

Title	Ambee : creating device interaction through ambient sound and location analysis
Sub Title	
Author	Ruiz Velasco Enriquez, Juan Pablo(Inakage, Masahiko) 稲蔭, 正彦
Publisher	慶應義塾大学大学院メディアデザイン研究科
Publication year	2012
Jtitle	
JaLC DOI	
Abstract	<p>While many researchers have experimented with visual Augmented Reality, development in the area of sound Augmented Reality is currently lacking. Using incoming Ambient sounds and location to generate interactive contents represents an unlimited source of entertainment that remains untapped.</p> <p>The Ambee project aims to create an interactive application that provides entertainment by analyzing incoming sounds and having an on-screen character react and generate music and other contents based on sound analysis.</p> <p>Ambee dynamically adjusts the contents being offered to the user, therefore no two locations are the same and no two users have the same experience. This in return creates a sense of Serendipity that if approached correctly, would have unlimited potential to generate in-app contents.</p> <p>This thesis analyzes the viability of this concept in mainstream electronics, consumer interest, as well as information on the design framework and the experiences that are be offered to our target user.</p>
Notes	修士学位論文. 2012年度メディアデザイン学 第215号
Genre	Thesis or Dissertation
URL	<a href="https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40001001-00002012-0215">https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40001001-00002012-0215</a>

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the Keio Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

# Master's Thesis

## Ambee

Creating device interaction through ambient sound and location analysis.

by

Juan Pablo Ruiz Velasco Enriquez

Submitted to the Graduate School of Media Design  
in partial fulfillment of the requirements for the degree of

MASTER OF MEDIA DESIGN

at the

KEIO UNIVERSITY

Academic Year 2012

© Graduate School of Media Design, 2012. All rights reserved.

# Ambee

**Creating device interaction through ambient sound and location analysis.**

by

Juan Pablo Ruiz Velasco Enriquez

Bachelor of Industrial Design

Submitted to the Graduate School of Media Design  
in partial fulfillment of the requirements for the degree of

MASTER OF MEDIA DESIGN

at the

KEIO UNIVERSITY

Academic Year 2012

© Graduate School of Media Design, 2012. All rights reserved.

Certified by .....  
Professor. Masahiko Inakage  
Professor, Graduate School of Media Design  
Thesis Supervisor

Certified by .....  
Professor Naohito Okude  
Professor, Graduate School of Media Design  
Thesis Co-Supervisor

Accepted by .....  
Professor Masa Inakage  
Professor and Dean, Graduate School of Media Design

---

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at the Keio University, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

(Signature) \_\_\_\_\_

Graduate School of Media Design

Keio University  
Hiyoshi, Yokohama, Japan

Date \_\_\_\_\_



# **Ambee**

## **Creating device interaction through ambient sound and location analysis.**

by

Juan Pablo Ruiz Velasco Enriquez

Submitted to the Graduate School of Media Design  
on August 7, 2012, in partial fulfillment of the  
requirements for the degree of  
Master of Media Design

### **Abstract**

While many researchers have experimented with visual Augmented Reality, development in the area of sound Augmented Reality is currently lacking. Using incoming Ambient sounds and location to generate interactive contents represents an unlimited source of entertainment that remains untapped.

The Ambee project aims to create an interactive application that provides entertainment by analyzing incoming sounds and having an on-screen character react and generate music and other contents based on sound analysis.

Ambee dynamically adjusts the contents being offered to the user, therefore no two locations are the same and no two users have the same experience. This in return creates a sense of Serendipity that if approached correctly, would have unlimited potential to generate in-app contents.

This thesis analyzes the viability of this concept in mainstream electronics, consumer interest, as well as information on the design framework and the experiences that are be offered to our target user.

Keywords: Ambient sound, entertainment, social interaction, location, music.

Thesis Advisor: Professor. Masahiko Inakage

Thesis Co-Advisor: Professor Naohito Okude

# Table of Contents

<b>Abstract</b> . . . . .	iii
<b>Table of Contents</b> . . . . .	iv
<b>List of Figures</b> . . . . .	vii
<b>Acknowledgements</b> . . . . .	viii
<b>1 Introduction</b> . . . . .	1
1.1 What is Ambee . . . . .	2
<b>2 Literature Review: Interacting with sounds, devices and locations</b> . . . . .	3
2.1 The importance of personal music . . . . .	3
2.2 Interacting with sounds . . . . .	6
2.3 Creating Synesthesia . . . . .	10
2.4 Exploring the world together . . . . .	12
2.5 Sounds create emotions . . . . .	14
<b>3 Implementations and related projects</b> . . . . .	16
3.1 Muscle fish . . . . .	17
3.2 Shazam . . . . .	17
3.3 Navinko . . . . .	19
3.4 Lumines . . . . .	20
3.5 Tamagotchi . . . . .	21
3.6 Seaman . . . . .	22
3.7 Milo . . . . .	23
3.8 State of Development . . . . .	24
<b>4 Ambee Concept</b> . . . . .	25
4.1 Scenario Analysis . . . . .	25
4.1.1 Crowded Trains . . . . .	25
4.1.2 Tapping on screens . . . . .	26
4.1.3 Waiting . . . . .	26
4.1.4 Moving around new areas . . . . .	26
4.1.5 Background Noise . . . . .	26
4.2 The ideal Ambee user . . . . .	27
4.3 Design Workshop: Creativity Box . . . . .	28
4.3.1 Character design analysis . . . . .	30

*Table of Contents*

---

4.4	Visualizing Background noise . . . . .	31
<b>5</b>	<b>Implementation: Ambee . . . . .</b>	<b>32</b>
5.1	Ambee Design Framework . . . . .	32
5.1.1	Early prototypes . . . . .	33
5.1.2	Dynamic Sound Interaction . . . . .	33
5.1.3	Gamification of music . . . . .	34
5.1.4	Location Aware benefits . . . . .	35
5.2	User Interface and functions . . . . .	36
5.2.1	Life . . . . .	37
5.2.2	Sound Generator . . . . .	38
5.2.3	World Map . . . . .	39
5.2.4	Social . . . . .	40
5.3	Application Framework . . . . .	40
5.4	Hardware requirements . . . . .	42
5.4.1	Sound analysis Methods . . . . .	42
5.4.2	Euclidean Distance . . . . .	42
5.4.3	Spectrogram . . . . .	42
5.5	Analysis Challenges . . . . .	43
5.5.1	Noise . . . . .	44
5.5.2	Distortion . . . . .	44
5.5.3	Data management . . . . .	44
5.5.4	Battery life . . . . .	45
5.6	Ambee as a gaming platform . . . . .	46
5.6.1	Restaurants . . . . .	46
5.6.2	Convenience Stores . . . . .	46
5.6.3	Parks . . . . .	47
5.6.4	Train Stations . . . . .	47
5.6.5	Karaoke Locations . . . . .	48
<b>6</b>	<b>Evaluation . . . . .</b>	<b>49</b>
6.1	Framework of this evaluation . . . . .	49
6.1.1	User Interface . . . . .	50
6.1.2	Music Interaction . . . . .	51
6.1.3	Location analysis . . . . .	52
6.1.4	Data analysis from Usability test . . . . .	53
6.2	Quantitative Research . . . . .	54
6.3	Discussion with Tetsuya Mizuguchi . . . . .	59
<b>7</b>	<b>Conclusions . . . . .</b>	<b>61</b>
7.1	Concept Feedback . . . . .	61
7.1.1	User Interface analysis . . . . .	62
7.2	Current revenue model . . . . .	62
7.3	Future Developments . . . . .	63
7.4	Revised Business Model . . . . .	63
7.5	The Story behind Ambee civilization . . . . .	64

*Table of Contents*

---

7.5.1 Planet Sonas . . . . .	64
7.5.2 City Domination . . . . .	65
7.6 Development continues . . . . .	66
<b>Bibliography</b> . . . . .	<b>67</b>

# List of Figures

2.1	Regency Tr1-G . . . . .	4
2.2	Walkman TPS-L2 . . . . .	4
2.3	S-SMP Sound Chip . . . . .	7
2.4	Dance Dance Revolution Competition . . . . .	10
2.5	Letter colors as perceived by Synesthetes . . . . .	12
3.1	Shazam for iPhone user interface . . . . .	17
3.2	Navinko navigation system user interface . . . . .	19
3.3	Lumines Puzzle fusion. Lights and sounds. . . . .	20
3.4	The first tamagotchi (Model V1). A portable revolution. . . . .	21
3.5	Seaman for Sega Dreamcast. Life simulation evolved. . . . .	22
3.6	User interacting with Milo trough a Kinect camera. . . . .	23
4.1	Our target user. . . . .	27
4.2	Desing form and tester inside creativity box . . . . .	29
4.3	Characters from Creative Box Workshop . . . . .	30
5.1	Ambee logo. . . . .	32
5.2	Printer triggered music generator . . . . .	33
5.3	Interactive Musicality . . . . .	34
5.4	Shake to scan. . . . .	35
5.5	Four menu user interface. . . . .	36
5.6	X-ray of an Alien life form. . . . .	37
5.7	Ambees live from sounds, feed them music. . . . .	38
5.8	Exploring the world together. . . . .	39
5.9	Sharing the Ambee experience. . . . .	40
5.10	Framework of the Ambee application . . . . .	41
5.11	Two sounds analyzed by Euclidean distance method. . . . .	42
5.12	Spectrogram analysis. . . . .	43
5.13	Follow the sequence (Simon says). . . . .	46
5.14	Energy supply at the store. . . . .	47
5.15	Hide and seek outdoors. . . . .	47
5.16	Item transportation game. . . . .	48
5.17	A new way of performing. . . . .	48
6.1	Tester using interface for the first time . . . . .	50
6.2	Interacting with sounds . . . . .	51
6.3	Testing sound based interface . . . . .	51
6.4	Clips from character creation video . . . . .	52

# Acknowledgements

This dissertation would not have been possible without the help of many people who helped me through the last two years.

First and foremost, my utmost gratitude to Prof. Inakage for his continued support and assistance during my time at KMD and providing me the tools to research the field of social entertainment and sound interaction. His experience and approachable personality have left a positive impact in my life.

I would also like to thank Mr. Tetsuya Mizuguchi for taking time out of his busy schedule and giving me the privilege of talking to him about this project. As one, if not the most important person in the field of sound entertainment, his opinion and feedback are invaluable to me. As a fan of his game Lumines, I am overwhelmed to be able to discuss my work with him. A true honor.

Special recognition goes to my family for letting me leave my hometown at such an early age and letting me explore my potential. Even though it often seems I never looked back, I appreciate everything they did and continue to do for me. Without the ongoing support of my sister Jessica, this project wouldn't have been completed. Thank you for always being there for me.

I feel it is important to dedicate part of this effort to Japan and Japanese society. Your work ethics, your strength and your example make me want to try harder every day. No matter how grim the future looks, Japan will always continue to innovate, to grow, to impress. It is with pride that I call Japan my home.

And last but not least, I would really like to thank myself<sup>[38]</sup> for believing in me back in the day when it seemed no one else did.

Today, I am happy being who I am and where I am. Things got better after all.

# Chapter 1

## Introduction

We are routinely exposed to ambient sound and background noise. As our lives evolve, we spend less time at home and more time wondering around the city and exploring an ever changing environment Train, home, restaurant, movie theater, shopping mall. However, the way we interact with the city has remained mostly the same.

As we move from point a to point b, our lives become routinary and commuting to work lacks excitement. In cities like Tokyo in particular, we start to complain about places being too crowded, about areas being too dirty or noisy and having limited leisure time. As we get frustrated by our surroundings, we usually look to portable devices to block ourselves from the stress of living in a big city.

Portable entertainment is a concept that we give for granted these days and most people carry some sort of portable electronic device, be it a smart-phone, a tablet or a portable gaming console, we can personalize our relationship with the city by controlling the soundtrack of our lives.

In a very simple manner, these devices allow us to control how we perceive our environment.

It is important to note however, that the experience offered by these devices has been designed mostly to replicate what we can already do at home. As we moved to the digital era, we focused on creating smaller devices with more capacity. The arrival of digital music meant that we could now carry our music collections at all times, an impressive feat in itself that completely changed the way we listen to music. It is now possible to buy a song anywhere from our internet connected smart-phone.

This is an impressive development, however, sales of pre-recorded media have been falling steadily over the last few years. It is then that the question arises, How can we evolve the way we create sound entertainment?.

As a gamer and music enthusiast myself, I have come to realize that the older I get, the more difficult it becomes to find time to enjoy this kind of hobby. And the rise of popularity of casual gaming as well as the gamification of many other services seems to be the answer to this trend.

With our hectic lifestyles, oftentimes we only have a few minutes to spare everyday between commuting from one place to the other or while waiting to meet up with someone.

This casual gaming or entertainment sessions have become an integral part of our lives and are re-shaping entertainment as a whole. The shift away from complexity is one that cannot be ignored.

With this in mind, we set up a team in Keio university to explore this area of research. How can we create an entertainment experience that cannot be replicated at home?

Analyzing incoming sounds to create music and dynamically generate contents we can do that, through the use of Ambee.

## **1.1 What is Ambee**

Ambee is a handheld digital pet application currently in development for the Android platform that allows the user to interact with a virtual character through analyzing background noise and sounds.

The application can analyze incoming sounds to create music accordingly and the character reacts to the beat of the song being generated in real time. By scanning the background sounds, the application can detect the current location from a group of pre-set environments (restaurant, train station, karaoke, etc.)to adjust the contents on the screen.

The character automatically changes costumes and prompts the user to try a specific mini-game related to the current location of the user. This creates a sense of wonderment and exploration that links the virtual world and the digital worlds to encourage the user to explore the city and try the application in different locations.



## Chapter 2

# Literature Review: Interacting with sounds, devices and locations

The Ambee application is based on two main ideas, expand on the concept of location and sound analysis and to introduce novel device interaction methods. We aim provide new forms of entertainment.

The research is mostly based on three areas: location, sound creation and emotional sound design, since the connection between the user and the device is one of the key elements for the success of our application.

We want to create a reliable and exciting companion for users who explore the city alone provides an opportunity to push the boundaries of both, location and sound interaction.

### 2.1 The importance of personal music

Even though the way we listen to music has changed and evolved drastically through history, one thing is constant, Music is an immensely important part of our lives. It can have a direct effect on our emotions and help us evoke valuable memories, create an appropriate atmosphere and even help us structure an activity.

In a way, Music provides a soundtrack to our lives, an extra layer that makes things come to life and also, a new way to express who we are and how we interact with each other.

Even though the concept of recording and playing music was introduced in the late 1800's by Edison, technology limitations and high cost of the units meant that it would remain a niche market for years to come.

