

# SoundShade

A creative tool designed to help users reduce the negative effects of unwanted sound in their working environment.



Scott McGregor  
Master of Arts - Sound Design in New Media  
Aalto University - Media Lab  
Supervisor - Antti Ikonen  
2021

# Contents

<b>Contents</b>	<b>1</b>
<b>Abstract</b>	<b>2</b>
<b>Chapter 1</b>	<b>3</b>
<b>1.1 Introduction</b>	<b>4</b>
1.1.1 Thesis Purpose	4
1.1.2 Team Composition and roles	5
<b>Chapter 2</b>	<b>6</b>
<b>2.1 Identifying a Problem</b>	<b>7</b>
2.1.1 Sound and Noise	7
<b>2.2 Auditory Processing</b>	<b>11</b>
2.2.1 Masking	12
2.2.2 Cocktail Party Phenomenon	12
2.2.3 Further Reading on Understanding Auditory Distractions	13
<b>2.3 The Effects of Nature on Stress</b>	<b>14</b>
<b>2.4 Moving Forwards</b>	<b>15</b>
<b>Chapter 3</b>	<b>17</b>
<b>3.1 Early Prototypes</b>	<b>18</b>
3.1.1 Prototype 2 - The Minimum Viable Product	20
<b>3.2 Seamless Looping</b>	<b>22</b>
<b>3.3 Capture &amp; Acquisition – Sound Recording</b>	<b>25</b>
<b>3.4 Audio Manipulation - A New Method</b>	<b>27</b>
3.4.1 Multi-band Compression	28
3.4.2 Multi-layered Frequency Emphasis.	29
<b>3.5 HRTF - Head Related Transfer Function</b>	<b>31</b>
<b>3.6 User Interface</b>	<b>33</b>
<b>3.7 Further Development</b>	<b>35</b>
3.7.1 Stereo Vs Mono	35
3.7.3 Healthcare Product	36
3.7.4 Artificial Masking	36
<b>Chapter 4</b>	<b>38</b>
<b>4.1 Conclusion</b>	<b>39</b>
<b>References</b>	<b>40</b>

---

**Author** Scott McGregor

---

**Title of thesis** SoundShade - A creative tool designed to help users reduce the negative effects of unwanted sound in their working environment.

---

**Department** Media Lab, Department of Media

---

**Degree programme** Master's Degree Programme in Sound in New Media

---

**Year** 2021

**Number of pages** 41

**Language** English

---

Modern day working environments see many people's work and study places being transformed into more cost efficient open-plan layouts. With this reduce in cost comes a reduction in privacy - specifically sound privacy, and as a result a degradation in working performance. In this paper we will look at how noise, later defined in this paper as 'unwanted sound', can negatively affect a person's psychological and physiological state and then offer a solution to help combat these negative effects via the use of a mobile application. The SoundShade app contains various sounds of nature along with other non-intrusive sounds that were designed in accordance with background research efforts, outlined in this paper, in an attempt to offer a person the ability to create their own preferable sound environment, hopefully resulting in an increase in their working performance and focus, with an increased ability to recover from stressors. The paper will look deeply into the design process of the SoundShade app, the sounds found within, and how certain problematic auditory characteristics were avoided.

---

**Keywords:** Noise, Sound, Workplace-performance, Nature-sounds, Sound Design, Mobile App, Ambience, HRTEF,

---

# Chapter 1

# 1.1 Introduction

Our lives are composed of the things we experience, the things we perceive. Our five main senses allow us to feel, see, smell, taste and hear the world in which we live. Due to urbanization and cities becoming increasingly populated, our senses are bombarded with information every minute of every day; and being able to give them a break is more difficult than ever. The only complete down-time our senses seem to get is when we fall to sleep. However, this is only partly true when it comes to our hearing, for our brain still listens while we sleep (Gondek 1998). The brain uses an additional, secondary area, the frontal lobe, to listen when we are sleeping.

*"The frontal lobe is believed to play a key role in vigilance functions, such as screening new stimuli and preparing the body to react. During sleep this part of the brain may analyse sounds to decide whether the person needs to be awakened to respond. This mechanism would allow a camper in the woods to sleep through non-threatening cricket chirps. But it might awaken the camper quickly to the growl of a bear."*

[S. J. Gondek (1998) - How Do We Hear While We Sleep?]

The perception of sound and noise, then, seems to be extremely high up on the scale of sensory priority for the brain. It has developed mechanisms of dealing with what it deems unimportant, in terms of noise, and is able to filter it out to some degree. Yet, there are disadvantages that arise in the systems it uses, especially when applied to modern day society. This is the core inspiration for what will follow throughout this document.

## 1.1.1 Thesis Purpose

The aim of this thesis was to design a mobile application, hereon referred to as SoundShade that would, in theory, aid a person in their ability to focus, by reducing their exposure to problematic noises during work hours or study time when their surrounding sonic environment was of an undesirable or distracting nature. This would be achieved by creating a mobile application consisting of various different ambient and atmospheric sound samples taken from the natural and urban worlds that could be individually mixed, combined and positioned in a virtual surround

environment with any other contained in the app. The sounds would be designed from a scientific standpoint and be based on existing background research on the topic of how noise and sound stimuli can have measurable negative effects on a person's psychology. Though separated into its own individual chapter of the thesis, the background research lead the way in identifying and understanding the problem that SoundShade would attempt to aid, for the information gathered would form a fundamental stepping stone in being able to achieve ambient sounds that were designed in such a way that they did not further emphasize the problem.

Releasing SoundShade into the mobile application marketplace for download by the general public would be the goal in regards to how 'complete' the application would need to be. With this in mind the mechanics of the user interface, sound design, and programming needed to be far beyond the prototype stage. Furthermore, the amount of content would need to be substantial and varied to provide any one user with unique customization options while designing their preferred working environment.

### 1.1.2 Team Composition and roles

The SoundShade project was developed by a team of three people. Here are their roles within, and contributions to, the project.

**Scott McGregor** - Sound designer, background researcher and analyser, thesis writer, developer.

Scott was the sole writer of the thesis and was responsible for undertaking the analysis of the background research that would lead to the specific design process of the sounds found within the project.

**Valeria Gasik** - Visual content designer and interaction developer.

**Valtteri Wikström** - Programmer, and interaction developer.

# Chapter 2

Background Research -

Is a person's working performance negatively affected by unwanted external audio stimuli, and is it possible to reduce these negative effects by providing people access to more preferable sound environments?

## 2.1 Identifying a Problem

### 2.1.1 Sound and Noise

Technically sound and noise are the same physical phenomenon, they are the result of kinetic energy causing air particles to compress and retract creating waves. These waves are what enter our ears. Though, they are separated by meaning in regards to how they are perceived. Noise is often defined to be ‘unwanted sound’ with sound simply being a word used to describe what one hears. The only place known to man where sound does not and cannot exist is in the confines of a vacuum. Even then sound can travel via other physical mediums such as the water or liquids in our own body. This can be experienced in the more advanced of anechoic chambers, such as the one located at Orfield Laboratories which absorbs 99.9% of all sound. The president there states that we can still hear sounds that are created from the internal workings of our own bodies.

*“When it’s quiet, ears will adapt. The quieter the room, the more things you hear. You’ll hear your heart beating, sometimes you can hear your lungs, hear your stomach gurgling loudly”.*

[Orfield, S. (2012)]

Regardless of one’s surroundings, sound is always present and is notoriously difficult to avoid. By utilizing specialized equipment offices and hotels are able to suppress the ever increasing noise produced by their city's traffic population. Triple glazed windows, heavy curtains and even special acoustically absorbent walls made with dense materials offer protection from high noise levels found in the streets outside. Though not all establishments are fortunate to have the opportunity, nor the funding, to apply such technologies, yet they must still endure noise produced by external sources.

The World Health Organization (W.H.O) published a report that stated that around half of European Union citizens are estimated to live in areas that have not been deemed acoustically comfortable for the population (W.H.O - 1999) These figures are from what W.H.O defines as “community noise” which is noise originating from any source, excluding that from industrial workplaces. It further

defines the main “*indoor sources of noise*” are caused by ventilation systems, office machines, home appliances and neighbours (W.H.O - 1999).

Once there was a time when an office worker would turn up for their nine-to-five shift, sit at their computer in the relative privacy of their own room/cubicle and get on with whatever it was they were hired to do. Today, however, sees more and more employers and business owners adopting something called the ‘open-plan office’. A main reason for this popularity is that it is cheaper. An example of this is Accenture, one of the world’s leading consultancy services, an early adopter of the open plan environment. They were able to reduce their spatial footprint from 23,000 square feet down to 8,600 (Puzzo, M. 2014).

