MAUU5900

Master Thesis

in

Universal Design of ICT

May 2018

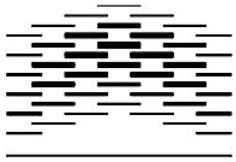
Universal Design of Intelligent Personal Assistants

Regont Kurtishi

OsloMet University, Formerly Known as HiOA

Department of Computer Science

Faculty of Technology, Art and Design



OSLO AND AKERSHUS
UNIVERSITY COLLEGE
OF APPLIED SCIENCES

1. Preface

This master thesis makes up the final sixty credits of a two-year study in Universal Design of ICT in OsloMet University, formerly known as HiOA.

First and foremost, I would like to thank my supervisor professor George Anthony Giannoumis for the remarkable assistance and direction with my work. My supervisor has always been helpful, supportive, available, energetic, and in good humour all the way. Also, he was extremely knowledgeable and fruitful even when it came to the tiniest details and always willing to share his opinions in regard to this thesis. During this one and a half year, I have come to realize the importance of walking this kind of journey with proper guidance and I was lucky to have such one.

I must also express my profound gratitude to thank my family for their support and for making it all possible for me to do my master studies abroad, respectively here in Norway. In all honesty, I wouldn't be here if it wasn't for them. On the other hand, I am happy that my wife was beside me during all times. She inspired and motivated me, game me strength to move forward, supported me, and more importantly made me believe in myself during all times, and I thank her infinitely.

Finally, I am also so grateful to the people who have participated in my thesis along with the associations and organizations that helped me in the recruitment process as well as their support. My sincerest gratitude is owed to them who took the time to be a part of my thesis. Lastly, I am pleased with my colleague friends whom every now and then made the time so we could all meet and work together by sharing our opinions and supporting each other.

Regont Kurtishi

Oslo, 16 May 2018

Abstract

As new ways of information and communication technology arise along with new forms of interactions, it is extremely important that these innovations are created, maintained, and further developed and improved under the particular rules of universal design that apply to the corresponding technology.

Since 2011, with the introduction of Apple's iPhone 4S with its intelligent personal assistant (IPA) Siri implemented, until today, major companies have announced their IPAs to make our lives easier. Day by day, they are continuously implemented in many other platforms, such as appliances and cars. The current statistics show a tremendous increase in usage and areas of development, on the other hand experts claim that an even more sophisticated era is yet to come.

This thesis aims to investigate the accessibility and usability of IPAs where the main research question stands: do intelligent personal assistants ensure accessibility to the technology they're present for the blind and non-verbal people? Further, this question is divided into four sub questions. Initially, the first sub question is: what barriers do fully blind people experience when using intelligent personal assistants, and how to overcome these barriers? The second sub question is: what barriers do non-verbal people experience when using intelligent personal assistants, and how to overcome these barriers? The third sub question is: are intelligent personal assistants useful to fully blind people in noisy environments? The fourth sub question is: do the user interfaces of intelligent personal assistants facilitate accessibility and usability?

In this thesis, the universal design of IPAs is approached by using qualitative methods of research. Moreover, the methods comprise of: semi-structured interviews with five blind participants and observational tasks used as a supplement to the interviews, usability testing using a persona to represent the non-verbal user, and a systematic literature review. Both interviews accompanied by observations and persona testing are used to gather data about the barriers that these two groups of users face and their experiences. The data from the interviews is then analysed using the content analysis technique. The interview schedule consists of eight questions, and there are four tasks part of the observations.

Lastly, the systematic literature review is designed with specific search terms and criteria and its main objective is to examine the state of the art of IPAs, incorporating the new ideas and features including guidelines and design methods. A systematic approach will be used to identify relevant papers and studies through methodical search of academic online databases. Although few researches exist regarding the universal design of IPAs, the previous research is limited and has mainly studied other areas, such as the privacy concerns or public usage of IPAs. However, limited as it is, there are themes of the current research that focus on the deaf and hard of hearing and have studied and tackled the impairments and disabilities that occur from situations. Nonetheless, existing research also focuses on the enhancement of IPAs, possibilities of integrating IPAs into computing environments in an Internet of Things context, and factors that affect the everyday use as well as suggestions for improvements.

Most importantly, this thesis is one of a kind and focuses on two groups of people with disabilities by exploring their usage experiences through semi-structured interviews to uncover barriers as well suggest solutions to disable these barriers. And, having in mind the concept of UD with the goal of providing the same benefits and opportunities to the broadest possible scope of users.

Keywords: intelligent personal assistants, universal design, accessibility, usability, blind, non-verbal, Siri, information and communication technology, barriers.

2. Table of Contents

2.1. List of Acronyms

Acronyms	Full Name
UD	Universal Design
IPA	Intelligent Personal Assistant
ICT	Information and Communication Technology
eISB	Electronic Irish Statute Book
ISO	International Organization for Standardization
NDA	National Disability Authority
NIH	National Institutes of Health
NOU	Norges Offentlige Utredninger
NTIA	National Telecommunications & Information Administration
OECD	Organisation for Economic Cooperation and Development
WHO	World Health Organisation
UDLL	Universal Design for Learning – License to Learn
WAI	Web Accessibility Initiative
PALS	Personal Assistant for Online Services
SDS	Spoken Dialog Systems
NSD	Norsk Senter for Forskningsdata
AI	Artificial Intelligence
NLP	Natural Language Processing
AT	Assistive Technology
IoT	Internet of Things
VUI	Voice User Interface
ASR	Automatic Speech Recognition
CA	Conversational Agents

Table 2.1. List of Acronyms

2.2. List of Figures

Table 2.2. List of Figures

2.3. List of Tables

Table 2.3. List of Tables

3. Introduction

Since 2007, we have seen a drastic change in our lives because technology has usurped major aspects of our daily living. Technology has become essential in fields like: business, communication, information, education, medicine, and governments to name a few. Today, technology is an all-encompassing entity of our society. The particular field that affects most of us globally is technology in information and communication (ICT). The universal design (UD) of ICT is defined as “the design of products and environments to be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design" (Story, Mueller, & Mace, 1998).

Universal design values diversity and inclusiveness. That’s because different users interact with systems in diverse situations and contexts, using diverse technology. According to Story et al. (1998) there are some vital factors that determine whether an application or service will be used to its greatest extent, and those factors are: (1) the user’s abilities (motoric, auditory, visual, cognitive, mental health), (2) the diversity among users (culture, language, age, ICT skills, education level, standard of living), (3) the situations (temperature, noise conditions, outside/inside, many/few people around, type of place, and level of tension in the situation), (4) diverse technology (different technological platforms), and (5) the application or service’s design (the interfaces). Moreover, ISO (1998) refers to usability as “the extent to which a product can be used by specific users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use".

The ICT products and services should be designed in a way that users shouldn’t face barriers when using a system (NOU, 1999). Barriers are referred as conditions (physical or social) that limit a person to participate in the community (NOU, 1999). The Nordic model of disabling barriers, also known as the GAP model, describes the cause of disabilities as a consequence of a mismatch between the users’ abilities and the requirements of the environment (NOU, 2003). Universal Design helps in reducing this gap, by taking into account all the characteristics of people and their abilities while attempting to accommodate to the broadest possible scope of users. In failing to do so, and by limiting people into participating in the community we

unfortunately build on the digital divide. The digital divide is a social and economic inequality and refers to the gap between demographics and regions that have access to modern ICT and those who don't or have restricted access (NTIA, 1995). The digital divide in households primarily depends on income and education level. To overcome the digital divide and ensure access for citizens, businesses, and regions, effectively and efficiently, it is important that governments have information on the nature and extent of the digital divide and on the kinds of measures that can help to overcome it (OECD, 2001).

Since 2011, with the introduction of Apple's intelligent personal assistant (IPA) called Siri, major companies have invested in artificial intelligence (AI) technology for various missions such as improve business processes, offer a better customer service, and so on. Since then, IPAs have been implemented in many platforms, the machine learning is continuously improving, and IPAs are rising day by day. Nowadays, there are plenty of IPAs in the market but the most distinguished ones are: Apple's Siri (Apple, 2017), Google Voice Search/Google Now (Google, 2017), Amazon's Alexa (Amazon, 2017), Samsung's Bixby (Samsung, 2017) and Microsoft's Cortana (Microsoft, 2017). Not to mention that, Facebook has also released the text based assistant within Facebook Messenger called the M assistant which is still not available to every country (Simonite, 2017).

The current research about IPAs mainly focuses on various areas, such as: how people in general experience IPAs, overviews and statistics about the current usage, first impressions, privacy concerns, comparing the task completion time with voice versus touch, etc. However, there are articles that suggest the enhancement of IPAs or the possible areas of improvements, and books about the design principles that are presented in this thesis. Anyways, the current research done about the universal design of IPAs remains extremely limited, especially the research targeted for the blind and non-verbal people.

In a systematic review, Bourne et al. (2014) estimated that globally 32.4 million people (60 percent women) were blind in 2010, and 191 million people (57 percent women) had a moderate and severe vision impairment. There are several definitions on who is called blind and visually impaired. Various definitions have been adopted by national governments, international agencies, and blindness organisations to determine qualification. Visual impairment includes

both low vision and blindness. Low vision is the inability to see clearly, even with corrective lenses, at a distance of six meters what persons with normal vision can clearly see at a distance of eighteen meters. Blindness is the inability to read the largest letter on a vision chart at a distance of three meters (Organization, 1968; Thylefors, Negrel, Pararajasegaram, & Dadzie, 1995). Over ninety percent of all blind people live in developing countries (Resnikoff et al., 2004). In a news report by WHO (2017), was shown that an estimated 253 million people live with vision impairment of which 36 million are blind. According to Norway's association of the blind: (1) 180,000 Norwegians are regarded as visually impaired, (2) more than a thousand Norwegians are completely blind, (3) every tenth person older than 70 years is blind of age, (4) ninety percent of the whole blind and partially sighted in the world live in poor countries (Norges Blindeforbund, 2016). Blindness and vision impairment are very important global health issues because they have an impact on decreased quality of life, increased morbidity and mortality, and substantial economic productivity loss, as well as are of high importance for the socioeconomics and public health of countries (Bourne et al., 2014; Frick & Foster, 2003; D. J. Lee, Gómez-Marín, Lam, Zheng, & Caban, 2005; McCarty, Nanjan, & Taylor, 2001). As far as possible, ICT systems, products, and applications or services should offer universally designed solutions. More importantly, the blind shouldn't feel as excluded from society by not being able to access and use technology that is essential for them (Hansen, 2017).

This thesis aims to show if IPAs ensure accessibility to the technology they're present for the blind and non-verbal people. In addition, this thesis aims to discover the usability barriers of the blind users when using IPAs as well as suggest mechanisms to disable those barriers. On the other hand, aims to discover the usability barriers of the non-verbal as well as how to overcome them. Furthermore, this thesis aims to find if IPAs are useful to fully blind people in noisy environments. Lastly, this thesis aims to find out whether the user interfaces of IPAs affect accessibility and usability to facilitate the blind people. By having in mind, the fact that, there are standalone products that are only operable and usable by using voice input, while providing no screen and no other alternative input. In those cases, voice control/input is not used as an assistive technology (AT). Burzagli and Emiliani (2013) concluded that the need of the involvement of users not only during the requirement gathering phase but also in the design

and development of possible solutions is essential for solving accessibility problems.

This thesis has six main sections. Initially, the introduction section poses the research questions. The literature review and background section provide an overview of: (1) the concept of universal design, (2) history of development of IPAs technology, (3) history of speech to text technology, (4) the industry, society, and the social trend, (5) current state of the art of IPAs, and (6) where is technology going? Then, the method chapter encompasses the research design of the techniques used to gather data: the systematic literature review, persona testing, and interviews and observational tasks. Furthermore, it continues with the processes of the data collection techniques. Then, it follows up with an analysis sub section which shows in what these processes resulted to. Lastly, the method section ends with the ethical considerations of the study.

The results section shows the results from the systematic literature review, the results of the persona testing, and the results from the qualitative studies – interviews and observations. Afterwards, the findings of this thesis are interpreted through an analysis and interpretation section which also consists of suggested possible solutions. Next is the discussion section which interprets the challenges of this study, and the thesis ends with a conclusion section about future research and practice.

3.1. Research Questions

Do intelligent personal assistants ensure accessibility and usability to the technology they're present for the blind and non-verbal people?

Of the main question, there are four sub questions:

1. What barriers do fully blind people experience when using intelligent personal assistants, and how to disable those barriers?
2. What barriers do non-verbal people experience when using intelligent personal assistants, and how to disable those barriers?
3. Are intelligent personal assistants useful to fully blind people in noisy environments?
4. Do the user interfaces of intelligent personal assistants facilitate accessibility and usability?

