offensive or abusive behaviour: contributions must be constructive and polite, not mean-spirited or contributed with the intention of causing trouble.
- No unlawful or objectionable content: unlawful, harassing, defamatory, abusive, threatening, harmful, obscene, profane, sexually oriented, racially offensive or otherwise objectionable material is not acceptable.
- No spamming or off-topic material.
- No advertising or promoting.

- No impersonation.
- No inappropriate (e.g. vulgar, offensive etc) user names.
- URLs (web site addresses) can only be posted if allowed under any relevant local house rules.
- Deliberate misuse of the complaints facility is not permitted. If you persist in doing this, action may be taken against your account.

If you have any concerns about this disclaimer, please email: cmydlarz@soundaroundyou.com

Appendix G

Analysis plots

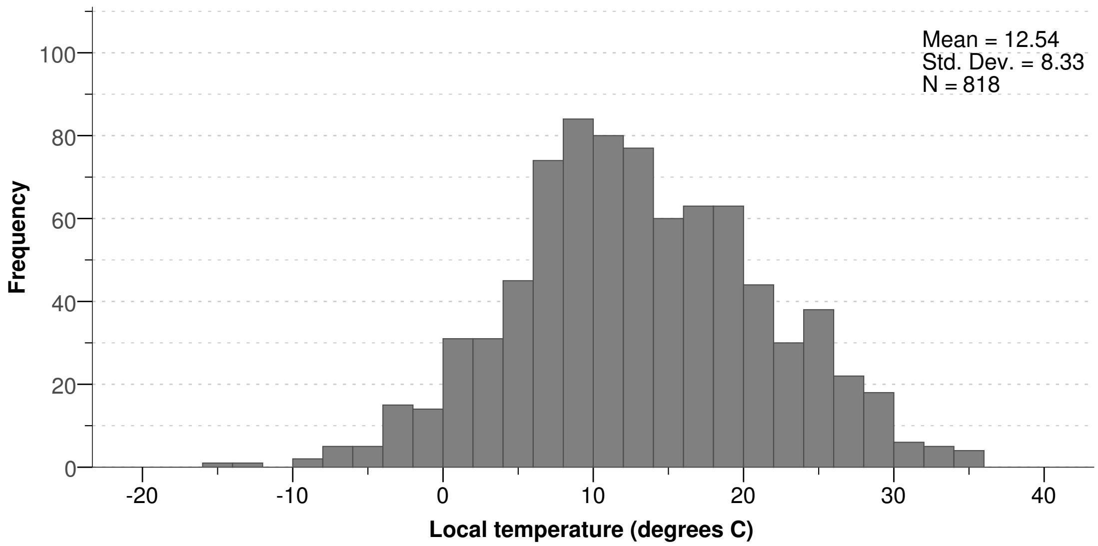

Figure G.1: Soundscape temperature distribution

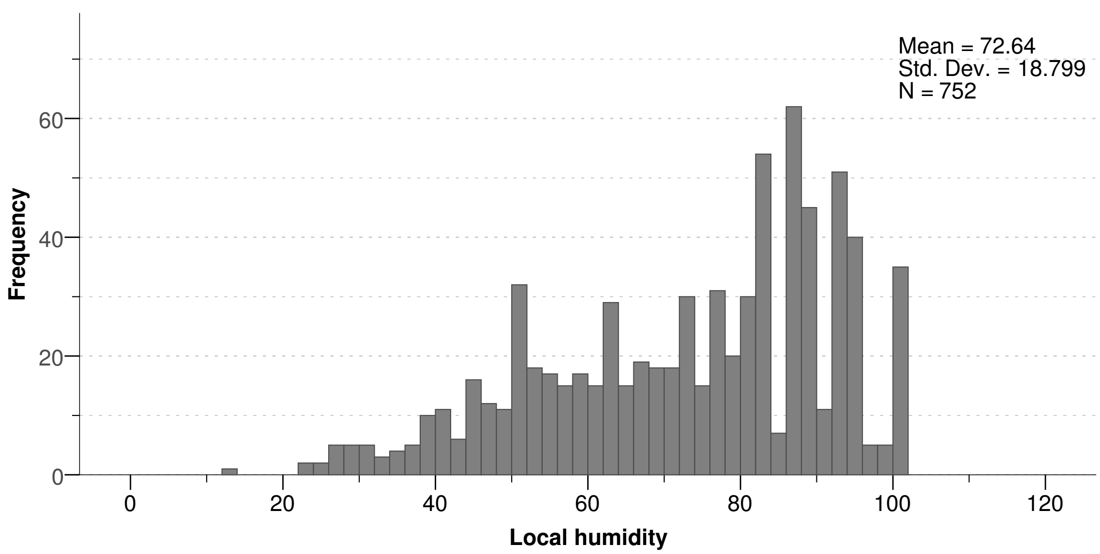

Figure G.2: Soundscape humidity distribution

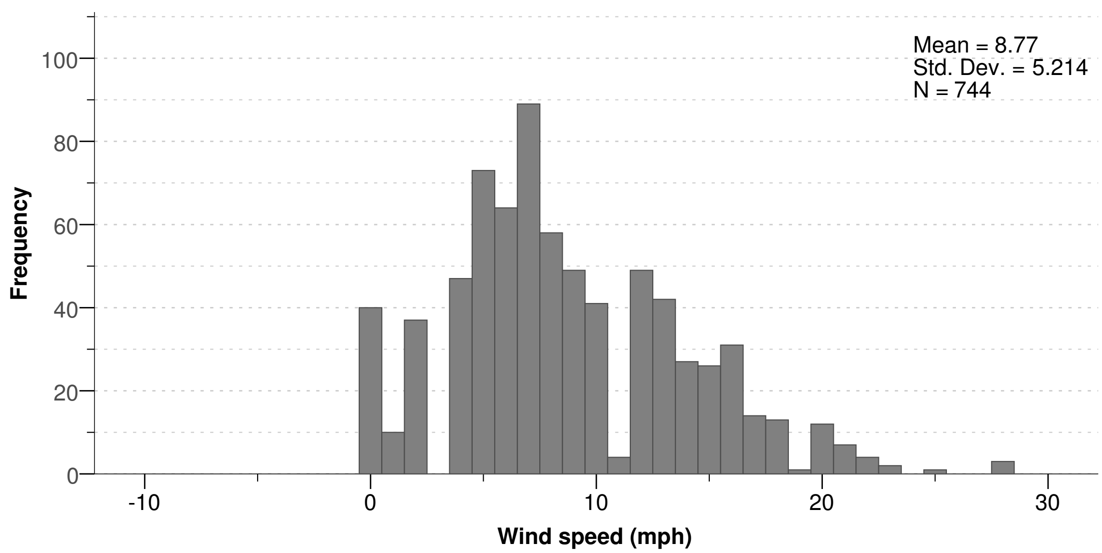

Figure G.3: Soundscape wind speed distribution

Bibliography

[1] S.R. Payne, P. Devine-Wright, and K.N. Irvine. People's perceptions and classifications of sounds heard in urban parks: semantics, affect and restoration. In *Internoise and Noisecon Congress and Conference Proceedings*, volume 2007, pages 2032–2041. Institute of Noise Control Engineering, 2007.

[2] G. Memoli, A. Bloomfield, and M. Dixon. Soundscape characterization in selected areas of central london. *The Journal of the Acoustical Society of America*, 123(5):3811, 2008.