After this huge saving on office space and, in turn, cost, what of the office workers’ personal space? It is unclear from reading the report whether or not a substantial amount or any of the workforce had been let go as a result of this space saving scheme. Let us presume that the company would still need the same amount of work to be attended to as before the transformation. Thus they would still need the same sized workforce that was previously present. If this is the case then each worker would see a dramatic drop in their own privately available area. The same journal from Heartford Business, also stated that a ‘CoreNet survey’ predicts that the average square foot of space per worker will drop from 225 in 2010 down to 152 by 2017 (Puzzo, M. 2014). With the same amount of people working in a smaller area designed to be open-plan, will see the amount of personal privacy of each worker somewhat diminished.

It is this lack of privacy that seems to be the main issue, reported from those working in open environments, that makes work life dissatisfying. It is not just the visual aspect of privacy, the fact that colleagues can freely see what one another are doing, but also the auditory aspect of having no privacy in terms of what can be heard. Fig.1. represents the percentage of people who are dissatisfied with different elements of working in a shared office space. The lack of sound privacy extends towards the 60% margin even when the office space is separated by high partitions.

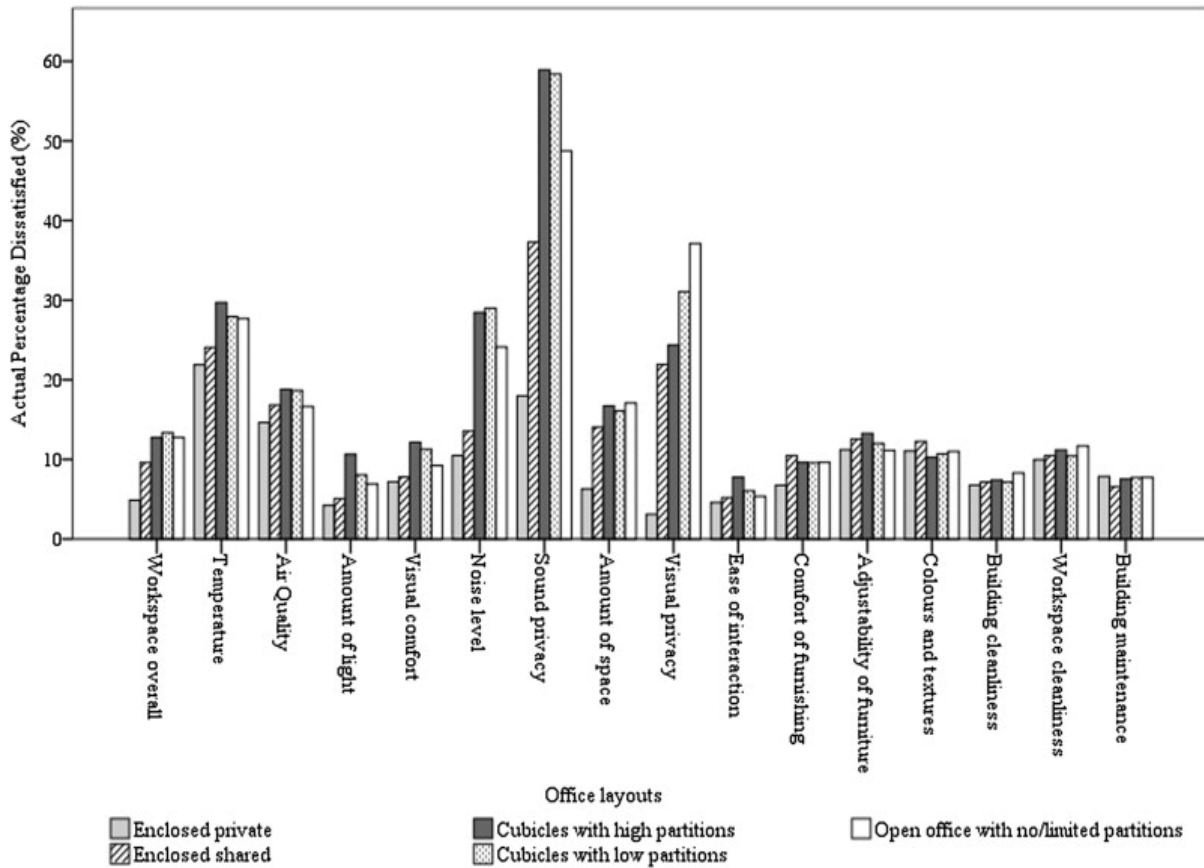


Fig.1 - Actual Percentage of Dissatisfied (APD) for IEQ questionnaire items by office layout configurations. Source: [sciencedirect.com/science/article/abs/pii/S0272494413000340](https://www.sciencedirect.com/science/article/abs/pii/S0272494413000340)

*“Our results categorically contradict the industry-accepted wisdom that open-plan layout enhances communication between colleagues and improves occupants’ overall work environment satisfaction. The open-plan proponents’ argument that open-plan improves morale and productivity appears to have no basis in the research literature,”*

[Jungsoo Kim and Richard de Dear (2013)]

Further research, such as the paper entitled ‘The Privacy Crisis’ (Steelcase) conducted by Steelcase’s ‘Workspace Futures’ team, reported that a mere 11% of 10,500 people surveyed across Asia, Europe and North America, were satisfied with their open plan workplace due to its distracting nature. This *“confirms that insufficient privacy in the workplace is an issue throughout the world”*.

This lack of workplace satisfaction leads to workers feeling disengaged and the cost of this disengagement in the US alone is approximated to be around \$450 – 550 billion per year. Moreover, Canadian Life Group Insurance research <sup>9</sup> found that people working in open offices took around 75% more sick days than those working from home.

---

<b>Cost of Disengagement</b>	
USA	<b>\$450 – 550B</b>
Germany	<b>€112 – 138B</b>
Australia	<b>\$54.8B</b>
United Kingdom	<b>£52 – 70B</b>

---

Fig.2 - [2013 State of the Global Workplace Report, Gallup](#)

With lower costs and other economical benefits, the chances of more open environments popping up across the globe is high. This could see exposure to unwanted noise rising and along with it the negative health issues that it promotes. If the common workplace is going to be a noisy one, solutions must be researched and implemented in a hope to negate or, at the very least, dramatically reduce the impact on the working population. Though, before one can hope to achieve such solutions we must first delve into why and how noise causes these issues.

## 2.2 Auditory Processing

Throughout history humans have had to adapt and become accustomed to what can be heard during everyday life. Once a sound is heard by the mechanical components of the inner ear, it is carried to the auditory cortex via the cochlear nerve (Landau & Barner, 2009) where it is then translated into meaning. It is, however unclear on what exactly happens to the auditory information once at the auditory cortex.

*“The cortex is so complex that the most we may ever hope for is to understand it in principle, since the evidence we already have suggests that no two cortices work in precisely the same way.”*

[James W. L. Beament - Biologist]

It is the meaning placed to the auditory information that causes a reaction. Mentally and physically. Hearing a threatening animal growl or a loud clap of thunder, for example, are things that may trigger a response to danger. On the other hand, noises that the brain deems less alarming can also have a measurable effect. For example, when one must work in the same room as an air-conditioning unit that produces consistent sound throughout the day may cause annoyance in those around it. This air-conditioning unit is a good example of sound being classed as noise. A study entitled ‘*Effects of Noise on Productivity*’ (Erret et al) explores the effects of noise in the office. The paper measured an adverse effect on the results from tests undertaken by their test subjects when they were exposed to external audio stimuli of an undesirable nature. If the subjects were annoyed by the sound then their test results were affected more. Moreover, imbalances in the external stimuli’s frequency content were also found to negatively affect a person's performance.

*For example, scores on the typing, math, and verbal reasoning tests tended to decrease when subjects were more annoyed by the noise. Perception of spectral imbalance, such as rumble, roar, and hiss, were also shown to impact performance. These findings support the idea that a more neutral-sounding spectrum is desirable in working environments.*

[Erret et.al. (2006)]

Most fascinating, regarding this topic, is that the study found that the noises were reported to be more annoying earlier in time. (Erret et. al. 2006) This supports a theory that the brain is able to apply a mechanism that reduces the presence of a noise that is not critical to a task currently at hand. This mechanism is referred to as masking.

### 2.2.1 Masking

The Erret et al study supports the theory, presented in the Szalma & Hancock Meta Analysis (2011), that there is some sort of active masking being applied by the brain. It does also prove that a person's work performance is drastically reduced by noise if the person finds it annoying. Though the mind is quite efficient at masking noises that are constant and unchanging, it struggles, when trying to ignore noises that it has evolved to pay attention to and those that are 'intermittent' - noises that are non constant, sudden, short, and louder in volume than that of the average noise around you. A printer suddenly coming to life, the squeak of a door hinge, laughter or words of a conversation from colleagues. These are all noises that the brain struggles to mask, because they are so short it doesn't have the time to apply its masking mechanism as theorized by Szlama and Hancock. Therefore, these are the most distracting types of noise, with speech having the most dramatic impact. One explanation as to why speech is the most distracting noise in the workplace was studied by Moray in 1959 - he coined the term the 'Cocktail Party Phenomenon'.