4. Literature Review and Background

There are various terms encountered within literature that are used interchangeably with the term IPAs, such as: voice assistants, virtual assistants, voice digital assistants, conversational assistants/agents, to name a few. Basically, they all refer to the same thing broadly known as voice assistants and academically referred to as IPAs, who are a part of the AI field. Intelligent personal assistants are software agents that can perform numerous tasks or services for an individual. In the field of computer science, a software agent is a computer program able to act on user's or another applications behalf (Nwana, 1996). In addition, Nwana (1996) describes that the term agent originally derives from the Latin word '*agere'* which means *to do*, and indicates an agreement to act on user's behalf (an action for which the user authorizes the decision if the action is appropriate).

Software agents are informally known as bots (from robot). Software agents can be embodied as a software like a chatbot that executes on a smartphone (like Siri). A chatbot is a computer program that conducts a communication via audio or text methods with humans (Mauldin, 1994). Today, chatbots are a part of IPAs. Intelligent personal assistants interact with users by using natural language processing (NLP) to match the input to executable commands. Natural language processing is a field concerned with the interaction between computers and humans particularly by programming computers to effectively process the human's natural language (Jurafsky & James, 2000).

4.1. The Concept of Universal Design

Universal design is defined by eISB (2005) as "the composition of an environment so that it may be accessed, understood, and used: (1) to the greatest practicable extent, (2) in the most independent and natural manner possible, (3) in the widest possible range of situations, and (4) without the need for adaptation, modification, assistive devices, or specialized solutions, by persons of any age or size or having a particular physical, sensory, mental health or intellectual ability or disability, and; Means, in relation to electronic systems, any electronic-based process of creating products, services or systems so that they may be used by any person". This definition is very important to understand the goals of this thesis because it particularly mentions two very important aspects of universal design which I had in mind about IPAs which

are: the natural manner, the situations, and to be used regardless of disabilities.

According to Story et al. (1998), there are seven principles of universal design and each principle has their corresponding guidelines, the principles are: (1) equitable use, (2) flexibility in use, (3) simple and intuitive use, (4) perceptible information, (5) tolerance for error, (6) low physical effort, (7) size and space for approach and use. The first principle has four guidelines which are: (1) provide the same means of use for all users: identical whenever possible, equivalent when not, (2) avoid segregating or stigmatizing any users, (3) provisions for privacy, security, and safety should be equally to all users, (4) make the design appealing to all users. Intelligent personal assistants should not segregate the blind or the non-verbal users and should be equal to all. The second principle's guidelines are: (1) provide choice in methods of use, (2) accommodate right- or left-handed access and use, (3) facilitate the user's accuracy and precision, (4) provide adaptability to the user's pace. The guidelines of the third principle are: (1) eliminate unnecessary complexity, (2) be consistent with user expectations and intuition, (3) accommodate a wide range of literacy and language skills, (4) arrange information consistent with its importance, and (5) provide effective prompting and feedback during and after task completion. Intelligent personal assistants may be the easiest to fulfil the guidelines of the second and third principles since they are flexible and simple to use. However, this thesis focuses to find if IPAs comply with particular guidelines such as: the fourth guideline of the second principle or the fifth guideline of the third principle. The guidelines of the fourth principle are: (1) use different modes (pictorial, verbal, tactile) for redundant presentation of essential information, (2) provide adequate contrast between essential information and its surroundings, (3) maximize "legibility" of essential information, (4) differentiate elements in ways that can be described, (5) provide compatibility with a variety of techniques or devices used by people with sensory limitations. It is important that the information that the IPAs present to the blind users is perceptible. The guidelines of the fifth principle are: (1) arrange elements to minimize hazards and errors: most used elements, most accessible; hazardous elements eliminated, isolated, or shielded, (2) provide warnings of hazards and errors, (3) provide fail safe features, (4) discourage unconscious action in tasks that require vigilance. The guidelines of the sixth principle are: (1) allow users to maintain a neutral body position, (2) use

reasonable operating forces, (3) minimize repetitive actions, (4) minimize sustained physical effort. Lastly, the guidelines of the seventh principle are: (1) provide a clear line of sight to important elements for any seated or standing user, (2) make reach to all components comfortable for any seated or standing user, (3) accommodate variations in hand and grip size, (4) provide adequate space for the use of assistive devices or personal assistance. For intelligent personal assistants, these three last principles might not be so relevant. However, the qualitative studies of this thesis, aim to uncover any uncertainties.

The traditional approach of UD lies in the inclusion in ICT environment (e-inclusion) where there is a problem of accessibility to systems and services (Burzagli & Emiliani, 2013). In addition, these accessibility problems are solved with adaptations where AT is used to allow access to ICT systems and services. According to ISO (2011), "an assistive product is any product (devices, equipment, instruments, technology, and software) especially produced or generally available for preventing, compensating, monitoring, relieving, or neutralizing impairments, activity limitations and participation restrictions". Assistive technology is technology utilized to give people with activity limitations abilities they do not have.

The term adaptation in UD refers to special solutions that are necessary when UD is not enough or when UD cannot be conducted (Story et al., 1998). Universal design is more than accessibility, it requires that persons with disabilities are taken into account when designing the main solution for products and the environment (NOU, 2005). Tollefsen (2011) states that UD when developing ICT systems and services is: (1) a goal – usable by all people, to the greatest extent possible without the need for adaptation or specialized design and efficient in use, (2) a tool – W3C/WAI Guidelines, heuristic analysis, UCD design methods, and (3) a process – needs analysis, design, prototyping, user testing, user experiments, implementation, and documentation. A universally designed product is accessible and usable to as many people as possible, and the best way to ensure this is to involve users when finding accessibility barriers.

4.1.1. Accessibility

According to WAI (2005), accessibility means that people with disabilities can use the Web. Moreover, it means that people with disabilities can understand, perceive, navigate, and interact with the Web, and that they can contribute to the Web. Web accessibility also benefits

others including older people with changing abilities due to aging. Web accessibility encompasses all disabilities that affect access to the Web: auditory, visual, speech, physical, cognitive, and neurological disabilities. Accessibility is also what makes products usable in a wide range of situations, contexts, and environments (Henry, 2007).

The web is an increasingly important resource in many aspects of life: education, employment, government, commerce, health care, and more. It is important that the Web is accessible so as to provide equal access and equal opportunity to people with disabilities (Henry, 2006).

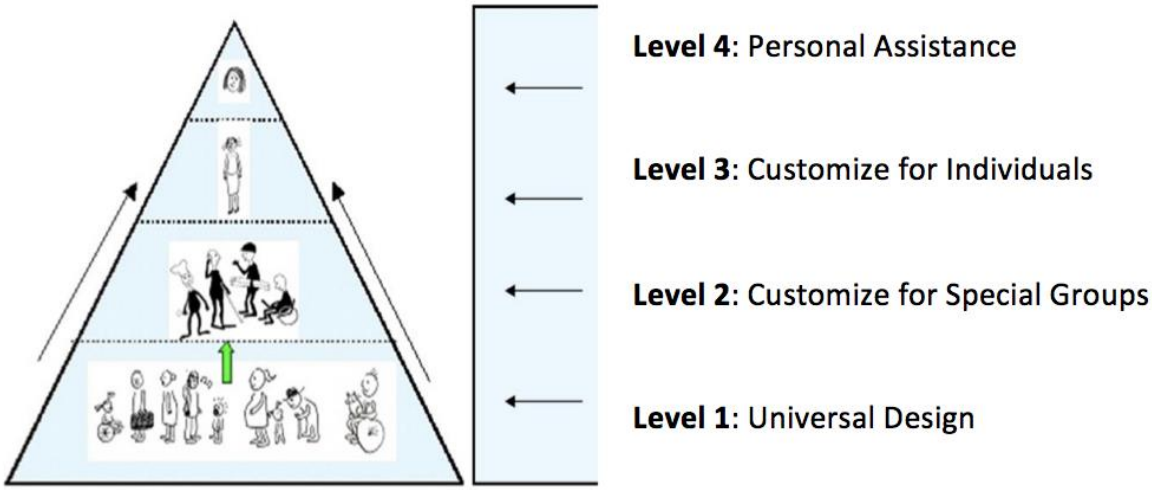


Figure 4.1.1. The Accessibility Pyramid (UDLL, 2016).

The figure above shows according to UDLL (2016), the levels of accessibility and how the market can be separated into four segments. Universal design being on the level one of the pyramid (bottom up approach) targeting all groups of people. Therefore, it is important to conduct a thorough accessibility evaluation to assess the accessibility of a software.

4.1.2. Usability

Usability is defined as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use" (ISO, 1998). According to Nielsen (2012), usability is a quality attribute that assesses how easy user interfaces are to use, and the word "usability" also refers to methods for improving ease-of-use during the design process. Moreover, Nielsen (2012) states that usability is defined by five quality components: (1) learnability – how easy is for users to accomplish tasks the first

time the encounter the design? (2) efficiency – once users have learnt the design, how quickly can they perform tasks? (3) memorability – when users return to the design after a period of not using it, how easy can they re-establish proficiency? (4) errors – how many errors do users make, how severe are those errors, and how easy can they recover from the errors? (5) satisfaction – how pleasant is it to use the design? Among many other quality attributes, utility is a key one and refers to the design's functionality: does it do what users need? Utility and usability are equally important and together determine whether something is useful (Nielsen, 2012). The definition of usability by ISO inspired me, is very important, and it is one of the many motives of this thesis because it emphasizes the specified users which in this case are the blind and non-verbal and the extent to which they can achieve the specified goals of a product. Likewise, by having in mind utility the interview questions and tasks were designed to find out does the system do what users need.

4.1.2.1. Usability Testing

Usability testing, known as "user research", deals with the research of the interface rather than researching the user with the purpose of making a specific interface better (Lazar, Feng, & Hochheiser, 2017). Moreover, usability testing is a broad term and includes elements of design and development such as personas, user profiles, card sorting, and competitive research that generally might not be considered "research" by those who consider themselves researchers (Kuniavsky, 2003).

Usability testing generally involves representative users attempting representative tasks in representative environments, on early prototypes or working versions of computer interfaces (Lewis, 2006). Usability testing is often used to test software that has already been implemented in existing systems (Lazar et al., 2017). In addition, Lazar et al. (2017) states that the interfaces being usability tested are typically screen layouts, lap-top, or tablet computers, as well as smart phones and other mobile devices. Except that the usability testing approach has one goal which is improving the quality of an interface by finding flaws of the interface (causing problems for the users) that need improvement, at the same time it discovers what is working well with an interface and makes sure to keep those features.

There are three types of usability testing: (1) expert-based testing, (2) automated

testing, and (3) user-based testing (Lazar et al., 2017). This thesis focuses on user-based testing, which is what most people mean when referring to usability testing, which particularly means the group of representative users that attempt a set of representative tasks.

4.1.3. Diversity

In the field of Universal Design of ICT, the term diversity includes: (1) diversity in user's abilities, (2) diversity among users, (3) diverse situations, and (4) diverse technology (Story et al., 1998). As mentioned earlier, UD values diversity. That is because different users interact with systems in diverse situations and contexts, using diverse technology. The diversity in user's abilities consists of: (1) motoric, auditory, and visual, (2) sensory – see, touch, taste, smell, hear, balance, (3) mental health – anxiety, mood, psychotic, and eating disorders, impulse control and addiction disorders, personalities disorders, obsessive-compulsive disorders, and post-traumatic stress disorders, (4) cognitive – think, concentrate, formulate ideas, reason and remember, read and write. The diversity among users includes: culture, language, age, ICT skills, education level, and standard of living. Meanwhile, the diverse situations include: temperature, noise conditions, outside/inside, many/few people around, type of place, and level of tension/stress in the situation. Lastly, the diversity in technology could be device or software related. Diversity in devices includes: PCs, laptops, tablets, smartphones, smart TVs, ATMs, payment terminals, ticket machines, queuing systems, internet kiosks, information boards, and real time transportation information. On the other hand, diversity in software includes: applications, operating systems (OS), browsers, assistive technology (AT), and new/old versions and devices.

This is what I had in mind when I decided to include in usability testing as many devices as possible. Also, having in mind this definition I found it important to include the blind and non-verbal when thinking about universally designed solutions.

4.1.4. ICT Barriers

According to NOU (1999), barriers are referred as conditions (physical or social) that limit a person to participate in the community. Marks (1997) states that there are the medical model and the social model of disability that are used to disable barriers. Some examples of barriers can be: **s**tereotypes and prejudice, organizational procedures and practices that are not flexible, inaccessible information, inaccessible buildings, inaccessible transport.

The medical model of disability views disability as a problem prone to the disabled individual, being an issue concerning not anyone else except the individual affected. In contrast, the idea of the social model of disability is that society disables people, by designing everything so as to meet the needs of the majority of people who are not disabled (Shakespeare, 2006). Shakespeare (2006) defines disability as the loss or limitation of opportunities to take part in the society equally with other people because of social and environmental barriers. Meanwhile, an impairment is defined as injury, illness, or a condition present from birth that causes or is likely to cause a long-term effect on the physical appearance and/or limitation of function within the individual that differs from others.