The concept of portable music started in 1954 when I.D.E.A. put on the market the first portable transistor radio with the release of the Regency TR-1. It was

## 2.1. The importance of personal music

---

an important change of direction for the music industry, because it was the first time that music was meant to be an "always on" service for personal use instead of pre-recorded tracks to be enjoyed together. The design was simple and music was played through a single lo-quality speaker.

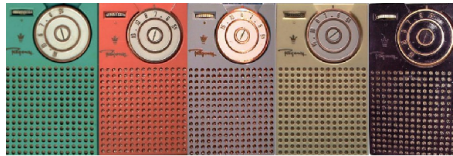


Figure 2.1: Regency Tr1-G

However, the personal music revolution arguably began with the introduction of the first self-contained portable cassette player, the Walkman TPS-L2 [3] by Japanese manufacturer Sony Electronics. At the time, Japanese companies were widely considered as the best in terms of hardware and transistor miniaturization[27] and this device was the culmination of the advancements of their technology.

The Walkman's real standout feature was its size. Being only slightly bigger than the tape itself and with an extraordinary battery life (on only 2 AA batteries) set the standard for generations to come. The use of headphones also meant that music was to be consumed by one user at a time, therefore creating a more private, personal experience.



Figure 2.2: Walkman TPS-L2

For the first time, Music stopped being the main action of the listener and instead became the background for many experiences outside the house. Tia DeNora [34] mentions the importance of music getting into our daily lives and experiences, and how it has become an important element that is part of the way "people make sense of a place". Music triggers memories.[17]

As part of DeNora's research, one user describes how she felt after listening to a specific music track:

"It [car radio, switched on as the speaker started the car] was playing the Double Concerto and I just had to stop, and some of my friends were coming behind, you know, and I was just in floods of tears, and they said, 'Why don't you turn it off'? And I said 'I can't' and that it was ages before I could listen to that, or anything like it without thinking of him".

Music can then have a direct influence in emotions and the mood we feel at any given time, that is why since its conception, the idea of personal music was based on giving the user ultimate control over what he would listen to, when he would listen to it, and eventually where to listen.

After the Walkman, the industry focused on incremental advancements to music technology, where most of the focus was placed on increasing the quality and the fidelity of music reproduction while the experience remained mostly unchanged for years to come. This saw the introduction of new formats like the Compact Disc that found mainstream success as well as less popular technologies like the Super Audio CD[43] and DVD Audio, both of which promised HD Audio quality but ultimately failed to find significant success.

It was with the arrival of the internet and the release of the iPod[6] that the way we consume music changed completely. For the first time, users had almost unlimited access to music libraries around the world, not constrained by location, language or availability. If a song was released somewhere in the world, it meant that a user could probably find it, store it and access it as desired on a device with enough capacity to hold literally thousands of songs at a time. Sound quality has since become a feature of less importance.

This important shift of direction should not be overlooked since it shaped the music industry that we know today. It meant that users were willing to share their music and at the same time, consume more of it while using it as an important way of personal expression.

iTunes[39] recently introduced the concept of sharing libraries in the same network, changing forever the concept of personal music. Based on research by Barry Brown and Abigail Sellen (3), since music collections are now visible to other co-workers or friends, people now feel the pressure to diversify their library and are also a lot more receptive to new music that is accessible through their network.

A new song or artist is a lot more likely to be successful in the market place if it is recommended by a familiar user.

As part of their findings, they mention how a similar music collection can lead to interest from people to learn more about each other and can also change the way we perceive a certain person, so even though the music is commercially available, our collection can represent who we are:

”There is an instant connection, like If I meet someone who listens to the early Verve stuff when I think there’s something really important going inside of them [...] I think it brings me a lot closer to people if you can share the exhilaration that music can bring to you.”

Music has stopped being a personal experience and expanded to a world where we can express ourselves through the artists and the kinds of music that we listen to, having access to nearly unlimited music libraries and bringing back memories at any given time means that the concept of personal music has evolved greatly from those early experiences with the transistor radio and cassette players, and it is now a social experience where we can share, create and express ourselves.

## 2.2 Interacting with sounds

While users had access to personal music players since the 60’s, it took more time for electronic gaming to become a mainstream solution. Early music players provided a access to previously generated contents, the concept of interacting with those devices in real time required far more processing power than was available at the time.

With the arrival of video game consoles, new game developers eventually realized that the best way to communicate and offer feedback to the player was through Visual and acoustic elements.

Modern video-games rely on sounds to enhance interactive experiences nowadays and include orchestral soundtracks and synthesizer sounds, however, they were basic and flat in the early stages of gaming. While the first home console available was the Magnabox Odyssey[8], It’s lack of a sound chip forced developers to move to the more advanced Pong[42] released shortly after. An article by Times Magazine [14]claims that the unique feature of Pong is the inclusion of a chip that can handle both Sounds and Images at the same time.

At an interview with IGN [7], Al Alcorn who worked on the game mentions that while finding many issues to move the project forward, one thing was clear:

You've got to have sound.' Oh okay, well I'm over budget and three months into this thing and Nolan said 'I want the roar of a crowd of thousands.' Cheers, applause. How do you do that with digital circuits? Ones and zeroes? I had no idea, so I went in there that afternoon and in less than an hour poked around and found different tones that already existed in the sync generator, and gated them out and it took half a chip to do that. And I said 'there's the sound if you don't like it you do it!' That's the way it was left, so I love it when people talk about how wonderful and well thought out the sounds are.

From that point, music became part of modern video game technology that has been recognized by users and gamers alike.

A few years later, Nintendo decided to introduce the Super Nintendo entertainment system, but given their lack of experience, they decided to ask rival Sony to develop the sound chip for their console, given their expertise in developing music players and recorders. The outcome of this collaboration was the Nintendo S-SMP, the chip that would become the only collaboration between the two Japanese companies.

The chip was a remarkable achievement in terms of sound quality allowing



Figure 2.3: S-SMP Sound Chip

for reproduction of 16 bit music. Designed by Sony engineer Ken Kutaragi[20], development allegedly started when Kutaragi, who was already working for the electronics giant, bought his daughter a Nintendo entertainment system. After listening to her playing with the system, he was so disappointed by the primitive sounds created by the console that he started researching in secret.[24]

He would later create the competing PlayStation[40] system for Sony that would dethrone Nintendo as the dominant player in the gaming Industry and give the Japanese giant it's most profitable brand.[30]

The Playstation was designed from the ground up to be more than a gaming system and integrate advanced graphical and sound capabilities. The inclusion of a CD-drive meant that for the first time, games could be fully orchestrated and offer high-quality stereo sounds without any problems with storage capacity or distortion. This was one big advantage that let PlayStation developers create more complex experiences when compared to those available in the competing Nintendo 64.

According to Julian Eggebrecht [9], CEO of social game design studio Touch-Factor mentions:

"The Super NES has a sound-chip, the N64 doesn't have a sound-chip [...] the more graphics you do on the N64, the less performance you have left over for sound. With the Super NES, you knew that you could do all this and then you still had a sound chip (S-SMP) to handle the music. On the N64, sound eats up performance[...] We work a lot with sound, so we value a game's audio highly – but what you expect from your game is something that the individual developers come up with by themselves. And if you don't have high expectations – the N64 makes it tough enough".

He mentions that the limitations of the system actually managed to hamper game innovation and presented an important challenge to in-game sound interaction.

Since the arrival of the PlayStation, music video games have become an important part of pop culture and their influence has forever reshaped the way we interact and socialize around music. Many games based on the concept of sound interaction have developed a very strong legion of fans that can offer models on how music can function in a digital media-scape.

One of the best examples would be the Dance Dance Revolution[10] series introduced by Konami in 1998. The game has the player following a series of on-screen commands that then have to be pressed in a timely manner on a special mat using feet and hands. During the game-play, big speakers will play music which turns the machine into an interactive digital jukebox of sorts.

The game developed a strong sense of community due to 2 main factors, original music and social entertainment.

While gaming had remained mostly an stay-at-home activity, Dance Dance Revolution encouraged users to go to the arcades, play the game and interact with each other. It has been referred as Karaoke for the feet. Karaoke has since the beginning been considered a Musical and Social experience[2], and Dance Dance Revolution was one of the first, if not the first game to merge both elements successfully.

Rob Drew's examination of Karaoke culture, Karaoke Nights [5], establishes that unlike other forms of entertainment, the performer and the audience create a special connection that enhances the feeling of community. Drew mentions that Karaoke performers often act as part of a team group members collaborate in appealing to the audience, synchronize their movements in relation to the audience, evaluate one another's conduct from the perspective of the audience.

At the same time, the creation of original music meant that users from around the world would collaborate and rate the music released with each new iteration. Most of the music and remixes are exclusive to the game and cannot be heard anywhere else.

The game doesn't limit itself to music however, and introduced other sound elements to enhance the gaming experience. As it made the transition to home consoles, it tried to emulate the social experience that was offered in the arcade version. The game has everything to emulate the arcade experience at home, including a prerecorded crowd that alternately used boos or cheers based on the player's performance.

If the gamers actions corresponds to the timing of on screen commands, the user will be greeted with Perfect, Great or Miss. This element is relevant because it drifts from the usual attention of on-screen information and also focuses on the emotional element of the game. The most hardcore players would learn the music and choreography's to dance them backwards without looking at the screen and based solely on sound elements and memorization. This also creates a spectacle for observants where the people are expected to dance for the crowd, not the machine.

This led to mainstream acceptance of music based gaming, and DDR enjoys one of the most diverse fan-bases on earth. It has successfully attracted Asian users with the use of Animation tracks, White fans via video gaming and African Americans through it's similarity with Hip-Hop and break dancing.



Figure 2.4: Dance Dance Revolution Competition

The success of Dance Dance revolution is in big part responsible for the creation of the music gaming genre. Some of the most important games of the genre are Parappa the Rapper created by Masaya Matsuura[41], Space Channel 5 created by Tetsuya Mizuguchi[45] and more recently Guitar hero and Rock Band.

## 2.3 Creating Synesthesia

While most recent music interaction experiences follow a similar basic design, advancements in technology have made it easier to create rich audiovisual experiences that intend to take the user to a parallel universe where time doesn't exist and perception is all that matters.

Synesthesia is a neurological condition that literally means senses together. This means that the stimulation of one sensory triggers an automatic reaction in a second. It is often described as hearing colors, seeing sounds.

Since the introduction of sound as part of the video game experience, it has been used mostly to guide players and guide them in a certain direction. According to Jesse Schell [32] sound helps the player look for something hidden, set the pace of the action, give hints of what direction to follow and so on. It is at this point that game music becomes a language, that speaks to the players to such a deep level that it can affect their moods, desires and actions.



One important player in the world of music games is Tetsuya Mizuguchi, often regarded as the Master of Synesthesia. A former Sega employee, he lead the United Game Artists studio, that placed special emphasis in pushing the boundaries of Sound Interaction Design. While at the helm of the studio, he created games like *Rez* and *Space Channel 5*. His games often do not fit any existing genres, however, one thing is constant: Every action triggers an audio-visual reaction in the game.

In Mizuguchi's own words Synesthesia can be described as:

"Synchro and aesthetics, everything together as one. These are not new words; these are very old words. We need that kind of a new concept all the time. 100 years ago, many artists described synesthesia, because they had no expression words, no good words to express. So, they made words, and Im just using their words. We need that kind of a concept in interactive experience."

It means that every action within the game has a reaction and more importantly, they are all connected to each other and the gamer to create an even deeper emotional impact. Research by the Massachusetts Institute of Technology (MIT) reached similar conclusions that seem to validate Mizuguchi's vision for Synesthesia as a game enhancer. According to Neurologist Richard Cytowic [4] Synesthesia tends to be involuntary and automatic and tends to be a reaction of the senses being stimulated. It is common that this will trigger a chain reaction that will cause two senses work together.

Synesthesia is also noted for having a spatially extented perception, meaning that they tend to have a sense of location. Synesthetes often mention the concepts of going to a place or looking at from a different perspective based on Cytowic's research.

Even more important however are his claims that Synesthesia is highly memorable and creates an emotional and affective connection.

Researcher Sean Day [21] carried a unique experiment where he gathered 43 Synesthetes and aksed them to assign colors to different letters of the alphabet to try to further understand how this individuals use their senses.

## 2.4. Exploring the world together

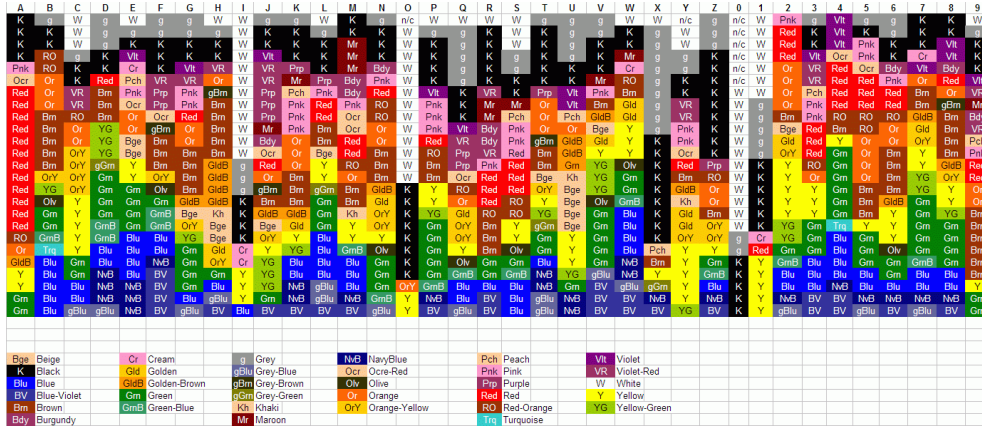


Figure 2.5: Letter colors as perceived by Synesthetes

It is interesting to observe that some letters have a very visible pattern. For example in column a, most of the participants agree that letter A has a red hue to it. At the same time, letters I, O and numbers 0,1 are related to white according to synesthetes.

The concept of Synesthesia in gaming is based on a similar ideology, synchronizing sounds with visuals to enhance and lead the gamer to a different place where interaction with the device has consequences and can create a more complex and affective emotional connection.

Among famous people, painter Wassily Kandinsky is notorious for combining four senses into his paintings: Color, hearing, touch and smell. At the same time, more recently singer Lady Gaga has mentioned in numerous interviews how she associates colors with her music and it is part of the song writing process.

## 2.4 Exploring the world together

In recent years and with the booming of mobile communications, it is easier than ever to be connected to the internet at all times. In the beginning, this was only used to offer communication services but has now extended to offer many different experiences including gaming, gps and location sharing.