### 2.2.2 Cocktail Party Phenomenon

The cocktail party phenomenon aims to answer why humans are so easily distracted to external speech outside the task they are currently attending to. Moray (1959) found that around 33% of subjects observed were able to recall or respond to their name being spoken in an unattended message whilst attending a primary one. Yet they were unable to recall simple lists or instructions given in the secondary message. He concluded that humans set up a "block" to reject incoming stimuli on a conscious level of the unattended message so that they can focus or attend to the primary message. Only stimulus of high meaning to the listener such as their name

can break through this block. However, this phenomenon does not hold equally true for the entire population as once thought back in 1959 when Moray first conducted his investigation. The expanded study of Moray's, made by Conway, Cowan and Bunting (1995), found that people with larger working memory spans noticed their name in the unimportant message much less often than those with a smaller span.

Working memory maintains goal-relevant information when a said person is performing a task (Baddley & Hitch 1974 & Miyake & Shah 1999). This means that persons with larger working memory capacities are able to focus on the task at hand without external distractions such as auditory stimuli becoming an issue.

### 2.2.3 Further Reading on Understanding Auditory Distractions

There are many models that have been put forward and studied over the years in regards to working memory and other memory functions such as short term and long term – Atkinson & Shiffrin, Baddeley & Hitch and Ericsson & Kintsch to name the ones used as a foundation to better understanding the brain's memory mechanics. Though their content is theoretical and each challenges or expands and in some cases argues with the others. For this reason I have not included dedicated sections to the different memory functions of the human mind as I feel it will distract and take a tangent from the main goal of this thesis. Nothing on the functions of memory had influence on the final project, rather they provided a wider understanding on how the human mind works in regards to task related performance.

With findings in how the general population are dissatisfied with their auditory surroundings during the working day, how sounds of a distracting nature such as speech and audio content of “a high importance”, and how external auditory stimuli can cause annoyance and inturn be detrimental to working performance, the focus of the background research at this point shifted to trying to discover weather or not sounds of nature could be used during working hours to remedy these negative effects.

## 2.3 The Effects of Nature on Stress

In 1984 Roger S. Ulrich conducted a study on hospital patients during their stay after surgery. The study saw two groups of patients recover from a surgical operation, the difference being that one group had a window that faced a brick wall, the second had a window that faced a natural scene. He found that those with a nature facing window had shorter hospital stays after their operation, evaluative comments from nurses were less negative and required less doses of moderate or strong analgesic doses. (Ulrich 1984).

In a later study Ulrich (1991) claimed that by producing positive emotional states nature sounds were able to increase the restorative states of a person. Since these studies others have been undertaken to better understand the effects of nature on stress recovery in some form or another. One very interesting study was that made by Jesper J. Alverson et. al (2010), entitled - '*Stress Recovery during Exposure to Nature Sound and Environmental Noise*' in which subjects were given a test specifically designed to increase psychological stress levels of both the sympathetic and parasympathetic systems. Skin conductance levels (SLC) were observed to measure the effects nature and noise audio stimulus had on the subjects. After each stress test was completed the subjects were placed in one of four auditory environments, listed below.

1. *Nature sound. A mixture of sounds from a fountain and tweeting birds. The average sound pressure level was set to 50 dB (Laeq,4min).*
2. *High noise. Road traffic noise recorded close to a densely trafficked road. The average sound pressure level was set to 80 dB (Laeq,4min).*
3. *Low noise. The same noise as (2), but set to a lower average sound pressure level, 50 dB (Laeq,4min).*
4. *Ambient noise. A recording from a quiet backyard, with a constant low level ambient noise, mainly caused by ventilation systems of the buildings surrounding the yard. The average sound pressure level was set to 40 dB (Laeq,4min).<sup>18</sup>*

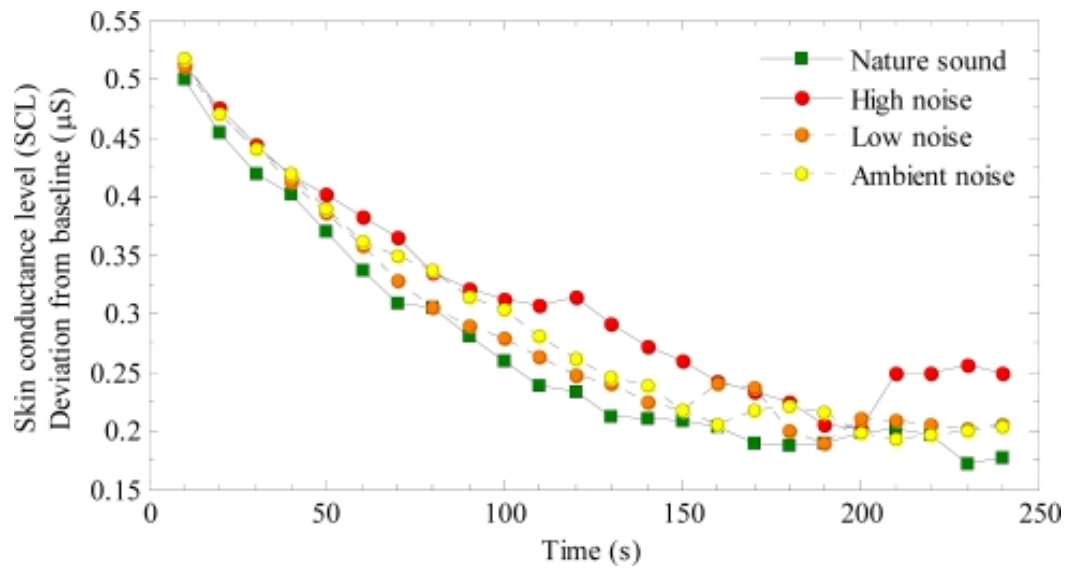


Fig. 3 – Skin Conductance Levels Reducing Over Time - Jesper J. Alverson et. al

The results showed that SCL reduced faster with exposure to nature sounds than any of the other noises. Although the numbers were only of significance between the nature sounds and high noise, the study states that -

*“Although this difference was statistically significant only between the nature sound and the high noise, detailed analyses of the recovery functions showed that half-life SCL recovery was 9–37% faster during the nature sound than during the noises. These results suggest a faster recovery of the sympathetic nervous system during the nature sound.”*

[Jesper J. Alverson et. al (2010)]

The authors conclude that nature sounds are able to reduce recovery time after a psychologically stressful situation.

## 2.4 Moving Forwards

All of the preceding background research formed an understanding of how nature sounds could be utilized to reduce recovery time from psychological stressors, and possibly act as a mask or block for unwanted external noises that have been proven to be a substantial problem in the modern day working environment.

The following part of this thesis will be a report of how SoundShade was designed with this background research in mind. We must pay close attention and

refer back to the findings of these studies as much as possible in an attempt to limit and, where possible, totally remove any of the unwanted audio characteristics that may hinder the effectiveness of the tool we are trying to create. As an example, the study '*Effects of Noise on Productivity*' (Erret et al 2006) found that audio stimulus that contained an imbalance of frequencies, specifically excessive low frequencies and high frequencies or as the study describes them as being rumbly or hissy, to be the most annoying to the subjects. However, some sounds of nature inherently contain these problematic characteristics, such as wind blowing in the trees. The main characteristic of the rustle of leaves is relatively hissy, so the design process must allow for the sounds to maintain their natural qualities whilst at the same time providing the user the ability to tone down the intensity of any one sound to suit their personal preferences. Furthermore, there must be nothing within the sound content that a user may find distracting or annoying based on the finding of the research. For instance, human speech is one sound that must not be present in any of the content, so clean, isolated recordings must be acquired. Failing that any speech must be removed during the editing process. Also, any sonic characteristic that could be deemed as intermittent must be addressed; all abrupt changes in frequency and intensity should be smoothed to a degree that these changes happen gradually, so as to not cause a distraction in the user.

# Chapter 3

Building SoundShade and it's Evolution

## 3.1 Early Prototypes

Before SoundShade was called as such, it was simply an idea. The idea of people having the ability to control the sonic environment surrounding them. With limitations set by the universal laws of physics, and indeed ethical behaviour, being able to adjust and manipulate the noises we must endure from day to day remains somewhat difficult.

With this idea in mind I began to experiment with a programming language called 'App Inventor' created by MIT to construct a basic user interface that would be used to control the sound level of ambient sound files, such as the call of seagulls, ocean waves and rain, for example, that a user could listen to on a mobile device via their headphones - thus creating their own personal sound world.

The original goal was to develop a working mobile prototype for a school course in which I was enrolled. The idea, at the time, was to create nothing more than a toy, something novel that would possibly raise topics of discussion regarding how noisy our world really is. This topic was of some interest to me as a sound designer, though it was not something I spent too much of my time thinking about. It would come to be that twelve months later the topic would be the sole focus of my studies.