[3] B. Berglund and M.E. Nilsson. On a tool for measuring soundscape quality in urban residential areas. *Acta Acustica united with Acustica*, 92(6):938–944, 2006.

[4] H.S. Becker and B. Geer. Participant observation and interviewing: A comparison. *Human Organization*, 16(3):28–32, 1957.

[5] Ö. Axelsson, M.E. Nilsson, and B. Berglund. A principal components model of soundscape perception. *The Journal of the Acoustical Society of America*, 128(5):2836–46, November 2010.

[6] J. Kang and M. Zhang. Semantic differential analysis of the soundscape in urban open public spaces. *Building and Environment*, 45(1):150–157, January 2010.

[7] R. Cain, P. Jennings, and J. Poxon. The development and application of the emotional dimensions of a soundscape. *Applied Soundscapes: Recent Advances in Soundscape Research*, 74(2):232 – 239, 2013.

[8] M. Weiser. The computer for the twenty-first century. *Scientific American*, 1991.

[9] M. Callon. Research in the wild and the shaping of new social identities. *Technology in Society*, 25(2):193–204, April 2003.

[10] N. Maisonneuve, M. Stevens, M.E. Niessen, and L. Steels. NoiseTube: Measuring and mapping noise pollution with mobile phones. *Information Technologies in Environmental Engineering*, 4:215–228, 2009.

[11] E. Kanjo, S. Benford, M. Paxton, A. Chamberlain, D.S. Fraser, D. Woodgate, D. Crellin, and A. Woolard. Mobgeosen: facilitating personal geosensor data collection and visualization using mobile phones. *Personal and Ubiquitous Computing*, 12(8):599–607, August 2007.

[12] A. Steed and R. Milton. Using tracked mobile sensors to make maps of environmental effects. *Personal and ubiquitous computing*, 12(4):331–342, 2008.

