Designing interfaces for the visually impaired

Contextual information and analysis of user needs

Stina Olofsson

Fall, 2017
Master's Thesis in Interaction Technology and Design, 30 credits
Supervisor at UmU: Shafiq Urréhman
Supervisor at North Kingdom Stockholm: Sonja Lakner

Examiner: Thomas Mejtoft

Umeå University
Department of Applied Physics and Electronics
SE-901 87 UMEÅ
SWEDEN

Abstract

This thesis explores how to design for the visually impaired. During the course of work, a literature study and interviews with blind and visually impaired people were conducted. The objective was to investigate what contextual information is wanted in new and unfamiliar spaces outside their home. The interviews also explored how they experience digital tools they are using today and what they think of the possibilities of voice and other user interfaces. The main finding from the study is that there are indications that multimodal interfaces are preferred. The interface should combine voice, haptic and graphics since the participants wanted to interact in different ways depending on functionality and context. Three main problem areas were identified, navigation, public transportation and shopping. Another result was that when developing for the visually impaired it should always be tested on people with a wide variation of vision loss to find the correct contextual information.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The World Health Organisation, WHO, estimated in 2014 that 285 million people are having a visual impairment worldwide [2]. One of the biggest obstacle in everyday life as a person with visual impairment is the decreased mobility and difficulties in discovering and getting around new and unfamiliar environments. This is something that hopefully will start to change. Along with the technological development of smartphones and their interfaces, it is opening up for a new and exciting challenge when it comes to user experience design and increased accessibility. The definition of accessibility is to which degree a product, device, service or environment (virtual or real) is available to as many people as possible [3].

One interface involved in this development, is voice user interfaces (from here on referred to as VUI). The interface allows a user to interact with a system through voice or speech commands and has gone from being a fictional encountering in countless science fiction creations to, now in 2017, becoming a part of everyday life in Sweden. Another interface, that is easier to overlook, is haptic interfaces. Haptic interfaces regards the sense of touch and it is common to underestimate the importance of it without the loss of another sense. For example, touch is the sense that facilitates or makes motor activities possible, permits the perception of nearby objects and spatial layout when viewing is not feasible. It can also inform us of object properties like temperature or material and draw our attention to events, signaled by vibrations, that is inaccessible to other senses [4].

The existing assisting technologies for the visually impaired generally focus on either voice or haptic, or both for alternative approaches for input and output. Some examples of this are devices like alternative keyboards and pointing devices, speech recognition, Braille displays, and screen readers [5] but since Braille literacy rates are low, voice interfaces have become the most common technology for users with visual impairment [6].

1.1 Problem Statement

Harper and Green says that most of the information provided to aid the independent traveller is in the form of visual cues [7]. These cues typically include graphical signs, coloured lights, road markings, printed public transport information etc. If you have a visual impairment the cues provided becomes next to zero. In the past years, attempts have been made of improving the technology and tools for the visually impaired. However, most of these systems take on a technology centric approach i.e., ignoring the user's actual requirement analysis. The effectiveness of both voice and haptics will ultimately be dependent on the extent to which contextual interactions are understood and accounted for in their design [8]. This includes being able to give the right information, at the right place, in the right way. The problem is therefore that there is a lack of research in the field of the users' need and which contextual information is necessary. That is why this thesis is focusing on filling the gap in what contextual information is missing from a user perspective and in which contexts there can be an improvement.

1.2 Objective

The objective of this thesis is to investigate how to design for people who have a visual impairment. This will be accomplished by a literature study and a qualitative study where the focus is to gather data and insights from people who are visually impaired and blind. This will then be used to look into how the technology available in a smartphone can be of use in the problem scenarios found, and in which context. The objective is divided into two research questions, focusing on the contextual information and user experience;

- What kind of contextual information is wanted from visually impaired in new, unfamiliar spaces outside their home?
- Based on the previous question, how, if possible, can an interface in a smartphone be used to provide such information?

1.3 Limitations

The study will not go into sensory substitution. The focus will instead be in identifying some of the most important contextual information necessary and prototype a concept from that scenario. It will not be a navigation aid but providing additional information necessary in the everyday life. The device for the VUI is constricted to use the current features supported by a smartphone device in the year of 2017. Another limitation is that the interface should be as intuitive as possible. That means that interfaces which may have a better longtime solution but require extensive training will be overlooked.

Chapter 2

Background

2.1 North Kingdom

North Kingdom is an award winning experience design company based in Stockholm, Skellefteå and Los Angeles with approximately 50 employees (2017). They have worked with companies such as Google, Disney, Volvo, Netflix, and Toyota, creating everything from mobile applications to marketing campaigns. North Kingdom describes themselves as:

> "We believe that new value can be created wherever people, business, and technology collide. We help our clients harness that value through the creation of experiences, products, and services that play a meaningful role in people's lives. Through human-centered design, we make the complex simple and relatable, no matter what medium or platform" [9].

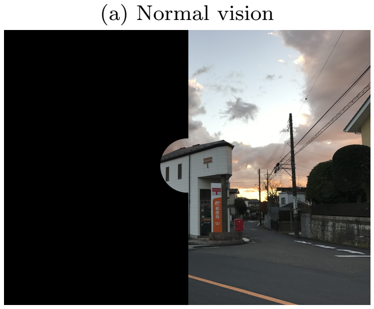
The author has conducted this master thesis at North Kingdom's office in Stockholm during the fall of 2017.

2.2 Definition visual impairment

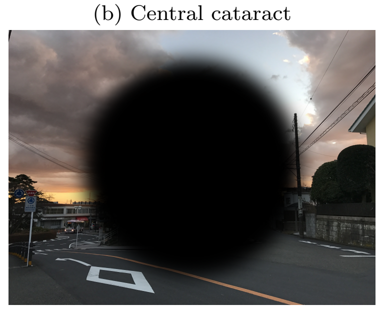
The estimation of 285 million people suffering from visually impairment worldwide are divided into two groups of 39 million blind and 246 million with low vision [2]. That means that someone in our world goes blind in every five seconds. Demographics of the blindness vary significantly in different parts of the world. In developing countries, about 0.4 % of the population is blind while in the rest part of the world this rate rises up to 1 % [2]. Consequently, 90 % of the visually impaired population are living in developing countries. About 82 % of the blind are over the age of 50 [2].

(a) Normal vision

(b) Central cataract



(c) Half vision



(d) Central loss of vision

Figure 2.1: Four examples of different kind of vision and visual impairment

It is important to distinguish between different kind of visually impairment, as each type are causing different problems which requires different kind of solutions and contextual information necessary [5]. In figure 2.1[1] it is shown some different kind of visual impairment and how it can effect the visual information available compared to normal vision. There are also several different degrees of visual impairment and according to WHO, vision function is classified in four levels as follows [2]:

- Normal (full) vision - no visual impairment
- Moderate vision impairment
- Severe vision impairment
- Blindness

Moderate vision impairment combined with severe vision impairment are grouped under the term “low vision”: low vision taken together with blindness represents all vision impairment. About 15 % of people who are having vision loss cannot see anything at all. The remaining 85-90 % may have residual vision or other types of low vision and may have difficulties with colour, light, form or movement perception [2].

Visual acuity is a number used by eye care professionals that indicates the

[1] Image source found in [10]

sharpness or clarity of vision. A visual acuity measurement of 20/70 means that a person with 20/70 vision who is standing 20 feet (approximately 6,1 meter) from an eye chart sees what a person with unimpaired (or 20/20) vision can see from 70 feet (approximately 21.3 meter) away [11]. Another way to refer to the definition of low vision is to describe low vision in terms of function, rather than numerical test results. So low vision is considered an uncorrectable vision loss that interferes with daily activities. Or in other words, low vision can be described as ”not enough vision to do whatever it is you need to do”, which can vary from person to person. Some eye care professionals can also be using the term ”low vision” to describe permanently reduced vision that cannot be corrected with regular glasses, contact lenses, medicine, or surgery [11].

2.3 Mobility

Mobility could be defined as the possibility of freely moving, without support of any accompanying person, at home, in public and private buildings, and in open spaces, as the streets of the town [12]. Travel and visit new places for blind people poses many more problems than it does for sighted people. Blind or visually impaired pedestrians can easily miss obstacles on the pavement in front of them or branches overhanging their path [7]. Neither can they look into the distance and see a landmark which they need to follow. Instead, visually impaired people tend to continuously process what is happening in their immediate personal space. This is to negotiate objects in surface level, whilst keeping orientation to find the final destination [13]. Each blind or vision impaired person are using their own strategies that works best for them. The information they are using to form their decisions may be gathered from residual vision. For example, very few vision impaired people in the world have no light perception. They can also use tactile information (e.g. underfoot surface texture), following building lines, auditory information (e.g. hearing landmarks or other cues) and other sensory information (e.g. smell from coffee shop or bakery).

The most commonly used primary mobility aids in many countries are a (long) white cane and a guide dog (”seeing eye dog”). These tools give assistance for the first group of problems, which constitute a part of the task of navigating through the immediate environment or micro-navigation [14]. Vision impaired might be using none, one or both of them depending on the environments they are travelling through. Generally, guide dog users are using a long cane, as good orientation skills are required to ensure that the guide dog is guiding in the correct manner and along the routes that the user are determining. In the countries which are using guide dogs, the dogs primary role is to walk in a straight line and avoid obstacles in the path of the user [13].

There are three different types of canes that may be used. They are called a long cane, an identification/symbol cane and a support cane. A long cane is used to help the user detect obstacles and hazards through providing information from the environment which assists the navigation. For example, the user can use the cane to detect changes in the surface textures between grass and concrete to be able to follow a footpath. The identification/symbol cane is a tool that is used to signal to the general public that a user is visual impaired. The cane is not designed or used to make contact with ground, but may instead be used to check the height of a step or drop-off/kerb. The support cane allows a person who requires balance and stability support to be identifiable as having a visual impairment [13].