For now the prototype began to take form and the functionality worked as desired, though the programming language was suppressing the development in certain areas. By the end of the course a basic functioning version of the application could be run on an Android mobile device. Fig. 4 shows the interface. The circles can be moved along the y-axis to increase or decrease the volume of each sound file independently from one another. Each sound was stored locally and was streamed from the mobile device's internal memory.

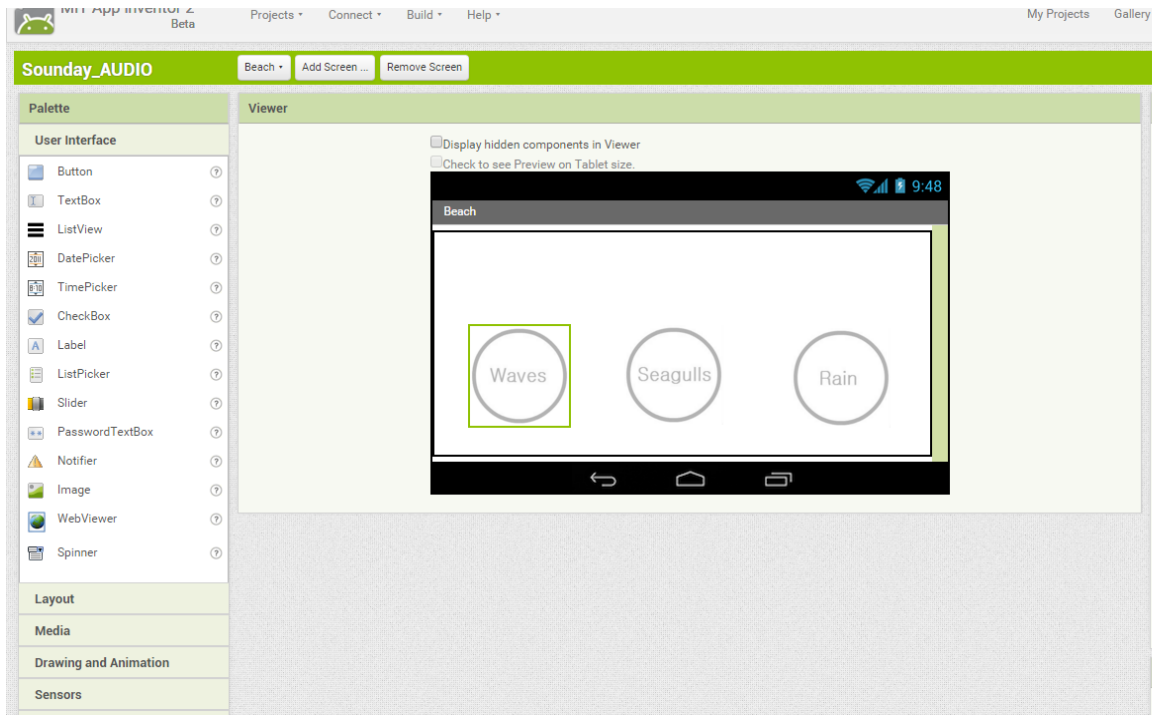


Fig. 4 - The first prototype of SoundShade Inside the App Inventor by MIT

Though the course was complete, I found myself drawn to the documentation that I had uncovered regarding how humans perceive sound and noise and how it has multiple effects upon us. I was transfixed. What had originally started out as a toy would become so much more. I began working on the prototype in my spare time adding performance updates and features to make the app more usable. Fig. 5 shows the location screen. Clicking a location will reveal different sounds that are available for that location. Fig. 6 shows the loading screen which plays a short musical jingle as the application loads its content.

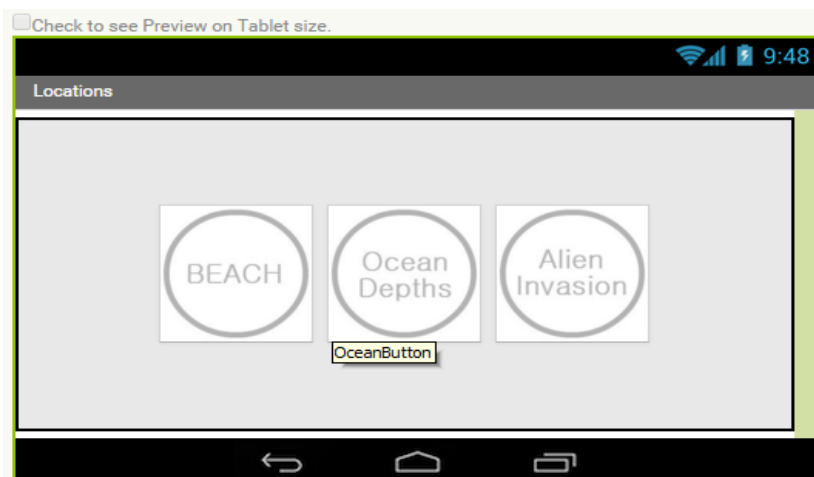


Fig 5 - Location Selection Screen



Fig 6- Loading Screen

Unfortunately, at that time, the progression of the development was halted by the limited control over audio files that the App Inventor language allowed. One of the limitations of the App Inventor language was that the ability to pan sound files in the stereo image was not possible. Panning the audio files would play a huge role in the user being able to craft their perfect sound world, and for it to sound realistic, not to mention any simulated movement automation would not be possible.

With the progression of the prototype slowly coming to a halt due to my basic understanding of different programming languages I needed to find a solution if I was to keep my idea alive and evolving so I set out to find a team that would help me take SoundShade to the next stage of development.

### 3.1.1 Prototype 2 - The Minimum Viable Product

After the SoundShade team grew from one to three, progress started on formulating the fundamental idea of what we wanted to achieve. We all agreed that we would create a mobile application for iPhone and other Apple devices. The team had a mutual agreement that applications developed for the iStore market place were of higher quality as the acceptance and review process was much more stringent than of the Android or Windows counterparts, which, at the time, there were almost no

quality standard requirements. This focused direction helped us pour all of our work into only developing for one platform and not have to worry about technical difficulties that would occur from cross platform development.

The next version of SoundShade would come in the form of a minimum viable product - a term used when developing a product to be market ready as quickly as possible with minimum features. To help increase the speed of our development we applied for a start-up accelerator funded by YLE and Nestholma who would push us to deliver our minimum viable product in just three months. With such a short turnaround time any problems that would arise in our programming, and the app overall, would need to be solved quickly. Something that fortunately fell into our favour was that the development platform we decided to use included the option of using a Head Related Transfer Function (HRTF) algorithm to place sounds in a virtual 3D environment. This feature was very important to us from the beginning of the development stage, as it was a feature that seemed to be missing from the existing products on the market similar to what SoundShade would become. We felt that this feature was a crucial part of the app and it must be included in the minimum viable product. The included HRTF proved to be somewhat effective and certain parameters were available for alteration such as the height of the sound source, and realistic distance roll off simulation - this simulated the sound moving further away from the listener as the volume decreased. This was a feature that added a believable and realistic quality to the sound worlds that would be created and provide a solution to those sounds that contained some of the problematic characteristics discussed in section 2.4. With this taken care of at such an early stage we could move on and focus on other critical problems that would not be acceptable when the time came to prove the viability of the project. One problem that arose almost immediately was getting the audio samples used in the app to loop seamlessly.

## 3.2 Seamless Looping

Normally when an audio sample is set to loop or repeat for continuous playback the end of the sample and the beginning have different frequency content and amplitude points. When the play header jumps from the end of the sample back to the beginning if the waveform does not contain identical data at both the end point and start point it can often cause an artefact to be audible at the point of looping. This artefact can be heard as a click or a pop. It was imperative that this was avoided to achieve our goal of SoundShade being used as a tool to help people focus. If the users were able to identify any artefact or patterns or hear something that sounded unnatural or disturbing within the audio it could lead to distraction or listening fatigue. Fig 7 shows a difference in amplitude in the waveform. The difference in amplitude can be incredibly small, visually, and still produce an unwanted audible artifact. This artefact can be avoided by making sure that the sample starts and ends at the zero crossing. The zero crossing is where the amplitude of a waveform is at zero. This can be done by simply adding a fade at both the end and beginning of the sample as this brings both amplitude points to zero, though the change in frequency content still makes the loop point abrupt and audible to the human ear.