This has led to the popularization of Geocaching, a new form of entertainment that relies on the concept of a real world treasure hunting game. With over 5 million active users worldwide, it has created a strong sense of community that like few other applications out there, encourages users to explore their surroundings alone or surrounded by other geocachers.

It is interesting in that it merges the digital world with the real world. It encourages it's users to hide objects worldwide and then have other players go out and find them.

Based on research by North Carolina State University [37], Geocaching is a great tool to enhance the way users interact and understand the possibilities of their electronic devices, as well as their multiple connectivity options. It has proven successful as a new form of entertainment as well as an innovative tool for education.

As part of the experiment by NCSU, users were given a GPS for the first time and explained the rules of Geocaching. The participants mentioned that the experience was provoking:

"I believe that geocaching is a grand way to incorporate technology, physical activity, and logical reasoning."

They mentioned that by being forced to go outdoors, the application provided them with a sense of excitement and wonderment, since they never knew what to expect once they started playing.

At the same time, the experiment was not without it's faults, and a few technical issues arose. One participant mentioned: "The microwave tower and three-story buildings messed up our GPS and distorted our coordinates. Sudden failure in technology (GPS units) was a hindrance, but it turned out to be a springboard for our learning."

However, even though GPS signals remain unreliable indoors, strides are being made to enhance the technology and offer a better experience for the users.

Google, creator of the popular GPS service Google maps[1] is aware of this situation and is currently working in a solution. To do so, it has created an innovative program that allows users to walk around an indoors location to then send the information to the server to help map the place. This is an interesting approach that lessens the burden of having to map each individual map on their own.

With this technology, if a mall, airport, stadium wants to offer a detailed map of their venue through their service, all they have to do is start the application, follow a few steps and start walking. The application will record the data and it will be available for other users in the future.

This is a breakthrough in mobile navigation, that also has business potential. According to Forbes magazine [19], the popularity of Android Devices[25] as well as the availability as a free application provides an excellent opportunity for companies wanting to offer their services. As soon as a user enters a building of any kind, he will have an instant view of the vendors that are nearby, real time promotions, directions on how to get there. The growth potential of such application remains mostly untapped, but if done correctly would turn in millions in revenue for Google.

This innovative approach of asking users to collaborate with the company to do indoor mapping provides a great insight of a successful collaboration between a company and the users to enhance a service. As it is now, Google has already partnered with Airports in Atlanta, Chicago, San Francisco, and Tokyo as well as some Malls, Ikea locations world-wide and other locations.

## 2.5 Sounds create emotions

The concept of emotional design is based in the book by Donald Norman[26] and it explores how emotions play a crucial role in understanding our surrounding and learning. To put it in a different perspective, attractive things work better.

According to Norman, humans react to design in three levels: visceral or how we perceive the object (the way it looks), behavioral (how the item performs) and reflective (the way we feel about the object or the memories it evokes).

This is important to understand if we are intending to create an emotional connection between our device or service and the user. We need to find the right balance between innovation and familiarity, between complexity and simplicity, all while making it visually appealing for the user.

When designing emotional sounds however, we don't have access to that visual element, so we must rely mostly in reflective reactions, meaning we must evoke memories from the user or to aim trigger a specific reaction or mood.

During an interview [33], the sound designers of the horror movie Nightmare on Elm Street give some insider information on how to excite the audience through the use of sound. They mention the importance of putting the audience in a setting that they can relate to:

”...During Kriss nightmare when she is outside looking for her dog, you hear ”real life” sounds in her backyard that put her and the audience on notice. We hear some wind, wind chimes, a night bird but if you listen carefully to the night bird you realize that the sound it makes begins with Freddy’s new signature sound of him ”scissoring his finger knives” and quickly morphs into the wing flaps of the bird. Its subtle, but effective. In general, we made an effort to have our design feel organic in nature. If anything felt ”processed” we usually would modify that element to keep things sounding more natural.”

The sound familiarity creates an emotional connection that drives the sound element as one of the most important to the audience and it sets the mood for the following sequences.

In the video game world, survival horror game Silent Hill has been renowned for its extensive use of sounds as a way to trigger emotions in the player. Designed by Akira Yamaoka, Akira states that a sound will reach the brain before a visual element, therefore changing the player’s expectations and emotions.

In his games, location background sounds provide a very valuable tool create emotions and control tension in a more subtle way than music does.

Sound design has greatly evolved from the early stages of device interaction and it is now a powerful weapon used by creators to enhance experiences, evoke emotions and feelings and transport the user to a new world. What was once considered an afterthought is now an important element of game design.

## Chapter 3

# Implementations and related projects

The Ambee application introduces a novel approach to offer interactive experiences to users through the use of their electronic devices. We aim to create engaging experiences for the user, and to analyze ambient sounds and location to provide a new form of entertainment.

While the idea of speech recognition is not a new one, the concept of sound identification was first researched around the late 90's when early pioneer StarCD introduced the idea of identifying music being played through radio stations. However, slow internet access hindered these early experiments. The user had to contact the radio station while the song was being played, then the radio station would automatically consult a third party play-list to match it with the song sample being sent by the user.

It was a novel idea that was however limited only to the songs being played by the station and being monitored by the play-list provider.

However, fast forward to today and internet access is mostly ubiquitous. At all times we are connected through all kinds of mobile devices, from laptops, music players, smart phones and tablets. With this kind of access, it became easier to use QBE (Query by Example)[12] to develop services and applications.

QBE is a lot more flexible than what StarCD originally offered, since it automatically takes a sample of the track and compares it with an existing database[15]. This led to continued development in the area of sound sampling. In the early 1980's a company called Broadcast Data Systems[29] developed a variety of methods to correlate sounds and waveforms.

In the next section we will analyze different applications and games that offer experiences in similar areas.

### 3.1 Muscle fish

Big advancements in sound analysis technology were made in the 90's, when Muscelfish[11] developed a new method that was based mostly on Euclidean distance metrics and multidimensional feature analysis.

Unlike other systems that use audio retrieval work, the system created by Muscle fish differs in its content based capabilities. In their program, a multitude of perceptual features such as loudness, brightness, pitch and timbre can be used to represent a sound. After the analysis is done, it is then possible to assign the sound to one of the sound classes in the data base.

### 3.2 Shazam

Shazam[18] is an application available for multiple platforms that allows users to scan music playing in the background and display the song information on the screen of a mobile device.

The application has gone through several different versions and benefited immensely from the soaring success of smartphones with constant internet access, specially on iOS and Android platforms.

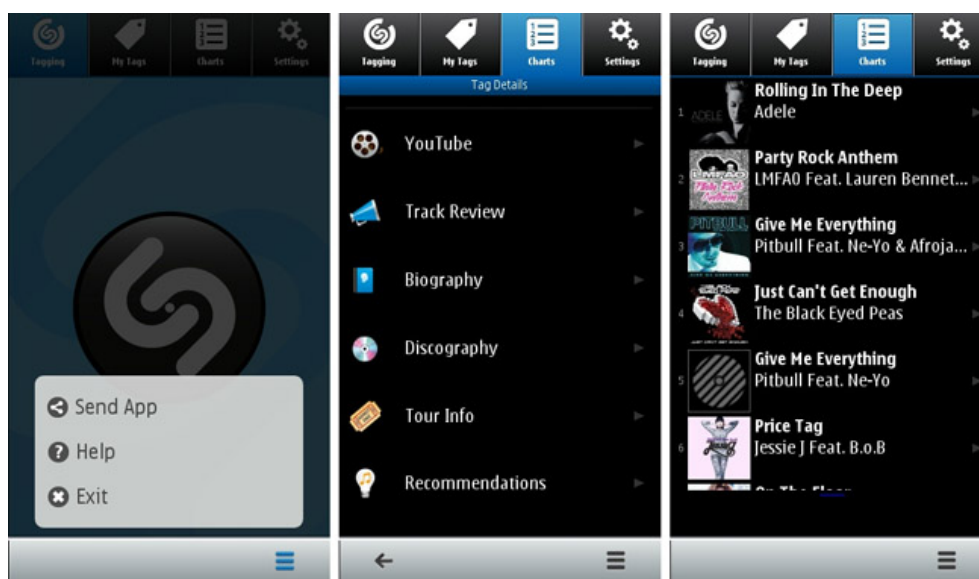


Figure 3.1: Shazam for iPhone user interface

### 3.2. Shazam

---

The first version released in the U.K. in 2002 allowed users to scan the sounds through the built-in microphone of their devices to take a small sample. The user would then have to send the file by SMS to Shazam (through the use of a special phone number) and Shazam would analyze the data and reply with an SMS containing the song information.

One of the main limitations of the technology is that it only works with pre-recorded music, since it has to match the sample with an existing data base, therefore limiting part of its potential.

The current version of the application offers a very straightforward and accessible design and is tightly integrated with social networking and shopping services. After a song has been analyzed, the application will show the track information and album cover, it is then possible to buy the track, ring tones, look for related videos on you tube or even share the track through Facebook, twitter and other services.

Several similar applications have appeared since the introduction of Shazam, amongst them:

Sound hound: Same functions as Shazam but also offers the option of displaying the song lyrics on the screen turning the device into a portable karaoke.

Midomi: This application can look for recorded songs when the user sings or hums the desired track to the phones microphone. It however relies on user-sent samples making it less reliable when compared to other services.

Gracenote Music ID: Gracenote analyzes the music from CD's and mp3 files to get the track information, artist, composer, etc. It is particularly popular with home entertainment systems. The main advantage is the database which contains more than 28 million tracks.



### 3.3 Navinko

Navinko[22] is a mobile GPS application available for iOS that uses an innovative sound based interface to provide information to the users. It also includes important social elements like friend tagging (assigning special sounds to specific users) or bike pooling, to organize social gatherings and explore the city.

The application was developed as part of the Urban Media Group of Keio University and was launched in 2009. Through the use of hands-free voice control, users can share locations with friends and record important information about their urban discoveries.

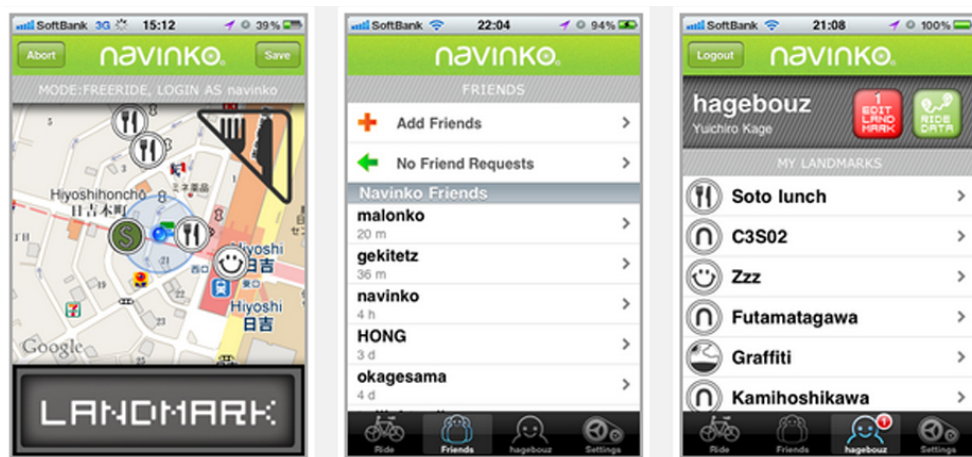


Figure 3.2: Navinko navigation system user interface

The main issue it had was the limited user base, since it was designed to cater to iPhone users with a 3g connection and a bicycle. This greatly limited the user base and at the same time made it complicated to get a sponsor for the project.

The other obstacle was that in some countries (including Japan) the use of headphones while riding a bicycle is illegal. At the same time, having the mobile device connecting to the GPS all the time proved to be a challenge for the battery life of current phones, and it was enough to drain a fully charged iPhone in about 40 minutes during internal tests.

At the same time, some successful things were achieved by the application, such as important advancements in designing a user experience that relies mostly on sound information. This is particularly important with a navigation application, as it gives the user more time to spend enjoying their surroundings and less time looking at the phone screen. The application is available for iOS devices.

### 3.4 Lumines

Lumines[13] is a video-game designed by music game developers, Tetsuya Miziguchi. It relies heavily in the concept of Synesthesia, often described as seeing colors, hearing sounds.

The game is a puzzle game similar to Tetris where blocks fall from the upper part of the screen and the user has to align them and trigger their elimination based on timing. Failing to do so, will cause the game to end.



Figure 3.3: Lumines Puzzle fusion. Lights and sounds.

However, the big difference with other puzzle games, is that sound and lights play a big role in the game. The music dynamically changes according to what happens on the screen and specific actions trigger explosions and music changes. The first game available as a launch title for the Sony Playstation portable was so well received that spawned a few sequels for different platforms.

However, given the success of the music created for the game, the developers decided to create a music band called Genki Rockets to continue making game music, concerts, and special collaborations with other companies like Sony and Panasonic.

### 3.5 Tamagotchi

Tamagotchi[44] is a portable digital pet introduced for the first time in Japan in 1996 by Bandai corporation. To date this portable companion has sold more than 78 million units. The Tamagotchi phenomenon is one that cannot be underestimated, as it is the perfect example of modest technology offering a robust user experience and is particularly impressive because of its sudden raise to popularity worldwide as a pop culture phenomenon.



Figure 3.4: The first tamagotchi (Model V1). A portable revolution.

The first unit (Model V1) consisted of a primitive black and white LCD and three buttons. However, this was enough to include life functions like feeding, health, playing, sports and so on. The many different interaction possibilities with the character guaranteed the lasting appeal of the device, not to mention the exceptional battery life.

However, two main features are usually attributed to the instantaneous success of the tamagotchi.

**Low cost:** At an introduction cost of around 34usd. It made it easy for families to buy them for their kids, and became an almost instant impulse purchase nation wide when it was introduced in Japan.

**Size:** The small size and long battery guaranteed that it was an always on device. Even if the user only had a few minutes it was enough to spend some time playing with the Tamagotchi and experimenting with its features and quirks. This created a special relationship within the device and the user as a reliable and dare we say affective form of entertainment.

**Surprises:** another element that made the Tamagotchi a success is that every time a user plays it, he gets a completely different experience. Random events happen that surprise the user every passing day and keep him coming for more. Given the success of the original Tamagotchi, it has expanded to multiple markets and newer devices offer color screens, social interaction elements and other advanced features.

However, the Tamagotchi craze seems to have slowed down in recent years, and current models are aimed mostly at a young female audience, failing to have the nearly global success it had at launch more than a decade ago. It has failed to adapt to a new generation of portable devices and connected users.

## 3.6 Seaman

Seaman[31] is a life simulation game released originally in 1999 for the Sega Dreamcast home console. It is similar in concept to the Tamagotchi in that it offers users a virtual pet to interact with. However, the main difference with other projects is that the Seaman was programmed to react to voice commands through the use of a special headset accessory released for the game.