However, these traditional primary aids give no assistance for the second group of problems, those related to navigating through the more distant environment, also called macro-navigation [14]. What is important to remember is that any navigation service should not replace the primary mobility aids, but rather it should be treated as an orientation tool used along with other skills to augment the user experience and reassure its users [13].

Landmarks and clues

Using landmarks and clues are two ways to use the environment to navigate. They play an important role to reassure a person that he is walking in the right direction as well as helping the person to locate himself in a space [15]. Primary landmarks are defined as objects which are difficult to miss at a particular location or environment. They can be things as kerbs, a change in the walking surface etc. Clues are more transient and may include sounds, smells, change of temperature etc. [13]. Vision impaired people tend to make use of different landmarks and clues than sighted people and sometimes more detailed ones [16]. The difference is that the visual impaired are using clues closer to the individual, in other words in areas that can be touched, for example through the white cane or felt through the soles of the feet.

2.4 Tools for the visually impaired

There are also other aiding tools aiming to support people with reading and in the use of digital products. Two of the most common tools that will be referenced in this study are Braille and screen readers. Braille is a tactile system of raised dots that can be read with the fingers by people who are blind or have low vision. It was invented in France by Louis Braille in the latter part of the nineteenth century. It is not a language in itself. Rather, it is a code by which many languages, such as English, Spanish, Arabic, Chinese, and dozens of others may be written and read. This means that it can be used by thousands of people all over the world in their own native languages, and provides a means of literacy for all [17]. With all of its benefits, the drawback of Braille is that it is hard to learn which means that the literacy of Braille is low and in Sweden there is about 1500 people using Braille to read and write on a daily basis [18]. That is to compare with the, in total, 120 000 persons who have a visual impairment in Sweden [19].

Screen readers are another tool that is common among blind and people with low vision. It is a software application that enables people with severe visual impairments to use a computer. There are two ways that a screen reader can provide feedback to the user, speech and Braille [20]. With speech the screen reader uses a Text-To-Speech (TTS) engine to translate on-screen information into speech, which can be heard through earphones or speakers. A TTS may be a software application that comes bundled with the screen reader, or it may be a hardware device that plugs into the computer. In addition to speech feedback, screen readers are also capable of providing information in Braille. Then an external hardware device, known as a refreshable Braille display is needed [20]. Screen readers also exists for smartphones and provides spoken descriptions of items on the screen and can read text items. The screenreader feature is called VoiceOver on Apple devices and TalkBack on most others (i.e., Android phones). The screen reader feature announces icons as the user touches or selects them on the screen. A user can also select a text-to-speech feature to read aloud long text items such as email or articles.

Blindsquare

One navigation application used as a reference in this study was the application Blindsquare. Blindsquare is an GPS application, specialized into giving better descriptions and directions of the environment to blind people. Blindsquare is self-voicing and are announcing points of interest, intersections and user-defined points through a dedicated speech synthesizer. On their website Blindsquare writes that the most important BlindSquare features can be accessed through an audio menu via any headset or speaker that supports Apple's music controller. This means you do not have the need to touch the screen of your phone or looking for tiny buttons when you travel with BlindSquare [21]. It is also possible to shake the device to hear the current address, as well as information about the location of the nearest street intersection and venues around.

Chapter 3

Theory

The theory section presents an introduction of things to consider when designing user interfaces for people with a visual impairment. It introduces the different possibilities with voice control and haptic interfaces as well as an introduction of the concept of multimodal and context-aware interfaces.

3.1 Designing for visually impaired

This section will give a short overview about things to consider when designing for visually impaired.

3.1.1 Improved Plasticity

Visual impaired and blind individuals are particularly dependent on their hearing and tactile senses. According to King [22] there is extensive evidence that they can develop superior auditory skills. Either as a result of plasticity within the auditory system or through the recruitment of other parts of the brain since those parts of the brain do not have to process visual information. But this is something that is debated. Some claim that this is a myth and that there is no evidence, but rather that the blindness makes vision impaired people pay more attention to their available senses in order to cognitively process the information from the environment [13]. Whether it is true or not, people who are visually impaired are often using more clues in their environment than sighted people. It can include non-visual clues such as wind direction, smell of the bakery, heat of the sun but these things are also more inconsistent and therefore also considered less reliable. In order to increase the reliability of the clues used, vision impaired often combine the use of different senses. For example they may use a tactile landmark followed by an auditory clue in order to confirm that they are approaching a landmark [16].

3.1.2 Workload for visually impaired

There are different ways or possibilities of compensating for loss of a sensory function. One way is to augment deficient senses by conveying missing information through an intact sense [23]. But there are many pitfalls and mistakes that can be done. Kristjánsson et al. [23] state that one common pitfall is how conveying extra information about the environment risks sensory overload. Related to this, the limits of attentional capacity make it important to focus on key information and avoid redundancies. Strothotte et al. [24] demonstrate how headphones used to transmit audio may mask/distort sounds in the environment which are necessary for avoiding hazards. This is because persons who have blindness or visual impairment are using and are depending on their hearing much more than sighted people. Devices should be task-focused since in many cases it may be impractical to convey too many aspects of the environment. Evidence for multisensory integration in the representation of the environment suggests that researchers should not limit themselves to a single modality in their design [23].

3.1.3 User preferences

Journal of visual impaired and blindness (JVIS) conducted a survey in 2004, which showed that the majority of blind and visually impaired people preferred a speech input and output interface [25]. The survey was performed to investigate preferences for a possible personal navigation device. The concerns expressed regarded cosmetic acceptability of a device and the use of a single or stereo headphone interface. In busy and noisy environments, hearing instructions from the smartphone speaker, particularly when in a pocket are a challenge. Apart from not being able to hear the instructions there is also a danger in blocking out auditory clues and warnings in the environment [13].

A new technology, partly solving this issue are bone conducting headphones. These headphones are constructed in a way so they do not obstruct the ear and are particularly suitable when the user wants to keep in contact with its surrounding environment [26]. Additionally, vision impaired people have reported that the ideally prefer using wireless headphones to avoid wires getting tangled [25].

3.2 Voice User Interfaces

VUI is what a person interacts with when communicating with a spoken language application. The elements of a VUI include prompts, grammars, and dialog logic (also referred to as call flow). The prompts, or system messages, are all the recordings or synthesized speech played to the user during the dialog. Grammars define the possible things callers can say in response to each prompt. The system can understand only those words, sentences, or phrases that are included in the grammar. The dialog logic defines the actions taken by the system. One example of this is responding to what the caller has just said or reading out information retrieved from a database [27]. One of the biggest and most impactful contributions of the latest developtment in VUI, is the vastly improved accessibility for a lot of people, including people who suffer from visual impairment or blindness [1]. According to Deepak [28], voice interaction is the best way to interact in a human-machine relation for people who have visual impairment.

3.2.1 Conversational User Interfaces

The latest development in VUI are Converstational User Interfaces, which differentiate itself from VUI by simulating natural language, rather than the command and answer conversation VUI have been in the past.

To understand conversational interfaces, it is important to understand what a conversation is, since it is a lot more than just a simple exchange of information. In a conversation, the participants normally share natural assumptions about the topic and know how a conversation should develop. There are also expectations about the quality and quantity of the contributions that each person should make. On top of that, natural conversations have predefined rules about politeness, consistency, and other natural rules of how to behave. Plus, there are a instinctively behaviour to look for and know how to disregard superficial meanings if they are unclear or as the expression says ”read between the lines”. And while we all do it naturally, conversation is actually a rather complicated process [29]. Grice said that to be understood, people need to speak cooperatively. He called this the Cooperative Principle. With this principle he introduces Grice’s maxims which consists of four basic rules to have a cooperative conversation [30]:

- Quality — Only say things that are true
- Quantity — Do not use more or less informative than needed
- Relevance — Only say things relevant to the topic
- Manner — Be brief, get to the point, and avoid ambiguity and obscurity

In conversations where these rules are not followed it is common to end up experiencing confusion and frustration. VUIs who are not following these maxims will experience similar issues. Pearl is giving some examples (see Table 3.1) of how this can manifest in badly designed VUIs [1].

Table 3.1: Examples of breaking Grices maxims [1].

Quality	Advertising things you can not live up to, such as saying "How can I help you?" when really all the VUI can do is take hotel reservations.
Quantity	Extra verbiage, such as "Please listen carefully, as our options may have changed" (Who ever thought, "Oh, good! Thanks for letting me know"?)
Relevance	Giving instructions for things that are not currently useful, such as explaining a return policy before someone has even placed an order
Manner	Using technical jargon that confuses the user

3.2.2 Accessibility in Voice User Interfaces

According to Nielsen, VUI have the best potential in the following cases [31]:

- Users with various disabilities, who cannot use a mouse and/or a keyboard or who cannot see pictures on the screen.
- Users who are in an eyes-busy, hands-busy situation. Whether or not they have disabilities, the keyboard-mouse-monitor combo fails users in these situations, such as when they are driving cars or repairing complex equipment.

Pearl [1] calls VUI a ideal, non-visual experience for the blind and visually impaired. She says that the constraints of designing for people with different disabilities can help to inform how we can solve challenges in not just VUI but in other areas as well. In her book, Designing for Voice User Interfaces [1] she has interviewed an expert: Chris Maury[1], for a few best practices for VUI design informed by accessibility. They suggest that VUI should prioritize personalizing over personality. For example, to let the user choose what text-to-speech voice they want to listen to in their app. Differences in voice can be things like their own specific personality or the user just might enjoy using some more than others. Many voices are built with high speech rates in mind, which means that they can sound more robotic but are also much more intelligible at higher words per minute [1].

It is also important to keep the information short. In an audio interface there can be no skipping around or skimming like in visual interfaces, where the user quickly can change their attention between different sections. Therefore, it is important to provide only the most vital information. The VUI should allow

[1]For more info see http://chrismaury.com/

the user to interrupt so that they can go back or end a reading of a list when they hear the one they were looking for.