4.1.5. The Disability GAP Model

Except the medical and social models of disability, there is also the relational model of disability or the disability GAP model. The GAP model is defined by NOU (1999) as, "a "disability" occurs when people's practical lives are significantly limited because of a gap or mismatch between the person's ability and the demands of the environment or society".

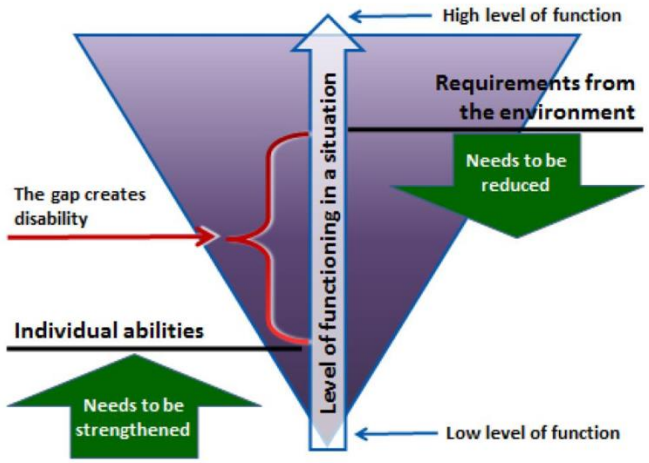


Figure 4.1.7. The Disability GAP Model (Fuglerud, 2014).

The picture above shows that, in order to reduce the gap that creates disability, the individual abilities have to be strengthened and on the other hand the requirements from the environment need to be reduced. That way, it results in a high level of functioning in a particular situation. The disability GAP model is one of the most important models of disabling barriers

and its worth mentioning that the research questions were designed by having this model in mind.

4.2. History of Development of Intelligent Personal Assistant Technology

The first tool developed that was able to perform digital speech recognition was IBM's Shoebox, presented to the public in 1962, after its launch in the market in 1961, and could only understand sixteen words in English and nine digits (IBM, 2012). Since then, IPAs have undergone more than five decades of development. The first two decades were like pushing back the arrow that was going to dominate today's lifestyle once shot.

4.2.1. The Foundation of Intelligent Personal Assistants

In the first two decades was built the foundation of today's IPAs. That includes the introduction of Harpy in 1970, which possessed a more efficient search strategy called 'beam search' and was more accurate. Harpy could understand 1,000 words and set a milestone in the development of IPAs (Waibel & Lee, 1990). In 1980, because of a new statistical method known as the hidden Markov model (HMM), voice recognition was able to predict the input by considering unknown sounds as being words, rather than using templates of words and looking for sound patterns (Juang & Rabiner, 2005). As late as 1990, another significant development was the release of Dragon Dictate. It was the first speech recognition product for consumers and it costed $6,000 (Mutchler, 2017).

In 1997, Dragon NaturallySpeaking arrived which recognized 100 words per minute by speaking naturally (Thompson, 2012). Through the years from 1997 to 2004, Microsoft introduced the Office assistant called Clippit which was included in the Microsoft Office (Cain, 2017). Before Clippy, in 1995, Bill Gates personally launched Microsoft Bob which was supposed to make the software for first-time users more friendly (Cain, 2017).

4.2.2. Intelligent Personal Assistants in the 21st Century

By 2001, speech recognition was found embedded in operating systems like Windows Vista and OS X or Office XP. The following year, Google released the application Google Voice Search, and used its servers of data to better predict what is being said (Pinola, 2011). In 2010, Google personalized the recognition of Voce Search app to better record users and produce a more accurate speech model. In February 2011, IBM's Watson won jeopardy. Watson is a

computer system that is capable of answering questions posed in natural language and was developed in one of IBM's projects called DeepQA (High, 2012). Today, Watson is AI platform for professionals and it can accelerate research and discovery and enrich interactions.

As late as 2011, IPAs reached another significant level in development. That was with the announcement of Apple's Siri that same year (Pinola, 2011). Originally, Siri was an application in the App Store in February 2010 and was acquired by Apple two months later. Then, it was integrated in the iPhone 4S which was released in October 2011. Today, Siri is an integral part of all the Apple products including the iPad, Apple Watch, MacBook, iPod Touch, HomePod, and Apple TV. In July 2012, Google launched Google Now first integrated in Android 4.1 Jellybean, which was a feature of Google Search, and it was first supported on the Galaxy Nexus smartphone.

In April 2013, Microsoft introduced Cortana at the Annual BUILD developer conference (Chris, 2014). In January 2015, it became available for Windows 10 and Windows 10 mobile. In December 2016, Cortana was added to mobile platforms: Windows Phone 8.1, the Invoke smart speaker, Microsoft Band, iOS and Android. In August 2016, Cortana was added to Xbox One. Today, Cortana can set reminders, recognize natural communication without keyboard input, and answer questions using information from Bing search engine.

In November 2014, Amazon introduced Alexa first used in Amazon's Echo speaker. Alexa is capable of voice interaction, playing music, set alarms, provide real time information like traffic, news or the weather. Alexa can control several devices by using itself as a home automation system. It is also possible to extend the capabilities of Alexa by installing skills (third party applications) which were introduced in June 2015. Amazon Echo officially launched in June 2015.

In March 2016 SoundHound launched HOUND, a virtual assistant application. Later that month, Amazon launched Amazon Echo Dot and Amazon Tap. In May the same year, Google introduced the Google Assistant as a part of the messaging app Allo. Unlike Google Now, the Google Assistant can engage in two-way communications and it was primarily available for mobile and smart home devices. In October 2016, Samsung acquired the virtual assistant Viv which was created from the developers of Siri (Kastrenakes, 2016).

In March 2017, Samsung introduced Bixby alongside Galaxy S8 during the Samsung Galaxy Unpacked 2017 event held on March 29, 2017. Actually, Bixby was introduced one week before the event but it made its first appearance on the event. In May 2017, Samsung announced that Bixby will come to its line of Family Hub 2.0 refrigerators and it will be the first non-mobile product to have an assistant. Bixby comes in three parts: Bixby Voice, Bixby Vision, and Bixby Home (Dolcourt, 2017). Bixby Voice can be activated by long pressing the Bixby button on smartphones. The Bixby Vision is built in the camera application, essentially an augmented reality camera that can identify objects in real time, search for the objects on different services, and offer the user to purchase them. Bixby Home can be activated by swiping to the right on the home screen of the smartphone and it consists of a list of information that Bixby can interact with such as weather and fitness activity.

In April 2017, Google introduced multi-user support for Google Home which could recognize six different voices. In April 2017, Amazon introduced the Echo Look, a camera with Alexa built-in. The device could make recommendations for outfit, take photos, or record videos. Additionally, it could take 360-degree photos or videos with built-in AI for fashion advice. It helps the consumers to create a catalogue for their outfits and rates the look based on machine learning algorithms with advice from fashion specialists. In May 2017, Baidu unveiled the AI device Xiaoyu in China. Days later, Amazon introduced new features for the Echo device – calling and messaging. Weeks later, Google announced that by summer 2017 Google Home will be available in Germany, France, Japan, Canada, and Australia. In June 2017, Apple introduced the HomePod which was going to ship as late as December 2017. In June 2017, Google Home launched in Canada. In May 2017, Amazon launched its speaker with a screen called Echo Show in the United States (Haslam, 2017a). In July 2017, Alibaba launched Genie X1 smart speaker (Chen, 2017).

4.3. History of Speech to Text Technology

Speech is the primary means of communication between people. It is worth mentioning that speech recognition is a sub-field in computational linguistics that develops methods and technologies in order to enable the recognition and translation of spoken language to text. The past five decades, the research in automatic speech recognition has attracted a lot of attention.

4.3.1. Automatic Speech Recognizers in Early Stages

To automate simple tasks is not a modern phenomenon, but it is one that goes back hundred years. In 1881, Alexander Graham Bell, Chichester Bell, and Charles Sumner Tainter invented a recording device designed to respond to incoming sound pressure (Juang & Rabiner, 2005). Based on this invention, Bell and Tainter created the Volta Graphophone Co. in 1888 to manufacture machines that could record and reproduce the sound in office environments. In 1907, the American Graphophone Co. acquired the patent and trademarked the term "Dictaphone". Later, it became the Columbian Graphophone Co. Approximately in that same time, Thomas Edison invented another phonograph and developed the "Ediphone" to compete with Columbia (Juang & Rabiner, 2005).

In 1773, Christian Kratzenstein, a Russian scientist and professor of physiology in Copenhagen succeeded to produce vowel sounds (Kratzenstein, 1782). Later in Vienna, Wolfgang von Kempelen constructed the "Acoustic-Mechanical Speech Machine" (Juang & Rabiner, 2005). In the mid of 18th century, Charles Wheatstone built a similar version like Kempelen's machine which could be controlled with a hand (Wheatstone, 1879).

In the first half of the 20th century, Fletcher with some others at Bell Laboratories documented how the power of a speech sound is distributed across frequency as well as the sound's characteristics and intelligibility, as perceived by a human listener (Fletcher, 1922). In 1930's, Homer Dudley influenced by Fletcher's achievements, developed a speech synthesizer machine named VODER (Voice Operating Demonstrator) which could synthesize a sentence through an operator who had to learn how to control it in order to produce the appropriate sounds of the sentence (Dudley, Riesz, & Watkins, 1939). The VODER was demonstrated in New York in 1939 at the World Fair and was considered to be a very important milestone in the evolution of speaking machines.

In 1952, Biddulph, Davis, and Balashek of Bell Laboratories built a system of digit recognition for a single speaker (Davis, Biddulph, & Balashek, 1952). During the 1960's, several other laboratories in Japan demonstrated the capability to build hardware with special purposes of performing a speech recognition task. Particularly, it was the vowel recognizer by Suzuki and Nakata at the Radio Research Lab in Tokyo (Suzuki, 1961). In another recognition system by Fry

and Denes at an University in England, their recognizer system could recognize four vowels and nine consonants by using statistical information which increased the recognition accuracy of words consisting of two or more phonemes, and marked the first use of statistical syntax in speech recognition (Denes, 1959).

4.3.2. Technology in 1980's and 1990's

The research in speech recognition in the 1980's was distinguished by a shift in the methodology from template based intuitive approach to a statistical and more rigorous approach. The Hidden Markov Model (HMM) mentioned in section 4.2, was early on known and understood in some laboratories like IBM the Institute for Defense Analyses (IDA), but it widespread after its publication of the theory and became the preferred method for speech recognition (Ferguson, 1980). Its popularity and the use of this model as the main foundation for speech recognition systems has remained the same these last two decades. This model is an essential component of a speech recognition system that follows the statistical pattern recognition approach and has its root in Bayes' decision theory (Theodoridis, 2003).

In the 1990's, a tool was developed named the FSM (finite-state machine) library that embodied the finite state network approach in a unified framework and since then has been a major component of almost all modern speech recognition systems (Mohri, 1997). Except the FSM tool, during these years, a lot of progress was made in the development of software tools and systems became more and more sophisticated. Other speech recognition systems include: (1) the Sphinx system (K.-F. Lee, 1988), (2) the BYBLOS system (Schwartz et al., 1989), and (3) the DECIPHER system (K.-F. Lee, 1988). The Sphinx system integrated the Hidden Markov Model and the strength of the Harpy system. The most distinguished and widely adopted software was the system made available by the Cambridge University team led by Steve Young and was named the Hidden Markov Model Tool Kit (HTK) as well as remains today the ground software tool for automatic speech recognition research (S. Young et al., 2002). In the picture below can be seen the milestones in research of speech and multimodal technology.