[13] E. Kanjo. NoiseSPY: A Real-Time Mobile Phone Platform for Urban Noise Monitoring and Mapping. *Mobile Networks and Applications*, 15(4):562–574, November 2009.

[14] WikiMedia. Mobile forecasts india - strategic planning. Online: http://strategy.wikimedia.org/wiki/Mobile/Forecasts/Indias, Accessed: June 2011.

[15] International Telecommunication Union - ICT Data and Statistics Division. ICT facts and figures. Technical report, International Telecommunication Union, 2012.

[16] P. Farago. iOS and Android adoption explodes internationally. Online: http://blog.flurry.com/bid/88867/iOS-and-Android-Adoption-Explodes-Internationally, Accessed: August 2012.

[17] Portio Research Limited. Mobile factbook - April 2012. Online: http://www.portioresearch.com, Accessed: April 2012.

[18] The European Commission. Directive 2002/49/EC of the european parliament and of the council of 25 June 2002 relating to the assessment and management of environmental noise, 2002.

[19] European Commission. Future noise policy: European commission green paper, 1996.

[20] R.M. Schafer. *The tuning of the world*. Knopf, 1977.

[21] B. Truax. *Acoustic Communication*. Ablex, 2nd edition, 2001.

[22] M. Zhang and J. Kang. Towards the evaluation, description, and creation of soundscapes in urban open spaces. *Environment and Planning B: Planning and Design*, 34(1):68–86, 2007.

[23] R.M. Schafer. The soundscape designer. In Helmi Jarviluoma, editor, *Soundscapes: Essays on Vroom and Moo*, pages 9–18. Department of Folk Tradition, 1994.

[24] K.N. Irvine, P. Devine-Wright, S.R. Payne, R.A. Fuller, B. Painter, and K.J. Gaston. Green space, soundscape and urban sustainability: an interdisciplinary, empirical study. *Local Environment*, 14(2):155–172, 2009.

[25] J. Punter, M. Carmona, and A. Platts. The design content of development plans. *Planning Practice and Research*, 9(3):199–220, 1994.

[26] R.M. Schafer. *The new soundscape*. Berandol Music, 1969.

[27] R.M. Schafer. *The book of noise*. Priv. print. by Price Print., 1970.

[28] R.M. Schafer. *The music of the environment*. Wien, 1973.

[29] J.B. Fritz, M. Elhilali, S.V. David, and S.A. Shamma. Auditory attention–focusing the searchlight on sound. *Current opinion in neurobiology*, 17(4):437–55, 2007.

[30] A.S. Bregman. *Auditory scene analysis: The perceptual organization of sound*. MIT Press, 2 edition, 1990.

[31] K. Wrightson. An introduction to acoustic ecology. *The Journal of Acoustic Ecology*, 1(1), 2000.

[32] M. Raimbault and D. Dubois. Urban soundscapes: Experiences and knowledge. *Cities*, 22(5):339–350, 2005.

[33] R. Cain, P. Jennings, M. Adams, N. Bruce, A. Carlyle, P. Cusack, W. Davies, K. Hume, and C.J. Plack. Sound-scape: A framework for characterising

positive urban soundscapes. volume 123, pages 3394–3394. [New York: Acoustical Society of America], 2008.

[34] W.J. Davies, M.D. Adams, N. Bruce, M. Marselle, R. Cain, P. Jennings, J. Poxon, P. Cusack, A. Carlyle, D.A. Hall, A. Irwin, I. Hume, and C. Plack. The positive soundscape project : A synthesis of results from many disciplines. In *Internoise 2009*, Ottawa, Canada, 2009.

[35] A. Irwin, D.A. Hall, A. Peters, and C.J. Plack. Listening to urban soundscapes: Physiological validity of perceptual dimensions. *Psychophysiology*, pages 1–11, June 2010.

[36] A. Irwin. How do listeners react to different urban soundscapes? an FMRI study of perception and emotion. In *Internoise 2009*, Ottawa, Canada, 2009.

[37] M. Raimbault. Qualitative judgements of urban soundscapes: Questionning questionnaires and semantic scales. *Acta Acustica united with Acustica*, 92(6):929–937, November 2006.

[38] C. Guastavino. The ideal urban soundscape: Investigating the sound quality of french cities. *Acta Acustica United With Acustica*, 92:945 – 951, 2006.

[39] S.R. Payne, W.J. Davies, and M. Adams. Research into the practical and policy applications of soundscape concepts and techniques in urban areas (DEFRA–NANR200). 2009.

[40] C. Guastavino. Categorization of environmental sounds. *Canadian journal of experimental psychology*, 61(1):54–63, March 2007.