To accommodate improved plasticity in VUI, Pearl says that interactions should be time efficient. This means that the VUI should have the ability to change the speed for the more advanced users who often are much better than normal users of listening. With time efficient, she also means that when designing visual experiences we try to limit the amount of clicks. The same should go for VUI, trying to keep the interactions as few as possible.

Another way of handling this in VUI is to make sure that VUI should provide context [1]. Providing context means guiding the users to what they can do. In graphical interfaces this is not so much of a problem but for voice interfaces the discovery of features is nonexistent. The design of the VUI should therefore be able to help informing the user how they should respond or what they can do, but in some cases this is not enough. In these cases, Pearl recommends that the user should be able to fall back to an explicit orientation action. This means that the user should be able to ask for help at any time and that the interface can guide them and help them reorient to their current context. Some common expression the user might use are for example:

- ”Help”
- ”What can I say”
- ”Where am I”
- ”Uhm, I am confused...”
- ... (silence)

3.2.3 Audio Signifiers

Auditory information plays an important role in the everyday life of a vision impaired person, whether this is environmental (e.g. a car passing by) or predetermined (e.g. any sound design component such as a keyboard sound or alert)[13]. Audio signifiers are also an example of the last kind. There are three types of audio-based signifiers or cues which can prompt user actions and inform users about possible commands [32]:

- Nonverbal sounds, or earcons (auditory icons), which are distinctive noises generated by the system, usually associated with specific actions or states. For example, Siri emits a 2-tone beep after detecting its activation phrase, to signal that it is now ‘listening’ for a command.
- Explicit verbal signifiers, when the system verbalizes a suggestion or request to let the user know what commands are available. For example, if you tell Google Home to “Set a timer,” it responds with “Ok, for how long?”

- Implicit verbal cues, when the system hints that an action is possible, without fully articulating the suggestion. For example, when Amazon's Echo detects its wake word while it is speaking, it pauses its own speech to let the user know that it is 'listening' for a new command. This behavior mimics human speech patterns, where people pause briefly to cue conversational partners that they are willing to stop speaking and listen.

3.3 Haptic Interfaces

Haptics, origin from Greek means "I touch", is referred to as any form of nonverbal communication involving touch. Perhaps the best asset of haptics is discretion - only the user feel the effect of the touch. The skin is ideal for tactile stimulation as the skin area is extensive, at 1.8m^2[33].

Without examples of temporary occlusion or permanent loss of a sense we tend to underestimate the role of touch in our perception of the world [4]. Most commonly haptic feedback is combined with other senses but an area where it can shine is when it comes to feedback. Haptic feedback, often referred to as haptics, is the use of the sense of touch in a user interface design to provide information to an end user. Tactual sensory information from the hand in contact with an object can be divided into two classes: tacticle information and kinesthetic information [34]. The difference between the two is complex, but at a high level: Tactual or tactile perception refers to every type of sensation related to the sense of touch, so for example things you feel with the skin of your fingers (pressure, vibration, temperature and texture). The Braille system is one example of a very successful tactile system used by blind and visually impaired all over the world. Kinaesthetic is the things you feel from sensors in your muscles or your position of the body [4]. Imagine holding a coffee-mug in your hand. Kinesthetic feedback tells your brain the approximate size of the mug, its weight, and how you are holding it relative to your body.

Haptic perception is fast and let the user response in a natural manner. While the sighted rarely recognize the importance of haptics, the blind have to make full use of it. The blind community is believed to gain the most benefits from haptic research. The limitations of haptics is its lower resolution capability than audio. For this reason, haptics alone is only suitable for limited information feedback. If more details needed, engineers should combine haptics with another sense such as audio to extend resolution [35].

A different study using tactile methods for providing information to visually impaired and blind was performed by McDaniel et al. [36]. They used a haptic (vibrotactile) belt to assist individuals who are blind or visually impaired by communicating non-verbal cues during social interactions. Results from their two experiments show that the haptic belt is effective in using vibration location and duration to communicate the relative direction and distance, respectively, of an individual in the user's visual field by using vibrations.

3.4 Context-aware design

Bradley and Dunlop writes in a study that visually impaired people will vary individually and collectively in their use of contextual information to navigate/orientate [37]. In their results they also present that there were differences in information presentation styles (verbal vs. visual) and clear preferences for control over contextual information.

The main usability implication/issue that they present is that the design of the application must allow an element of user control in order to present contextual information that is appropriate to a user's task and situation. For example, some scenarios speech output would be better for reaching a destination promptly. Whereas, visual presentation involving additional information may be preferred (or used in conjunction with speech output) when touring a city for the first time in order to provide a greater spatial orientation and awareness of surrounding environmental features/landmarks. Lastly, they state that another usability implication may be that more contextual information needs to be provided for speech output than for visual presentation [37].

3.4.1 Contextual information

In an experimental wayfinding study Bradley and Dunlop [10] indicate that instructions formed from visually impaired people resulted in a lower weighted workload score, less minor deviations, and quicker times for visually impaired participants. In contrast, these instructions were found to cause a higher weighted workload score for sighted participants. This suggests that a different approach for sighted and people who have visual impairment is needed to be able to design a good user experience. Bradley and Dunlop [10] give two examples of this;

1. The presentation and notifications needs to be prioritized with respects to the user's focal activity and high level goal (e.g., a blind person may not want to be pushed with information about a friend in a nearby cafe as he or she crosses a busy pedestrian crossing). Factors such as cognitive workload would need to be considered.

2. Users who wish to send messages at specific locations for others to retrieve may need to prioritize their messages to suit a fitting audience. For instance, recommending a bookstore to a community of visually impaired travelers would be given low priority, whereas warning of a dangerous hazard would be given high priority (and possibly pushed be to them).

3.4.2 Navigation

When looking into context for visually impaired and their way of navigating through an area or open space there is an difference between how technology can aid in the dealing with the space near to the user and the space further away. When it comes to outdoor navigation, Petrie is distinguishing between the two by calling them micro- and macro navigation technologies [14]. Micro navigation technology is providing information from the environment closest to the user and macro navigation technologies are providing information about the distant environment. Examples of micro navigation developed to assist visually impaired travellers are for instance, Electronic Travel Aids, such as obstacle avoidance systems (e.g. Laser Cane and ultrasonic obstacle avoiders). Global Positioning Systems (GPS) and Geographical Information Systems (GIS) are examples of technologies that have been used to assist macro-navigation.

The context information that is provided is vital and the most important thing is to provide contextual information from a multidisciplinary viewpoint. Harper and Green [7] says that most of the information provided to aid the independent traveller today is in the form of visual cues, including graphical signs, coloured lights, road markings and printed public transport information. These variations in mobility markings for sighted users suggest that travel should not be thought of as a single activity but a sequence of different activities because there has evolved a preferred means of conveying the useful information associated with each of the different travel activities to a user [7]. Harper and Green further regards to the act of travelling as a ”flow of travel”, rather than a series of interrelated tasks.

3.5 Multimodal interfaces

Graphical User Interfaces (also called GUI) are a standard in Human-Computer interaction but it is important to develop alternate modes of interaction for visually impaired or “sight occupied” users. Since sight is the most rapid and precise way to convey graphical information, it appears essential to replace it not by the use of one method but many [38]. The term used to describe that a user is interacting with a system using different senses or modalities such as speech, pen, touch, gestures, eye gaze, head and body movements is multimodal interaction [39]. Systems that is designed for people with an visual impairment often uses this kind of approach to combine the different senses. Many different sources [38][23][37], are stating that the best way to design for the visually impaired is to use a multimodal approach. In a study performed by Dufresne et al. [38] they performed a test where a visual modality in a website was replaced by one or two modality(ies). Their research showed that multimodality was associated with better performance for blind and sighted subjects and that it was ranked as the best interface, with the haptic ranked as the second best.

Chapter 4

Methods

The methodology is presented in this section and consists of five main sections: 4.1 Introduction, 4.2 Literature Study, 4.3 Interviews, 4.4 Data Analysis and 4.5 Workshop

4.1 Introduction

The method was conducted in two phases: Understanding and exploration.

Understanding

In the first phase, which includes section 4.2 and 4.3, the focus was to understand the research area and the current user situation. This was done by conducting a literature study resulting in extensive knowledge of areas touched on in this thesis. That was followed by a series of interviews with persons who either had an visual impairment or a blindness. This was done to understand the usage, needs and challenges living with an visual impairment and how they were using digital tools and voice interfaces in their daily life.

Exploration

The goal of the second phase was to use the knowledge and research found during the first phase to explore new areas. This was done through a workshop at the office of North Kingdom. The goal of the workshop was to explore some problem areas and come up with an idea or concept on the bleeding edge of how interaction with voice and a haptic interface could work and a concept proposal (presented in section 6) was made.

4.2 Literature Study

In the beginning of the work a broad variety of literature was evaluated. The type of literature varied from books, articles, previous interviews, videos, magazine posts to blog posts in order to get a wide understanding of the field. The reviewed research mainly consisted of articles found through Umeå University's online library[1], by searching using keywords and phrases such as "voice user interfaces", "designing for the visually impaired", "visually impaired", etc. The conclusion of the literature study covered most areas of design for VUI, haptics and multimodal interfaces and as well as guidelines in designing for blind and the visually impaired. This included workload, plasticity, experience design, etc. and served as a foundation for the theoretical framework (described in Section 3).

4.3 Interviews

A qualitative study consisting of interviews with people with a visual impairment was conducted. The goal was to investigate and identify the biggest obstacles and problem in the everyday life in Sweden. Another goal was to investigate what tools are most common and what the thoughts and feelings the subjects have towards VUI and haptics. According to Rowley [40], semi-structured interviews are good for collecting data when:

- The research objectives centre on understanding experiences, opinions, attitudes, values, and processes.
- There is insufficient known about the subject to be able to draft a questionnaire.
- The potential interviewees might be more receptive to an interview than other data gathering approaches.