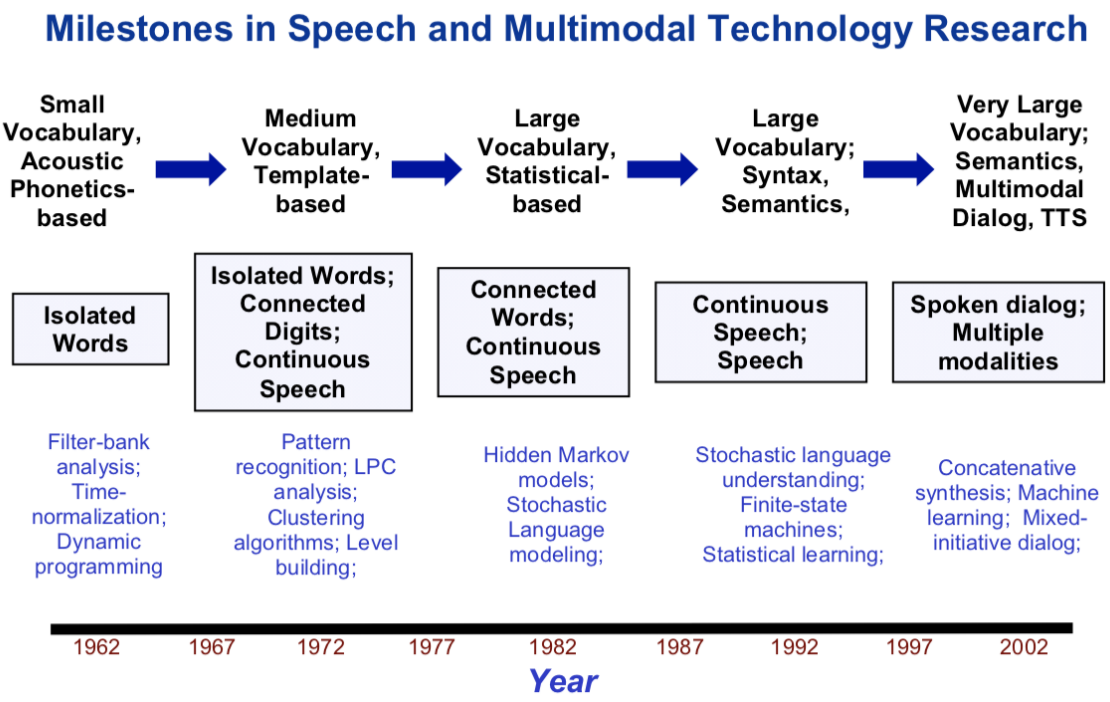


Figure 4.3. History of Speech Recognition Technology (Juang & Rabiner, 2005)

4.4. Industry, Society, and the Social Trend

According to Bergen and Moritz (2017), 500,000 Google Home units were shipped in 2016. Meanwhile Osborne (2016) shows that, Cortana has 133 million monthly users. It is estimated that Amazon sold approximately 2 million units in the first nine months of 2016 (Lewin & Lowitz, 2016). Marchick (2017) concluded that, eleven percent of the respondents who already own an Amazon Alexa or Google Home device will also buy a device from their competitors, on the other hand the growth of Amazon Alexa has been over five hundred percent in the second half of 2016.

In a survey conducted by Miller (2017), the results indicated that almost sixty-three percent of the respondents would want a more intelligent Siri. Phil Schiller (Apple's SVP of Worldwide Marketing) expressed his opinions regarding IPAs and suggested that they are better with a screen (Haslam, 2017b). The research report by J. Smith (2017) from Business Insider's Intelligence shows the consumer usage and interest in using IPAs in the following figure:

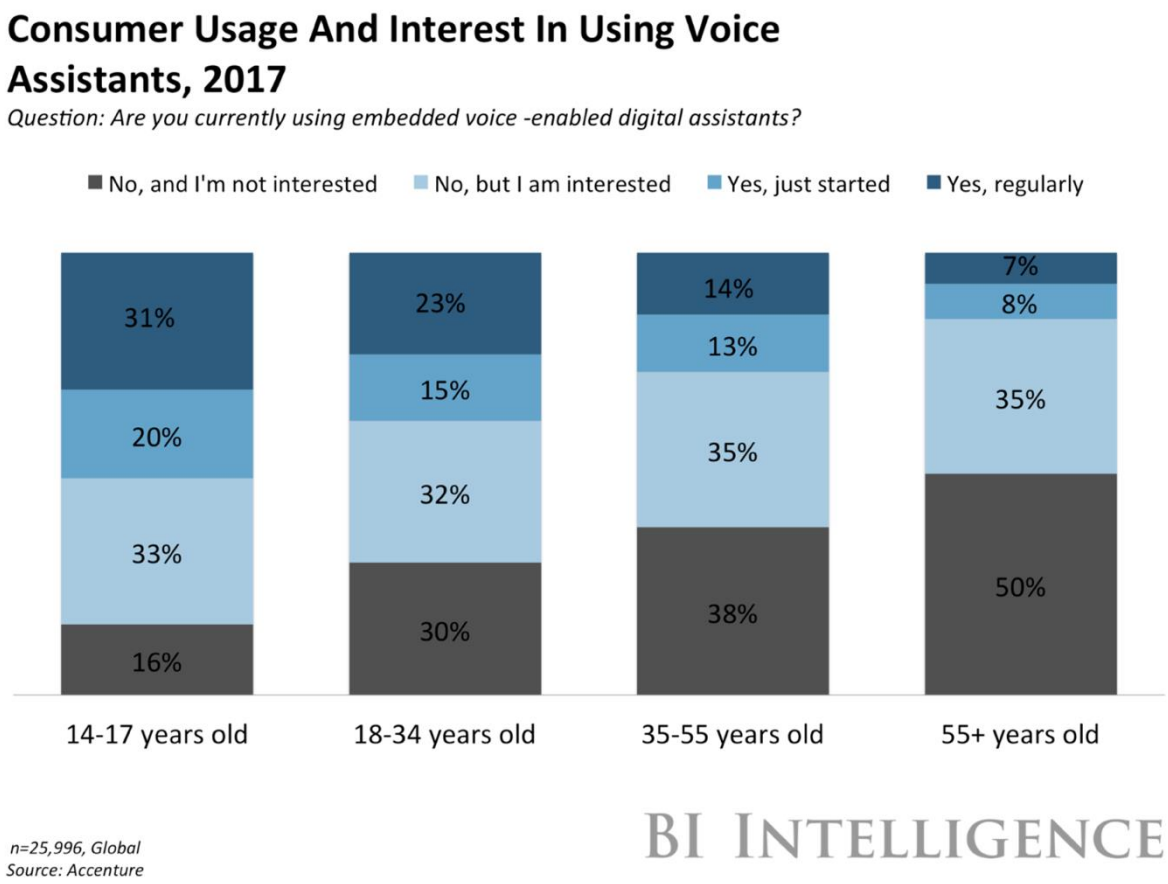


Figure 4.4. Consumer Usage and Interest in Using IPAs in 2017 (J. Smith, 2017).

Usage statistics according to Pilon (2016) show that, nine percent of users said they have used IPAs like Siri or Cortana in the past day, and forty-five percent of those using IPAs are users of Siri, thirty-three percent have used Google Now, twenty-seven percent used Cortana, ten percent have used Amazon Echo or Alexa. Meanwhile Yurieff (2017) presented some of the main reasons why people prefer voice search over typing and those were: (1) people can speak 150 words per minute versus type forty words per minute, (2) twenty-eight percent think voice is more accurate than typing, (3) forty-two percent say that using voice while driving is a very strong reason to, (4) twenty-one percent say they don't like typing on the phone and use voice instead.

According to Sterling (2016), the statistics of searches made using voice show that one in every five searches made with Google Android app in the U.S is a voice search, and in the future companies will have to optimize their voice search queries to fit long queries. Moreover, Olson

(2016) concluded that forty percent of adults use voice search at least once per day. Meanwhile, Huffman (2014) shows that, more than half of the teens and forty-one percent of adults in U.S use voice search on a daily basis, and its use is continuing to grow every day. In addition, K. Young (2016) did find out that twenty-five percent of people aged 16-24 use voice search on mobile. On the other hand, Sentance (2016) found out that voice searches on Google in 2016 are up thirty-five times over than in 2008.

4.5. Current State of The Art

The interaction between users and IPAs is done in three different ways. The current ways include three forms of input: (1) text – such as in messaging or email applications, (2) voice – which is the usual interaction, and (3) images – by uploading photos or by using the camera for real time interaction. Nowadays, IPAs are integrated in numerous platforms, such as: objects (smart speakers), in instant messaging apps (Facebook M), built into mobile OS (Siri), independent from the smartphone's OS (Bixby), on smartwatches (Apple Watch), appliances (cars or refrigerators).

Fluent.ai (2018a) brings voice interface software products with their goal being to disable the barriers that lead to high adoption of voice user interfaces. Furthermore, on the 8th of March 2018 Fluent.ai presented in the Mobile World Congress in Barcelona the Universal Speech Recognition (Fluent.ai, 2018b). What is unique about FluentAI Core is that it aims to universalize voice user interfaces. It means that it seeks to solve the problem of accents by offering an alternative of Alexa and Google Assistant which allows everyone to be understood. They have created a voice interface that delivers high accuracy in recognition in any language or accent or mix of languages, even in noisy environments.

4.5.1. Intelligent Personal Assistants Accessibility

Since the most famous and most used IPAs are Siri, Google Now, Alexa, and Bixby, this section shows the current state of IPAs regarding accessibility. Siri is available in all Apple's devices, including: iPhone, iPad, iPod, MacBook, iMac, Apple Watch, Apple TV, and HomePod. Except the basic features, Siri has the following features: (1) maps – standard, satellite, directions, turn-by-turn navigation, traffic, lane guidance, speed limits, transit, flyover, nearby, indoor maps airports, indoor maps malls, (2) Siri – sports, Twitter integration, restaurants

(information, reviews, reservations), movies (information, reviews, showtimes), dictionary, calculations and conversions, (3) QuickType keyboard – language support, predictive input, dictation, multilingual typing, contextual predictions, emoji predictions (Apple, 2018a). Each of these features is only available in selected countries, particular airports, and malls.

Google's has two assistants, Google Now and Google Assistant. Google Now works within iOS and Android with the Google Search application and is not specified to a platform. Also, it doesn't get better and cannot get to know the user. The hardware controlling capabilities are limited in iOS. Meanwhile, Google Assistant is integrated in Google Home and Google's Pixel phones. The Google Assistant is an enhancement of Google Now, it has the same capabilities plus more features and an interface and it provides information more like in the form of a conversation. The Google Assistant works with more than 1,500 smart home devices and more than 200 popular brands (Google, 2018). The Google Assistant is available in: phones (Google Pixel 2), speakers (Google Home), wearables (Guess Connect), laptops (Google Pixelbook), TVs (Android TV), cars (Android Auto), smart displays (Lenovo Smart Display), and other (Nest Cam IQ Indoor) (Google, 2018).

Some of Alexa's supported devices are: smart speakers (Echo), TVs (Amazon Fire TV), phones and tablets (Amazon Fire Tablet, Huawei Mate 9), laptops and desktop computers (Asus ZenBook, HP Pavilion Wave), smart home (Amazon Dash Wand, LG InstaView smart refrigerator), wearables (Pebble Core), and automotive (all 2018 models of BMW or Toyota). Similar to other assistants, Alexa can answer questions or play music. Also, it has a wide range of skills that can be installed which are produced by third party developers or it is even possible for the users to create their own skills with a new skill called Blueprints (Haslam, 2018). More importantly, since Amazon is mostly focused on Smart Speakers, Alexa can control a wide range of smart home devices, such as: switches, lights, TVs, security cameras, thermostats, door locks, iRobot's vacuum cleaner, ceiling fans, etc. On the other hand, Bixby can be found only in Samsung's Galaxy range of smartphones like the Galaxy S9 and the Galaxy Note tablets. The table below shows an overview of the current state of IPAs regarding accessibility, including: activation methods, supported languages, voice characteristics, and other distinguishing features.

Activation Methods	Languages	Voice	Other Features
Siri			
Press and Hold Home Button, or Press and Hold the Power Button in iPhone X	20 Languages: Arabic, Chinese, Danish, Dutch, English, Finnish, French, German, Hebrew, Italian, Japanese, Korean, Malay, Norwegian, Portuguese, Russian, Spanish, Swedish, Thai, Turkish.	Gender: Male, Female.	Maps, Apple CarPlay, Apple TV App, Sports, Twitter integration, Apple Pay, QuickType Keyboard, News, iTunes Store, Apple Music, App Store, Spotlight Suggestions.
"Hey Siri"		Accent: English in American, Australian, and British.	Add owners information.
Hold Down the Microphone Button in Apple's Headphones		Voice Feedback: Always On, Control with Ring Switch, and Hands-Free Only	Notes: Searchable Handwriting.
Raise to Speak (Older Generations)			Incoming Call Announcements
Double Tap the Airpods			Type to Siri
CarPlay: Voice Command Button on the Steering Wheel			Access while iPhone is locked.
Google Now/Google Assistant			

"OK Google" command	8 Languages: English, French, Italian, Japanese, Korean, Spanish, Brazilian Portuguese.	Gender: Male, Female.	Now on Tap
Press and hold the Home Button in Pixel phone.		Accent: English in American, British, and Australian.	My Day briefing
"Hey Google" command			Routines
Open Google Assistant App in iOS or Android			
Squeeze the Screen in Pixel 2			
Alexa			
"Alexa" default wake word in Echo (can be changed in user preferences).	3 Languages: English, German, and Japanese.	Gender: Male, Female.	Can install "Skills" from Alexa Skills Kit.
Activate Echo speaker using Alexa app or web interface.	Available in 38 Countries.	Accents: English (Indian, Australian, British, Canadian),	Can be paired with a great variety of devices.
Activate using Tap			
Activate using a button in smartphones.			
Bixby			

Long press the Bixby button on Smartphones to activate Bixby Voice.	2 Languages: English and Korean.	Gender: Male, Female.	Bixby Vision
Swipe to the right on the home screen to activate Bixby Home.	Available in more than 200 countries.	Accent: American English only.	
"Hi Bixby"			

Table 4.5.1. Intelligent Personal Assistants Accessibility

4.6. Where is Technology Going?

Amazon has launched its Alexa for business initiative. It is designed to provide a framework for Alexa to enter the workplace and not just the home. Aimed to add value to the corporate environments. "Alexa in Business is the next big step in making voice UI ubiquitous. Using Alexa at work adds to the continuum of experience, which now extends from the home, to the car, to the workplace. Further, the ability to create custom private skills, which live in specific organizations, will springboard a new market for Alexa development moving forward" (Metrock, 2017).