[41] J-J. Aucouturier, B. Defreville, and F. Pachet. The bag-of-frames approach to audio pattern recognition: a sufficient model for urban soundscapes but

not for polyphonic music. *The Journal of the Acoustical Society of America*, 122(2):881–91, 2007.

[42] R.J. Mooney and R. Bunescu. Mining knowledge from text using information extraction. *ACM SIGKDD explorations newsletter*, 7(1):3–10, 2005.

[43] S. Imai. Cepstral analysis synthesis on the mel frequency scale. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'83.*, volume 8, pages 93–96. IEEE, 1983.

[44] C.M. Bishop. *Neural networks for pattern recognition*. Oxford University Press, 1995.

[45] L. Ma, D.J. Smith, and B.P. Milner. Context awareness using environmental noise classification. In *8th European Conference on Speech Communication and Technology*, 2003.

[46] J. Ge and K. Hokao. Applying the methods of image evaluation and spatial analysis to study the sound environment of urban street areas. *Journal of Environmental Psychology*, 25(4):455–466, December 2005.

[47] B. Schulte-Fortkamp. The quality of acoustic environments and the meaning of soundscapes. In *Proceedings of the 17th International Conference on Acoustics*, volume 3, 2001.

[48] T. Hashimoto and S. Hatano. Effects of factors other than sound to the perception of sound quality. In *Proceedings of the International Congress on Acoustics, Rome*, 2001.

[49] A. Zeitler and J. HellBruck. Semantic attributes of environmental sounds and their correlations with psychoacoustic magnitudes. *Proceedings of the 17th International Congress on Acoustics, Rome*, 2001.

[50] S. Kuwano and S. Namba. Dimensions of sound quality and their measurement. In *Proceedings of the 17th International Congress on Acoustics*, 2001.

[51] B. Berglund. Perceptual characterization of soundscapes in residential areas. In *Proceedings of the 17th International Congress on Acoustics*, 2001.

[52] W.W. Hsieh. Nonlinear principal component analysis. In Sue Ellen Haupt, Antonello Pasini, and Caren Marzban, editors, *Artificial Intelligence Methods in the Environmental Sciences*, chapter 8, pages 173–190. Springer Netherlands, Dordrecht, 2009.

[53] S. Kolenikov and G. Angeles. The use of discrete data in pca: Theory, simulations, and applications to socioeconomic indices, 2004.

[54] M. Candel. Recovering the metric structure in ordinal data: Linear versus nonlinear principal components analysis. *Quality & Quantity: International Journal of Methodology*, 35(1):91–105, February 2001.

[55] C. Guastavino, B.F.G. Katz, J-D. Polack, D.J. Levitin, and D. Dubois. Ecological validity of soundscape reproduction. *Acta Acustica united with Acustica*, 91(2):333–341, 2005.

[56] S. Viollon, C. Lavandier, and C. Drake. Influence of visual setting on sound ratings in an urban environment. *Applied Acoustics*, 63(5):493 – 511, 2002.

[57] B. Berglund, P. Hassmén, and A. Preis. Annoyance and spectral contrast are cues for similarity and preference of sounds. *Journal of Sound and Vibration*, 250(1):53–64, February 2002.

[58] S. Kuwano, S. Namba, and T. Kato. Auditory memory and evaluation of environmental sounds. *The Journal of the Acoustical Society of America*, 123(5):3159, 2008.

[59] H. Mller. Fundamentals of binaural technology. *Applied Acoustics*, 36(34):171 – 218, 1992.

[60] C. Guastavino and B.F.G. Katz. Perceptual evaluation of multi-dimensional spatial audio reproduction. *The Journal of the Acoustical Society of America*, 116(2):1105, 2004.

[61] F. Rumsey. Spatial quality evaluation for reproduced sound: Terminology, meaning, and a scene-based paradigm. *J. Audio Eng. Soc*, 50(9):651–666, 2002.

[62] K. Shirima, O. Mukasa, J.A. Schellenberg, F. Manzi, D. John, A. Mushi, M. Mrisho, M. Tanner, H. Mshinda, and D. Schellenberg. The use of personal digital assistants for data entry at the point of collection in a large household survey in southern tanzania. *Emerging Themes in Epidemiology*, 4(1):5, 2007.

[63] S. Lane, N. Heddle, E. Arnold, and I. Walker. A review of randomized controlled trials comparing the effectiveness of hand held computers with paper methods for data collection. *BMC medical informatics and decision making*, 6(1):23, 2006.