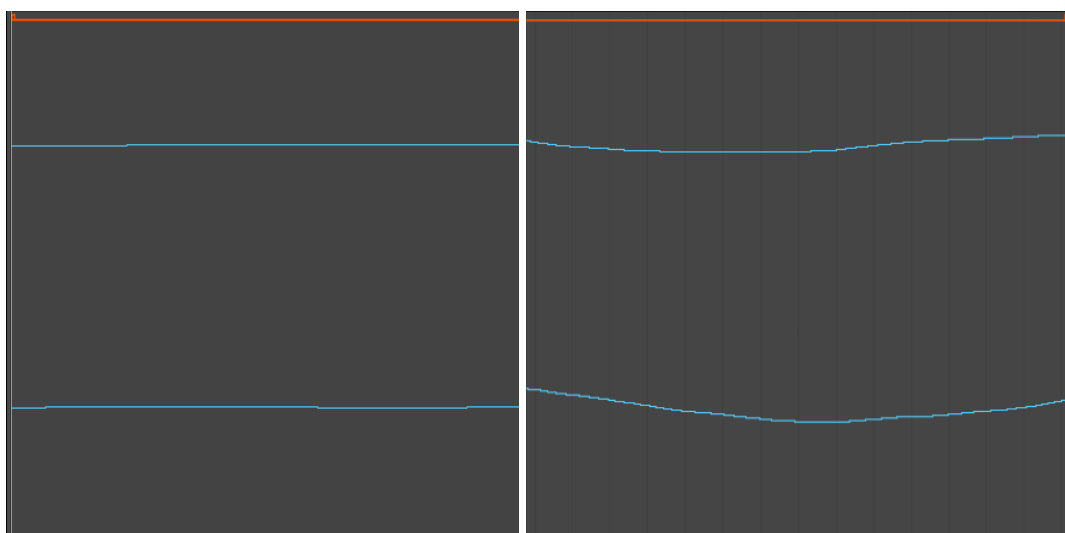


Fig 7 - Differing amplitude points of the end (left) & beginning (right)

The technique that was adopted to seamlessly loop the audio files in SoundShade was one called 'tail-to-front looping'. Fig 8 illustrates how this is achieved. The black highlighted region of audio in fig 8 (top image) indicates an edit cut being made in the sample. Then the block of audio being brought to the front of the sample (middle image). This will create a seamless continuation of the audio wave amplitude points as the play header jumps from the end of the sample back to the beginning. This technique is heavily used in the game audio industry as many sounds such as atmospheres and sound effects which have indefinite lengths, caused by events in the game, are called on a regular basis.

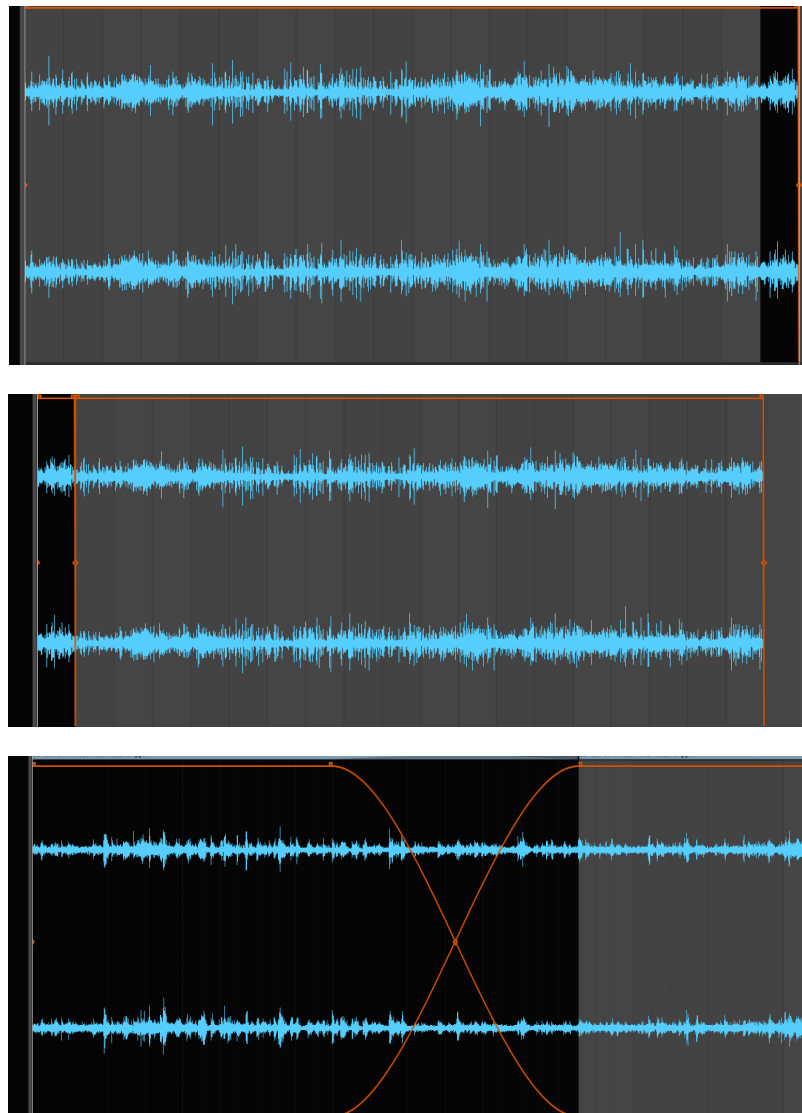


Fig 8. Tail-to -front looping technique

With the samples now seamlessly looping, the next task was to make sure the new start point of each sample would cross fade naturally into the rest of the audio sample that followed it. For example, often it would be the case that the end of a sample would be louder, quieter or have different frequency content than the original beginning location of the sample. Simply cross fading over a long period of time would help to make this transition more natural fig 8 (bottom image). In addition small, but crucial, amplitude changes, seen in fig 9, were automated onto the sample at transient points to ensure maximum smoothness during the cross fade.

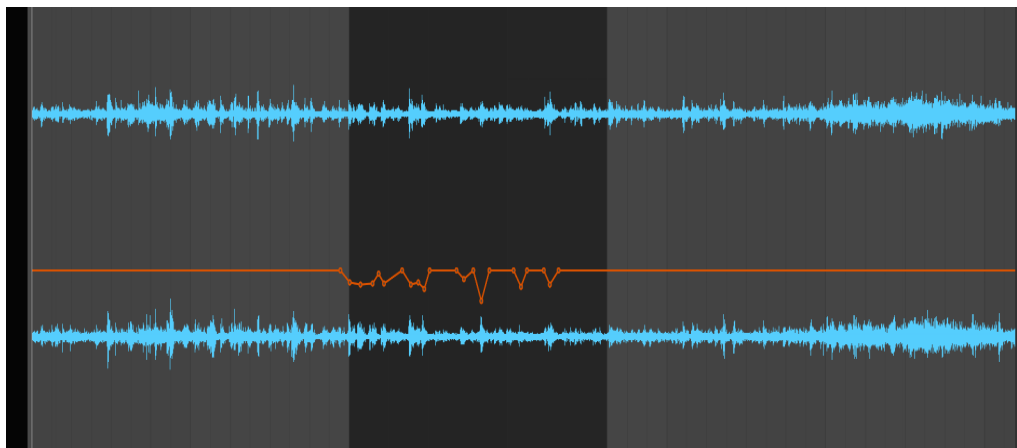


Fig 9. Transient amplitude automation over the crossfaded region

Now that the audio files looped seamlessly they were exported and converted to .MP3 file types ready to be implemented into the current version of the SoundShade application. Upon testing the audio in the application we could hear that the looping point was actually audible with a click or a pop despite when played back in the audio workstation they did not produce this. The files were double checked to ensure they looped seamlessly - and they did. It wasn't clear why the files wouldn't loop without a problem. After a considerable amount of research it would turn out to be an issue caused by MP3 encoders adding a delay, more specifically - a certain number of samples of silence, to the start and end of the converted audio files.<sup>19</sup>

This delay can be reduced by changing parameters in an encoder's settings, though these vary from software to software and for each setting tried in various encoders the silence added was impossible to eradicate.

The original reason for choosing the .MP3 file type for SoundShade was due to the fact that extremely small file sizes could be obtained. The team were a little anxious

about releasing an app that would equate to around 200 megabytes or more in size, this would be the result if we were to use the full lossless audio file type .WAV. Though the .WAV file did loop without issue we decided against using it. Upon researching further into suitable file types that could be used the team stumbled upon a feature found in Apple's QuickTime software that would convert .WAV file types into something known as Core Audio format (CAF) file types. The .CAF conversion results in lossless compression which was optimal for our needs. Furthermore, the file type allowed for seamless looping without adding any silence to the samples. We were still a little concerned about the overall size of the application because though .CAF is a compressed file format it is lossless which means that file size reduction is limited. Upon building the app in a simulator we found that the application only reached a total size of 70MB. The team decided this was adequate as storage on smart devices in the modern time far exceeds one hundred times this amount.

### 3.3 Capture & Acquisition – Sound Recording

With the behaviour of our audio samples now totally under control, it was time to set out and build the real content for the application.