Gartner predicted that by 2019, twenty percent of all user interaction with the smartphone will be done by using IPAs (Van Der Meulen, 2016). Moreover, Olson (2016) states that, by 2020 half of all searches will be voice searches. According to Forni (2016), the two top predictions for users and IT organizations are: (1) by 2020, one hundred million consumers will shop in augmented reality and 30 percent of web browsing sessions will be done without a screen, (2) by 2021, twenty percent of all activities an individual engages in will involve at least one of the top-seven IPAs.

Overmyer (2016) in a report shows that, by 2019 the voice recognition market will be a $601 million industry. Moreover, based on a new study by Voice Labs, by the end of 2017 will be shipped twenty-five million devices and the total number of voice-first devices will reach to thirty-three million in circulation (Marchick, 2017). According to Gurman and Stone (2018), Amazon is working on a robot for homes powered by Alexa. Furthermore, Britt (2018) presents

the seven key predictions for IPAs: (1) personalized responses with contextual understanding, (2) individualized experiences, (3) development across multiple platforms and channels, (4) changes in search behaviours, (5) voice notifications, (6) touch interactions, (7) focus on security. Overall, seeing the industry of IPAs and their integration, usage statistics and predictions, the most important thing is that they are accessible and usable by all the people.

5. Method

In this thesis, I used the qualitative approach to answer the research questions. Qualitative research is concerned with the meanings people attach to their things in lives, it is used to understand people from their own frames of reference and experience reality as they do (Taylor, Bogdan, & DeVault, 2015). This approach is inductive, and the researcher develops concepts, insights, and understandings from patterns in the data (Taylor et al., 2015). Furthermore, the researcher looks at people and settings holistically in which case people, groups, or settings are viewed as a whole (Taylor et al., 2015). The researcher is concerned with how people think and act in their everyday lives. In qualitative interviewing, it's possible for the researcher to model the interview like a normal conversation rather than a question-and-answer exchange (Taylor et al., 2015). In qualitative research, all perspectives are worthy of study and the assumption that the perspectives of powerful people are more valid than those of the powerless is rejected (Taylor et al., 2015). In this kind of research, the researcher emphasizes the meaningfulness of their research and the researcher has something to learn in all settings and groups (Taylor et al., 2015). The researchers are flexible in how they conduct the studies.

Berg and Lune (2004) states that the qualitative approach is usually done in interviews, open-ended questions, or focus groups. In most of the cases, the number of participants in this kind of research is small in number and the research requires a lot of resources and a lot of time. Interviews and focus groups can be used for both formative and summative purposes (Lazar et al., 2017, p. 187).

According to Lazar et al. (2017), the three areas that researchers need to pay careful attention to when people with disabilities are involved are: (1) participants, (2) research methodology, (3) logistics. Also, Lazar et al. (2017, p. 495) emphasizes the fact that "the research that leads to improved interface and design experiences for people with disabilities may eventually lead to interfaces that are better for the general population!". This thesis has five fully blind participants for the interviews and observations, and Lazar et al. (2017) mentions that this number of participants is acceptable when the research focuses on users with disabilities.

Usability testing is known as “user research” in which case the research is about the interface and not the user (Lazar et al., 2017, p. 263). This kind of technique is used to find out how to make an interface better. Usability testing involves representative users (personas, user profiles) attempting representative environments in working versions of computer interfaces (Lewis, 2006). I have included this method because mobile devices need usability testing more frequently because the interaction approaches are newer (more and more content on smaller screen sizes) and features can easily be activated accidentally. The goal of this technique is to improve the quality of an interface.

In short, the techniques used to gather data in this thesis include: a systematic literature review, semi-structured interviews, observational tasks, and persona testing. Although the main approach used to answer most of the research questions is the semi-structured interviews, the persona was created and put to test especially for representing the non-verbal user. Meanwhile, the observations consisting of specific tasks are used as a supplement to the interviews with the fully blind users. On the other hand, the goal of the systematic literature review is to find sources that can support or help answer the research questions or uphold the findings of the thesis and examine the current state in the development of IPAs.

5.1. The Research Design

5.1.1. Systematic Literature Review

The systematic literature review focuses on these research questions:

- What barriers do people, including people with disabilities, experience when using intelligent personal assistants?
- Are intelligent personal assistants useful to people with disabilities when impaired by the situation, context, or the environment?
- Do the user interfaces of intelligent personal assistants facilitate accessibility and usability?
- Which are the methods of enhancing or improving intelligent personal assistants?

In order to find the relevant literature about the UD of IPAs, the data collection was based on specific search terms and inclusion criteria. The search terms included the following keywords: intelligent personal assistants, Siri, barriers, blind, non-verbal, useful, user interface, disabilities,

accessibility or usability, alternative access, assistive technology, machine learning, and artificial intelligence.

5.1.1.1. Literature Search and Database Selection

The main objective of this analysis is to:

1. Carry out a systematic literature review to examine the current state of the art of IPAs, incorporating the new ideas and features including:
    a. Guidelines
    b. Design methods.
2. Which are the most common barriers that people, including people with disabilities, experience when using IPAs?
3. How does the physical environment (noisy situations) affect the usefulness of IPAs?
4. Does the user interface of IPAs facilitate people with disabilities regarding accessibility and usability?

A systematic approach will be used to identify relevant papers and studies through methodical search of academic online databases and search engines, such as: Google Scholar, Springer, ACM, IEEE Xplore, or Web of Science.

5.1.2. The Qualitative Data

To answer the research questions, I conducted semi-structured interviews with five blind users. In addition, I also included some observational tasks.

5.1.2.1. The Interview Schedule

The following are the questions I decided were best suited to find out or understand the barriers that fully blind people experience when using Intelligent Personal Assistants:

1. How long have you been using a voice assistant? What do you use it for? Complete Tasks, Make Searches?
2. Which type of a voice assistant are you currently using? Have you considered switching to another voice assistant? Why?
3. Tell me about your general opinion about voice assistants based on your experience?
4. What do you like about your voice assistant? How often do you use it?

5. What do you dislike about your voice assistant? Can you mention any specific problems encountered? How did this affect you?
6. How important is it for you that voice assistants are easy to access? Tell me more about this?
7. What do you think about the layout (the screen) when activating your voice assistant? Would you want to add or remove anything? Any reasons?
8. Tell me about your suggestions for how to improve the voice assistant you are using? Is there anything you would like to change? Are there any features you are missing?

5.1.2.1.1. Why this Initial Suite of Questions?

There are several reasons as to why I ended up with these questions. When deciding how to conduct the interviews, I considered several different interview structures. Interviews can usually be divided into structured, semi-structured, and open-structured/unstructured (DiCicco-Bloom & Crabtree, 2006).

With my goal being to find out the barriers that the blind people experience, learn more about the usefulness of IPAs, and see how the user interface affects these users, an approach that produces qualitative data was the right choice. I decided against a structured interview, since they often result in quantitative data (DiCicco-Bloom & Crabtree, 2006).

According to Holloway and Galvin (2016), semi-structured interviews are the most common interviews in qualitative research, were the researcher asks the participants a series of pre made questions that limits the response categories (Qu & Dumay, 2011). This method, with both structured and unstructured questions would best suit to answer these questions. That's because, the open questions would be used to really go into the depth of the participant's answers, meanwhile by including questions that are easy to answer I would learn more about the participants relationship with IPAs and their assistant's user interface. Doody and Noonan (2013) points out that by using only open questions, the researcher needs to be a skilled interviewer, and it is advisable to start with easier questions before moving onto harder ones. This is what I had in mind when creating the questions, since I am in the infancy of research, and I don't see myself as a very skilled interviewer and don't want to make the interview too complex.

Another popular method in qualitative research that inspired me was the unstructured or in-depth interview. An unstructured interview can be considered similar to a guided conversation (DiCicco-Bloom & Crabtree, 2006; Lofland, 1995). You can say that the interview is a conversation with a purpose, and it is therefore important to have a clearly defined purpose when conducting a qualitative study (Ritchie, Lewis, Nicholls, & Ormston, 2013). Since I believe that my research has a clear purpose I drew inspiration from this. Ritchie et al. (2013) continues that there is still a clear difference between normal conversation and an unstructured interview. The roles of the researcher and participant, and their objective does not bear much resemblance to a "regular" conversation (Ritchie et al., 2013).

As mentioned earlier, when designing this interview, I decided to use mostly open questions. According to Ritchie et al. (2013), a good interview usually involves open questions, and one of the strengths of qualitative research is that it makes it possible for me to explore unexpected issues as they appear. When designing the questions, I used both content mapping questions and content mining questions as a manner of raising issues, while content mining questions are used to explore the issues in detail. These types of questions also involve probes which are used to follow up on the answers I received. As researchers, we have a responsibility to encourage the participant to talk freely when answering questions and use techniques like "follow up questions" to achieve deeper answers and to fully understand the participant's meaning.

I avoided using leading questions that could influence the answers, and focused on creating short and clear questions. Ritchie et al. (2013) emphasizes that short and clear questions, where the participant has no uncertainties, are the most effective questions. Doody and Noonan (2013) mentioned an important aspect to have in mind when creating questions: "The responses to the interview questions should provide the researcher with an opportunity to answer the research question" (Doody & Noonan, 2013, p. 3). This helped me with avoiding structured questions that provided only yes and no answers. Jacob and Furgerson (2012) states that the phrase "tell me about" is a strong way to start a question because it invites the participants to tell a story and also makes it harder to create a question that is too complicated to answer. The question number seven, which might look unusual to ask a fully blind person,

was included in order to understand the relationship of voice control and the IPA and whether they can be used together for easily completing tasks in the phone or, quite the opposite, cause confusion. That would help to see if the fully blind users need a better interface with more suggestions about actions.

Another important aspect in mind when creating the questions was that I don't want to take up too much of the participant's time. I don't want the interview to last too long and therefore developed an interview with 8 main questions that would last about 20-30 minutes. Having in mind that my participants are fully blind people, I believe the interview could take another five more extra minutes.

This interview can also be done online, using Skype, if it is not possible to meet face-to-face with the participants. Skype is the best option when it is not possible to meet because I can also observe the participants. Lazar et al. (2017) mentions that when conducting an online interview by only chatting, the researchers lose some control over the process and participants in comparison with doing a face to face interview, and this kind of research is more appropriate for quantitative research. Further, Lazar et al. (2017) also states that the contextual feedback you get is much more limited, since the research doesn't let you observe the user to the same extent. During the process of choosing the questions, I practiced the process with some classmates in order to come up with the best questions. And while doing that, I came up with new questions and figured out some of the questions I previously had were not clear enough, and could be misunderstood, so they were revised.

5.1.2.2. The Observations

Interviews have their shortcomings and challenges. They involve data collection that is separated from the task and context under consideration, and they suffer from problems of recall (Lazar et al., 2017). The participants tell the interviewer what they remember. When asking participants about software related questions, the answers I will get during the interviews can be very different from the answers I might be able to find by sitting in front of the user and actually seeing them using the IPAs.

To avoid these potential detachments, as Lazar et al. (2017) advises, I combined the

interviews with observations. The observations took place before the interview sessions. The following table shows the series of tasks and what the participants were asked to do.

Task No.	About the Tasks
1.	Activate their intelligent personal assistant.
2.	Complete a search using the default search engine.
3.	Accomplish a task that serves as a shortcut of another application, such as setting an alarm, creating a note, or choosing a similar task of this nature of their own preference.
4.	Exploring the suggestions of commands that their IPA offers.

Table 5.1.2.2. The observational tasks.

This helped me see the difference between what my participants said and did. “Look at behaviour, listen to perceptions” (Crabtree & Miller, 1999).

5.1.2.2.1. Why These Series of Tasks?

When choosing these specific tasks, there were some good reasons considering the great variety of possibilities. One of the main reasons was to create as shorter and simpler tasks as possible because I wouldn’t want the energy of my participants to go towards these tasks since the most important is the interview. The participants should be as fresh as possible for the interview and I wouldn’t want something very complicated and long to distract them before the interview.