[64] M. Tomlinson, W. Solomon, Y. Singh, T. Doherty, M. Chopra, P. Ijumba, A.C. Tsai, and D. Jackson. The use of mobile phones as a data collection tool: A report from a household survey in south africa. *BMC medical informatics and decision making*, 9(1):51, 2009.

[65] J. Anhøj and C. Møldrup. Feasibility of collecting diary data from asthma patients through mobile phones and sms (short message service): response rate analysis and focus group evaluation from a pilot study. *Journal of medical Internet research*, 6(4), 2004.

[66] S. Gaonkar, J. Li, R.R. Choudhury, L. Cox, and A. Schmidt. Micro-blog: sharing and querying content through mobile phones and social participation. In *Proceeding of the 6th international conference on Mobile systems, applications, and services*, pages 174–186. ACM, 2008.

[67] H. Lu, W. Pan, N.D. Lane, T. Choudhury, and A.T. Campbell. Soundsense: scalable sound sensing for people-centric applications on mobile phones. In *Proceedings of the 7th international conference on Mobile systems, applications, and services*, MobiSys '09, pages 165–178, New York, NY, USA, 2009. ACM.

[68] N. Eagle, A.S. Pentland, and D. Lazer. Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences*, 106(36):15274–15278, 2009.

[69] M. Azizyan, I. Constandache, and R.R. Choudhury. Surroundsense: mobile phone localization via ambience fingerprinting. In *Proceedings of the 15th annual international conference on Mobile computing and networking*, MobiCom '09, pages 261–272, New York, NY, USA, 2009. ACM.

[70] G. MacKerron. Mappiness. Online: http://www.mappiness.org.uk, Accessed: August 2012.

[71] E.G. Guba. *The paradigm dialog*. Sage Publications, Incorporated, 1990.

[72] J.M. Morse. Approaches to qualitative-quantitative methodological triangulation. *Nursing research*, 40(2):120–123, 1991.

[73] L.A. Curry, I.M. Nembhard, and E.H. Bradley. Qualitative and mixed methods provide unique contributions to outcomes research. *Circulation*, 119(10):1442–1452, 2009.

[74] N.L. Leech and A.J. Onwuegbuzie. A typology of mixed methods research designs. *Quality & quantity*, 43(2):265–275, 2009.

[75] J.L. Kincheloe. Describing the bricolage: Conceptualizing a new rigor in qualitative research. *Qualitative Inquiry*, 7(6):679–692, 2001.

[76] K. Nyunt. Waterfront soundscape of auckland. In *18th International Congress of Acoustics*, 2004.

[77] W. Yang and J. Kang. Soundscape and sound preferences in urban squares: A case study in sheffield. *Journal of Urban Design*, 10(1):61–80, 2005.

[78] Wunderground. Wunderground weather API. Online: http://www.wunderground.com/weather/api/, Accessed: November 2012.

[79] D. Schreckenberg. The association between residential quality of life and aircraft noise annoyance around frankfurt airport. In *Acoustics 08, Paris*, 2008.

[80] Nestoria. Nestoria API. Online: http://www.nestoria.co.uk/help/api, Accessed: October 2012.

[81] C.E. Osgood, G.J. Suci, and P. Tannenbaum. *The measurement of meaning*, volume 47. University of Illinois Press, 1967.

[82] G.R. Kidd and C.S. Watson. The perceptual dimensionality of environmental sounds. *Noise Control Engineering Journal*, 51(4):216–231, 2003.

[83] P.N. Dokmeci and J. Kang. Objective parameters for acoustic comfort in enclosed spaces. In *20th International Congress on Acoustics, ICA 2010*, 2010.

[84] D. Dubois, C. Guastavino, and M. Raimbault. A cognitive approach to urban soundscapes: Using verbal data to access everyday life auditory categories. *Acta Acustica united with Acustica*, 92(6):865–874, 2006.

[85] H.K. Park T.K. Lee S-W Kim, G.G. Song. Effects of transportation noise exposure time on the subjective response. In *20th International Congress on Acoustics, ICA 2010*, 2010.

[86] B. De Coensel and D. Botteldooren. The quiet rural soundscape and how to characterize it. *Acta Acustica united with Acustica*, 92(6):887, 2006.

[87] T. Choudhury, S. Consolvo, B. Harrison, J. Hightower, A. LaMarca, L. LeGrand, A. Rahimi, A. Rea, G. Bordello, B. Hemingway, P. Klasnja, K. Koscher, J.A. Landay, J. Lester, D. Wyatt, and D. Haehnel. The mobile sensing platform: An embedded activity recognition system. *Pervasive Computing, IEEE*, 7(2):32 –41, april-june 2008.