The game would basically transform the television into a water tank and give the user the responsibility of looking after the Seaman as well as analyzing its life.

It is a very unique game in terms of design because of the limited actions that the gamer can actually do. It provided basically no cues on what would trigger any sort of reaction within the game, encouraging users to experiment to make progress.



Figure 3.5: Seaman for Sega Dreamcast. Life simulation evolved.

The developers were praised for including many Easter eggs (surprise elements) into the game that kept the most hardcore fans entertained for years to come.

Seaman received an Excellence Award for Interactive Art at the 1999 Japan media Arts festival and received the Original Game Character of the Year award at GDC 2002. In 2008, Game Informer named the game one of the top ten weirdest of all time.

In a way, it could be said that Seaman is the silly and primitive version of Apple's newly released Siri personal assistant service for the iPhone.

### 3.7 Milo

Also known as Milo and Kate[23], it is a tech demo introduced by Microsoft at the 2009 Electronic Entertainment Expo in Los Angeles, CA.

The technical-demo was used mostly to introduce the concept of a controller-free gaming platform where sounds and images would be used to interact with a virtual character.

It offered the first glimpse at what the new Kinect camera for the Xbox 360 console could offer in terms of advance interaction with users through video and sound recognition technologies. The demo was well received by the gaming press and helped Microsoft sell millions of Kinect devices based don the promise of advanced interaction.

There were multiple reports that intensive research was being carried out at Microsoft to turn the demo into a retail game, however, this rumors never materialized and led to the departure of Peter Molyneux, the main developer behind the concept, from Microsoft Game Studios.



Figure 3.6: User interacting with Milo trough a Kinect camera.

One of the main reasons for the cancellation of the game was the apparent pressure that Microsoft had on the development team to turn the life simulator into something more akin to a traditional game experience, things that went against the vision of the game designer.

## 3.8 State of Development

As we discussed in the previous chapter, improvements in technology and lower production costs have led to many innovative products and services over the last few years. However, many interesting concepts fail to make it to the marketplace due to lack of support or an inappropriate concept behind them.

Innovative ideas for AR and location based services have existed for some time now, and the technology available is both reliable and comparatively inexpensive. It is however important to find a core idea that attracts both, users and companies to create a commercially successful product or service.

In the next chapter we will describe the scenario and the inspiration behind the ambee concept.

## Chapter 4

# Ambee Concept

The creation of the Ambee project is the culmination of extensive research in the area of sound and location based interaction experiences. It marks the beginning of a collaboration between the Japanese branch of a leading Mobile chip maker and the team members of the Around.Hear group at Keio University.

While pitching the project to our sponsor, the application went through several different revisions in multiple areas until a final agreement was reached between the two parties. It was decided to develop a digital pet mobile application that would rely on their background sound analysis technology to dynamically generate contents for users. In return we would design an experience that would help them expand their ambient sound data base by obtaining the sound samples recorded by users. Increasing the sound database is a key element to guarantee the reliability and accuracy of their new technology.

In this chapter we will analyze the most common scenarios where the Ambee application will be used, as well as a brief description of our target audience based on the request from our business partner.

### 4.1 Scenario Analysis

The Ambee concept started while analyzing the way people live and interact with the city and their devices.

As our lives in a Metropolis like Tokyo become busier and more chaotic, we have less time to relax and enjoy casual entertainment experiences. We spend little time at home, so the need to create new and novel forms of mobile entertainment arises. The inspiration behind the Ambee concept came to exist based on the analysis of the following scenarios:

#### 4.1.1 Crowded Trains

It is a reality that we live every day. Trains are getting more and more crowded and it is not always easy to have access to our devices (phones, tablets, game

consoles). However, one look around and we realize that most users are either listening to music through their headphones or struggling to look at their screens.

Ambee aims to create a form of entertainment that can be experienced without relying on visual information.

### 4.1.2 Tapping on screens

While listening to music, we tend to react to the beat of the sound we are listening to. Part of this was noticed by co-researcher Yupeng Pan, when he noticed how many users tap the screen of their devices with their fingers while they listen to music in different locations. This random movement that is often involuntary can be analyzed to change the way we interact with our devices.

### 4.1.3 Waiting

In a big city, there is many situations where we have to wait for a few minutes. Restaurants, Train stations, Convenience Stores, Karaoke. The waiting time even though variable, is always enough to create some sort of frustration, yet not long enough to start a full fledged game or start reading a book. During these short periods of time, I often find myself bored and wondering what to do. The purpose of Ambee is to create an experience that can be had instantaneously after launching the application.

### 4.1.4 Moving around new areas

Exploring new areas of the city is a big part of our lives, specially when we travel or move to a new city. Many times we visit locations for the first time and wonder around to see what we can find. This is often limited to the real world, but has successfully been expanded to the virtual world through gaming applications and Augmented Reality. Visiting different places provides an opportunity to provide different contents and forms of entertainment to users that is often overlooked in a design process.

### 4.1.5 Background Noise

People who move from a small to a big city like myself, often complain that it is too noisy and that there is always background noise wherever we are. To block ourselves from that noise, we use headphones to create our own virtual environment. However, we can use this noise to create a bridge between the real world and what we are doing with our electronic device. Ambee analyzes this sounds and creates music and other related contents (costumes, locations, accessories) for the character related to our current location.



## 4.2 The ideal Ambee user

Our target user is described briefly in the next section. The application is being designed for local distribution first, with a possible global launch in the near future. With this in mind, we aim to design an application that can be used in as many markets as possible without deep modifications

Our target user falls in the following categories:

- 1) People who like to have fun with technology and like to experiment with novel concepts and application. This is the kind of early adopter that can test our service from the start and help enhance the experience before an open release.
- 2) Urban citizens who have to spend time in the city, time commuting and who have a few minutes a day to spend gaming.
- 3) Music and Gaming enthusiasts.
- 4) People with a Smart phone.
- 5) Ages 15-35.



Figure 4.1: Our target user.

### 4.3 Design Workshop: Creativity Box

To analyze how users react to sounds and the concepts to be included in the Ambee application, we invited a group of people to participate in a character design workshop.

The workshop was designed to see how users were influenced by being exposed to multiple familiar background sounds. The users did not know what sound they were listening to.

Users were invited to use a desk that was covered by a black cubicle. Once inside, they would be blocked from any external light and sound. Through the use of a small LED light inside the cubicle, only a piece of paper and some markers were visible to the testers.

Once the test started, the user would listen to a sound sample for 3 minutes. Using noise-canceling headphones, the participants were blocked from external sound as well. Each track consisted of ambient sounds recorded on the following locations:

Track 1: Fast food restaurant Yokohama Bay Quarter

Track 2: Inside Train toyoko line between Yokohama and Tanmachi

Track 3: Party inside a family restaurant

Track 4: Yoyogi park

Track 5: Coffee shop in Tsunashima Station

### 4.3. Design Workshop: Creativity Box

During their listening time, testers were asked to design characters that they related to the things they were listening to. In our early internal testing, users were confused and often blocked and didn't know what to draw, since they couldn't look anywhere to look for inspiration or copy design elements.

For that reason, on the day of the test we provided a worksheet that included a brief explanation of the activity as well as some elements that they can use to design their characters.



Figure 4.2: Design form and tester inside creativity box

In total, 16 participants designed characters during the workshop. Their designs were scanned and transferred to digital format by the team members of the Around.hear team of Keio University.

### 4.3.1 Character design analysis

The characters obtained after the workshop were in general, unexpected to our team. It was interesting to note that some sounds that were part of the recordings made an impact and influenced the participants design, even though we hadn't heard them previously. Sounds like cars in the background or thunder helped the designers get more creative while inside the dark cubicle.

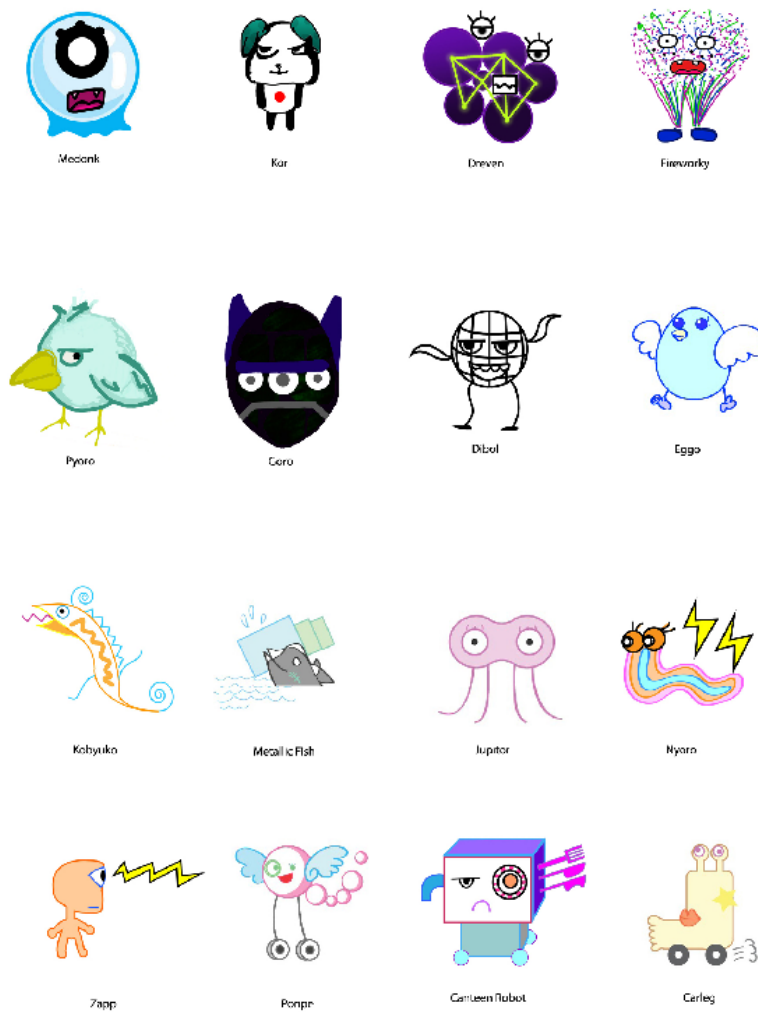


Figure 4.3: Characters from Creative Box Workshop

## 4.4 Visualizing Background noise

Sounds ultimately had a direct effect on how users created their characters. With this came the realization that using background noise and sounds to dynamically generate contents is an unexplored method that has unlimited potential.

Moving around the city we are always exposed to different locations and sounds that provide a novel context to enhance our service in a meaningful way. With the use of modern electronic devices like smart-phones we can analyze sounds to create contents in real time.

The service is designed to be used by the consumers described in the scenarios at the beginning of this chapter, as well as having users upload sound samples recorded through their devices to our server, one of the requests of our business partner. In the next chapter we will streamline the design framework for the Ambee application and how it creates new forms of entertainment. The focus is on the sound and basic interface elements of the application.

# Chapter 5

## Implementation: Ambee

### 5.1 Ambee Design Framework

When development of the virtual pet application began, we aimed to design a novel interactive experience that would rely on location and sound analysis. The outcome of that research is AMBEE: Sound Interaction.

The name is based on the importance of Analyzing Ambient sounds for entertainment.

AMBEE: Ambience.

The application is aimed at young users who live in big cities who have a few minutes a day to play with their devices and we can offer them an innovative experience as they move around the city. If you are somewhere new, you can play something different with AMBEE.



Figure 5.1: Ambee logo.

The Aesthetics of the logo and the color palette for the project were chosen to target a young adult user base. The colors were chosen to represent an exciting sunrise in a big city, where the day is starting, the user is leaving home, the exploration has begun. It also includes the silhouette of Tokyo Sky Tree, a new building that is likely to be emblematic of Tokyo, our host city, in coming years.

### 5.1.1 Early prototypes

As part of the design process, early prototypes and demos of the application have been developed for early usability tests as well as pitching to our collaborator during the negotiating process. The prototypes were developed using the Flash platform and are in the process of being translated for the Android SDK. The aim of Ambee is to analyze sounds and their respective locations and generate contents for the user through sounds and on screen animations of the character. To implement this goal, the application has been designed with the following concepts in mind.

### 5.1.2 Dynamic Sound Interaction

We are always surrounded by different kinds of noise. However, what we consider annoying also has information that can be transformed into a more pleasant sound by the Ambee application.

Analyzing the locations that we visit more often we can come up with a list of sounds that we are often exposed to.

Say for example we spend the day at the office. Through the day, we will very likely at some point listen to the air conditioning system working, a printer printing a document, the fan of a computer working, even flushing the toilet. These sounds that we often ignore all carry the elements that can be analyzed and dissected through a logarithm and can be transformed into music.

The goal of Ambee is to create an immersion application that lets the user interact with an on-screen character while it analyzes incoming sounds and location to generate music and entertainment contents. With this in mind an early demo of the application was created where the character can detect constant sounds (printer, air conditioner, airplane engine) and replace them with a song.

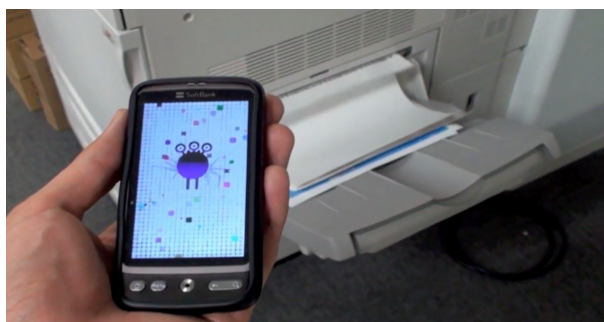


Figure 5.2: Printer triggered music generator

Through the use of an Android phone's microphone, the logarithm of the application can detect and analyze the incoming sound in real time to create a track that matches the beat and rhythm of the sound. On top of that, the character reacts accordingly and dances to the rhythm of the song being created, offering a new form of entertainment. This is a novel approach to generate contents based on something we would otherwise consider background noise.

### 5.1.3 Gamification of music

Over the last decade, the way we experience music has not changed in any significant way. Users are less likely than ever before to buy physical media or to pay for contents. This has been noticed by some innovative media creators that have decided to provide tools rather on top of recorded music to their fans. Basic applications or services are now provided (often free of charge) that let them interact with their music in any way they see fit. With Ambee users can interact with their music by tapping on the screen and using different multi-touch gestures

In an interactive Demo, the character is displayed in the center of the screen while music plays in the background. The user is encouraged to interact with the music by using a different gestures that can be recognized by the device's touch screen.