A semi-structured interview opens up to more detailed and insights that might be missed in a questionnaire [40]. According to Rubin this model also allows the interviewer to ask follow-up questions that deviate from the predefined set of structured questions in order to get a more clear and explanatory answer [41]. This structure was chosen since it opens up for discussion and allows for further understanding of the interviewees answer.

4.3.1 Participants

The total amount of participants included in the study was ten people and their information can be seen in table 4.1. Their age ranging between 11 - 84 years of

[1]For more information, see http://www.ub.umu.se/

age. In total there were two females and eight male subjects. The participants were recruited through different platforms and organizations for the visually impaired in Sweden, as well as through common acquaintances. The eye sight were differing from being completely blind to having a visual impairment around with around 10 percent vision. It was also a variation with some people being visual impaired from birth and some getting the eye condition in their teens or as a result of getting an eye disease after retirement. To remain anonymous, the letter in the name column are used to reference the subjects in the Results (see section 5).

Table 4.1: Gender and age group of interviewees.

Name	**Gender**	**Age**	**Vision**
A	Male	10-20	Blind one eye, little vision at the other. From birth
B	Male	20-30	Blind. From birth
C	Male	20-30	1% eyesight. Have gradually been losing sight since 4 years back
D	Female	20-30	Less than 10 % eyesight in the best eye. Have gradually been losing sight since age of 12
E	Male	40-50	Blind. From birth
F	Male	40-50	8% eyesight in the best eye. From birth
G	Male	40-50	Less than 10% eyesight. From birth
H	Male	70-80	Losing sight since age of 55. Can today only see sun and bright lamps
I	Male	70-80	Losing sight since age of retirement
J	Female	80-90	Losing sight since age of retirement

The interviews were all taking place during the fall of 2017. Two of the interviews were conducted in person. One at the office of the company North Kingdom in Stockholm and the other interview was conducted at a public library in central Stockholm. A part from these two, the other interviews were conducted through telephone. All of the interviews were recorded by using a wire connecting the phone to the computer and recording with the software QuickTime. Choosing telephone as the main interview method was based in the problem of finding subjects in the near area. Since travelling to different parts of Sweden was not an option in this study, telephone was chosen as the most fitted alternative. There is not so much research into conducting qualitative interviews by telephone since it is often regarded as a down prioritized option, but there is little research into the actual effects. Novic [42] states that:

> *"The absence of visual cues via telephone is thought to result in loss of contextual and nonverbal data and to compromise rapport, probing, and interpretation of responses. Yet, telephones may allow respondents to feel relaxed and able to disclose sensitive information,*

and evidence is lacking that they produce lower quality data" [42].

Regarding the lack of visual cues, research has been done by a group from the University of Geneva. Their result shows that visual learning seems not to be necessary in order to produce the same pattern of facial expressions in real emotional contexts but that blind subject have a difficult time producing expressions when solicited. As well as controlling the degree of intensity [43]. Which could make it harder to interpret potential visual cues in the right way.

4.3.2 Interview themes

The interviews was based upon 4 different main themes with one extra, gaining extra information if time was found excessive. The four themes consisted of background, tools used today, new places/environments and contextual information. The length of the interviews was around 30 minutes and the extra theme was general questions to further get the subject to think freely about possible solutions to obstacles in their life if time was enough. The order of the themes and questions was ordered to first be easy things to answer to make the subjects feel comfortable before moving on to harder or potential more delicate questions. The questions were designed to be open-ended and a scenario was also chosen to make it easier to set a frame without priming the subjects. The themes and questions can be found in Appendix A.

4.3.3 Interviews with expert

As a part of the prestudy, before the other interviews were conducted, a pair of questions was sent to Léonie Watson. Watson is Communications director and Principal engineer at The Paciello Group (TPG), and also works with Government Digital Service (GDS) on the GOV.UK platform. She is a member of the W3C Advisory Board, and co-chair of the W3C Web Platform WG (working on specs like HTML5), Watson is closely involved with the web standards community and frequently asked to talk about web standards and/or accessibility at conferences. She also writes her own blog[2]. Apart from all this, Watson also have personal experience as she began to lose her eyesight in her twenties and is today completely blind. The purpose of the questions was to get an opinion from an expert about the possibilities and future of using voice for better accessibility. As well as what to watch out for when designing good user interfaces. The questions and answers can be found in appendix B. The answers were answered via email in the fall of 2017.

[2]For more information, see https://tink.uk/

4.4 Data Analysis

The goal after the interviews was to analyze them to find the users needs and goals. Since the interviews were using a semi-structured form, the data analysis started with a transcription of the recordings and then highlighting sections with all the concepts, ideas and findings in each interview. The findings were put onto post-it notes, with each point on its own paper and a small note on who had said it. After all the post-its were written down they were grouped into similar themes that appeared through the interviews. After the grouping were done the notes in each group were further analyzed and divided into ”nice to have”, want and required before the results were compiled.

4.5 Workshop

After the interviews were conducted and analyzed a workshop was performed. The theme and goal of the workshop was to come up with new ideas and work on a possible solution based of the findings from the interviews. This was made to gain the expertise and creativity of the people working at North Kingdom. In total 6 persons joined the workshop and the schedule of the workshop was divided as shown in table 4.2. The main frame of the workshop was based after a toolbox with processes for idea and concept development of the school Hyper Island [44] were brainstorming were used as a tool to generate as many ideas as possible. The idea was to generate quick ideas that later on could be built upon to come up with a concept.

Table 4.2: Schedule Workshop.

Time	**Activity**
10.00	Introduction to workshop (5-10 min)
10.10	Individual brainstorming (5 min)
10.15	Presentation and Cluster ideas (10 min)
10.30	Group brainstorming (15 min)
10.50	Presentation (10 minutes)

In the individual part everybody got time to think about the issues themselves and after that was done everybody got to present their ideas and the ideas were clustered. That part ended with a vote were each team member got three votes each, voting for their favourite idea to continue developing it further in a group using inspiration from everyone.

Chapter 5

Results

The results are divided into two different sections. The first one, section 5.1, is presenting the results from the interviews. Section 5.2 presents the result from the workshop held.

5.1 Interviews

During the interviews there were certain things that were repeated by several person and were therefore considered to be a common problem and things to take into consideration. In the section 5.1.1, the answers from the expert are presented and in section 5.1.2, thoughts and opinions from the visually impaired participants are presented. In section 5.1.3, the identified main problems are concluded. In section 5.1.4, conclusions from the interviews and what to move forward with are presented.

5.1.1 Expert interview

Watson was positive about the future for VUI and thinks that the technology and experience will become more ubiquitous. She says that she is already using tools in her life which are using voice in different contexts. Her screen reader uses synthetic speech to announce what is on-screen, the Echo tells the weather, what is on her calendar, what time it is, and much more. Siri is used to make calls and sends texts and the navigation apps tells which directions to take en route to the end destination. She believes the future can only be better but that there are some issues that needs to be addressed. Privacy, security, and trust are among them. On other thing that Watson is emphasizing when it comes to designing accessible interfaces is the verbosity. Watson says that:

> *You need information in manageable chunks, and that will vary depending on the context. Directions for a walking app need to be short and to the point, descriptions for an app that guides you around an art gallery may need to be more verbose.*

The full responses of Leonié Watson can be found in the appendix.

5.1.2 Thoughts about VUI

All of the participants had, at least once, tried using Siri or some other kind of VUI in their phone. Most of them were in general positive to the VUI, as well as positive in general towards the idea of using VUI on a daily basis, like a lot already were. The most common tasks performed with a VUI in the mobile (in no particular order) was:

- Set an alarm
- Call someone
- Dictation for SMS or email
- Check the weather
- Book a meeting (check calendar)
- Shorter fact search
- Open up an app
- Check the time
- Ask where I am
- Calculate (math)

The issues towards VUI was foremost in shortcomings in the technology. Difficulties in understanding what was being said by the user and the VUI not being able to answer the questions. An example of a shortcoming was that if asking Siri to google on a topic - she answers that she found this on Google and then asks the user to read or explore it by themselves. If asking Siri to read a Google search she do not understand what is being said or is including it in the search, see figure 5.1. This is done according to a guideline in VUI [1] that longer lists should be presented in lists or for the user to explore themselves. The problem is that, for a user that is visually impaired, the only option left is to use TTS to read it out loud, making the shortcut of googling it through a VUI not that useful. One person put it like this:

> *The annoying thing about Siri is when I ask her to search for something and then that she does not say what she has found - A*

On the question in which context they were using VUI most of the participants commented that they use VUI in their mobiles when they are at home or when

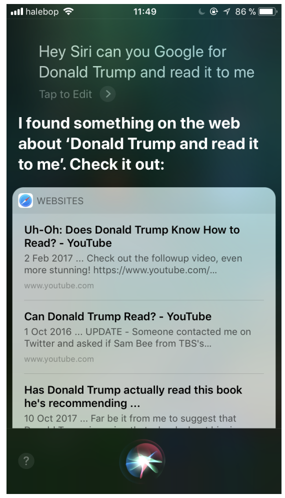


Figure 5.1: Example of Siri not being able to read a Google result out loud

being around people that they know, but that they still feel a bit uncomfortable to do it in public spaces like the bus.

Voice and tonality

The voice and tonality was a factor where the participants had very different opinions. There were in particular one participant, G, who refrained from using VUI just based on the fact that he considered the sound of the voice was too robotic to be used. Other than him, all participants said that they generally did not care about the voice in their VUI except if they were using the voice to read books or doing similar things that was more for leisure than practicality. In those scenarios, it was crucial that the VUI was using a pleasant voice. A few persons, who were well accustomed with their smartphones, mentioned that they had downloaded specific voices that they used in their TTS. Other settings mentioned was pitch and speed. The settings of speed was both to be able to speed it up - for the experienced participants this was very important. The older participants, who had a little bit difficulty with their hearing, mentioned on the other hand that they wanted to be able to slow down the voice to be able to hear what was said.

5.1.3 Main problems

Under this section, there are listed three main problem or areas in the participants daily life which were identified after analyzing the interviews. The main problems identified were navigation, public transportation and shopping.