There are a lot of articles that I have read online where blind people describe their experience using IPAs. One of the most repeated questions/wonders risen by blind users was if only sighted people knew how the blind actually use the accessibility features of assistive technology how different the developers thinking would be. From a blind person’s point of view, us as researchers or developers should be willing to undertake the development of an assistive application by asking real users with disabilities about the areas of daily life that a vital application for them could really help them (Hansen, 2017). This topic should be well researched and these people with disabilities should be asked about their opinions. This has been one of my motives for choosing to do the interviews and observations in order to learn more from fully blind users and offer them the same experience as sighted users have by

offering them fully accessible application interfaces rather than simplified “blind-friendly” ones.

The first task is a basic task and is chosen to see which option the user chooses to activate their IPA having in mind that there are many other alternatives. Also, this task could open the way for the participant to talk about accidentally activating IPAs in the interview. The second task was chosen since the literature review shows a great increase in searches and I want to see how it goes with a fully blind user and are they able to get the information they need. Frankly, to see if they actually get the information they’re searching for. The third task has to do with the usefulness of these possibilities. With this technique, by creating a very simple task and seeing how it goes the participants will be more prepared about the interview and they will consider this as something to talk about regarding other applications they use and not only the alarms or notes.

When using technology that is always updating and improving, we often see tips or short tutorials when opening applications for the first time or after an update which are very important to understand how to use it and get the most out of it. The last task was created in order to understand how blind users deal with this, can they use the suggestions that their IPA provides, are they enough, are these suggestions accessible, and so on.

5.1.3. The Persona Testing

I used a persona to add user diversity, but most importantly I need valuable information coming from the non-verbal users’ point of view in order to answer the research questions. The usability testing technique will be used to represent, by using this persona hearsay, the non-verbal people. In this case, the non-verbal people were imaginary and the interfaces of the IPAs in different devices were put to testing regarding accessibility and usability.

The persona was not created by me, I have taken the liberty and fitted a persona created by others in my evaluation. Fortunately, I have experience using this technique since I practiced it in a previous course. The persona used is taken from the online course Digital Accessibility: Enabling Participation in the Information Society from FutureLearn (2016) that I followed last semester.

5.1.3.1. Lily – Non-verbal

Lily is 21 years old and is non-verbal since birth. She communicates using sign language. Lily is a big fan of the artificial intelligence field and she studies physics. On her free time, she likes to read about robots or watch fantasy movies. Her hobbies are tennis and hiking. Even though she receives social benefits, she likes working and is employed in a service centre where she usually sits and works in the computer.

5.1.3.1.1. Challenge 1: interact with the smartphone using the IPA

Even though she can complete all tasks on her smartphone using touch interactions, she also wants to use the IPA in her smartphone to quickly look and find what she's searching for to save time. Lately, she has heard about this new trend of smart speakers and she wants to own one to control her house. Sometimes, she feels left aside, and wonders what it is like to have a virtual assistant to communicate and she is challenged by the new opportunities and trends that technology offers.

5.2. Data Collection

5.2.1. The Process of the Systematic Literature Review

To select appropriate papers for inclusion in the review, some search terms were decided and the following criteria were used:

- Articles published between 2009 and 2018.
- Articles must involve IPAs.
- Articles must address one of the followings:
    - How to enhance IPAs.
    - How to disable barriers of IPAs.
    - How to design for people with disabilities, especially involving the blind and non-verbal.
    - Guidelines used when developing IPAs or design methods.
    - The physical environment in relation with the usefulness of IPAs.
    - The IPAs' user interface in relation with accessibility and usability.
- Articles must include at least one combination of these search keywords.:

- Intelligent personal assistants OR voice assistants AND Siri AND barriers AND blind OR non-verbal OR disability AND useful AND user interface AND accessibility OR usability.
- Artificial intelligence AND Machine Learning AND Siri AND barriers AND accessibility OR usability AND user interface AND blind OR non-verbal OR Disability.

- Articles must be from peer-reviewed publications.

5.2.2. The Process of Persona Testing

The testing using the persona started, during the first phase of this thesis, with Apple's iPhone 7 running iOS 10.3.1 (the latest at that time). Siri in this iOS was the same in all Apple's devices of that time. The same procedure was continued in the second and final phase of this thesis by gathering other devices such as Samsung's Galaxy S9, and also by testing the latest version of iOS. As mentioned earlier, universal design values diversity and inclusiveness. Therefore, I decided to put in this testing the other devices paired with IPAs too.

The IPAs tested and their operating system are put on the table below:

Brand	**Device Type**	**IPA Name**	**Operating System**
Apple	iPhone 7	Siri	iOS 10.3.1 and iOS 11.3
Samsung	Galaxy S9	Bixby	Android 8.0
Samsung	Galaxy S5	S Voice	4.0
Apple	MacBook Pro 13 with Touch Bar	Siri	macOS High Sierra Version 10.13.3
Google	Pixel 2	Google Assistant	Android 8.0
Amazon	Echo Show	Alexa	601481220
Amazon	Echo (2nd Generation)	Alexa	608490620

Table 5.2.2 The list of Devices Tested using Persona

The process of testing these devices is relatively simple, the only thing that needs to be done is for Lily to pick up the device, activate the IPA and find information online or complete tasks using only touch interaction. The goal is to successfully interact without using voice.

In order to have these devices on hands to make this test, I went in Telenor or Telia stores in Oslo and used the ones available there. Meanwhile, since I own the Apple products in the list they were tested at home. For the other devices, such as the smart speakers that I couldn't get hands-on I searched for videos online of full reviews. The full reviews also helped me discover things I might have missed.

5.2.3. The Processes of the Qualitative Data

5.2.3.1. The Process of Conducting the Interviews and Observations

"Asking questions and getting answers is a much harder task than it may seem at first" (Fontana & Frey, 2000, p. 642). As I mentioned earlier, about my lack of experience, I used a guide to help me through the process of creating and conducting the interview. Jacob and Furgerson (2012) talks about the difficulties students often have when conducting qualitative research for the first time and points out the importance of using aids when going through the interview process. The guide consisted with not only the questions of the interview that I asked, but also what I said before the interview (including the informed consent), prompts that helped me under the interview, and what I said after the interview. Purpose of the research, time line, confidentiality, and motivation of work are some of the most important parts in the guide. Doody and Noonan (2013) states that the researcher should learn the guide by heart before the interview, because the participants may be distracted by it, and in worst case ruining the flow in the interview. I worked this out by making something in between, I let the participants know before the interview that I am going to use a guide to help me remember the questions I will ask. A guide is commonly used and is not intended only for those lacking experiences.

When taking notes of the participants answers I mostly used keywords and not fully sentences. Doody and Noonan (2013) also talks about the use of "why" questions, and how this question may cause participants to get defensive, and in worst case hamper their response to the question and future questions. In my interview, I tried not to give this word too much significance while believing that the participants will not be distressed. For us, as researchers, the participants wellbeing during the interview is very important. Ritchie et al. (2013) talks about the importance of creating climate of trust, where the researcher shows a real desire to understand from the perspective of the participant, and where the researcher appears

comfortable and professional. For this interview, I made a written contract to receive the participants' consent to use their answers for this thesis. An important point, besides getting the answers, was to act professional and make the participants feel comfortable. This also helped me gain experience.

Ritchie et al. (2013) states that the researcher's ability to listen is an essential part of interviewing, and how a researcher comprehends the participant's answers affect the interview. Further, Ritchie et al. (2013) also points out that a researcher has to think quickly in order to extract the most important information out of what the participant is saying, knowing what to pursue, and formulate relevant follow up questions. Prior to conducting the interview, I made some pilot-testing and I felt more comfortable when really conducting it. I tried to look original, have faith in myself, and not follow the prepared questions in a submissive manner. I consider myself to have solid knowledge about this field of research so I don't get myself in a position where the participants might ask me questions that I might not be able to answer, and thus seem unprofessional and unprepared.

Interviews can be extremely flexible, based on interviewee responses the interviewers can choose to reorder questions or invent completely new lines of inquiry (Lazar et al., 2017, p. 188). When I conducted the interviews, I considered the possibility of reordering the questions. I tried to increase the understanding by having in mind the possibility of opportunistic interviewing, taking an interesting idea and running with it (Lazar et al., 2017, p. 188).

Overall, the process of the interviews and observations went very well. Firstly, I contacted Norges Blindeforbund which is a vision-impaired organisation focusing on the interests of the blind and those with low-vision. The organisation has approximately nine thousand members and their overall goal is social equality for the blind, the partially sighted as well as other groups of people with disabilities (Norges Blindeforbund, 2018).

This organisation contacted Statped and made it possible for me to interview three people there. Statped is the special education service in the sector of education for municipalities and county municipalities that provides resources and expertise for customized and inclusive education for children, adolescents and adults with special educational needs in collaboration with educational psychological service (Statped, 2018).

The first three interviews took place in Statped's building in Statped sørøst Hovseter. When I arrived there, we met at the reception, had a coffee, and I followed them to the room they had reserved for the interviews. The environment was accessible, and they moved around like they knew every inch of the place since they've been working there for quite some time.

First of all, during the analysis of the data, I realized two of my mistakes in the design of the interview schedule and observational tasks while having in mind the process and how it went. I wouldn't exactly call these mistakes because they required from me to be flexible.

As I mentioned earlier about the flexibility of the interviews, in three interviews I had to skip the first part of the second question because I knew beforehand which type of IPA they used from the observations and didn't have to ask because it was obvious. I included that question just in case I wouldn't be able to recognize which one it is and thus have to ask. Basically, I had to clarify and add "except the one in your smartphone". By adding that, it would also include other devices they own such as a second phone or a smart speaker at home.

The second mistake was in the first interview, in the third task of the observations, where the first participant was asked to complete a task of another application. I shouldn't have mentioned as an example the alarm because the participant preferred to set an alarm five minutes from the moment that was asked to. In my opinion, the participant chose to set one in five minutes because we would both know that the command worked and that would make me happy. And when that happened, the alarm rang as soon as I asked the first question of the interview and it distracted us a little bit. However, that didn't pose a serious problem because it was funny and actually made the interview friendlier. After the first interview, I removed the alarm as an example for the task.

Another case that required from me to be flexible and not follow up on the questions submissively was when I asked about having considered to switch to another IPA. There were cases when the participant had considered to switch to another one and was eager to tell the reason why and I didn't interrupt, I let them continue. That reason was, of course, something that they didn't like about their IPA and hence I didn't have to ask the first part of question five, instead I went deeper with the second part of that question.

Since I was in contact with Blindeforbundet, the general secretary there invited me for

an interview. After I was interviewed, I was given the opportunity to interview two more people there, in their building in Majorstuen. The interviews went very well, and this time the whole process was easier for me, I was calmer, and more confident. In the end, I managed to possess interview data from five participants in total.

5.2.3.2. The Process of Analysing the Gathered Data

The collected data from the interviews was analysed as soon as possible while the interviews were fresh in my mind. This helped me remember details that I might not have written in my notes, as notes became less and less useful over time. When analysing the data, I used my notes rather than audio or video recordings. I didn't use any computer aided software for analysing the interview data because the number of interviews is low and can be dealt with manually.

The data from the interviews underwent five key stages of analysis. The stages are listed below:

1. Read the Transcripts:
    a. The transcripts were put as a whole,
    b. Notes were created about the first impressions,
    c. The transcripts were re-read again, very carefully and line by line.
2. The Labelling of Relevant Information:
    a. Were labelled: words, sentences, phrases, or sections regarding actions, activities, concepts, differences, opinions, processes (known as coding). Based on: repetition, explicitly stated important by the participant, or other reasons.
    b. Focused on: description of things, and conceptualisation of underlying patterns.
3. Decided about the most important codes and created categories by bringing the codes together:
    a. Created new codes if needed by combination of two or more codes,
    b. Conceptualised the work by keeping only the necessary codes and grouping them together in categories.
4. The labelling of categories:
    a. Decided which were the most relevant ones,

b. Described the connections between categories.

5. Other options considered:
    a. Whether there was hierarchy among categories,
    b. Whether one category was more important than the other.

5.3. Analysis

5.3.1. The Analysis of the Articles from Systematic Literature Review

To evaluate the relevance, the gathered data was then put in screening process according to the objective of the literature review. Only the articles written in English were included, and abstracts were reviewed for the articles that met the specified inclusion criteria. If the article met the objective, it was downloaded and thoroughly reviewed.

It was a total of 122 articles, using both search terms, that Google Scholar presented. Those were as well of other databases like ACM or Web of Science. Many of the articles that were found by Google Scholar directed to other webpages that required a membership in order to access the full text. However, I used the databases that the university provides to search again and again in each one to access the content.