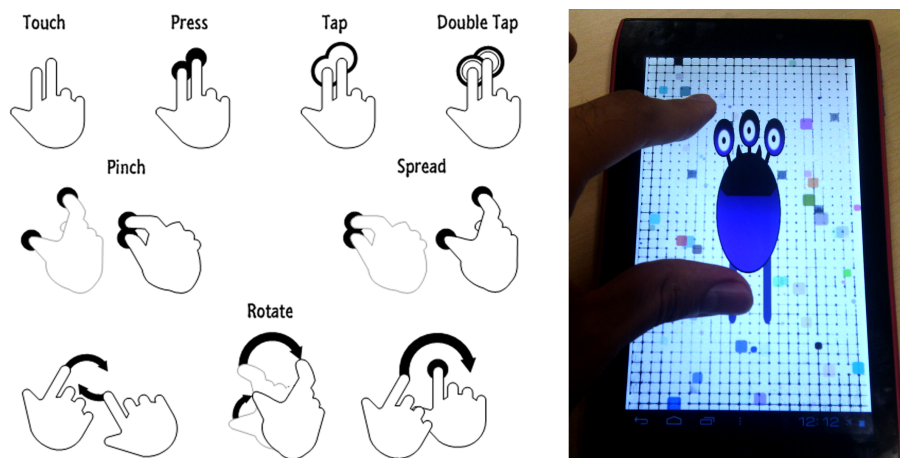


Figure 5.3: Interactive Musicality



This was designed to warrant that the user can directly interact with the music without having to look at the screen and based mostly on aural information.

The sudden increase of popularity of smart phones has meant that music players as we know them have mostly ceased to exist and users now have in their pockets powerful devices with advanced microprocessors, multi-touch touch screens and constant internet connections. Ambee uses this hardware to enhance the music experience. This can be used with pre-recorded media or with contents being generated on the from the application itself through background noise analysis.

### 5.1.4 Location Aware benefits

Offering automatic location detection, we can guarantee that the experience will always be different and fresh. While many forms of entertainment in the gaming, music and movie industries try to replicate the experience offered at home, they often fail to enhance it or change it in any meaningful way.

With Ambee we aim to solve that issue by dynamically adjusting the contents by analyzing the location and the background sounds.



Figure 5.4: Shake to scan.

To activate the location scan it is necessary to shake the phone for a few seconds. This activates the sound scan in the device for a few seconds to analyze the background sound information.

The current SDK from our business partner allows us to detect the location and change the contents accordingly. The character changes clothes depending on the place where the application is being used.

It is possible to analyze the background sound in-app or if the device has an internet connection, information is sent to and from server for better quality location matching. Handling the analysis within the code helps us improve battery life, as well as guarantee that users can continue using the service overseas, where internet roaming access is expensive.

## 5.2 User Interface and functions

The user interface is designed around the concept of Musicality meaning that every element has a special sound that the user can recognize after spending sometime with the application.

My goal when designing the main elements of the app was to make sure that some basic functions can be accessed when the user is not looking at the screen.

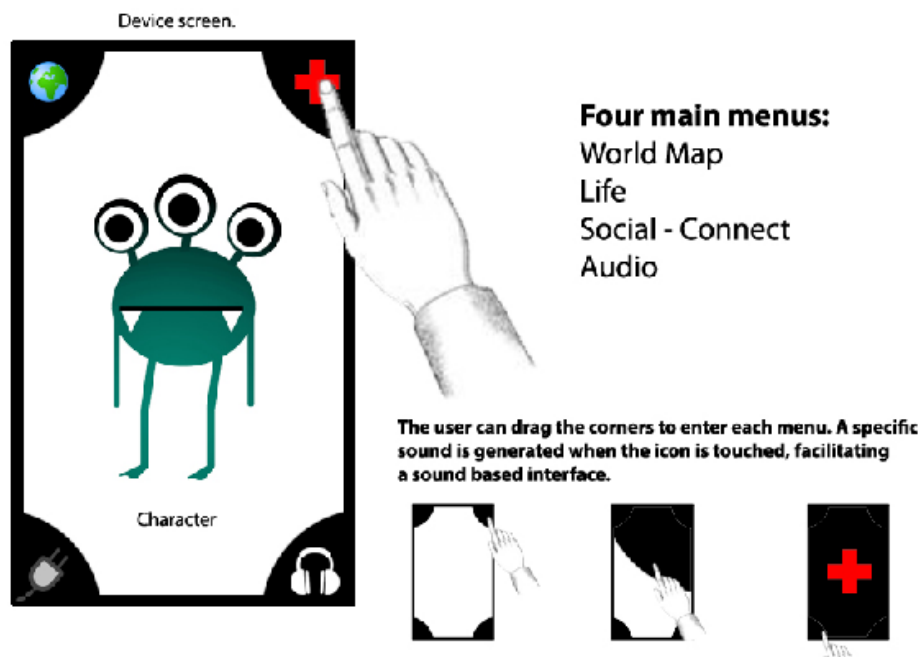


Figure 5.5: Four menu user interface.

The main screen features the character prominently in the center, as it is the main element of the application.

Four bubbles placed on the corners of the screen provide access to the main

4 elements of the game. Life, Music Generator, World Map, and Friend list. Each bubble is assigned an individual original sound.

When a bubble dragged to the opposite corner, the menu is selected. By doing this, we grantee that once the user is familiar with the location of the menus or the sound they generate, they can access and move around the menus without the need of visual clues.

### 5.2.1 Life

The life section of Ambee covers all the vital signs and the ways you interact with the virtual character. It is one of the most important sections as it needs to engage the user and create an emotional connection between the user and his Ambee.

In this section, the user can access infomation about the character like age, favorite sound, etc.



X-ray view

#### Life

**Character history:** check the vital signs of the character, life expectancy, etc.

**Food level:** Ambees live of sounds if you don't play with them, they will lose body parts and eventually fade.

**Clothes:** When you visit locations, you can obtain clothes and uniforms for your Ambee. You can wear them later.

**Mood:** Your Ambee can develop his own personality based on the attention you pay to him. Share him on facebook? He will be proud.

Figure 5.6: X-ray of an Alien life form.

As part of the interaction element of Ambee, each character has a growing pattern. It starts as a simple basic life form, but will continue to evolve in surprising ways as the real time goes by. This creates a surprise element for the players and aims to keep them engaged longer.

### 5.2.2 Sound Generator

In this section, our Ambee becomes a music interaction tool. It is possible to access our music collection and interact with the sounds through the use of the touch screen. Our intuitive interface makes it possible to play with the music without looking at the phone. It is possible to save our remixes and then share them with fiends.

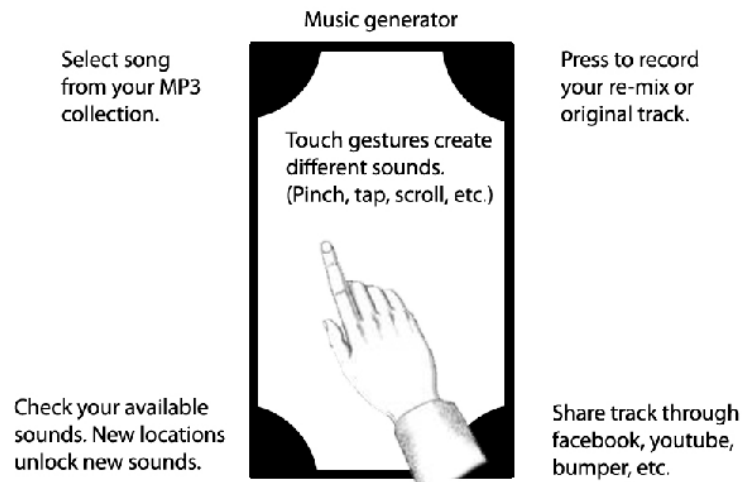


Figure 5.7: Ambees live from sounds, feed them music.

Each gesture generates different sounds that can be used to enhance our existing music collection. By pinching, zooming, tapping the screen it is possible to play with our tracks in real time all without necessarily looking at the screen. This tracks can then be shared with your friends through the application or social networks. This creates a platform not only for interaction but for artistic expression.

### 5.2.3 World Map

The World map section shows the user the locations where he can obtain special items for this Ambee as well as a list of our sponsors and collaborators for future versions of the application.

When the user accesses the world map, he can spin with his finger to look at the completed areas and locations.

The map also shows unlocked locations and finished challenges. It is possible to see a percentage of the World Map that the user has conquered. With a colored coded design, the map can be divided by location type or POI (points of interest).

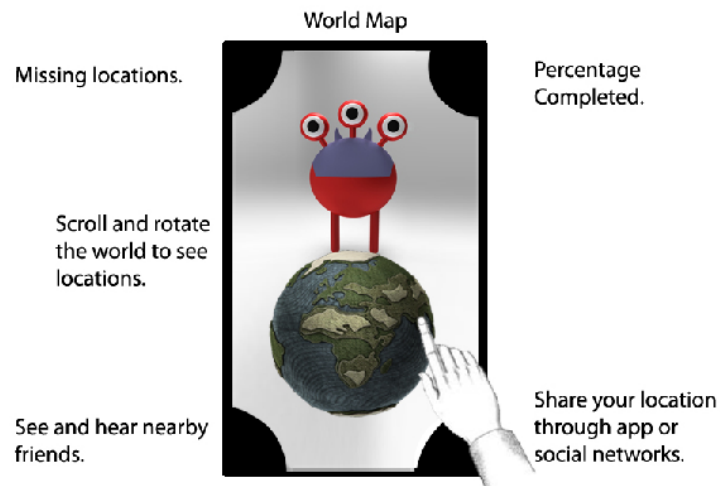


Figure 5.8: Exploring the world together.

As the user walks around the city with our Ambee, he will run into hidden items and treasure that can be unlocked automatically, triggering a sound on our device.

For example, the user is walking near a Clothing store, the Application can detect the location and give a digital version of the clothes on display at the window. The user can select this object in the life section at any time to dress his Ambee with this branded item.

### 5.2.4 Social

In the social section of the application it is possible to hear other friends playing Ambee or other users nearby using the service. It is possible to tag friends with an specific sound. That way if the user decides to do so, he can use Ambee to meet new people as well as getting notifications when friends are nearby.

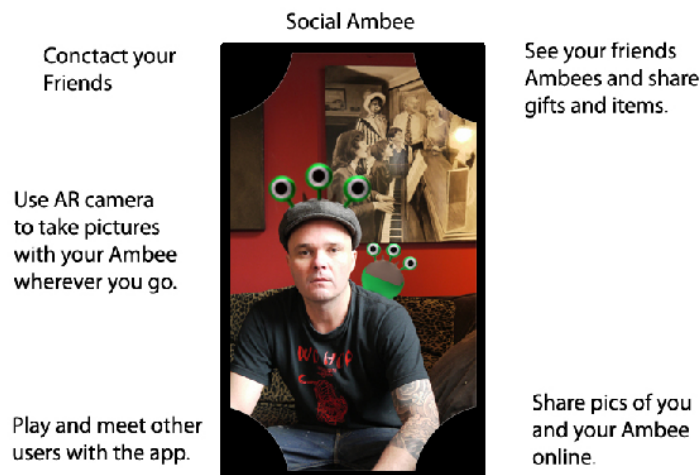


Figure 5.9: Sharing the Ambee experience.

As part of the social element it is possible to take snapshots of the character at any time that can be share through social networks. This is an important element designed to create a deeper connection between the gamer and his pet.

By using the built-in front facing camera available in many electronic devices today and with the help of basic AR technology the user can blur the line between real world and digital world by taking pictures with his Ambee wherever he goes. This kind of interaction has been designed to give the users more incentives to use the application outdoors and when traveling.

## 5.3 Application Framework

The current demos of the application have been developed in the Flash platform, since it is possible to export the code to an .apk file that can be recognized on Android devices. Since the core element of the experience is the sound recognition technology, the following framework has been selected for the location matching function of the application.

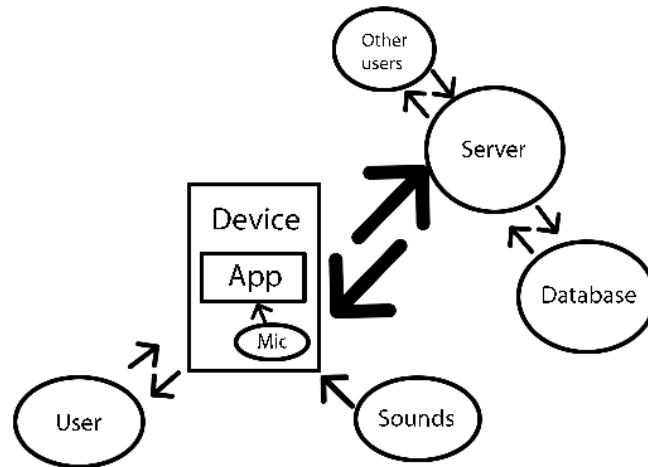


Figure 5.10: Framework of the Ambee application

When the user activates the location scan function of the application, incoming sounds can be recorded and analyzed by the built-in microphone. After a short sample has been recorded (5-10 seconds) the information is analyzed in app to create a small image of the sound. This information of around 30kbytes is then sent to the server through the internet connection on the device (wifi or 3g).

Once the file has been sent, the server will compare the sample with the existing data base to match the sound and assign the best match. The information is then sent back from the server to the device letting the application know the kind of location where the application is being used.

This in return opens an opportunity for our business partner in that it allows the database to be extensively updated and expanded with the help of the users sending their sound samples through their server.

This is specially convenient or indoor locations where GPS signal is limited and also provides a new possibility for travelers, since a basic data-base of samples could be stored in the application itself to allow basic matching of the most common locations without the need of an internet connection.

## 5.4 Hardware requirements

The current SDK is based on the Android platform and is compatible with a wide variety of hardware configurations. In order to make the system work properly, a noise canceling microphone and an internet connection is recommended for a more accurate sound analysis. However, any phone with Android 2.3 (Gingerbread) or later is compatible with the required SDK.

### 5.4.1 Sound analysis Methods

Sound analysis is based on the concept of extracting information of an audio file. This information can then be used for different purposes, such as comparison, storage, classification, synthesis, etc. These are the methods used to analyze the background noise within recorded with the device.

### 5.4.2 Euclidean Distance

Euclidean distance[36] is an analysis technique where a spectrogram is measured to classify a sound. The image is analyzed and special points in the sound wave are distanced. By doing so, it is possible to compare with an existing database to establish a correlation of the sound either individually or as part of a context.