Navigation

As a normal sighted person you can walk different roads. But as a visual impaired person you normally memorize and learn to walk A - B and B - A. H, who lost his sight in his fifties said that he still spent a lot of time walking around in the places were he still remembers what it looks like. What he is avoiding is new streets and roads. All participants mentioned navigation as a troublesome area and the most common aspects in the field of navigation and mobility was:

- Where am I and how do I get to where I want to go?
- Confirmation - how do I know if I am walking in the right direction?
- Getting information about places around me
- Avoiding obstacles

Addressing navigation, there was a difference between people with low vision and the ones who were blind. All participants who were blind were still very capable of travelling and moving around by themselves but still said that there were certain places they would never go without anyone to accompany them. This was because of the risk of getting lost, or walking in the wrong direction at the wrong place. E, who were blind said that no matter how good a navigation application are, it will always just be a complementary aid. When talking to B about travelling by train, and if there would be any information who could aid him in travelling independently in Sweden he put it like this:

> *Well, the problem is not that you ca not find information. It is on the internet. Just as it is for everyone else. And reading things on the internet is not hard. Or I am not going to say that it is not hard, it is available. But on the other hand, it is not really the same as for anyone who sees. They can look at a map, and from that map understand how they should walk. The information is there. On the other hand, I could use a navigation app, if I would now be willing to do it at a train station. Which I can answer, I am not. That is because they are not always precise and I do not want to end up on a track. And therefore, as well as the transport between the station buildings, I would always call first and use a guiding companion. - B*

B said that to navigate by himself in other places he was using two applications at the same time. He was then using his white cane and walking with Bluetooth headphones, listening to the instructions. One of the applications being the Blindsquare and the other one a normal GPS-application like Google maps to give directions to a specific place.

Two persons stated that they had tried using the Blindsquare app to navigate, both of them being blind. In general they both liked the application and the features in it, like the option of giving directions according to the clock. (Giving directions according to the clock model is for example to say "The royal castle is at 2 o'clock"). None of the persons had tried the voice feature of Blindsquare since it was a premium feature that had to be paid for every month. They had not been using the feature of changing between the buttons in the headphones either. B thought that the application was giving a little bit too much information sometimes, with the comment that it could probably be fixed, but he was a bit to lazy to do it. With that noted, B still wanted more information at for example crossings.

The ones who had low vision expressed the difficulty of seeing signs, fear of walking over roads or bike lanes. Everyone with low vision was using some kind of map application in their phones and the ones who still were using their eyesight mentioned that it was especially good when it is dark outside and the vision gets even poorer. Most participants were using the application by zooming a lot and making sure that they were following the dotted line. A few said that if they were going to a completely new place they always called guide service to take them there or going with the subway and asking someone to meet them at the gate. Still if being able to travel around by themselves it takes a lot of energy. G said that:

> *It takes a lot of effort to walk around and be totally observant of your surroundings all the time - G*

The problem with GPS and its inaccuracy[1] was mentioned as a big trouble. As well as normal navigation apps only giving information as street name when asking the question for directions or asking "Where am I?". All of the participants also talked about the issue of walking into obstacles like poles, signs, bikes and road constructions blocking the road. Regarding the information wanted about the environment all participants preferred practical and objective descriptions before getting impressionistic descriptions of the environment. Another common problem was settings of transportation means. Some GPS applications have a default of car as transportation, which is bad option for a visually impaired user. Changing those settings was also considered difficult, especially for the participants with lower vision.

Public Transportation

The majority of participants with low vision were using public transportation and memorized routes to travel within their community. One problem mentioned was how to know that you are entering the right vehicle. In that specific case, announcements were mentioned as extremely good, but not always reliable. All participants who were living in Stockholm were preferring the subway, as the easier way of transport, since it always stopped at exactly the same place. Buses were harder because of the exact opposite reason. They could

[1] For more information see: https://www.gps.gov/systems/gps/performance/accuracy/

stop at different places depending on external things like the bus driver, traffic, road constructions or if other buses where stopping at the same bus stop. H summarized it like this:

> *Yes, they say that there will come a train but then when a train is coming it is another one. The same applies to the buses. I am using a bus stop where five different buses stop, and it is not so easy to know which one to take. And the difficulty is also to find the right bus stop. - H*

Half of the participants said that if they were going some place completely new they prefer ordering a car service taking them there over commuting.

Shopping

Shopping is a task that is really hard for persons who are blind or visually impaired. 7/10 participants, both blind and with low vision said that this was a problem in their lives. Some said that they had started to shop online instead to be able to get better and more information about what they are buying. But shopping online was not considered as an option for the older participants.

The trouble of shopping begins already at the entrance. H mentioned that he sometimes could feel a bit unsure walking into a store, not sure if it was the right store he was walking into. Other questions in the store were:

- Where is my product
- What does it cost
- Does not know or easily miss special campaigns and offers

The issue of not being able to see special prices or offers as "Take 3, pay for 2" was mentioned to be an issue because it was much harder to buy cheaper food or other products. Only one participant mentioned that knew of an application in his phone to make shopping easier. This was an application to scan bar codes to see which grocery they he was holding.

Further research shows that when organizations are giving tips about how people better can plan and do their shopping most of them involve asking for help from either friends, family or staff at the store [2]

.

[2] For more info: `https://sandysview1.wordpress.com/2015/04/16/how-do-people-who-are-blind-or-visually-impaired-shop-independently/` and `http://www.visionaware.org/info/essential-skills-2/shopping/25`

In general

The most common method used by the younger participants with low vision was to use their smartphone to convey visual information was to take pictures of objects, signs or posters and then zoom into the picture to read the text. This was also done on objects like street signs or menus in restaurants.

One thing that became obvious during the interviews was that there was a big difference between being completely blind and having low vision - even if the person with low vision only could see a few percent. The difference showed itself in how the persons could use their smartphones and how much they still where dependent of and were using visual clues in their daily life. Using a VUI in his home, B, who are blind were noting that he wanted to have audio feedback of visual actions like "Turning down the lights". This was something that was considered unnecessary for the people who had low vision since they still can notice the difference.

Asking people what they wanted most, the most common answer was just to be able to be independent. To be able to do whatever they they wanted to do, whenever they want to do it, without asking for help from friends, family or strangers.

> *I want to be free and able to do what I want, whenever I want to. - D*

5.1.4 Conclusion

All were in general very positive to be using voice as a way of interacting, especially in their homes and in private places. Using voice interactions in public spaces was not considered ideal by most participants but still a possible way if done in the right way, being able to use headphones. Only one person mentioned haptics during the interview. E said that he thought that vibrations and tactile methods was something that generally could be used more and in a more effective way to improve the usability.

One subject talked about in almost all interviews was about audio description and the importance about it. Asking if people wanted more audio descriptions of public places and their environment people were in general more skeptic but they liked the feature of having audio description of for examples images.

D talked about the importance of assuming that people who have a visual impairment wants everything with voice or tactile element but that it should be a mix of them all to be able to accommodate for different functionality. She put it like this:

> *It is important to know that some people who cannot really see still uses their eyes a lot. So it is important with images and text and to*

> *not only use Braille. I think it is good with Braille but personally, I do not even know it - D*

All participants with low vision mentioned that they used some kind of zooming tool, either in their phone or with a magnifier glass to be able to see and read visual information. F said that he was actively not choosing the special tools provided by the help companies in Sweden. But wanted to use normal apps and tools in his smartphone to make things more accessible. This is an opinion of making the app work for a larger group of people, not only the visually impaired or blind. But foremost it is about being able to deliver a service that is needed.

5.2 Workshop

Going into the workshop with the knowledge from the interviews and using the expertise of people at North Kingdom who are good at thinking creatively, the workshop generated a lot of different ideas based of the most common main problems. The ideas produced were clustered in the themes of public transportation, shopping, navigation, geofencing and general ideas looking at the problem at a higher level.

One main concept was chosen as the most promising one was a concept which combined the technology of geofencing and beacons in a smartphone application. The idea was chosen as the best one since it could address the problems in a more general way and also provide contextual information in different places. After the workshop a benchmarking was made and the concept was further inspired by an application called Wayfindr[3]. Wayfindr are also using similar technology and are doing research for a system for the blind using beacons to provide information at subway stations in London and Sydney.

[3]Read more on: https://www.wayfindr.net/

Chapter 6

Concept Proposal

The concept proposal is based of an idea from the workshop - incorporating the insights from the interviews and the literature study. It involves technology of beacons and geofencing to provide an application platform were contextual information, which often are displayed visually, can be shared and accessed by visually impaired people. The goal of the concept is to start designing better and more inclusive spaces to help visually impaired people live more independent lives.

The concept is a platform which aims to be of assistance in all problem areas identified in the interview process. It means the platform should be able to be applied in different contexts, such as subway stations, bus stops, stores or other public places. In this thesis only one context, a grocery store, have been chosen to be used as a prototype presentation to explain the concept.

6.1 Concept Technology

In this section the technology of geofencing and beacons are briefly descibed. Figure 6.1[1] is showing how the geofence is active around the specific location of the store and how the beacons are working within a closer radius.

6.1.1 Geofencing

Geofencing is a feature in a software program that uses the global positioning system (GPS) or radio frequency identification (RFID) to define geographical boundaries[2].

[1] Image source: https://www.plotprojects.com/
[2] More information: http://whatis.techtarget.com/definition/geofencing

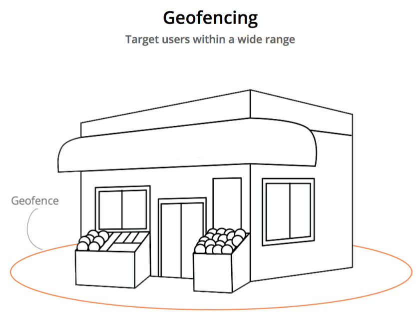


Figure 6.1: Image showing the proximity of geofencing and beacons.