After reading their titles and abstracts, only thirty-two articles and two books were downloaded and thoroughly reviewed that met the criteria. Sometimes, when in doubt whether to download and read or not, their introduction or research questions were skimmed in order to be sure that the article didn't meet the criteria.

5.3.1.1. Analysing of Selected Articles

The search results were combined and condensed. Eighty articles were excluded due to not being relevant, and seven for not being peer reviewed. During the review process of the thirty-two articles and two books that were downloaded it was found that twenty-three articles were irrelevant and one book was not so related to this thesis and thus were ignored. At this stage, only nine articles and one book remained. Meanwhile, at a later point, three more articles that were found in the citations of the other articles were downloaded and added to the review because important and relevant information was cited from them. At the very ending point, it was a total of twelve articles and one book that remained that are discussed in the results section.

5.3.2. The Analysis of the Data from Persona Testing

The collected data from persona testing was analysed based on certain aspects. Firstly, the same IPA of the corresponding device was compared with the older OS of the same device. By the way, since from before attempting on the challenge, it was obvious that the collected data in the end will be 'yes' and 'no' answers. In other words, it would either be successful to interact with the IPA by typing or not. In order to avoid results that would only describe this aspect, I decided to compare based on OS.

The goal of this comparison was to see if there were any connections, patterns or any common ground about how these giant companies decide to approach the design of their IPA. Moreover, I compared the old IPA of Samsung which was S Voice with the new one Bixby. Then, there was compared Alexa of Echo with Echo Show. I got this idea in the first place because I had to test again Apple's Siri in iOS 11 since the test of Siri in iOS 10 during the first phase of the thesis became old and I noticed very big changes.

5.3.3. The Analysis of the Qualitative Data.

There are a variety of techniques used to analyse interview data. The approach known as content analysis, which is the one that I used, involves examination of the text of the interview patterns of usage. It includes frequency of terms, co-occurrences, and other structural markers that may provide indications of the importance of various concepts and the relationships between them (Lazar et al., 2017). This approach, builds on the assumption that the structure of an interviewee's comments provides meaningful hints as to what the interviewee finds important and why (Robson, 2002).

Holsti (1968, p. 14) defined content analysis as "… a technique for making inferences by objectively and systematically identifying specified characteristics of messages". Kulatunga, Amaratunga, and Haigh (2007) mentions the main types of content analysis methods: (1) word count, (2) conceptual content analysis, (3) relational analysis, and (4) referential.

Word count involves counting the frequency of words in the text. The assumption behind word counting is that the words mentioned most often indicate the important concerns. In conceptual content analysis the text is scrutinized to check the existence and frequency of a concept or theme (Colorado State University, 2006). The relational analysis examines the

relationships between the previously mentioned concepts (Palmquist, Carley, & Dale, 1997). The referential content analysis is suitable to analyse the complexity of language to produce a meaning (Michael, Bryman, & Liao, 2004).

What makes content analysis rich and meaningful is the use of codes (Stemler, 2001). According to Ryan and Bernard (2003, p. 780) “Coding is the heart and soul of whole content analysis”. According to Weber (1990, p. 37) “Category is a group of words with similar meanings or connotations”. The text itself can generate categories, concepts, or codes, therefore codes can also be identified after data collection (Ryan & Bernard, 2003). Having too few codes can produce invalid and unreliable conclusions (Palmquist et al., 1997). What is known as *in vivo* coding recommends to develop the categories and codes closer to the original text by using actual phrases or words in the text (Neuman, 2002).

5.3.3.1. The First Stage

Initially, in this stage, I took the transcripts that consisted with only the keywords I wrote during the interviews and started writing in each one, by also keeping the interview schedule near to help me remember, as much details as I remembered from the interviews, such as: sentences, phrases, actions, opinions, and so on. One transcript represented one participant. After that, I read each transcript very carefully and tried to visualise the processes of each interview in my head in order to remember if I had missed something. I read them a couple of times to make sure everything was on place. Then, I put the transcripts as a whole and moved to the second stage.

5.3.3.2. The Second Stage

In this stage, I took markers with different colours and started coding by marking the repeated words, each word with a different marker. After that, I labelled phrases that expressed opinions, represented activities or processes, differences, and concepts as well as statements that the participants explicitly declared important. The most repeated words in this stage were: dictating, Norwegian, search, functions, problems, recognition, English, internet, available, understand, use, important, screen, know, always, errors, something, correctly, communication, issues, pronounce, answer, things, information, Siri, type, and difficult.

After marking the words, I started creating new codes by using two or more codes, and

ended up with: “I couldn’t interact with the phone”, “written correctly”, “issues with the keyboard”, “available in Norwegian”, “punctuation marks”, and “search for things in the internet”. When taking notes, I focused mostly on keywords and gave my full concentration into remembering every important phrase and sentence.

5.3.3.3. The Third and Fourth Stages

At this point, I dropped the words: ‘always’, ‘screen’, and ‘something’ because they were irrelevant in most of the cases and I used them to help me remember the sentences. Furthermore, I took words, phrases and sentences of the same meaning and created the categories. I took two codes and made one code. I used *in vivo* coding to name the categories based on actual words and phrases. When I wrote everything I remembered, the text itself produced concepts and helped me create the categories. Each category consists of the corresponding keywords, phrases, and sentences. The seven categories are listed and discussed in section 7.3. Meanwhile, the information gathered from the interviews and observations that wasn’t related with the categories and didn’t need any representation i.e.: what the participants liked about their IPA, or how long they were using IPAs, is showed in section 6.3.

5.4. Ethical Considerations

Ethical codes or principles express how we should behave as a society and as individuals. They are moral judgements that can be applied to particular situations to guide our behaviour and help us make decisions. Ethics requires respect to the research environment as well as attention to the appropriate use to human participants (National Institutes of Health, 2016). Research ethics requires that participants must be treated fairly and with respect (Lazar et al., 2017). Furthermore, Lazar et al. (2017) states that to the participants must be provided information about the study, in order for them to make a meaningful decision as to whether or not they really want to be involved. The notion of informed consent is a critical component of modern research on human participants.

Research that involves people with disabilities is very important because it discovers issues that require attention, inform policy, evaluate programmes and services, and track how social and economic change affects people with disabilities (NDA, 2009). The National Disability Authority (2009) presents six key principles as a guidance for good practice in research with

people with disabilities: (1) to promote the inclusion and participation of people with disabilities in research, (2) ensuring that the research is accessible to people with disabilities, (3) to avoid harm to research participants, (4) ensuring voluntary and informed consent before participating, (5) to understand and fulfil relevant legal responsibilities, (6) to maintain the highest professional research standards and competencies.

Participants of this research were voluntary. Moreover, the participants were fully informed about the procedure of the research, and it only happened based on a consensus between us. There were not presented any burdensome demands to the participants. The environment where the interviews were conducted was physically accessible. There was diversity in participants including age, education level, gender, ICT skills, to name a few. The participants were included in the dissemination of this thesis findings.

There was no harm for the participants, neither physical nor psychological. According to National Disability Authority (2009) possible sources of harm include: breaching the confidentiality, manipulation to participate in research, the nature of the interaction between the researcher and participants. Confidentiality was guaranteed and the promise was kept for the participants and they were assured that the information gathered here will only be available to those directly involved in this study.

The benefit of their participation was emphasized. Openness and honesty were key components in which case this thesis was explained in details, including the motives of the research, objectives, goals, and how their participation is good for society and for other people who might find themselves in the same situation at some point in life. This research didn't ask a vulnerable human to participate, such as mentally disabled individuals unable to give an informed consent. I was sincere with my participants, showed them my work in advance, and assured them that I would not make up any of the data, and wound not publish misleading information.

Lastly, this research was notified to the relevant authority here in Norway which is NSD (Norsk Senter for Forskningsdata) and was done in Norway. The notification receipt can be seen in Appendix B in the list of appendices. NSD (2017) is one of the largest resource centres for archiving research data and also assists researchers and students with data gathering. NSD

(2017) also states that the same legal rules apply if a research was to be conducted abroad as long as the student/researcher is of an institution in Norway.

5.4.1. Informed Consent Form

There was presented the consent form for the research to the fully blind participants. The informed consent document was also provided in electronic format (accessible PDF) and was sent by email, to be used for reading. The form was verbally instructed to the participant regarding what to do and asked them to respond verbally while recording the conversation of the consent, and Lazar et al. (2017, p. 516) states that this is appropriate.

Each participant was paid after the session with 200kr cash or by using Vipps (an easy way of sending money through the Vipps app provided by DNB bank), a likeable sum from a student that shows gratitude and respect for the participants. This also helped me in the recruitment process. The informed consent can be seen in Appendix A in the appendices section.

6. Results

6.1. Systematic Literature Review

An overview of the topics from the systematic literature review is presented in the following table. The table consists of four columns: (1) the reference, (2) the method used in the study, (3) the aim of the paper, and (4) the results from the research. Each source is then further discussed in the topics section. The table consists of only the articles, the book is discussed at the very end.

References	Methods/Participants	Research Aims	Research Results
1. (Nagata, Oostendorp, & Neerincx, 2004)	An exploratory study with 24 participants to determine user requirements. The features of PALS were presented on storyboards viewed on a pocket PC.	Enhance user interaction with mobile devices and web services by using an intelligent interface to facilitate quick and accurate web task performance	Assistance for the mobile Web interface is needed to support a user's attention and memory for handling interruption.
2. (Milhorat, Schlogl, & Chollet, 2014)	The implementation of a spoken dialog system (SDS). The base components are specialized agents who communicate with each other using a server.	Make the human-machine dialogue more flexible and adaptable to the user's requirements and bypass the limitations of the current technological capabilities.	For some time, it will be improbable to develop any standard for designing effective dialogue systems.
3. (Santos, Rodrigues, Casal,	A literature review to determine the current state of the art of IPAs	Review the state of the art of IPAs based on IoT solutions.	The IoT technology will enable the creation of

Saleem, & Denisov, 2016)	based on IoT solutions.		omnipresent communication scenarios where almost all the devices in the environment will be able to communicate.
4. (Cowan et al., 2017)	20 participants from a university community. The method: Online questionnaire, Tasks while reporting any issues and observations, and semi-structured focus groups.	Understand the experience of people who use IPAs occasionally, but not frequently. Address why people currently don't use IPAs or use them infrequently by presenting findings from a focus group study.	Identified barriers that were relevant to infrequent users of IPAs more widely, such as: the difficulties caused by interruptions to hands free interaction, the tendency to imbue Siri with human-like qualities and the lack of trust in Siri performing tasks.
5. (Abdolrahmani, Kuber, & Hurst, 2016)	Interviews with 8 blind participants. Snowball sampling. Scenarios created and presented to inspire discussion.	Examine the issues faced by blind individuals using a mobile device when impaired by the situation, context, or environment.	Blind individuals are interested in using voice interfaces to complement touch screen interaction when situationally impaired.

6. (Glasser, Kushalnagar, & Kushalnagar, 2017)	Five participants who evaluated ASR software.	Assess the usability of the applications in face-to-face spoken language interaction by providing a visible text representation of speech.	ASR has been and continues to be focused on hearing speakers. Deaf and hard of hearing individuals find it difficult to be fully comfortable with ASR and cannot dictate to ASR services reliably. All ASR services that were tested do not reliably provide accurate or usable transcripts.
7. (B. Smith, 2017)	Semi-structured interviews with 19 participants.	Provide suggestions of design guidelines for information architects when approaching smart speaker user-interaction. A critical look at designing information architecture based on users' reported experience.	The users must: (1) feel understood and heard by the device, (2) trust the device, (3) have minimal breakdowns in understanding, (4) be involved in the process of designing information architecture.

8. (Tundrea, 2017)	Online survey with 18 participants with 24 questions. A focus group with 7 participants.	Evaluate IPA usage motivations. Address various privacy issues and concerns related to IPAs.	The users are concerned and aware of their personal data collection. Users feel trapped in the network society. The denial of technology would bring unfavourable consequences from a social perspective.
9. (Luger & Sellen, 2016)	Semi-structured interviews with 14 participants.	Aims to understand user experience of Conversational Agents (CA).	Without the humanlike cues and affordances relied upon by multimodal systems, CAs have particular challenges.
10. (Moorthy & Vu, 2014)	120 participants for an online survey.	Gain a basic understanding of the usage patterns of IPAs in public spaces.	The participants reported that they use the IPA and the smartphone keyboard to enter private vs. non-private information depending on location/place.
11. (Easwara Moorthy & Vu, 2015)	76 participants for an online survey.	Gain an understanding of the	The participants were more cautious

		usage patterns of IPAs in public spaces.	of disclosing private information than non-private information.

Table 6.1. The overview of the sources from the systematic literature review.

6.1.1. The Topics

Each subject matter found in the systematic literature review is further discussed into more details in the following sections, and the most important information is extracted.