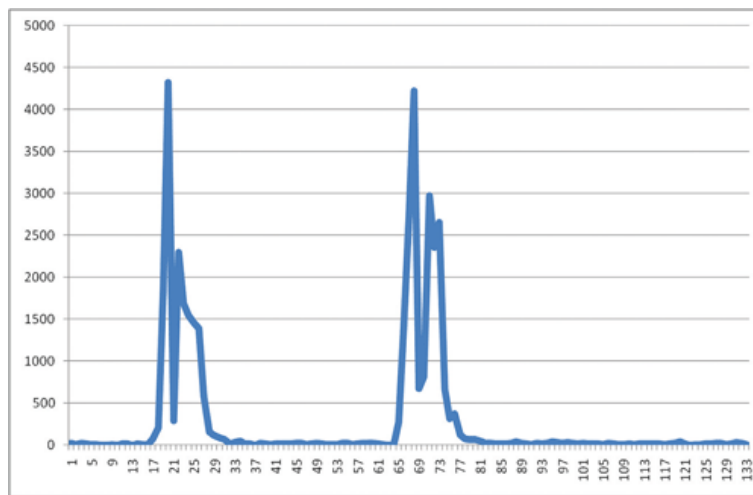


Figure 5.11: Two sounds analyzed by Euclidean distance method.

### 5.4.3 Spectrogram

Also known as voice prints, a spectrogram[28] is basically a graph that shows a sound's frequency on the vertical axis and time on the horizontal axis. Different speech sounds create different shapes within the graph. Spectrograms also use



colors or shades of grey to represent the acoustical qualities of sound.

Modern voice print recognition is advanced and reliable enough for some companies to offer access to information based on voice commands to unlock the database.

Basically, the concept of voice identification states that every voice has specific characteristics that are different enough to distinguish from each other through a process called voice-print analysis. It takes into consideration voice uniqueness factors such as vocal cavities (throat, shape and length of vocal cords) that make the likelihood of two people having an identical voice-print appear very remote.

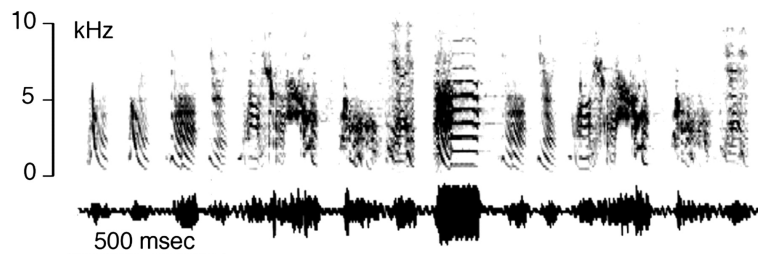


Figure 5.12: Spectrogram analysis.

The other element that can affect voice analysis is the way in which the speaker manipulates speech muscles during the speech. These include the lips, teeth, tongue, jaw muscles and so on.

Many studies have been published on the reliability of using voice-print as a means of voice identification and the results have been encouraging. According to an investigation by the Federal Bureau of Investigation [16], based on the analysis of 2,000 forensic cases, the error rate is under .31 percent.

However, when we analyze voices and sound, a sound spectrograph can create a complex speech wave that can be used to differentiate between two or more sounds. They can also help us identify the source of the sound successfully.

## 5.5 Analysis Challenges

The existing methods of sound analysis have improved exponentially over the last five years with the mainstream popularity of the internet. This meant that now, information didn't have to be stored in our devices, but could also be obtained

in real time from a server. This however created new issues. In the next section we will analyze some of the issues that current sound analysis technologies are trying to overcome.

### 5.5.1 Noise

In some environments, noise[35] will compete with other, more relevant ambient sounds. While being in a car for example, we can perceive traffic or engine noise to identify the location, but other elements such as rain or street noise.

The algorithm should be robust enough to detect the background noise and analyze it properly. For example, the average background noise in a classroom is about 50-60dB, a level that is also considered similar to the sound produced by a person speaking.

In the world of Android Development, the usual common denominator to start testing is the Nexus One manufactured by Taiwanese company HTC. This phone includes a noise canceling system developed by telecommunications company Audience that successfully suppresses background noise during a conversation.

Noise reduction is usually achieved through a combination of hardware and software analysis. The use of a directional mic can help in some conditions, but given the different hardware configurations and segmentation specially in the Android platform, we can not give for granted that the device will have such kind of microphone. Therefore, it is important to do it via software. Usually this is done by using a 1-2 second sample before recording begins to allow the application to adjust the levels to detect any noise.

### 5.5.2 Distortion

Distortion can arise from a variety of sources, such as background noise reverberation (echo). In this case, this elements could affect the way a mobile device takes in the sound data, since telephony equipment reduces sample response to an average of 300Hz-3,400Hz. It is important to note, that many modern phones include functions such as noise canceling and zoom in the built-in mic, elements that can either enhance or decrease the quality of our recordings.

### 5.5.3 Data management

The system must be able to receive and index the sound samples from millions of locations in the server database. For this reason it is important to analyze the sound sample in-app and then send only the relevant data to the cloud. Otherwise, it would incur in intensive processing load on the back end search engine.

A statistical scale within the system guarantees that even if we add a big number of samples, the possibilities of matching the correct one will not decrease exponentially. Another issue is that an application that relies exclusively in an online server or a cloud server, would be unusable should any issues with the network arise.

### 5.5.4 Battery life

Based on previous experience by applications developed at Keio University such as the Navinko application, having a constant internet connection can drain the battery very fast. It is important then to limit server access as much as possible.

Based on similar technology developments, the most battery effective method for sound analysis and sampling is to take a short recording of 5-10 seconds that is then analyzed within the application. Once a simple spectrogram is created, a very small sample file of approximately 30kbytes is sent to the server to compare with an existing database.

## 5.6 Ambee as a gaming platform

The sound analysis technology provided by the advanced SDK provided by our collaborator allows us to not only utilize the incoming sound wave, but also to detect through the microphone the kind of location where the device is being used without having to rely on a GPS signal or an internet connection.

Based on our early user test feedback, the following locations were chosen as ideal to develop the application around use in these areas. A visual example is provided for each location for future developments. Ambee offers mini-games that are related to the place where the application is used.

### 5.6.1 Restaurants

Visiting a fast food restaurant provides a comfortable yet noisy environment where the user has a few minutes to spend. Proposed game themes: Food Making, Guest management, food eating, food fighting.

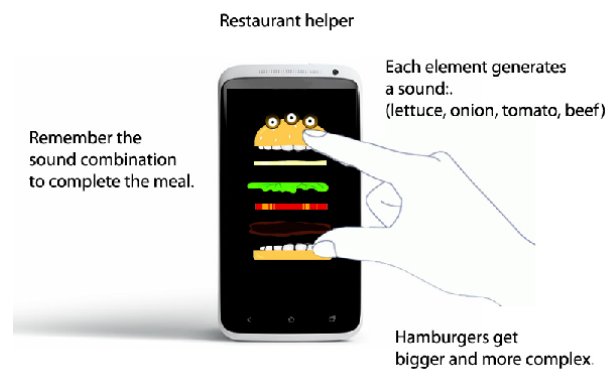


Figure 5.13: Follow the sequence (Simon says).

### 5.6.2 Convenience Stores

In a country like Japan, most people visit a convenience store almost on a daily basis so it is important to offer them in-game benefits for doing so. Companies like Lawson already give points for every time we visit their establishment. We can take that element to the next level and getting branded items and food for our Ambee on every visit. Proposed game themes: Queue management, stock arrangement, disaster prevention.

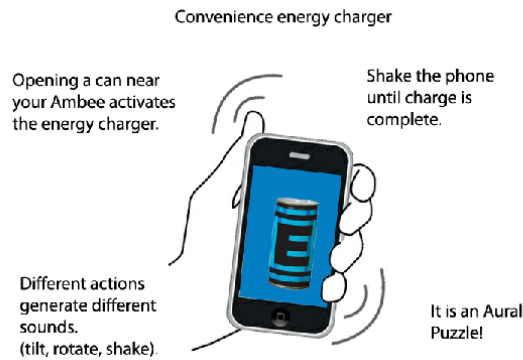


Figure 5.14: Energy supply at the store.

### 5.6.3 Parks

Parks provide a good location for mobile gaming. One of the main reasons is that the users usually spend extended amounts of time there and they at all times have a clear signal and strong internet connection. It is also often done in groups, so it can help users interact with each other through the application in big open areas. Proposed game themes: Hide and seek, Ambee hunt, Sports.



Figure 5.15: Hide and seek outdoors.

### 5.6.4 Train Stations

Waiting in a train station is a common experience in most big cities. In a city like Tokyo, users often have to wait for up to 15 minutes when waiting to the desired train. Because of crowding, it is very normal for train riders to wait next to the track long before the train arrives. We can use this waiting time to offer them a train related gaming experience. Proposed game theme: Crowd breaking, Balance in a crowded space, driving train.

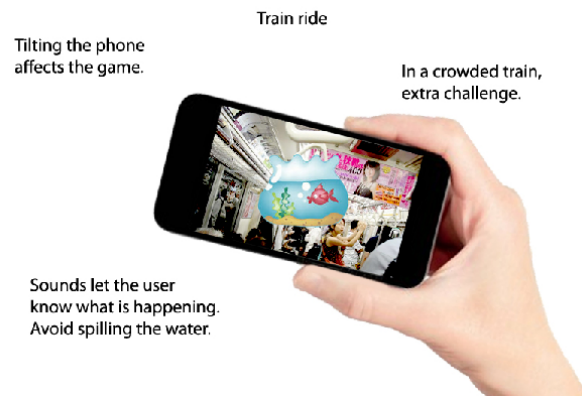


Figure 5.16: Item transportation game.

### 5.6.5 Karaoke Locations

One of the most popular forms of entertainment in Japan. And while in the west it is mostly a social activity enjoyed in open spaces, in Tokyo it is common for people to go to Karaoke on their own. With Ambee the user can take the character and interact with it through the night. The device records our performance and has the Ambee character perform it on screen. It analyzes our vocals and boo or cheer at the end of our song. With the company of our Ambee we are never alone.

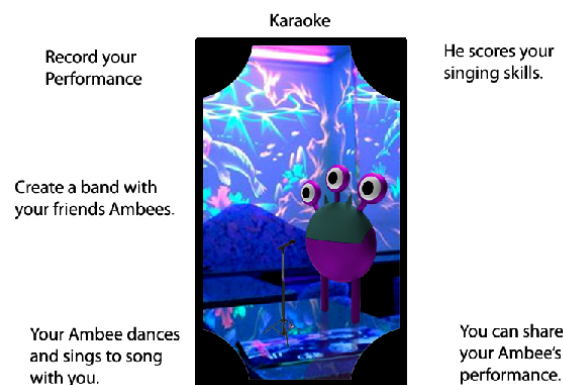


Figure 5.17: A new way of performing.

## Chapter 6

# Evaluation

### 6.1 Framework of this evaluation

In this evaluation, the concepts and demos of the Ambee application were tested in different conditions. Trying to get a better understanding of how the application would work in real world conditions. A basic usability test was carried out where 5 users were asked to test the user interface to see whether or not the sound based interface provides an important leap over existing designs.

Given that the application is designed to be used in different locations, users were asked to use the service under different levels of background noise. At the same time, they used the music interaction demo to see if they could activate the commands without using the screen of the device.

To test the dynamic character creation element, a video was prepared to be shown to a bigger group of users (64). The users were explained the basic functionality of the Ambee application and were asked to answer a few questions related to the most important elements.

Given that a global launch in the future is being considered, it is important to measure the level of interest in the overseas markets.

Finally, a QA session was conducted with the help of game Designer Tetsuya Mizuguchi to introduce him to the concepts proposed by Ambee.

### 6.1.1 User Interface

The aim of this test was to allow 5 users to try the user interface of the application and measure the accuracy of their actions and if users could recall the location of the menus after a few minutes of usage.



Figure 6.1: Tester using interface for the first time

The testers were shown the application for the first time and were explained that by moving their fingers around the screen they would get aural information that would help them know if they were activating a menu or getting closer to a button in the touch screen of the device. Testers were given 3 minutes to play with the interface to analyze their reactions and the accuracy of their commands and then they were asked a few questions about their experience with the application.



### 6.1.2 Music Interaction

In the second evaluation, users tested the interactive music demo where they can play a song and interact through multi-touch inputs on the touch screen. Users were not explained how to interact with the application. They were only explained that with their commands, the character would react accordingly. Users were encouraged to try any touch gestures they remembered using in other applications.

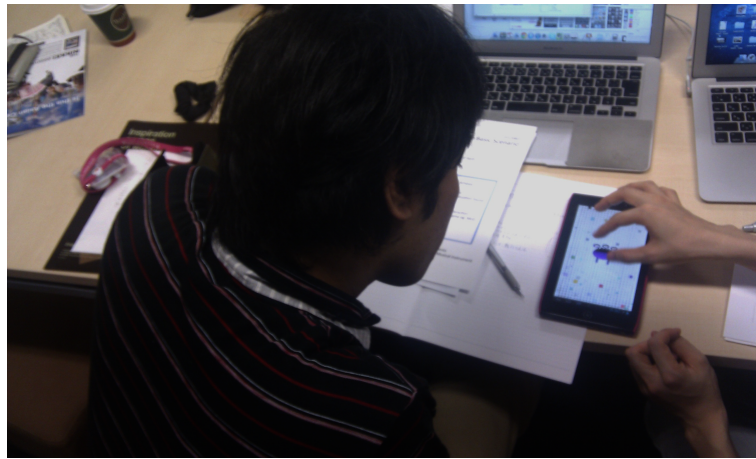


Figure 6.2: Interacting with sounds

The users were then encouraged to try using the application with the mobile device inside their pockets, to see if it was possible for them to activate the commands they already knew, but this time without looking at the screen.

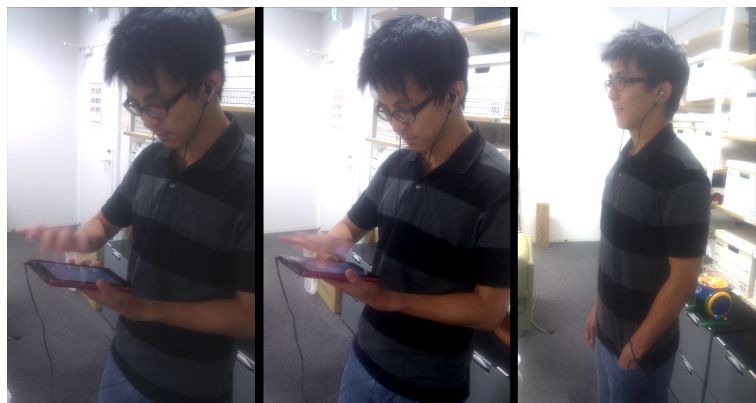


Figure 6.3: Testing sound based interface

### 6.1.3 Location analysis

For this section of the test, users were shown a video of characters being generated based on the background sounds on a specific location. The video contains sections recorded at the beach, in a park, in a restaurant, inside a train and in a busy street. The reactions of the users were recorded to analyze their reactions.