By using geofencing it allows an administrator to set up triggers so when a device enters (or exits) the boundaries defined by the administrator, an alert is issued. Many geofencing applications incorporate services like Google Earth, allowing administrators to define boundaries on top of a satellite view of a specific geographical area. Other applications define boundaries by longitude and latitude or through user-created and web-based maps.

Geofence virtual barriers can be active or passive. Active geofences require an end user to opt-in to location services and a mobile app to be open. Passive geofences are always on; they rely on Wi-Fi and cellular data instead of GPS or RFID and work in the background. This is the kind of geofencing which will be used in this concept.

6.1.2 Beacons

Beacons are one-way transmitters that can be used to mark important places and objects. Typically, a beacon is visible to a user's device from a range of a few meters, allowing for highly context-sensitive use cases. Beacons are small, battery-powered, always-on devices that use BLE (Bluetooth Low Energy) technology to transmit signals to devices, such as smartphones and tablets, within a range of about 90 meters[3].

Multiple beacons can be positioned around an area, such as inside a store or airport, to broadcast relevant information to portable devices within their proximity. Mobile device owners can then react to, engage with, or use the information for indoor, turn-by-turn navigation and store discounts, among other things.

[3]More information: http://www.ibeacon.com/what-is-ibeacon-a-guide-to-beacons/

6.2 Concept Guidelines

Below are four areas which have been used as basic guidelines for the concept. They are based of the results from the literature study and interviews.

Contextual Information

To be able to provide contextual information, the location of the user is used. It is accessed through GPS and Bluetooth in the user's smartphone. This is to locate the closest geofences and be able to access the information available.

Multimodality

The multimodality in this concept, is that the user have the opportunity to choose to interact with the application either with voice or the graphical interface. The interface have a graphical interfaces so the user can be using the keyboard and pressing the buttons if wanted. Maps and other visual information is also provided to accommodate those who prefer to use GUI in their smartphones. Text-to-speech is used to provide information on the screen.

Feedback

Feedback is provided both through audio, graphics and vibrations. The user is getting feedback through haptic actions like vibrating of the phone or sound confirming actions or signaling when a geofence is active. Since many of the participants in the interviews were saying that they were using their screens a lot, feedback should also be provided in the GUI. The blind participants were saying that they were either holding the phone in their phone or having it in their pocket while walking.

Workload and Context

The idea of the concept is that there should only be available information about the area closest to the user. This is to minimize the workload and avoiding the pitfall of conveying to much information. It is also for making it easier for the user to navigate the application and know what they can do in each specific area. The available actions are displayed at the home screen. When the user is entering a geofence the application is vibrating and making a sound making the user aware that he is entering an active geofence. To gain more information, the user have to search or ask for it to not provide too much information in wrong or busy situations. To minimize the workload the user should also be able to easily filter which contexts he/she want to access.

6.3 User Journey

The user journey describes a scenario of how the user could be using and interact with the service. For a visually impaired user the journey of going shopping differs from a sighted person. In figure 6.3, the seven steps describes a journey from the user's home, realizing he needs to buy some groceries. Walking to the store and getting his groceries. The figure shows in which parts of the journey the application can aid the user in his tasks.

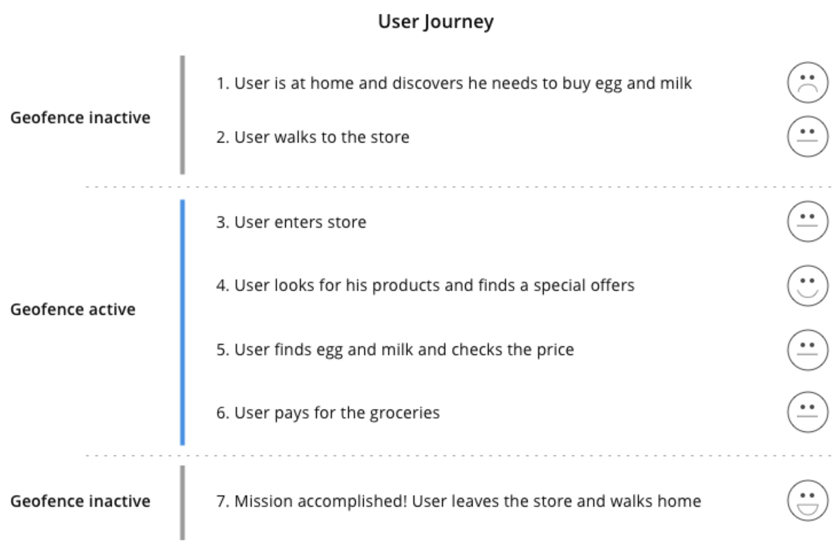


Figure 6.2: Figure showing the seven steps of a user going shopping.

The application is active around the actual location of the store and that is where the user can access the available information. This is to provide the contextual information in the right context.

User enters store

When the user walks into the geofence the application vibrates and lets the user know he/she is now in an active geofence. When the user enters the store he receives another vibration and message on the screen that he/she is entering the store. The contextual information wanted in this scenario is a confirmation walking into the right store and which information the user can access inside the store. Beacons located at the entrance can provide feedback that the user are entering the right store.

User looks for his products and special offer

If you are visiting a new or not so familiar store visually impaired people are often advised to make a list and ask someone of assistance. Either a friend or a staff member to find the right spot of the product and help to pick the right item off the shelf. This is something that is wanted to be done independently. Which can be done through a indoor map and beacons with information where products are stored.

If it is a store you know well, you might have memorized the route to find a specific product. But if the special offers only are presented through signs the user might miss the special offer that if you buy two packages you get the eggs cheaper. The special offers can instead be found in the application.

User finds egg and milk och checks prices

When the user have found the eggs and milk he/she might want to check the price to see how much the products costs. Checking the price is hard to do without someone telling you or if you have low vision you might be able to take pictures of the tag and then zoom in to see the price. The application can be of aid through scanning the product and having the application reading the information and providing the information on the screen for the user to zoom into.

User pays for products

When all products have been found the user pays for his products at the cashier. Visiting a big store for the first time it might not always be that easy to find the cashier. The application can provide a indoor map and confirmation that the user are approaching the cashier and what payment methods are available.

6.4 Application

This section describes the concept's application, how the smartphone communicates with the physical world, and how the user interface is structured. The application is a minimum viable product and the flow is designed to be as easy to navigate as possible.

6.4.1 Application flow

Non-trivial screens of the application are described in this subsection. The flowchart in figure 6.3 show the structure of the application. The site map describes the main screens and the pop-up to allow the application to use GPS. The application also need the user to have Bluetooth active to be able to access the beacons.

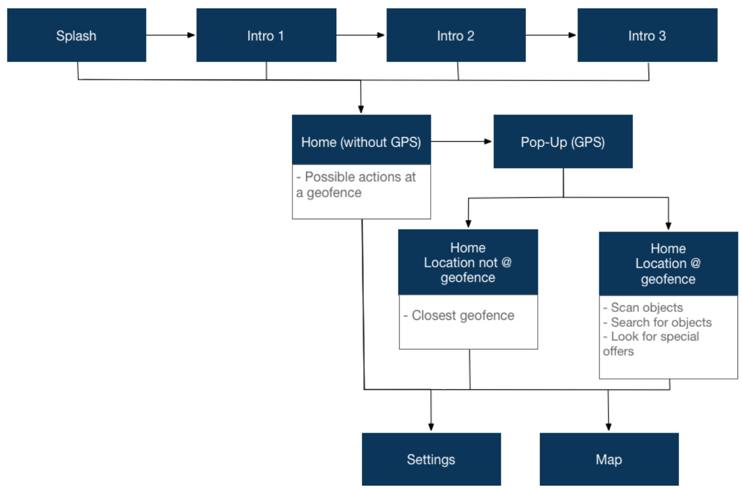


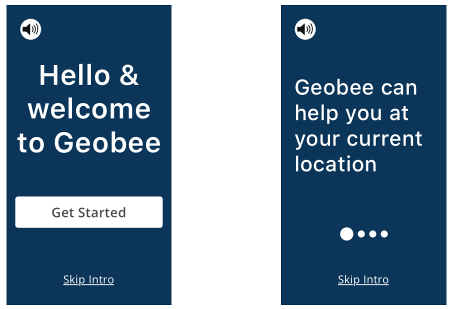
Figure 6.3: Image showing the site map and flow of the concept.

6.4.2 Onboarding

The onboarding have the purpose of introducing the application and what it can do for the user. The default of the onboarding is to have the text read out loud but the voice can be turned off. The onboarding flow can be seen in figure 6.4 and figure 6.5. The first screen, also referred to as a “splash screen”, is the very first thing the users will see when they open up the application. After that a flow of four screens will follow. The user can however choose to skip these at any time.

6.4.3 Home

The home screen is the main screen of the application. It can be seen in figure 6.6. The first version that the user are exposed to is a version with no GPS allowed. There is a short introduction at the top. Below the heading there are some possible actions the user can access, increasing the users curiosity of what the application can do. All of the buttons trigger a pop-up to allow GPS since it is necessary to make the application work properly. The purpose of showing

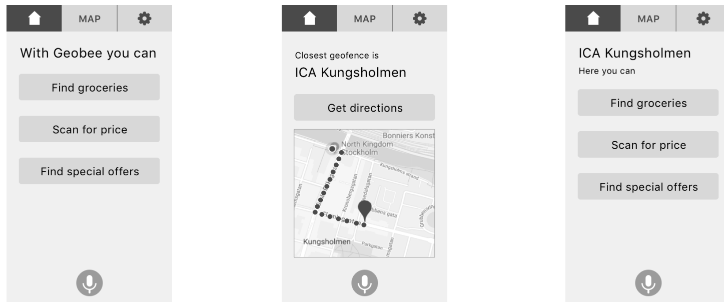


(a) Splash screen (b) Intro - first screen

Figure 6.4: Two screens showing a splash screen and the first introduction screen.

(a) Intro - second screen (b) Intro - third screen (c) Intro - last screen

Figure 6.5: Three screens showing the rest of the introduction screens.