The aim for SoundShade was to have totally unique sound samples recorded 'in-house' so that we eliminate the chance of the same sample being found in other similar applications. Though, with this said, for the minimum viable product we did resort to using some royalty free sounds due to ease of access, with one or two making it into the final product, though the post production process that was applied would make sure that the audio sounded unique and original in terms of frequency content and texture. Initially the sound recording was performed with the following equipment:

- *Tascam DR100 & Tascam DR60D MKII Portable Recorders*
- *Rode NTG 2 Shotgun Microphone*
- *Samsung HD MKII Headphones.*

During the recording of the 'easy-to-acquire' sounds in early winter, such as traffic and rain, the equipment performed without complaint. The street in which I live provided much of the sound content found in SoundShade, such as wind, rain and traffic. There is even a railway line close by which I was able to capture with little effort. However, when the time came to record the more delicate sounds the equipment's performance began to fall off somewhat dramatically. By more delicate sounds I almost exclusively refer to bird song. The NTG 2 microphone, though a popular microphone among amateurs and professionals, requires a lot of power in order to record a strong signal, whether it be power in terms of volume from the sound source or power in terms of the pre-amplifier driving the microphone. In this case, as ambient and distant bird song isn't the loudest of sounds the Tascam DR 100's pre amps would need to be set to a very high gain level to capture a strong signal. This proved to be a problem when the sounds were taken into the studio for post processing. The preamps of the Tascam DR100, when set to high gain, produced a very noticeable hiss caused by the electrical components being driven almost to their maximum. This forced me to make the decision to purchase a new recorder, one with more powerful preamps with less operational impedance. Until the new recorder was acquired, and for the minimum viable product, I would search the internet for bird sounds that I could purchase or find freeware versions. In fact, the final Japanese Birds sample was created from a royalty free sound file combined with one of my own early recording of birds, one that was not suitable to be used as the main file, but served well as a low level noise floor to add depth and texture to the sample.

The new recorder would be the Tascam DR60D MKII which is considered an entry-mid level recorder with a price tag of only 200€. However, the reason for choosing this specific model and make was the reviews on how it performed in practice. The Tascam was compared to the Zoom H6 by numerous outlets and professionals and the results were consistent. The Tascam's preamps out performed the Zoom H6's in every test. The operational noise was far lower than the Zoom and indeed any of its competitors. The only recorder bettering it in tests was the Sound Devices 702 field recorder, which is considered a professional high-end piece of gear. There were a few negatives to the Tascam, though they did not pose a problem

to my specific needs. I needed the quietest preamps I could afford. The sound sample 'Lakeside Birds' was recorded with the new Tascam DR60D MKII and it performed remarkably. I was able to get a good strong level with minimal post production in regards to the dynamics of the sample.

Further expanding on the recording process I also began to use a second NTG 2 shotgun microphone to record stereo images of the sounds I wanted to capture. Usually set up in a wide stereo pair, when space was no issue, or just pointing in a different direction completely. This allowed me to capture more material that could be used to layer the sound samples, making them more believable and sonically varied. The English Kitchen sounds were recorded in this manner, to capture the room reverberation which resulted in them sounding well placed in a virtual space. At this time winter had set in and would remain for some months and the bird song available was minimal and I could not find a suitable birdsong sample to use from the web. This meant I would use the recordings I made with the Tascam DR 100 MKI and manipulate it in post production so that it became usable. I would have to rethink my entire understanding of editing audio to complete this task with the best results.

### 3.4 Audio Manipulation - A New Method

The audio design process was a research-then-apply procedure. It saw individual sounds undergo a rapid evolution as new information regarding auditory characteristics, such as frequency content for example, were discovered to actually induce a negative effect on the human psyche. Excessive high frequency hiss or low frequency rumble, for example, could affect working performance when listening to sounds that had more balanced frequency content.

*For example, scores on the typing, math, and verbal reasoning tests tended to decrease when subjects were more annoyed by the noise. Perception of spectral imbalance, such as rumble, roar, and hiss, were also shown to impact performance. These findings support the idea that a more neutral-sounding spectrum is desirable in working environments.*

[Erret et.cal. (2006)]

Thus some of the original sounds were no longer viable to be applied as a countermeasure to distracting noises. This learning curve of what sounds, and indeed their characteristics, were favourable in a working environment required new and more advanced methods of audio manipulation to be utilized during the design process. This encouraged me to find new ways in which to manipulate the audio files that were to be included in the sound library of the SoundShade application. Simple equalization fell short in achieving an overall low noise floor whilst boosting and cutting frequencies that were deemed to be unwanted or important to the natural characteristics of each sound. For example, a sound file named 'Bird Song' was the sound of Autumn forest birds. The originally recorded sample was a relatively clean recording with minimal artefacts that needed to be edited out such as distant traffic, human speech etc. With the desire to isolate each primary sound for each sound file it would prove difficult to achieve a result that was satisfactory with simple E.Q. The 'Bird Song' sample for example contained an unwanted level of hiss from the recording device itself, the Tascam DR-100 MK I, and that of general city atmosphere buzz. By using standard equalization many frequencies contained in the bird's calls were compromised and began to sound unnatural. The first step would be to experiment with multiband compression.

### 3.4.1 Multi-band Compression

Multi-band compression is a form of audio processing that allows you to compress different frequencies independently from one another rather than the whole audio signal. By isolating specific frequencies I hoped to raise the signal amplitude of the necessary frequencies needed to maintain a natural sound for each specific audio file. After using this process on the entire audio library for a new update for the SoundShade App I received some feedback that the sounds were perceived to be very "condensed and flat sounding". I went back to the workstation and did some closer analysis of the audio files. I found that although I had reduced the level of unwanted noise and frequencies I had been too focused on that specific goal to be able to see, or hear in this case, the bigger picture. What had indeed happened was that the wanted frequencies of each sound had been boosted to such an extent, in order to raise their signals, that they had lost much of their natural dynamic range

and the average signal volume was extremely high with little to no headroom between the peaks of the signal and the digital 0db maximum. Fig 9 shows the multi-band compressed audio file of 'Rain Pavement', an audio sample of heavy rain in a city street (top) compared to the more recent version of the same audio file (bottom) that was manipulated using a technique I like to call Multi-layered Frequency Emphasis (MFE).

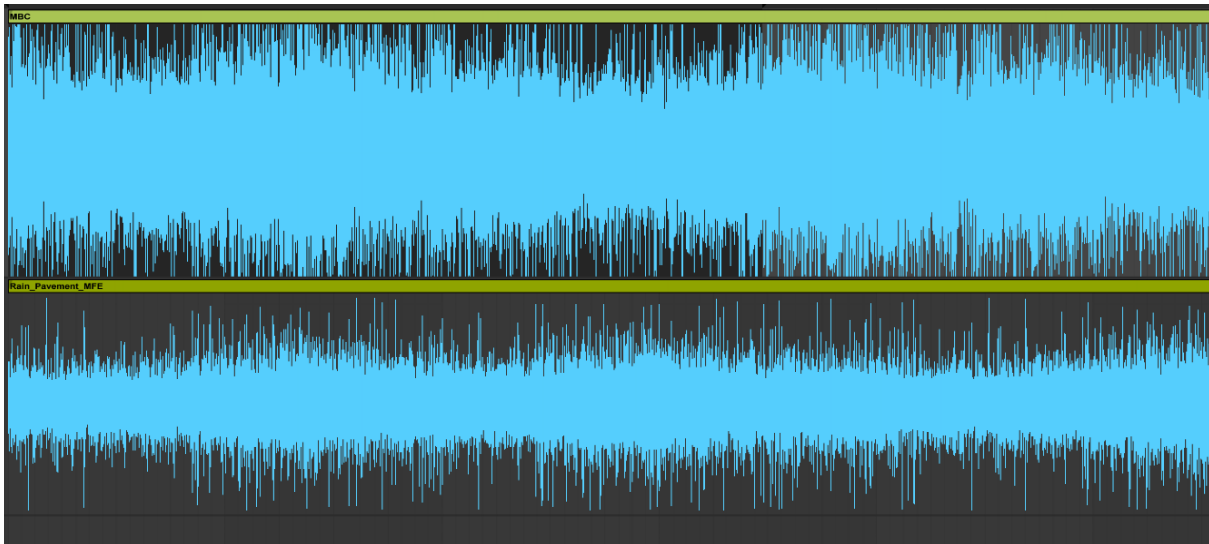


Fig. 10 - Multi-band Compression (top) MFE (bottom)

Note the lack of dynamic range in the upper sample (MBC), and how its peaks are pressed right up to the 0dB digital maximum, compared to that of the lower(MFE). This lack of dynamic range was what I identified to be causing the lack of naturalness in the audio files. Some of the other files that were less constant and that had fewer important frequencies such as the bird song, which had a narrow area of importance on the frequency spectrum, were less affected by the multi-band compression method. With this said I still made the decision to discontinue using multi-band compression in the project to keep a standardized and efficient workflow throughout the project.