[88] T.E. Starner. *Wearable computing and contextual awareness*. PhD thesis, Citeseer, 1999.

[89] mobiThinking. Global mobile statistics 2012 part A: Mobile subscribers, handset market share, mobile operators. Online: http://mobithinking.com/mobile-marketing-tools/latest-mobile-stats/a#phone-shipments, Accessed: June 2012.

[90] Microsoft. Windows phone. Online: http://www.microsoft.com/windowsphone, Accessed: July 2012.

[91] Research In Motion. Blackberry OS. Online: http://www.blackberryos.com, Accessed: July 2012.

[92] Apple. Apple iOS. Online: http://www.apple.com/uk/ios/, Accessed: July 2012.

[93] Google. Android. Online: http://www.android.com, Accessed: July 2012.

[94] Java. JSR 135 mobile media API - final release. Online: http://www.jcp.org/aboutJava/communityprocess/final/jsr135/, Accessed: February 2010.

[95] M. Marinilli. *Java deployment with JNLP and WebStart*. Sams Publishing, 2002.

[96] Nuance Communications, Inc. T9 text input: The global standard for mobile text input. Online: http://www.nuance.com/for-business/by-product/t9/, Accessed: September 2011.

[97] Google. Google maps API family - Google code. Online: http://code.google.com/apis/maps/, 2011.

[98] A. Barth. HTTP state management mechanism. *Internet Engineering Task Force, University of California*, (2070-1721), 2011.

[99] J. Zeman, S. Dance, and D. Abraham. Acoustic uses for the iphone. In *Proceedings of the Institute of Acoustics & Belgium Acoustical Society*, volume 3. IOA, 2010.

[100] BS EN 24869-1. Acoustics. hearing protectors. sound attenuation of hearing protectors. subjective method of measurement, 1991.

[101] BS EN 60268-4. Sound system equipment - part 4: Microphones, 2010.

[102] Thawte. Thawte code signing certificates, 2010.

[103] D. Collins. Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research*, 12(3):229–238, 2003.

[104] M. Goodchild and J.A. Glennon. Crowdsourcing geographic information for disaster response: a research frontier. *International Journal of Digital Earth*, 3(3):231–241, September 2010.

[105] J. Silvertown. A new dawn for citizen science. *Trends in ecology & evolution*, 24(9):467–71, September 2009.

[106] C.J. Goodwin. *Research in psychology: Methods and design*. John Wiley and Sons, 2009, 6 edition, 2009.

[107] D.C. Hildum and R.W. Brown. Verbal reinforcement and interviewer bias. *The Journal of Abnormal and Social Psychology*, 53(1):108, 1956.

[108] S.T. Cavusgil and L.A. Elvey-Kirk. Mail survey response behavior: A conceptualization of motivating factors and an empirical study. *European Journal of Marketing*, 32(11/12):1165–1192, 1998.

[109] U.D. Reips. Standards for internet-based experimenting. *Experimental Psychology*, 49(4):243–256, 2002.

[110] A. Furnham. Response bias, social desirability and dissimulation. *Personality and individual differences*, 1986.

[111] A. Frick, M.T. Bächtiger, and U.D. Reips. Financial incentives, personal information and drop-out rate in online studies. *Current Internet Science-Trends, Techniques, Results. Online Press, Zurich*, 1999.

[112] MATLAB. *Version 7.9.0 (R2009b)*. The MathWorks Inc., Natick, Massachusetts, 2009.

[113] O. Lartillot and P. Toiviainen. A matlab toolbox for musical feature extraction from audio. In *Proceedings of the 10th International Conference on Digital Audio Effects, Bordeaux, France*, pages 1–8, Bordeaux, 2007.

[114] K. Rajagopal, P. Minnick, and C. Leider. Automatic soundscape classification via comparative psychometrics and machine learning. In *Audio Engineering Society Convention 131*, 10 2011.

[115] B. Berglund. Relationship between loudness and annoyance for ten community sounds. *Environment International*, 1990.

[116] J. Kang. *Urban Sound Environment*. Taylor & Francis, 2006.

[117] E. Zwicker. A method for calculation of loudness. *Acustica*, 10:304–308, 1960.

[118] E. Zwicker and B. Scharf. A model of loudness summation. *Psychological Review*, 72(1):3–26, 1965.

[119] E. Zwicker and H. Fastl. Method for calculating loudness level, 1975.

[120] B. Berglund. Scaling loudness, noisiness, and annoyance of community noises. *Journal of the Acoustical Society of America*, 1976.

[121] E. Zwicker and Hugo. Fastl. *Psychoacoustics: Facts and Models*. Springer, 2 edition, 2001.