(a) No GPS (b) GPS - no geofence (c) GPS - geofence active

Figure 6.6: Three examples of the home screen. Without GPS and within and outside of an geofence.

this screen first is to lower the drop-out rate by showing the GPS pop-up in a context.

The second version of home is when the GPS is allowed but the user is not in any active geofence. The closest geofence is shown in a map and the user can get direction to the geofence.

The third version is when the user has allowed GPS and is in an active geofence. The possible actions in that geofence are shown as call to actions as buttons. The user can access the actions through the buttons or by the microphone button which opens up to voice commands. This is something that is in common through all screens.

6.4.4 Settings

Settings available to the user is filtering on geofences and being able to set more specific accessibility settings. In figure 6.7, the settings screen can be seen. The application has for example an option of inverting the colors to make the contrast better. The app should overall be designed to only be using high contrast colors and making it easy to use TTS. The font size can also be increased in the application.

6.4.5 Map

In figure 6.8, the map screen is shown. The main action available here is to search for geofences. The map is showing the user's location and nearby geofences. If the user have not allowed GPS a random location is shown in Sweden. The user can also search for geofences in the map section. Entering both generic words or specific locations.

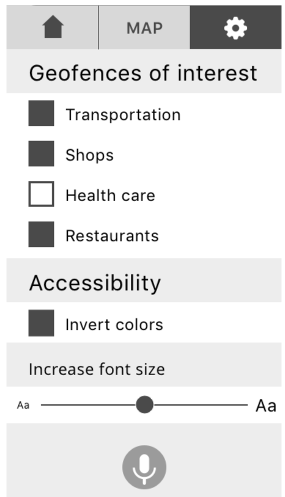


Figure 6.7: Screen of the settings of the concept application.

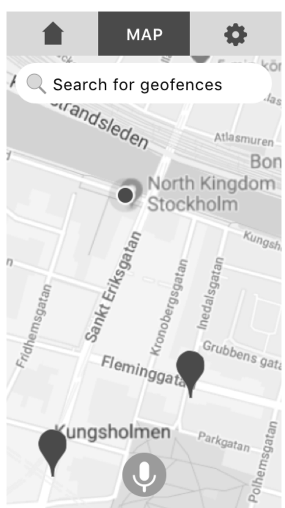
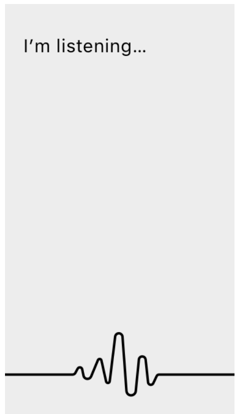
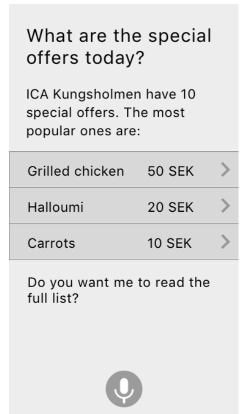


Figure 6.8: Screen of the map section, showing nearby geofences.



(a) Application listening



(b) Application have registered a question and the user is receiving an answer

Figure 6.9: Two example screens of the VUI in the concept application.

6.4.6 Voice

The application can be controlled with voice and all actions available with buttons can also be accessed by voice. In figure 6.9, the two main screens of how the voice works are shown. The first one, in a, is the first screen when the user have pressed the voice button and the application vibrates and makes a listening sound. The screen is having the screen showing the text ”I am listening” and the line in the bottom of the screen signaling that listening mode is on.

In the b screen the user have asked a question. In this case the application understood the user saying: ”What are the special offers today?”. The application have the question in the top of the screen and should start by telling the user how many offers the store has in total. Then instead of telling the whole list, breaking it up and only giving the three most popular ones before asking the user: ”Do you want me to read the full list?”. This way the AI is not getting tiring listening to. The user can get more about each offer by pressing it or asking for more info about the halloumi for example. The information is in different layers to only provide the most important information first.

Error handling

If the application have not registered any question or command by the user, it moves on to the screen shown in figure 6.10. This example shows an example of when the user is not in any active geofence and he/she is given some examples of things the user can ask for. The suggestions of the things a user can ask should be adapted according to context. For example, if the user would have been in a active geofence, the most central actions of that geofence would have been shown instead.

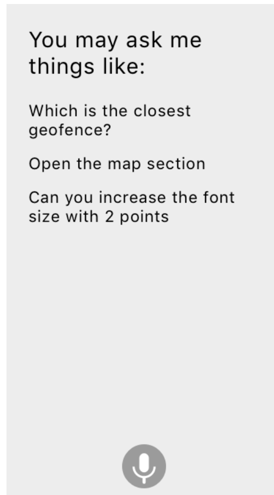


Figure 6.10: Example of error handling if the user is quiet or if the application cannot register a question or command.

6.5 Concept Limitations

The entire concept needs testing and validation of users with a wide range of different visual impairment to evaluate it. The technology have not been a focus in this study and to be able to perform a fully functioning prototype, further research into the technology would be necessary. Providing a proper setup of beacons and the geofence to make an accurate and precise prototype. This is a vital part since this is a prerequisite to be able to provide a functional system. The Wayfindr are a company using the same technology of beacons and have been used as a reference in the technology but the author have not found any other references in the area.

6.6 Future Work

The concept proposed in this study are a platform from which other services can be deployed and it would be interesting doing further work, trying different contexts based on the same site map. It would have been interesting to proceed with interviews with a deeper depth, focusing just on the main scenarios. Future work would therefore involve more prototyping, making more user stories and a fully functional prototype and testing it on users with different visual impairment. Doing user tests in the wild, as well as including more contexts, the author believe can bring valuable insights into how to move forward with the concept. Testing in the wild might also give valuable insights into other contextual information which the participants in the study might not have thought of since the questions in the interviews were asked out of context.

Chapter 7

Conclusions

The objective of this thesis was to investigate how to design for people who have a visual impairment. The two research questions were what kind of contextual information is wanted from visually impaired in new, unfamiliar spaces outside their home and how, if possible, can an interface in a smartphone be used to provide such information.

The method was divided into two parts. The first one was understanding and the second one exploration. The understanding consisted of a literature study, which investigated existing guidelines of designing for the visually impaired. The literature study also involved a brief benchmarking of applications and tools designed specifically for visually impaired. The result of the literature study was that it is important to be aware of the workload, not providing too much information. The information should also be context-based and a multimodal interface is preferred. The literature study was followed up with ten interviews. The interviews were conducted with people in different ages, ranging from the youngest person being 11 years old to the oldest person of 84 years of age. The participants had different degrees of visual impairment, grading from completely blind to about 8 % vision. After the interviews the answers were analyzed by finding common themes in issues expressed and contextual information asked for.

The results of the interviews showed a few common themes, consisting of the participants thoughts about voice interfaces and how the participants liked to use their smartphones. The users were in general positive to the tools they were using today and all of them were in some context using their smartphone as a tool. The results also showed that there was a difference in how the users wanted to interact with their smartphones, depending on their visual functionality and the context. Their satisfaction with smartphone applications can be put in context with the quick development and general improvement of accessibility in smartphones and technology in the last few year. Especially Apples products were mentioned by many participants.

Going into the exploration phase, three main problem areas were discovered. They were navigation, public transportation and shopping and the contextual information wanted in each space was identified. The information the participants were asking for when navigating was "Where am I and how should I get to my destination?". The existing applications were considered decent but could be improved in how the transportation settings were made. The contextual information that was being provided could also be improved. In public transportation the main issues was knowing when the vehicle was arriving and where it was stopping, both when entering and leaving the vehicle. Subway was considered to be a easier option in this context. Stores and shopping were considered to be a difficult task in the sense of finding the right product, checking prices and not being able to find information of current special offers.

A workshop was held with the purpose of idea generating, based of the issues concluded above. An idea was chosen as promising from the workshop and a conceptual proposal was made. The proposal was addressing the issues by using a smartphone application and the technology of geofencing and beacons. The concept was using a multimodal interface, contextual information, feedback and paying attention to the workload. The concept was presented through a user journey of a person going shopping for groceries and explained through wireframes.

As a final conclusion, the author believes that the most important thing, to be able to make inclusive design, is to start making all places and tools, virtual or real, as accessible as possible no matter functionality. The author strongly believe that this could be made in a much better way but that it requires more knowledge and research. One way of achieving this would be if visually impaired people could be included in more user testing of new products, services and environments.

7.1 Limitations

The objective of this study was reached but some limitations and improvements are discussed below.

The small amount of participants in the study makes it hard to draw fully supported conclusions and generalizations. The amount was limited by the time constraint and the difficulty in finding participants. The semi-structured approach was considered to be good since it, during the interviews, became obvious that depending on the participants vision loss, age and social situation the discussion developed into different directions. The disadvantage of the approach was that it was hard to go deep into all aspects. At the start of the thesis the main focus was on VUI but as the knowledge in the area grew, the author understood that the interface approach needed to be more open. That meant that with each interview the follow up questions could better penetrate how the participants experienced different situations. An improvement would have been to have more time to research and find a more definite direction before the interviews started.

A user journey touches many areas of possible developing, the wide spread research can always be made deeper and in different fields. The actions available was questions mentioned by the participants in the study. With deeper research in each field stronger conclusions and more numerous finds might emerge. Previous research used in the theoretical section are mostly about developing and designing for visually impaired people with focus on mobility and navigation aid. Since the proposed concept are applying a different context, the concept would benefit from more studies and tests from similar applications.

The concept proposed requires further testing in order to fully evaluate the system, both in interaction and information level. Due to lack of time and because of the time consuming task of creating and testing a multimodal interface this was not possible. Especially to test the prototype and make it work with a TTS, would have required an almost fully developed product.