### 3.4.2 Multi-layered Frequency Emphasis.

Having discarded multi-band compression I began playing around with other techniques and found that layering the same audio files on top of one another, by multiplying the tracks in the audio workstation, I was able to manipulate each

track's frequencies independently of one another. This allowed me to have total control over the volume of the frequency bands I wanted to emphasise and those which I wanted to soften without affecting the dynamic range of the sample I was creating. As an example the Rock Fountain audio sample alone was a clean recording, although somewhat quiet, by duplicating the sample three times I was able to increase the signals output level on the group's master channel. Though the sample was boosted in volume to a sufficient level, the desired gentle sounding texture was lacking as the high frequencies were too dominant for long listening periods. You can listen to the track Rock Fountain Gain Boost and Rock Fountain MFE for comparison on the USB provided. By utilizing an individual equalizer plug-in on each separate track I could create three tracks with three different and specific bandwidths of frequency. This would allow me to adjust volumes and frequency content of the final sample without using compression, thus maintaining a varying dynamic range. Expanding on the possibilities of the multiple layering I found that by time stretching one of the audio tracks to be a few semitones lower in pitch I could further simulate the gentleness I was trying to achieve by creating more mid-tone frequencies. Though on their own this bandwidth would sound muted and unrealistic. However, when layered with the higher frequencies it would result in the gentleness I was looking for. The third track was pitched five semitones down and a low/mid-pass EQ filter was applied. I wanted this track to replicate the sound of the water passing under itself as it trickled over the edge of the rocks and back into the main water source. This layer of the sample would sit most dominant in the mix as it replicated the gentleness I was aiming for. The results of this process were extremely fruitful though quite time consuming.

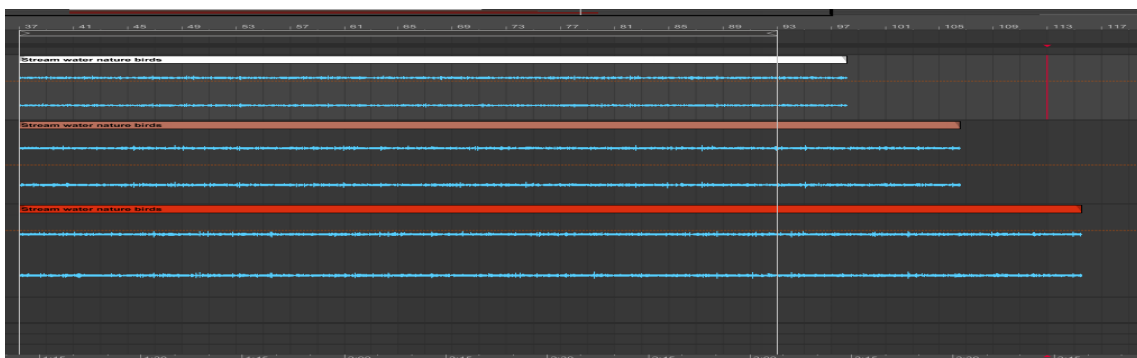


Fig. 11 - three layers of different time stretching amounts of the MFE technique



Fig. 12 - EQ curves used in the MFE technique

Once I had a mix I was happy with I found that the signal level was still rather low, so there would be a need for some signal boosting. Though, rather than raising the volumes of each individual track I simply recorded the mix to a new track, consolidating all three tracks onto one and then increased the output gain of this new sample. This saved time and eliminated the risk of altering the mix of the three tracks that had been achieved. I would repeat this process for many of the sounds contained in the SoundShade app. Examples of some of the sounds included in the SoundShade app will be provided via a storage drive along with this paper.

### 3.5 HRTF - Head Related Transfer Function

From the very beginning it was planned to make use of the Head Related Transfer Function. The function is an algorithm that simulates 3D sound within a stereo playback system. As discussed briefly at the beginning of the part of the thesis it was taken care of almost immediately by the development platform we chose. Though by

choosing to make use of the HRTF we created the need to use exclusively mono sound files. This was due to the fact that as a sound is moved around the listening position by the user the algorithm adds equalizing and minute delays to the sample which produced a phasing artefact to be audible on the samples. When the sound was stationary however this phasing was no longer present, though the frequency content of the sample was altered. When using stereo samples this altering of the frequencies was more dramatic than when mono files were used in their place. This is a point of research that could be covered in the future development of SoundShade.

On a side note, though we had positive feedback from our customers and test subjects, from a professional point of view I found the HRTF algorithm included in the Apple coding software to be lacking when the sounds were placed to the rear of the listener. There was something happening sonically but I was sceptical about whether it actually made the sounds feel like they were physically behind you. This was another point I would have liked to investigate and develop further.

Along with the 3D placement of sounds the HRTF also allowed us to decrease the intensity of a sound. When the sounds were moved away from the listening position on the sound stage there were things happening to the sound other than a simple volume decrease, such as subtle frequency roll. This gave the sense that the sound was actually further away and not just quieter. This was the answer to the problem that some sounds inherently contain problematic characteristics such as the hiss or the wind blowing in the treetops. By placing this sound further away a user could reduce the intensity of the sound enabling them to listen for long periods of time, whilst at the same time keep a sense of immersion and realism that we hoped to portray.

### 3.6 User Interface

When it came to the user interface we aimed to design something that would compliment the nature of the sound content. This meant no overly saturated or bright colours, as little clutter as possible, with soft, easy to look at and an intuitive UI and buttons/ sliders etc. Below you will find multiple screenshots from the most up to date version of the app.



The main page that the users would spend most time on was the ‘mixing stage’. After they choose the sounds they wish to listen to they can return to the mixing stage and adjust the positioning and proximity (volume) in relation to the listening position, shown by the head in the bottom third of the image here. The radial patterned hexagons that fade the further from the listening position were used to portray the volume diminishing the further from the head the sounds were placed. The HRTF algorithm did a fantastic job of emulating distance when moving the sounds further from the listening position until the sounds were placed at the rear, at this point, as discussed earlier I felt the HRTF’s effectiveness lessened.

Fig. 13 Mixing stage of SoundShade

This was the reason why the listening position indicator was placed in the bottom third of the screen. It limited the user to placing the sounds closer to the listening position when the sounds were placed in the rear arc as we felt that the effect sounded better at closer proximity.

The hexagon shaped sound icons were originally circular, but we wanted something more original and unique to SoundShade. The hexagon is also similar to the shape of an umbrella when looked at from a birds eye view and the umbrella was one of our inspirations for the name SoundShade. A sunshade is basically an umbrella used to

protect you from the harmful effects of the sun, and so a SoundShade would be a sonic umbrella to protect you from the harmful effects of unwanted noise.

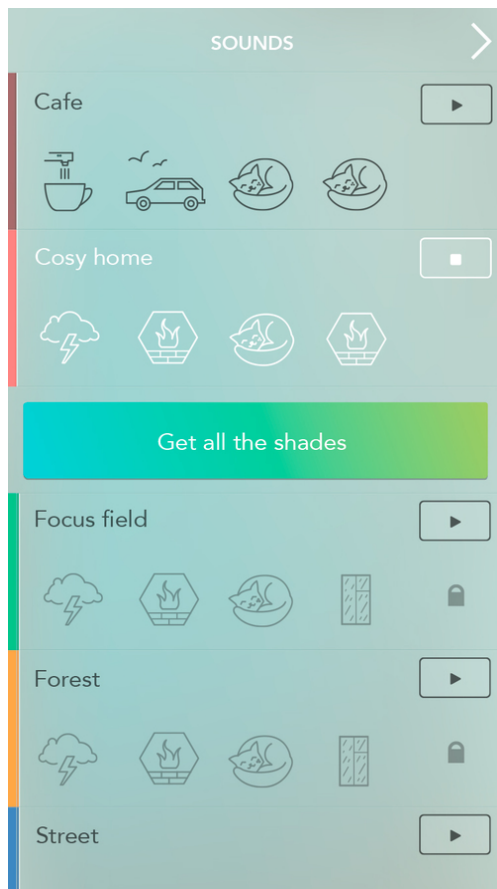


Fig. 14 Sound Library of SoundShade

Next is the sound library or the ‘sound drawer’ as we called it. The working project title for SoundShade was actually called My Sounday, a play on words for an ice cream dessert - Sundae. The idea was we would have a ‘drawer’ with different ‘flavours’ of sounds and the user would create their desired ‘Sounday’ from them. This name was unfortunately taken by a rock band so we had to reconsider. However, we still implemented a drawer of sounds that the user could browse and select from a wide range of sounds. The user would simply press one of the individual sound icons and it would illuminate white and begin playing whilst also being added to the mixing stage in the background.

All the sounds were arranged in themed ‘Shades’ such as Cosy Home for instance, the sounds are illuminated in fig 14, this allowed for easy playback of a pre-mixed shade using all of the sounds in that category. The user could then remix and reposition the sounds from the mixing stage as normal. The user could add as many sounds from the library as they wished, though multiples of the same sound could not be added. This was something I would have liked to experiment with as this would have allowed me to create stereo images of certain sounds, for example the rain on a window sound. By positioning two of the same sounds in different locations the effect of rain hitting two different windows could have been simulated adding another layer of variation to the experience.