[122] G. Von Bismarck. Sharpness as an attribute of the timbre of steady sounds. *Acustica*, 1974.

[123] B. Gygi and V. Shafiro. Development of the database for environmental sound research and application (DESRA): Design, functionality, and retrieval considerations. *EURASIP Journal on Audio, Speech, and Music Processing*, 2010, 2010.

[124] G. Tzanetakis and P. Cook. Musical genre classification of audio signals. *Speech and Audio Processing, IEEE*, 10(5):293–302, 2002.

[125] J. Saunders. Real-time discrimination of broadcast speech and music. In *Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings*, volume 2, pages 993–996. IEEE, 1996.

[126] F. Guyot, M. Castellengo, and S. Viollon. Perceptive characterisation of the acoustical quality of real complex sounds-validation with synthesis. *ACUSTICA*, 82:78–78, 1996.

[127] P. Vary and R. Martin. *Digital speech transmission: enhancement, coding and error concealment*. John Wiley and Sons, 2006, 2006.

[128] B. Logan. Mel frequency cepstral coefficients for music modeling. *International Symposium on Music Information Retrieval*, 2000.

[129] B. Lee. A new algorithm to compute the discrete cosine transform. *Acoustics, Speech and Signal Processing*, 32(6):1243 – 1245, 2003.

[130] R. Plomp. Tonal consonance and critical bandwidth. *The Journal of the Acoustical Society of America*, 38(4):548–560, 1965.

[131] K. Jensen. *Timbre models of musical sounds*. PhD thesis, University of Copenhagen, 1999.

[132] C. Shannon and W. Weaver. The mathematical study of communication. *Urbana, IL: University of Illinois Press*, 1949.

[133] I. Paraskevas, S.M. Potirakis, and M. Rangoussi. Natural soundscapes and identification of environmental sounds: A pattern recognition approach. In *Digital Signal Processing, 2009 16th International Conference*, pages 1–6. IEEE, 2009.

[134] J.A. Ballas. The niche hypothesis: implications for auditory display design. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 44, pages 718–721. SAGE Publications, 2000.

[135] J. Sueur, S. Pavoine, O. Hamerlynck, and S. Duvail. Rapid acoustic survey for biodiversity appraisal. *PLoS ONE*, 3(12):e4065, 12 2008.

[136] Y. Osada. An overview of health effects on noise. *Journal of Sound and Vibration*, 127(3):407–410, 1988.

[137] T.H. Park. *Introduction to digital signal processing: Computer musically speaking*. World Scientific Publishing Company Incorporated, 2010.

[138] O. Lartillot. Mirtempo: tempo estimation through advanced frame-by-frame peaks tracking. 2010.

[139] H. Rubenstein and J.B. Goodenough. Contextual correlates of synonymy. *Communications of the ACM*, 8(10):627–633, 1965.

[140] G. Nunberg. The non-uniqueness of semantic solutions: Polysemy. *Linguistics and Philosophy*, 3(2):143–184, 1979.

[141] S. Deerwester, S.T. Dumais, G.W. Furnas, T.K. Landauer, and R. Harshman. Indexing by latent semantic analysis. *Journal of the American society for information science*, 41(6):391–407, 1990.

[142] G.W. Furnas, S. Deerwester, S.T. Dumais, T.K. Landauer, R.A. Harshman, L.A. Streeter, and K.E. Lochbaum. Information retrieval using a singular value decomposition model of latent semantic structure. In *Proceedings of the 11th annual international ACM SIGIR conference on Research and development in information retrieval - SIGIR '88*, pages 465–480, New York, New York, USA, 1988. ACM Press.

[143] D. Zeimpekis and E. Gallopoulos. TMG: A MATLAB toolbox for generating term-document matrices from text collections. *Grouping multidimensional data: Recent advances in clustering*, pages 187–210, 2006.

[144] A. Field. *Discovering statistics using SPSS*. Sage Publications Limited, 2009.

[145] Flurry. Flurry blog. Online: http://blog.flurry.com, Accessed: October 2012.

[146] J. Stott. Earthtools.org - Sunrise & sunset times API. Online: http://www.earthtools.org/webservices.htm, Accessed: November 2012.

[147] G. Stasser and S.I. Vaughan. *Understanding Group Behavior: Consensual Action by Small Groups*, volume 1, chapter Models of Participation during face to face unstructured discussion, pages 165 – 192. Routledge, 1996.

[148] ComScore. iPhone gains share among affluents, women, older consumers. Online: http://www.emarketer.com/Article/iPhone-Gains-Share-Among-Affluents-Women-Older-Consumers/1008635, Accessed: June 2011.