Chapter 8

Acknowledgements

First I want to thank Sonja Lakner for being a great supervisor at North Kingdom, giving inspiration, feedback and pushing my thesis forward. Thank you Shafiq Urréhman, my supervisor at Umea University for providing great insights and constructive feedback. Big thanks to Malin Jofjärd Lövgren and Pontus Henstam for giving me the best peer-reviews one could possibly have. Last, but not least, thank you North Kingdom and the Stockholm office for all ping-pong games and friendly athmosphere at the office.

Bibliography

[1] Cathy Pearl. Designing voice user interfaces - principles of conversational experiences. O'Reilly Media Inc, 2016.

[2] Visual impairment and blindness - fact sheet no.282. http://www.who.int/mediacentre/factsheets/fs282/en/, October 2017.

[3] Telecommunication Standardization sector of ITU. Accessibility terms and definitions. http://www.itu.int/itu-t/recommendations/rec.aspx?rec=12624, 2015.

[4] Lederman S. J. Loomis, J. M. Tactual perception. *Handbook of perception and human performance*, Vol 2:pp 1–41, 1986.

[5] Michael Paciello. Web accessibility for people with disabilities. https://goo.gl/gjpa9D, 12 2006.

[6] Plaisant C. Shneiderman B. Lazar J Zhao, H. A framework for auditory data exploration and evaluation with geo-referenced data sonification. http://hcil.cs.umd.edu/trs/2005-28/2005-28.htm, 2006.

[7] Peter Green Simon Harper. A travel flow and mobility framework for visually impaired travellers. https://goo.gl/9eqH7V, 2000.

[8] N.A. Bradley and M.D. Dunlop. Investigating context-aware clues to assist navigation for visually impaired people. https://strathprints.strath.ac.uk/2573/, 09 2002.

[9] North Kingdom. http://www.northkingdom.com/, 2017.

[10] Mark D. Dunlop Nicholas A. Bradley. An experimental investigation into wayfinding directions for visually impaired people available. https://link.springer.com/article/10.1007/s00779-005-0350-y, 6 2005.

[11] Maureen A. Duffy. Low vision and legal blindness terms and descriptions. http://www.visionaware.org/info/your-eye-condition/eye-health/low-vision/low-vision-terms-and-descriptions/1235.

[12] Serbanati A. Biader U.C Panizzi E D'Atri A. D'Atri E., Medaglia C.M. A system to aid blind people in the mobility: A usability test and

its results. http://ieeexplore.ieee.org.proxy.ub.umu.se/document/4196337/?reload=true, April 2007.

[13] Wayfindr. Open standard for audio-based wayfinding. https://www.wayfindr.net/open-standard, 2017.

[14] Petrie H. User requirements for a gps-based travel aid for blind people. In Gill JM, Petrie H (eds) Orientation and navigation systems for blind persons, 1995.

[15] G. L. Allen. Spatial abilities, cognitive maps , and wayfinding: Bases for individual differences in spatial cognition and behavior. *Wayfinding behavior: Cognitive mapping and other spatial processes, ed. R. G. Golledge*, pages 46–80, 1999.

[16] Louise Fryer, Jonathan Freeman, and Linda Pring. What verbal orientation information do blind and partially sighted people need to find their way around? a study of everyday navigation strategies in people with impaired vision. *British Journal of Visual Impairment*, 31(2):123–138, 2013.

[17] American Foundation for the Blind. What is braille? http://www.afb.org/info/living-with-vision-loss/braille/what-is-braille/123, 2017.

[18] SRF Swedish association for the Blind. Louis braille - punktskriftens uppfinnare. http://www.srf.nu/leva-med-synnedsattning/att-ha-en-synnedsattning/att-lasa/punktskrift/louis-braille---punktskriftens-uppfinnare/, 2016.

[19] Funka. Design för alla - statistik. https://www.funka.com/design-for-alla/tillganglighet/statistik/, 2017.

[20] Leonie Watson. What is a screen reader? https://www.nomensa.com/blog/2005/what-screen-reader, 2005.

[21] Blindsquare. What is blindsquare, 2017.

[22] Andrew King. Crossmodal plasticity and hearing capabilities following blindness. https://link.springer.com/article/10.1007/s00441-015-2175-y, 3 2015.

[23] Jóhannesson Ómar I.a Balan Oanab Spagnol Simonec Valgeirsdóttir Vigdís Valaa Unnthorsson Rúnarc Kristjánsson Árnia, Moldoveanu Alinb. Designing sensory-substitution devices: Principles, pitfalls and potential. https://content.iospress.com/articles/restorative-neurology-and-neuroscience/rnn160647, September 2016.

[24] Michel R Raab A Petrie H Johnson V Reichert L Schalt A Strothotte T, Fritz S. Development of dialogue systems for a mobility aid for blind people: Initial design and usability testing. https://dl-acm-org.proxy.ub.umu.se/citation.cfm?doid=228347.228369, 1996.

[25] James R. Loomis Jack M. Klatzky Roberta L. Golledge, Reginald G. Marston. Stated preferences for components of a personal guidance system for nonvisual navigation. *Journal of Visual Impairment and Blindness*, 98:135–147, 2004.

[26] Antonino Cotroneo. Bone conduction: what it is and its benefits. http://blind.tech/bone-conduction-three-simple-questions-answered/, 2016.

[27] James P Giangola Michael H Cohen, Michael Harris Cohen and Jennifer Balogh. Voice user interface design. Addison-Wesley Professional, 2004.

[28] Siddharth Pathak Deepak Begrajka, Shekhar Jain. Hci guidelines for designing website for blinds. http://research.ijcaonline.org/volume103/number11/pxc3899403.pdf, 10 2014.

[29] Google. Be cooperative...like your users. https://developers.google.com/actions/design/be-cooperative, 2017.

[30] Kylie Martin Heather Bowe. Communication across cultures: Mutual understanding in a global world, 2007.

[31] Jacob Nielsen. Voice interfaces: Assessing the potential. https://www.nngroup.com/articles/voice-interfaces-assessing-the-potential/, January 2003.

[32] Kathryn Whitenton. Audio signifiers for voice interaction. https://www.nngroup.com/articles/audio-signifiers-voice-interaction/, September 2017.

[33] I. Fridman. Surface area of human skin, 2001.

[34] Nathaniel I. Durlach and Anne S. Mavor. Virtual reality, scientific and technological challenges. https://www.nap.edu/read/4761/chapter/8, 1995.

[35] Chang Nguen. Haptic obstacle detector for the blind. http://www.diva-portal.org/smash/get/diva2:818827/FULLTEXT01.pdf, 2014.

[36] Vineeth Balasubramanian Dirk Colbry Sethuraman Panchanathan Troy McDaniel, Sreekar Krishna. Using a haptic belt to convey non-verbal communication cues during social interactions to individuals who are blind. http://ieeexplore.ieee.org/abstract/document/4685291, 2008.

[37] Mark D. Dunlop Nicholas A. Bradley. Understanding contextual interactions to design navigational context-aware applications. http://dx.doi.org/10.1007/3-540-45756-9_37, 2002.

[38] Martial O. Ramstein C. Dufresne, A. Multimodal user interface systme for blind and.

[39] S. Oviatt. Ten myths of multimodal interaction. *Communications of the ACM, 42(11):74–81*, November 1999.

[40] Jennifer Rowley. Conducting research interviews. *Management Research Review*, Vol 35:pp. 260–271, 2012.

[41] Herbert J Rubin and Irene S Rubin. Qualitative interviewing: The art of hearing data. Sage, 2011.

[42] Gina Novick. Is there a bias against telephone interviews in qualitative research? *Nursing Health*, Vol 31:pp. 391–398, 2008.

[43] Edouard Gentaz Dannyelle Valente, Anne Theurel. The role of visual experience in the production of emotional facial expressions by blind people: a review. *Psychonomic Bulletin Review*, pages pp. 1–15, 2017.

[44] Hyper Island. Idea concept development. `http://toolbox.hyperisland.com/idea-concept-development`, 2014.

Appendix A

Interview Script

Hi! Thank you for taking time to talk to me. My name is Stina Olofsson and I am writing my master thesis about user interfaces for persons who have a visually impairment and I am therefore curious about your experiences and how you experience challenges and technological possibilities in your everyday life.

Is it okay if I record our conversation so that I go back to it later? The recordings will only not be published anywhere and if you want to I can send you the transcript after the interview. The interview will take about 30 minutes and we can stop at any time.

Below is the different themes used during the interview. Follow up questions was asked depending on the subject, as well as the order of the questions differed depending on the subjects background and visual condition.

- Background
  - Could you tell me some about your background?
  - Visual impairment?
- Tools used today
  - What kind of tools are you using today?
  - Are you using any voice interfaces, like Siri or Alexa for example?
  - (If using) When are you using it and what do you think about using that kind of interface? Why?
- New places/environments
  - Scenario: If I where to become visually impaired or blind and would start to visit new a place - say come to visit you in ... Could you

describe for me how I would be able to do it and what I should watch out for?

- Contextual information
    - What kind of information do you want when you are visiting a new place? Why?
- General
    - What are the biggest obstacles in your everyday life?
    - If you could dream - what kind of tool would you like to use or have in the future?

Appendix B

Interview Expert

1. What are the most common things that designers and developers overlook when it comes to designing accessible user interfaces?

> That verbosity is an important factor. You need information in manageable chunks, and that will vary depending on the context. Directions for a walking app need to be short and to the point, descriptions for an app that guides you around an art gallery may need to be more verbose.

2. What do you think about voice interfaces that exist today - is there anything that they are missing when it comes to accessibility?

> They leave out people who are Deaf or hard of hearing, and people who are unable to speak or articulate properly.

3. What do you think are the future of voice interfaces?

> I think they will become ubiquitous, but we have lots of problems to solve - privacy, security, and trust for example.

4. If the world could become more accessible through a voice interface - where would you like to start and what kind of information would it provide?

> It's already happening. My screen reader uses synthetic speech to announce what's on-screen, my Echo tells me the weather, what's on my calendar, what time it is, and much more besides, Siri makes calls and sends texts for me, navigation apps tell me which directions to take en route to my destination.