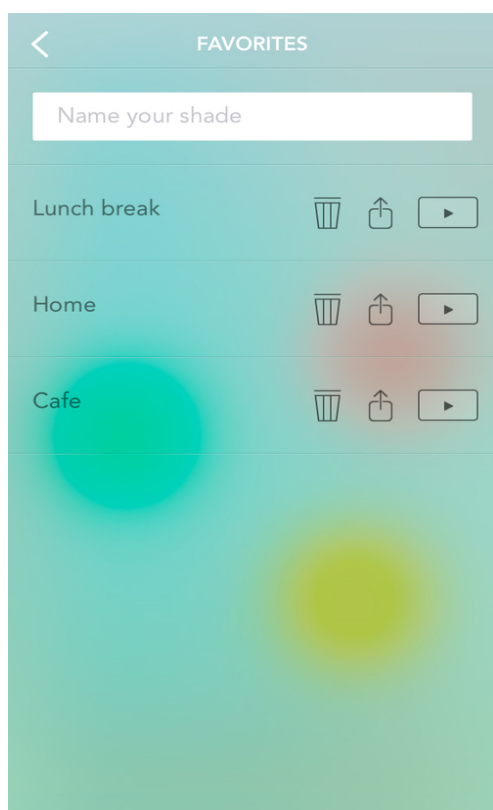


Fig. 15 Favourites page

Just as we had pre-mixed some of the sounds into categories, or themes, we implemented a feature to allow our users to do the same so that they would be able to quickly load up their preferred shade. Fig 15 shows the Favourites page. When loaded all of the sounds saved in that shade would be added back to the mixing stage in the positions they were in when the shade was saved.

Whilst the user interface would not be the focal point of the app it would be the method of interacting and manipulating the content of it, and for this reason we decided to spend more time on fine tuning and refining the usability of its interactive elements throughout the development of the project.

## 3.7 Further Development

### 3.7.1 Stereo Vs Mono

During the mixing and editing process I found the sounds to be much more sonically interesting and had a far more natural feel to them whilst listening in stereo. Due to the reasons explained above in regards to the HRTF we couldn't use stereo files. This was something that I sacrificed in order to benefit from the unique HRTF feature in the App as it was so popular, even if it did conflict with my artistic desires.

### 3.7.3 Healthcare Product

With SoundShade being based almost entirely on existing research on how noise and nature sounds affect the human mind it would be of great interest to see how well SoundShade would perform in lab tests relating to the working performance of office workers and indeed if our sounds can have a measurable effect on reducing the time it takes to recover from psychological and physiological stress. This was an early talking point of the team and our investors. If we could prove our product worked and had numbers and results to show for it, the potential of being able to provide SoundShade as a health care product would be incredibly high.

### 3.7.4 Artificial Masking

Chapter 2 of this paper looks into how people are affected by external audio stimuli. It was found that the brain is believed to apply a masking mechanism to sounds it deems unimportant to task related goals. Furthermore we see that it is also believed that the brain is able to set up “blocks” to reject unwanted auditory messages. Something that is of great interest to me is if SoundShade could be used as an artificial mask or block, so the brain would not have to apply its own mechanisms. I theorise that if an external device was to apply a method of reducing the perception of the unwanted external auditory stimuli that have been proven to be detrimental to working performance, this could possibly reduce working fatigue of the brain. If we assume that these mechanisms use up the brain's resources, and if the need for the use of these resources were eliminated, could this increase a person's ability to focus and inturn increase their workplace performance? Combining this theory with the results found that nature sounds are able to reduce recovery time from psychological stressor, could we simultaneously increase working performance whilst reducing the rate at which stress levels rise?

The ability to block external sounds with personal devices such as headphones is greatly reliant on the type and quality of the headphones. Passive noise-reducing inner ear buds are relatively cheap and more accessible compared with active noise-cancelling over-ear headphones but the latter are far more

effective at blocking external sounds. Tests would need to be undertaken with varying forms of listening devices to get a more accurate set of results as to whether this theory proves true.

It would be very interesting to see the results of tests designed around this theory as significant positive results could be extremely beneficial to the population who are exposed to unwanted noise during their working life.

# Chapter 4

## 4.1 Conclusion

The SoundShade mobile application was published to the iStore and quickly gained over fifty thousand downloads and holds close to a 5 star rating from those who reviewed the app. Though, the original purpose of SoundShade - to be designed as a tool for the work or study place, aiming to increase a person's ability to focus on workplace tasks by masking unwanted noises around them - did not seem to be apparent to our target audience. The majority of the user base uses SoundShade as a sleep aid. Still, SoundShade's content is being used to raise the noise floor of a room chosen for sleeping which masks unwanted, intermittent noises allowing the user to sleep undisturbed.

The SoundShade team believes we achieved what we set out to accomplish. The app works very well. The sounds are of a high quality and of a varying nature, that loop seamlessly and are undisturbing to the users. From my individual and artistic point of view, though, I did make some compromises regarding the audio content. I strongly feel that my team and I have given the users of SoundShade the opportunity to create their own perfect world using sounds. As a sound artist my hobby within the profession was, and is still, to create small soundscapes that I can retreat into for moments of respite and relaxation. The actual process of creating these worlds is as calming to me as the effect of the final creation. If SoundShade has enabled other people outside of the professional sound industry to become sound artists in their own rite and create their own perfect sound worlds, then I consider the project a success, regardless of the monetary income or any "*star-rating*" it achieved.

Overall this project did not simply result in a market-ready product. Rather, it is the result of a professional sound artist completely overhauling their understanding of the material they have chosen to work with for the past decade. The new found insights have revealed sound to not be something we simply physically hear, but something we experience on a conscious, subconscious, physiological and a psychological level, with substantial effects measurable in both the body and in the mind.

## References

1. Johns Hopkins University (1998). "How Do We Hear While We Sleep?" *Science Daily* [Electronic], Available: <[www.sciencedaily.com/releases/1998/04/980430044534.htm](http://www.sciencedaily.com/releases/1998/04/980430044534.htm)>.
2. Orfield, S. (2012) 'We all crave it, but can you stand the silence? The longest anyone can bear Earth's quietest place is 45 minutes', *Daily Mail* [Electronic], Available: <[www.dailymail.co.uk/sciencetech/article-2124581](http://www.dailymail.co.uk/sciencetech/article-2124581)>, paragraph 6.
3. World Health Organization, Geneva (1999). 'Guidelines for Community Noise', page 7 (vii)
4. Puzzo, M. (2014) 'Greater Hartford office market gaining momentum', *Hartford Business* [Electronic], Available: <[www.hartfordbusiness.com/article/20140106/PRINTEDITION/301029943](http://www.hartfordbusiness.com/article/20140106/PRINTEDITION/301029943)>, paragraph 9.
5. - Kim, J. & de Dear, .R (2013), 'Workspace satisfaction: The privacy-communication trade-off in open-plan offices' *ScienceDirect* [Electronic], Available: <[www.sciencedirect.com/science/article/pii/S0272494413000340](http://www.sciencedirect.com/science/article/pii/S0272494413000340)>, page 25 -Conclusion.
6. Steelcase - 'The Privacy Crisis Taking a Toll on Employee Engagement', *Steelcase* [Electronic] Available: <[www.steelcase.com/insights/articles/privacy-crisis/](http://www.steelcase.com/insights/articles/privacy-crisis/)> , paragraph 9
7. Paul Robertson (2014) - Open Plan Offices are a Health and Productivity Risk - Canada Life, *Cover Magazine* [Electronic], Available: <[www.covermagazine.co.uk/cover/news/2344756/open-plan-offices-are-a-health-and-productivity-risk-canada-life](http://www.covermagazine.co.uk/cover/news/2344756/open-plan-offices-are-a-health-and-productivity-risk-canada-life)>

8. Landau & Barner (2009) - Vestibulocochlear Nerve
9. Jessica Errett et al (2006) - 'Effects of Noise on Productivity: Does Performance decrease over time?', University of Nebraska, Page 7, Paragraph 1.
10. J.L. Szalma & A- Hancock (2011)- Noise effects on human performance: a meta-analytic synthesis.
11. Moray (1959)- Attention in dichotic listening: Affective cues and the influence of instruction.
12. Conway, Cowan & Bunting (1995)-The Cocktail Party Phenomenon Revisited
13. Baddley & Hitch (1974); Miyake & Shah (1999)
14. Roger S. Ulrich (1984)- View through a window may influence recovery from surgery
15. Ulrich et al. (1991) Stress recovery during exposure to natural and urban environments
16. J. J. Alvenson et. al. (2010)- Stress Recovery during Exposure to Nature Sound and Environmental Noise.
17. Mp3 encoder delay times -  
[http://mp3decoders.mp3-tech.org/decoders\\_lame.html](http://mp3decoders.mp3-tech.org/decoders_lame.html)