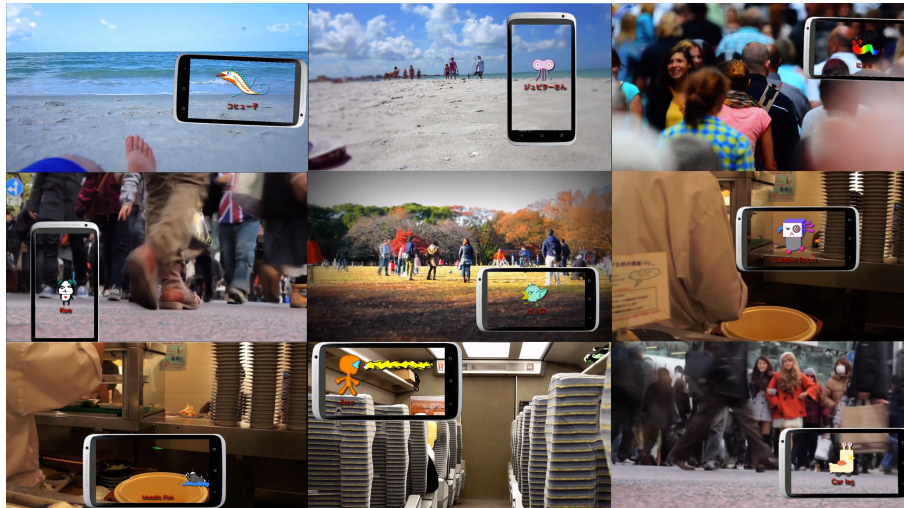


Figure 6.4: Clips from character creation video

This is the Ambee system, a system to adjust the contents to encourage users to explore different areas of the city looking for new sounds that trigger actions in the Application.

#### 6.1.4 Data analysis from Usability test

In the first evaluation, users seemed to understand the concept on the four corners after a few seconds of use. It is important to note that all users tried to touch the character first, instead of the menus in the corners. Within a few seconds of using the device they knew how to trigger the menus. However, when asked to close their eyes and try to navigate the screen, all users had issues knowing where the end of the touch-screen was (since the device is covered completely with a glass panel).

2 of the users had Android phones, so the application was installed in their devices and they were asked to try the process again. The familiarity with the size of their own phone instead of using a 7" tablet fixed the issue. Users expressed that the aural information provided through the application was helpful, but sometimes not clear if they were already touching a command.

In the second evaluation, users started using the application and enjoyed exploring the possibilities of creating new sounds with different commands. However, probably influenced by the familiarity with the iOS ecosystem, users were mostly using the "pinch to zoom" and "spin" commands. None of the testers tried scrolling down their fingers. 3 tried with only one finger at first, before being explained that 2 fingers were required to input commands.

Similar to the first evaluation, the most successful tests were carried out with the 2 users that had compatible headsets. After they tried the application on their devices, the users said that they felt more comfortable with the sensitivity of the screen of their phones.

About the sounds, users said that even though it was interesting to "interact" with the music, they wished it would go deeper into the song and instead of just adding elements on top of the song, they wished they could slowdown the beat or the tempo of it. The third evaluation was the most important to verify the concept behind Ambee and provide information on how to enhance the concept.

Users expressed that in the beginning of the video, they didn't understand the relation of the locations and the characters, but the more they saw, it became clearer to them that the characters were specific to each place.

The response of the testers to the sounds from the characters in the video were particularly positive and even more so that their visual designs. During the test with this 5 users, we obtained a good amount of information that was analyzed to validate the design of the application.

## 6.2 Quantitative Research

The concepts designed for the application as well as early demos were shown to 64 people that fall in line with our target audience. The goal was to get a better understanding of the market situation and decide which elements of an application are more important to our target user.

On the next section, we will look at the results, with a summary and conclusions at the end of this chapter. The survey was taken by people of the following demographics:

34 of the participants were men, 30 were women. In terms of age, 43 of the participants were aged between 25-30 years old, making for the vast majority of this test, while 10 participants were 20-25 and older than thirty. A single tester was below 20.

The testers were asked the following questions:

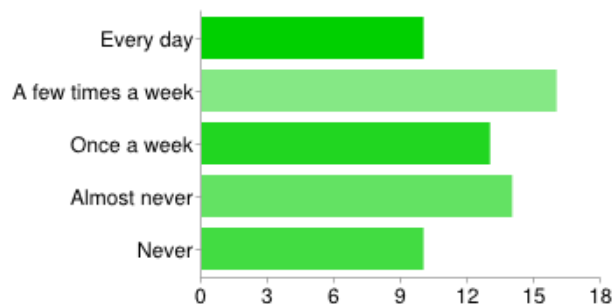
Do you own a Smart Phone?:

47 of the users currently own a smart-phone, 14 do not own one at the moment. However, out of those 14, 3 of them are planning on getting one soon.

Does your smart phone have a constant internet connection (GPS, 3G, LTE)?

43 users have a constant internet connection, while 17 do not have one. It is important to notice that of all the smart-phone users, only 4 don't have access to the internet at all times.

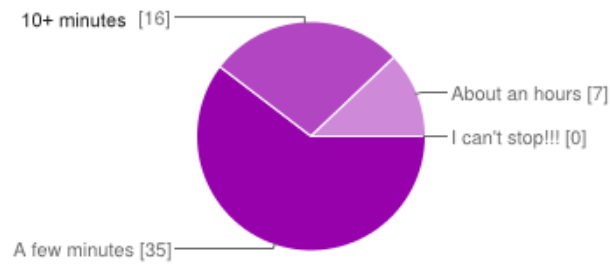
How often do you play games on a mobile device? (Phone, console, tablet)?



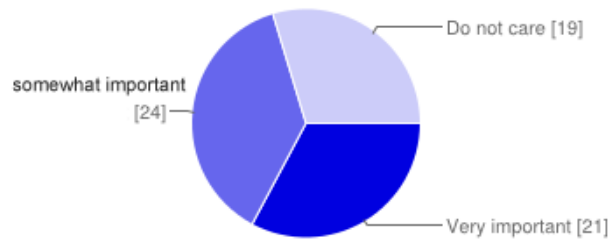
## 6.2. Quantitative Research

---

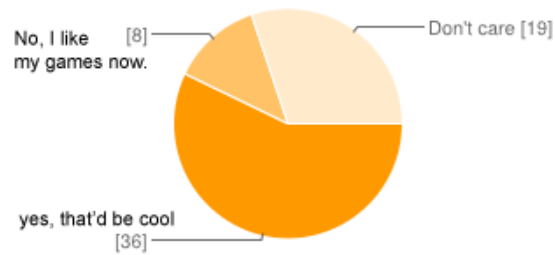
How long is your usual game session on a mobile device?



How important are the application sounds to you?



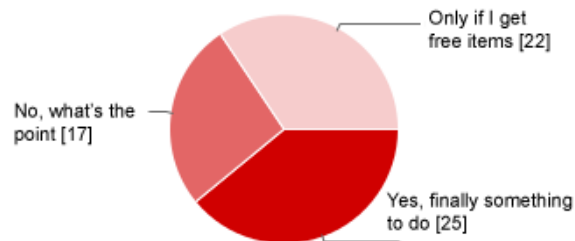
Would you like your location to affect the experience?



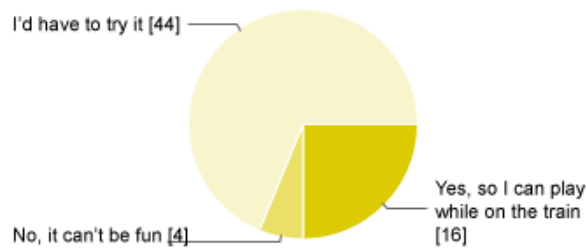
## 6.2. Quantitative Research

---

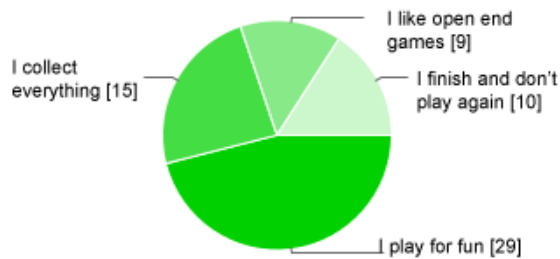
Would you play a sponsored minigame (McDonalds, Starbucks, etc.) if it was free?



If you could use the application without looking at the screen (based on sound reactions), would you do so?



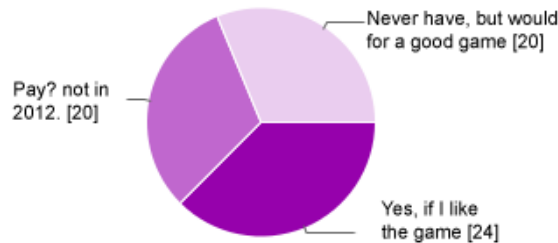
Do you play games for the entertainment of the moment, or do you think you NEED to accomplish something (level up, finish, etc)?



## 6.2. Quantitative Research

---

Have you ever payed for a mobile game? To unlock full version, new characters, new stages....



What is your favorite mobile game?

"I really enjoy puzzle games because you don't need a huge element of game control with a touch screen, you can jump in and out on a whim, and progress per level is usually saved."

"angry birds, easy to understand and fun to play. Funny sounds as well."

"Don't have a particularly fave one but I normally play mobile games which are either puzzles or quizzes. Reason: It's difficult to see the small screens on mobile devices so I tend to play games which do not consist heavily of graphics."

"angry birds . . It is Quick , fun, you can play for a short period while you are on a queue and the levels are easy to pass nothing will happen if you just have 2 min to play it"

"Where's my water? Attractive Graphics, good sound effects, nice challenge and short stages to progress"

"LUMINES....time just flies playing that thing!!!"

"Bike Race. It's short, fun – yet challenging – and great replay value. Like Excite-Bike, but you blow up!"

"Jet pack Joy ride: Simple gameplay, good unlocking curve. Great music and sound effects, just a dash of humor"

"Where's Wally?. It's an old game, but It's really good to kill time on the crowded train"

"Little Things, it is a game where you have a bunch of objects and have to find certain items, I get addicted to finding things."

"little things forever. iphone game. cute illustration, and sound"

"Triple Town....Alex got me addicted...it's all his fault. I like puzzle games and having cute growling bears is a plus!"

"mew mew tower because it's simple, cute and addictive"

"Blendr, It's fun to talk to people around you, but kind of mysterious because you don't know who/where they are. Sometimes the people around you can be surprising!"



## 6.3 Discussion with Tetsuya Mizuguchi

The following article is the result of an interview we conducted with game designer and music producer Tetsuya Mizuguchi who has extensive experience in the sound and entertainment fields. He worked for Sega for many years before starting his own company Q Entertainment. Mizuguchi is known for creating video games that incorporate a notable emphasis on interactive sound design.

When asked about the main consideration for strong interactive design mechanics, Mizuguchi San mentions that the application should feel rewarding. This can be achieved by musical and visual elements with the main goal of moving people emotionally. Every action has to trigger a reaction to convince the user that he is in control of what is happening. Similar to the human body, interaction design needs bones and muscle. So if your core idea or bones are strong enough, you can work on the muscle later in the design process.

The spirit of newer applications is to be rewarding immediately even for short periods of time. And while just 5 percent of the people pay for contents, we have to offer the other 95 percent something to do without any cost. The Power of free should not be ignored. Google is proof of that.

People use the app many times, and maybe through marketing or offering extra contents they might pay for it. Maybe. This is a new scheme that has not been proved yet, but the industry is moving in that direction. When asked about the importance of providing people with the tools to create their own contents, he mentions that people by nature enjoy feeling creative and expressing things. So the Democratization of these services gives the user the feeling of being a god and it is one of the new fundamentals of Social Application design. Technology is mature enough to offer these platforms through cloud technologies to let people create their contents and share them.

It is a good moment to be a game designer, music designer or both. Game mechanics are disappearing and evolving towards the concept of gamification where the user doesn't have to choose between different activities as long as the user is having fun. Maybe we can merge music with basic gaming or even reading a book. If the users have fun, the experience is justified.

The basic scheme is finding the fundamentals of how the concept of a game will evolve into something else.

The goal now is finding ways to link our lives and our daily activities with other forms of entertainment. He was asked how we can decide what is fun and

what is not and he explains that you should always follow your intuition and that of those around you. It relates to creating a connection with either previous experiences or what we are experiencing right now. Does it feel good? Recognizing how I or others feel is important. Asking them how they feel but also observing them. If I am the director or the Producer, I have to make decisions based on my own knowledge, and I need a very clear image of the tone.

It is a reality that applications or games sometimes don't live up to their expectations, something that in his own words is just "excuses". The best applications are the ones that are more creative and have the best business model. A game is not art, we can have an artistic approach but it is all about entertainment. And it is a business. So we have to benefit in some way. Sometimes companies know a service won't profit, but it will be a good marketing tool or something missing in the current portfolio of the publisher, so it can be a good success without selling.

Media design as a whole is fuzzy and very abstract and therefore we can focus on what we as humans want and dream and find ways to create it. Humans use media as a mirror or a dream.

After discussing sound interaction and gamification design elements with Mizuguchi san, he was introduced to the Ambee concept.

He mentions that many people have already explored the concept of location based gaming but many applications are expanding on the idea. We need to find ways to create serendipity and make it personal.

About the concept of using sound input to generate concepts and making them interactive by analyzing the sounds, he mentions that it is an excellent concept, and that partnering with a company who can provide the technology is a good idea, since the bone or core concept as it is, is very strong in his opinion.

Based on his experience, he says he is not convinced if it should be turned into a full-fledged/deep gaming experience and probably would work better if kept as a music generator with some basic gamification elements. The next step is starting a demo as soon as we get access to our collaborator's SDK to create a basic functionality demo since many companies have very good technology, but can't offer an experience like the one we are creating. Before closing the Interview he mentions the possibility of starting a collaboration with Keio Media Design to work on new applications for sound interaction design.

# Chapter 7

## Conclusions

Offering new contents through sound and location analysis is what Ambee proposes. When the user starts the application, it can create music from incoming sounds or detect the kind of environment where the user is to adjust contents dynamically.

The main goal we had when we started the project some months ago was to develop a new way of sound interaction and find a business partner to collaborate with the project. Both goals have been accomplished and the contract with our collaborator has now been approved after extensive negotiations. The areas of location and sound analysis to create interactive entertainment experiences remains the main goal of the team. Based on the feedback we obtained through usability testing, quantitative research and an interview with an expert in the area of sound interaction, the following evaluation has been made.