[149] W.J. Davies, M.D. Adams, N.S. Bruce, R. Cain, A. Carlyle, P. Cusack, D.A. Hall, K.I. Hume, A. Irwin, P. Jennings, et al. Perception of soundscapes: An interdisciplinary approach. *Applied Acoustics*, 2012.

[150] J. Feinberg. Wordle.net. Online: http://www.wordle.net/create, Accessed: November 2012.

[151] C. McNaught and P. Lam. Using wordle as a supplementary research tool. *The qualitative report*, 15(3):630–643, 2010.

[152] D.A. Hall, A. Irwin, M. Edmondson-Jones, S. Phillips, and J.E.W. Poxon. An exploratory evaluation of perceptual, psychoacoustic and acoustical properties of urban soundscapes. *Applied Soundscapes: Recent Advances in Soundscape Research*, 74(2):248 – 254, 2013.

[153] M. Raimbault, C. Lavandier, and M. Bérengier. Ambient sound assessment of urban environments: field studies in two french cities. *Applied Acoustics*, 64(12):1241–1256, 2003.

[154] S. Viollon and C. Lavandier. Multidimensional assessment of the acoustic quality of urban environments. In *Proceedings of The 29th International Congress on Noise Control Engineering*, August 2000.

[155] K. Kawai, T. Kojima, K. Hirate, and M. Yasuoka. Personal evaluation structure of environmental sounds: experiments of subjective evaluation using subjects own terms. *Journal of Sound and Vibration*, 277(3):523 – 533, 2004.

[156] J.D. Guilln and I. Lpez-Barrio. The soundscape experience. Madrid, 2007. International Congress on Acoustics.

[157] B. Schulte-Fortkamp and A. Fiebig. Soundscape analysis in a residential area: An evaluation of noise and people's mind. *Acta Acustica united with Acustica*, 92(6):875–880, 2006.

[158] B. Schulte-Fortkamp. How to measure soundscapes, a theoretical and practical approach. *Acoustical Society of America Journal*, 112:2434–2434, 2002.

[159] M. Nilsson. Acoustic indicators of soundscape quality and noise annoyance in outdoor urban areas. In *Proceedings of the 19th International Congress on Acoustics, Madrid*, 2007.

[160] A.J. Torija, D.P. Ruiz, B.D. DeCoensel, D. Botteldooren, B. Berglund, and Á. Ramos-Ridao. Relationship between road and railway noise annoyance and overall indoor sound exposure. *Transportation Research Part D: Transport and Environment*, 16(1):15–22, 2011.

[161] H. Abdi and L.J. Williams. Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(4):433–459, 2010.

[162] J.P. Stevens. *Applied multivariate statistics for the social sciences*. L. Erlbaum, 2001.

[163] D.G. Altman and P. Royston. Statistics notes: the cost of dichotomising continuous variables. *BMJ: British Medical Journal*, 332(7549):1080, 2006.

[164] G. Brambilla and L. Maffei. Responses to noise in urban parks and in rural quiet areas. *Acta Acustica united with Acustica*, 92(6):881–886, 2006.

[165] K. Foale and W.J. Davies. A listener-centred approach to soundscape evaluation. In *Acoustics 2012 - Nantes*, 2012.

[166] B.K. Kahn, D.M. Strong, and R.Y. Wang. Information quality benchmarks: product and service performance. *Communications of the ACM*, 45(4):184–192, 2002.

[167] J. Hunter, A. Alabri, and C. Van-Ingen. Assessing the quality and trustworthiness of citizen science data. *Concurrency and Computation: Practice and Experience*, 2012.

[168] D.R. Sanqunetti. Implementing geo-fencing on mobile devices, April 13 2004. US Patent 6,721,652.

[169] I. Garcia-Mart, L.E. Rodrguez, M. Benedito, S. Trilles, A. Beltrn, L. Daz, and J. Huerta. Mobile application for noise pollution monitoring through gamification techniques. In *Entertainment Computing - ICEC 2012*, volume 7522 of *Lecture Notes in Computer Science*, pages 562–571. Springer Berlin Heidelberg, 2012.

[170] K. O'Hara. Understanding geocaching practices and motivations. In *Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 1177–1186. ACM, 2008.

[171] G. Memoli, G. Licitra, and M. Cerchiai. Perspectives for a strategical mapping of soundscapes. *Journal of the Acoustical Society of America*, 2008.

[172] B. De Coensel. 1/f noise in rural and urban soundscapes. *Acta Acustica united with Acustica*, 2003.

[173] D. Botteldooren, B. De Coensel, and T. Demuer. The temporal structure of urban soundscapes. *Journal of Sound and Vibration*, 292(1-2):105–123, 2006.