### 7.1 Concept Feedback

During our early testing we got encouraging results with most users showing interest in the concept of a sound interaction experience that dynamically changes based on the location of the user as well as incoming background sounds. This goes in line with our previous expectations and can be considered a success from a design point of view.

The technology challenges are being ironed out at the moment, so technical limitations should not hold back development of the application. Since the framework design of the application prompts users to scan and record their surrounding, our partner benefits with an exponential increase of their sound database once the application is commercially available.

During testing, users expressed genuine excitement over the possibility of a easy to access and quick companion for short periods of time that would let them interact with sounds with the possibility of mini-games being included in future versions of the application. This was a very interesting discovery when showing the demos to our testers. Many of them mention that they like an application that is easy to explain to others, since most of they apps they use were rec-

ommended by a friend, and it was possible to do so with a few words. If an application is too complex, they mentioned they think it might be too challenging and therefore won't try it.

About paying for the service, the users were divided, with some saying they have never paid for a mobile app, but that they would if the experience was engaging enough. However, it remains difficult for most phone applications to make a profit this way. Probably the optimal solution it is then to either rely solely on negotiations with our Business partner or to create partnerships with other game companies.

### 7.1.1 User Interface analysis

In terms of user interface, users had no problems with the four corner design. However, when asked to use the device without looking at the screen and relying solely on the aural information, some mentioned that it was complicated to use in a device other than their own since they were not familiar with the dimensions of the screen. This can be fixed by adding some basic haptic feedback when using the application.

Users had a tendency to touch the character at the center of the screen before reaching for the menus on the corners. Touching the character had no reaction to the user, something that should be addressed immediately for the next versions of the application.

The interactive music player as well was well received with the one limitation being difficult to use multi-touch with one hand only in bigger devices like iPads or tablets, something that is not an issue with smart-phones.

## 7.2 Current revenue model

The team was excited to create a platform for creating and sharing music. However, it was a feature not well received during our tests. Most users mentioned that they usually do not dedicate the time to create in-game contents and would rather keep things simple.

Similarly, as a form to obtain extra revenue we considered partnering with companies to create "branded" mini-games or items for the application. This way, we could partner with shops, restaurants, etc. to offer users exclusive contents. It is a novel concept that however was not well received by our testers. Most of

them mentioned that they would be excited to play branded mini-games, only if they received some sort of reward for doing so or otherwise they would feel used for marketing purposes.

At the same time, the area that is currently lacking the most is the integration of a deep social element or multi-player experience. As it stands today, the application is mostly targeted towards single users that can share their experience with other people, but that is not encouraged as part of the current design.

## 7.3 Future Developments

Now that the team is obtaining the Software development kit from our partner to start full-on development of the application, there have been internal discussions on how to utilize the feedback obtained during the writing of this thesis and how it will affect the direction of the project.

## 7.4 Revised Business Model

After users expressed that they are not likely to play for branded mini-games unless they get some sort of real award in the way of free items from the sponsors, we began analyzing how to obtain revenue if future developments required it. One option is linking the virtual and the real world. Collaborating with real world retailers to offer digital versions of some of their products that can be obtained with every visit and then exchanged for items.

This would be an interesting alternative to the popular point card system used in most Japanese shops. With this concept, scanning the same place for a certain number of times would give the user access to a special coupon that can be exchanged for a real world item.

Also important to mention is that through the use of a sound analysis software, it is possible to use the sounds recorded by the users and increase the existing sound database of our sponsor, an element that is important to guarantee the quality and accuracy of the sound matching. That way, even if the profits are limited we could benefit with the potential future developments that a rich sound database can provide, be it for internal use, or licensing the technology to third parties.

Outside Gaming, the development of the sound based interface that relies on the "four-corner" sound design can be used for basic applications for users with

sight problems or aid elderly populations when they interact with electronic devices.

Social applications such as Skype would benefit from offering a basic version of the service that has a more accessible and simple design created for special population, something of special importance in countries like Japan, where the population is aging rapidly.

## **7.5 The Story behind Ambee civilization**

As part of the team effort to further enhance the connection between the users and the characters, a new story element is being written. In our tests, some users expressed interest in finding out more about the characters. Why are they here? Why do they need sounds?. This is an early draft of the story that is meant to be used in future versions of the application.

### **7.5.1 Planet Sonas**

Ambees live on the planet of Sonas. Sound is their only energy source. Sonas is a planet that is far away from the sun and the moon, thus there is no changing from daylight to darkness. Ryk Star is the closest star to the planet of Sonas. It is a star which emits brightness that is similar to our candle-light. It everlastingly illuminates Planet Sonas. Thus, Ambees live in cozy lighting all-year-round. The planet does not evolve like the Earth does. One side of the planet is always dark. There is no gravity on Planet Sonas. Because of that, sound does not take place at random. Ambees live on a scarce energy source.

The planet of Sonas bounces itself up and down five times a year. And that is when Ambees get together at the Sonas Plaza to celebrate the day. With the planet's every bounce, they jump as high up as they can. They group into teams and run into each other's team to see who can create the biggest sound.

Princess Suara and her father lived on the bright side of the planet. Hundreds years ago, Suara's father, a crazy scientist, managed to invent a device which could transfer sound into energy that could be used for various purposes. It changed the lives of Ambees, so he became the king of Planet Sonas. Lord Zvuk lived on the dark side of the planet and he wants to make a machine to take away all the sound on Planet Sonas in order to make himself the most powerful Ambee. He kidnapped a group of Ambees from the bright side of the planet and put them in a concentration work camp in order to supply the energy need on the dark side. Suara's father wants to save the kidnapped Ambees and light up the entire planet. Therefore, Lord Zvuk and Suara's father are sworn enemies.

Because sound resource is scarce on the planet they resided in. Suara was then sent off to a journey to outer space on a quest to locate abundant sound resource. She tries to store sound from Earth in attempt to use it to defeat Lord Zvuk and bright up the entire planet. Suara landed on Earth and found it to be the perfect place for sound collection. Many Ambees came to Earth after to participate in the same mission of sound-collecting. When Suara did not return, Lord Zvuk followed Suara to Earth. He wants to take Suara home and take over the sound that Suara collects.

Now Ambees are creatures that live among us humans. They do not have a voice like us human beings. They feed on sound and they communicate with each other in sound that they capture. They are invisible through human eyes. However, by the sound feeder(which comes with smartphone application designed by the ARound.hear group), we can establish relationship with them. Your smartphone is a device which allows you to see them and hear them if they choose to reveal themselves to you. Ambees enjoy interacting with each other. Their favorite pass-time is challenging each other in a Sound Match. Winner of the game takes away a piece of sound from the loser.

### **7.5.2 City Domination**

As part of our goal to offer a deeper integration as a social game, the concept of turning the game in the future into a tower defense kind of game has been discussed. If this happened, users would have to complete the single player portions of the game to strengthen their Ambees. Based in the concept of having two races fighting for the same territory, users would have to chose sides from the beginning and play the game with the goal of collaborating with others to take different areas of the city in real time.

Users have to avoid users from other species from entering an area has been claimed as theirs. To do so, users will have to collect different Ambees and leave them in different parts in the city. The gamer can visit them every day as they move from one location to the other and feed them specific sounds as they are within their range. This is an exciting development that has not been done in the past but that proved too ambitious to be developed by our team in KMD. If the first open test of the application is successful, it is planned that in the future, this version of the game will be pitched to our sponsor to take the application one step further in terms of social interaction.

## 7.6 Development continues

With the signing of the contracts with our sponsor, my contribution to this project comes to an end. Having started the project from start as well as forming the current team of collaborators, it is with pride and excitement that I wait to see what further developments come out of the Ambee application.



# Bibliography

- [1] Google Maps Blog. A new frontier for google maps: Mapping the indoors. <http://googleblog.blogspot.jp/2011/11/new-frontier-for-google-maps-mapping.html>.
- [2] Edited by Kenton O'Hara and Barry Brown. Consuming music together: Social and collaborative aspects of music consumption technologies. pages 37–45. Springer, 2006.
- [3] Walkman Central. Tps-12. <http://www.walkmancentral.com/products/tps-12>.
- [4] Richard E. Cytowic. ynesthesia: A union of the senses. page 122. MIT Press, 2010.
- [5] Rob Drew. Karaoke nights: An ethnographic rhapsody. page 450. Ethnographic Alternatives, 1997.
- [6] Apple electronics. ipod music player. <http://www.apple.com/ipod/>.
- [7] IGN entertainment. Al alcorn interview. <http://retro.ign.com/articles/858/858351p2.html>.
- [8] IGN entertainment. Classic game museum. <http://classicgaming.gamespy.com/View.php?view=ConsoleMuseum.Detail&id=1&game=12>.
- [9] IGN entertainment. What's wrong with music on the nintendo 64. <http://ign64.ign.com/articles/061/061845p1.html>.
- [10] Konami Digital Entertainment. Dance dance revolution. <http://www.konami.jp/bemani/ddr/jp/>.
- [11] Muscle Fish. Muscle fish. <http://musclefish.com/>.
- [12] E. Herre J, Allamanche and Helmuth. Robust matching of audio signal using spectral flatness features. *In proceedings of the 2001 IEEE Workshop on Applications of Signal processing to Audio and Acoustics*, pages 127–130, 2001.
- [13] Q Entertainment Inc. Lumines. <http://lumines.jp/lumines/>.

## Bibliography

---

- [14] Time Inc. A history of video game consoles. <http://www.time.com/time/interactive/0,31813,2029221,00.html>.
- [15] L. Brown L. Kenyon, S. Simkins and Sebastian. Broadcast signal recognition patent and method. [www.uspto.gov](http://www.uspto.gov).
- [16] Koenig. Spectrographic voice identification: A forensic survey. pages 79:2088–2090. Journey of the acoustical society of America, 1986.
- [17] Jeremy Hsu Live Science. Music-memory connection found in brain. <http://www.livescience.com/5327-music-memory-connection-brain.html>.
- [18] Shazam Entertainment Ltd. Shazam. <http://www.shazam.com/>.
- [19] Forbes Magazine. Google indoor mapping at malls, airports creates new market. <http://www.forbes.com/sites/greatspeculations/2011/12/05/google-indoor-mapping-at-malls-airports-creates-new-market/>.
- [20] Time Magazine. Ken kutaragi. <http://www.time.com/time/magazine/article/0,9171,994014,00.html>.
- [21] Musanim. Synesthesia analysis. <http://www.musanim.com/synesthesia/>.
- [22] Navinko. Navinko. <http://navinko.com/>.
- [23] Eurogamer Net. Milo and kate. <http://www.eurogamer.net/articles/bbc-milo-and-kate-is-a-tech-demo-after-all>.
- [24] CBS news. Great entrepreneurs in business history. [http://www.cbsnews.com/8301-505125\\_162-51196888/great-intrapreneurs-in-business-history/](http://www.cbsnews.com/8301-505125_162-51196888/great-intrapreneurs-in-business-history/).
- [25] Cnet News. Google: 200 million android devices now active worldwide. [http://news.cnet.com/8301-1023\\_3-57326649-93/google-200-million-android-devices-now-active-worldwide/](http://news.cnet.com/8301-1023_3-57326649-93/google-200-million-android-devices-now-active-worldwide/).
- [26] Donald Norman. Emotional design: Why we love (or hate) everyday things. page 84. Basic Books, 2004.
- [27] Lee Pei-Ling. A study on the miniaturization of a product. [http://www.idemployee.id.tue.nl/g.w.m.rauterberg/conferences/CD\\_doNotOpen/ADC/final\\_paper/232.pdf](http://www.idemployee.id.tue.nl/g.w.m.rauterberg/conferences/CD_doNotOpen/ADC/final_paper/232.pdf).
- [28] Nature publishing group. Acoustic and neural reconstructed spectrograms for speech from a single speaker or a mixture of speakers. [http://www.nature.com/nature/journal/v485/n7397/fig\\_tab/nature11020\\_F1.html](http://www.nature.com/nature/journal/v485/n7397/fig_tab/nature11020_F1.html).

## Bibliography

---

- [29] Interactive radio system. Nielsen broadcast data systems. <http://www.interactive-radio-system.com/docs/Nielsen%20BDS.pdf>.
- [30] Sony Financial Report. Financial quarter 4, 2011. <http://www.sony.net/SonyInfo/IR/financial/fr/viewer/11q4/>.
- [31] Sega Retro. Seaman. <http://segaretro.org/Seaman>.
- [32] Jesse Schell. The art of game design. page 292. Morgan Kaufmann Publishers, 2010.
- [33] Designing Sound. a nightmare on elm street exclusive interview with andrew decristofaro, michael payne, and david farmer. <http://designingsound.org/>.
- [34] De Nora T.. How is extra-musical meaning possible: music as a place and space for work. pages 84–94. Sociological theory, 1986.
- [35] Audacity team. Reducing noise. [http://wiki.audacityteam.org/wiki/Reducing\\_noise](http://wiki.audacityteam.org/wiki/Reducing_noise).
- [36] Hiroko Terasawa. Stanford University. Determining the euclidean distance between two steady state sounds. [https://ccrma.stanford.edu/~hiroko/timbre/Terasawa2006\\_ICMPC9.pdf](https://ccrma.stanford.edu/~hiroko/timbre/Terasawa2006_ICMPC9.pdf).
- [37] North Carolina State University. Using gps and geocaching engages, empowers and enlightens middle school teachers and students. <http://www.ncsu.edu/meridian/win2007/gps/04.htm>.
- [38] Juan Pablo Ruiz Velasco. Juan pablo ruiz velasco facebook. <http://www.facebook.com/jpanese/>.
- [39] Wikipedia. itunes. <http://en.wikipedia.org/wiki/iTunes>.
- [40] Wikipedia. Playstation. <http://en.wikipedia.org/wiki/PlayStation>.
- [41] Wikipedia. Playstation. [http://en.wikipedia.org/wiki/Masaya\\_Matsuura](http://en.wikipedia.org/wiki/Masaya_Matsuura).
- [42] Wikipedia. Pong. <http://en.wikipedia.org/wiki/Pong>.
- [43] Wikipedia. Super audio cd. [http://en.wikipedia.org/wiki/Super\\_Audio\\_CD](http://en.wikipedia.org/wiki/Super_Audio_CD).
- [44] Wikipedia. Tamagotchi. <http://en.wikipedia.org/wiki/Tamagotchi>.
- [45] Wikipedia. Tetsuya miziguchi. [http://en.wikipedia.org/wiki/Tetsuya\\_Mizuguchi](http://en.wikipedia.org/wiki/Tetsuya_Mizuguchi).