6.1.1.1. The Personal Assistant for Online Services

Nagata et al. (2004) presented the Personal Assistant for online Services (PALS). PALS aims to enhance user interaction with mobile devices and web services by using an intelligent interface to facilitate quick and accurate web task performance. PALS facilitates a user to complete a web task by using supportive concepts like interactive displays.

To determine user requirements for PALS, it was conducted an exploratory study with 24 participants on specific features of PALS presented on storyboards and were viewed on a pocket PC. Each participant viewed 22 PALS features and had to rate each feature on a scale from 1-5 for usefulness, attractiveness and necessity of the feature. They concluded by proposing that mobile assistance can improve usability of the mobile Web for a wide range of users.

6.1.1.2. The Implementation of a Spoken Dialog System

Milhorat et al. (2014) proposed the implementation of a Spoken Dialog System (SDS), to approach four possible areas of improvement for IPAs: (1) Extended Dialog History, (2) Improved Context Awareness, (3) Dynamic System Adaptation, and (4) Supported Task Hierarchy Design. They argue that the current technology is not able to provide algorithms for processing and understanding free-forms of conversations. It is crucial for the user to have the feeling of a naturalistic interaction in order to ensure an effective and efficient use of the system.

The aim of this paper is to make the human-machine dialogue more flexible and adaptable to the user's requirements. They concluded that for some time it will be improbable to develop any kind of standard for designing effective dialogue systems in the hopes of these

ideas helping in the future as an attempt to progress towards better and more flexible natural language user interfaces.

6.1.1.3. Intelligent Personal Assistants in an Internet of Things Context

Santos et al. (2016) in an Internet of Things (IoT) context, discuss the possibilities of the integration of IPAs with IoT objects as a promising solution to offer people the perfect personal assistant that has abilities to act, manage, and interact autonomously with the environment and suggest suitable solutions to problems that arise in daily life. The aim of this paper is to review the current state of the art of IPAs based on IoT solutions. They contribute by: (1) overviewing IoT technology and reviewing the available enabling technologies and protocols, (2) showing the main features of IPAs providing some examples of their approaches that involve the features, (3) describing how IoT networks might improve the functionalities of IPAs.

Application	Usability
Healthcare [92], [93]	IPAs can be used on ambient-assisted living scenarios, where sensors monitor a person that is under care of a caretaker (medical staff or not). The caretaker uses an IPA that receives information from the body sensors. It allows a person to be monitored anytime and anywhere. The sensors can communicate through low-energy protocols (IEEE 802.15.4 or ZigBee).
Smart home [94]–[98]	IPAs can be used to control and manage smart home appliances through mobile devices, such as smartphones. They can be used to turn off the lights inside a house, for example. Besides that, an IPA must be able to notify its user if some unexpected event is detected, such as a fire in the kitchen, or an abnormality on the air conditioner temperature.

Figure 6.1.1.3 Usability of IPAs on Healthcare and Smart home scenarios under an IoT context (Santos et al., 2016, p. 7)

In conclusion, this paper talks about how the IoT technology will enable the creation of omnipresent communication scenarios. In which case, all the devices in the environment are going to be able to communicate. Internet of Things technology offers new opportunities, such as creating IPAs that can assist their users while communicating with other smart objects in the environment. Furthermore, increase the knowledge of IPAs as they would learn the behaviour of their users through direct communication with them and by communicating with other smart

objects in the environment. To achieve this, there is the necessity of developing new mechanisms to increase the intelligence of IPAs, such as new machine learning algorithms and speech recognition mechanisms.

6.1.1.4. Intelligent Personal Assistants and Experience

Cowan et al. (2017) argues that other studies investigating IPAs about the experiences of people who use IPAs on a frequent basis (daily), and the use of IPAs by distinct groups including children and older adults, are invaluable. They cannot explain the experiences and barriers to use of all potential users. In this article, they adopt a distinct approach aiming to understand the experience of people who use IPAs occasionally, but not frequently. Other surveys cited in this paper have suggested that this reflects the most common pattern of use, seventy percent of all users, and if companies wish to extend the use of IPAs they should focus on the infrequent users as they represent the key target group which is not studied in detail.

There are twenty participants grounded in a series of focus groups most of whom are infrequent users, and discuss their experience using an IPA on a smartphone. The analysis of this paper focuses on users' views, everyday practices and barriers to use. The findings of this paper from focus groups resulted in six core themes: (1) issues with supporting hands-free interaction, (2) problems with performance with regards to user accent and speech recognition more widely, (3) problems around integration with third party apps, platforms and systems, (4) social embarrassment being a barrier to using mobile IPAs in public, (5) the human-like nature of IPAs, and (6) issues of trust, data privacy, transparency, and ownership. This study aims to address why people currently don't use IPAs or use them infrequently by presenting findings from a focus group study investigating how infrequent users experience and interact with IPAs.

The paper concludes that the research around user experience of IPAs remains in its infancy. Furthermore, they identified barriers that were relevant to infrequent users of IPAs more widely, such as: the difficulties caused by interruptions to hands free interaction, the tendency to imbue Siri with human-like qualities, and the lack of trust in Siri performing tasks. In addition, they mentioned a potential drawback of using humanness as a metaphor for interaction with IPAs. These were significant barriers for further using IPAs by infrequent users meanwhile the frequent users were less impacted by these issues.

6.1.1.5. Mobile Devices and Situationally-induced Impairments and Disabilities

Abdolrahmani et al. (2016) talk about situationally-induced impairments and disabilities that may be encountered when using mobile devices while in motion or in noisy environments. In which case, the situation itself places demands on the user's attention, vision, and motor ability. Examples of these include ambulatory mobile devices that are used when certain people attempt to navigate through an environment, like crossing the street, and attempt to divide attention between the mobile interface and the path ahead.

The situationally-induced impairments and disabilities are experienced by any user regardless of ability who might not be able to use their mobile device as expected. The aim of the paper is to examine the issues faced by blind individuals using a mobile device when impaired by the situation, context, or environment. The study conducts interviews by phone and video conferences, has eight blind participants who are users of mobile devices, and uses snowball sampling. The paper presents strategies and workarounds that are commonly used by their participants to address these barriers. All their participants had stated that they could not rely on their residual vision. The interview protocol consists of scenarios created by pilot studies that are shown to participants to find out whether they found themselves in some of the situations presented. These situations were used as a means for inspiring discussion.

They found out that individuals who are blind are interested in using voice interfaces to complement touch screen interaction when situationally impaired. The participants felt augmentation of voice interfaces would be beneficial if the situationally-induced impairments and disabilities being experienced were particularly limiting. This investigation revealed difficulties using phones while on the move, and other situations which might not impact sighted people. The participants expressed concerns about social factors like privacy, security, and not wanting to look different from others or to appear rude.

6.1.1.6. An Experience Report about the Accessibility Challenges by the Deaf

The study by Glasser et al. (2017) is an experience report and describes the accessibility challenges by two deaf, one hard of hearing and two hearing participants, including the authors, in using Automatic Speech Recognition (ASR) applications on personal devices for commands and group conversation. The participants all have different challenges and accessibility needs in

mixed group conversation in most settings, including academic and workplace settings.

The five participants of this study used and evaluated some ASR software: DEAFCOM, Dragon Dictation, Siri, Virtual Voice, Ava, Google Assistant, and Amazon Alexa. The apps were chosen because they were available for free and had high rating.

The study concludes that ASR has been and continues to be focused on hearing speakers. Deaf and hard of hearing individuals, even those who use voice on a regular basis, find it difficult to be fully comfortable with ASR. Furthermore, they cannot dictate to ASR services reliably, because there is a big variance in their speech, even if their speech is understandable by their hearing peers. Although the deaf person can be understood by hearing persons, all ASR services that were tested don't reliably provide accurate or usable transcripts.

6.1.1.7. Improve the User Experience of the users of Smart Speakers

The master's thesis of B. Smith (2017) focuses on smart speakers, respectively on Google Home and Amazon's Echo. Moreover, the thesis focuses on finding what the smart speaker users experience when managing and finding information and what might an information architecture designer do to improve user experience.

The study is approached by conducting semi-structured interviews with nineteen users of smart speakers. The study results in four main themes: (1) the users feel that they should be understood by their speaker, (2) trust their speaker, (3) minimize misunderstandings, and (4) be involved in the process of designing information architecture.

6.1.1.8. The Usage Motivations and Privacy Concerns of IPAs

The master's thesis of Tundrea (2017) focuses on the usage motivations and privacy concerns of IPA users. The research questions are: (1) to what extent IPA users agree to personal data collection in order to gratify their needs? and (2) what are the motivating criteria that determine the usage of IPA?

To data collection techniques in this study are an online survey with eighteen participants consisting of twenty-four questions, and a focus group with seven participants. This study, while assessing the motivations of IPA usage at the same time addressed different obstacles that were discovered in the process. The main conclusions from this study was that: (1) IPAs are a rather new element for the general public, (2) the participants from both the

survey and focus group revealed that they are very sceptical when it comes to data collection of different applications including IPAs, (3) the participants admitted that they feel 'forced' to accept in many cases the data collection in the exchange of services, (4) even though the participants didn't use their IPAs very often they are still sharing a lot of personal information with their devices, (5) the participants would love to use the IPAs in the future if they were more efficient.

6.1.1.9. The Factors that Influence Acceptance and Success

Luger and Sellen (2016) as two Microsoft researchers in a very interesting paper state that the current existing research about IPAs tends towards technical papers about architecture, systems created for specific contexts (guiding users around space), delivering information, or supporting language learning. Moreover, they state that we fail to truly understand the dynamics such as how and why these IPAs are used and which are the factors that influence acceptance and success in these scenarios. Out of this, they pose two research questions: (1) what factors currently motivate and limit the ongoing use of IPAs in everyday life, and (2) what we should consider in future design iterations?

Through semi-structured interviews the study explored how fourteen existing users used (who considered themselves 'regular' users) IPAs, where they used them, and the emotions elicited by their use. The interviews were conducted through telephone or skype. Siri was their most used IPA (10 out of 14).

Luger and Sellen (2016) concluded that the operation of the conversational agent (CA) systems failed to bridge the gap between user expectation and system operation. Users had poor mental models of how their IPA worked and that these were reinforced through lack of meaningful feedback regarding system capability and intelligence. Their work also found out that the interactions with the agent was generally seen as a secondary task. There was found a deep 'gulf of evaluation' demonstrated through the extent to which users were consistently unable to ascertain the simplest tasks, and their reluctance to use the CA for complex or sensitive activities. In addition, while the key use for IPAs are 'hands free', the handling of errors by the system raised the question of the design goals. The users didn't trust the system to do complex tasks (like writing emails) with the thought that the system might not get the task done

correctly. Moreover, the users used strategies like complex words, reducing the number of words used, speaking more clearly or slowly, changing the accent when interacting with IPAs. More importantly, these users saw IPAs as an interface that had to be learned, and whether it changed over time with more interaction.

In conclusion, Luger and Sellen (2016) made the following suggestions for improving the challenges faced: (1) consider ways to reveal system intelligence, (2) reconsidering the interactional promise made by humorous engagement, (3) consider how is best to indicate capability through interaction, and (4) rethink the system feedback and design goals.

6.1.1.10. Intelligent Personal Assistants and their Context of Use

Both studies by Moorthy and Vu (2015; 2014) have explored the question of context of use. Their studies have shown a clear need to consider the public nature of the context of use and the information being transmitted in such interactions. In a quantitative study of imagined scenarios, the participants rated themselves more highly likely to use an IPA in a private place as well as when disclosing non-private information. Also, they rated IPAs as more acceptable to enter non-private information and to use it at home. Additionally, there was no difference in acceptability rating for private and non-private information in home than public contexts.

Comparatively, usability research on voice enabled multimodal tools in another study by Jöst, Häußler, Merdes, and Malaka (2005) found that people prefer to use these types of interfaces alone rather than in social situations.

6.1.1.11. Designing Voice User Interfaces

The book by Pearl (2016) focuses on designing voice user interfaces (VUI) for mobile apps and devices, and includes : (1) the VUI principles, (2) personas, avatars, actors, and video games, (3) speech recognition technology, (4) advanced VUI design, (5) user testing for VUI, (6) voice-enabled devices and cars. A multimodal interface means allowing the user to interact both via voice and by using a screen (Pearl, 2016). I will only focus on the design principles of this book because of the goal of the systematic literature review and because the other sections are not relevant. Also, this section covers accessibility as well. The principles according to Pearl (2016, pp. 13-61) are:

1. The visual capabilities of mobiles are essential to create a rich VUI experience because they allow the user to have a leisurely pace in interaction. VUI and visual are essential components of the system and it is vital to design together;
2. Conversational Design – to interact with a VUI system beyond one turn (to imagine what the user might want to do next), meaning except answering queries, the VUI should also consider what happens next without forcing the user to take another turn but by anticipating and allowing it. Additionally, it is vital that the system keeps history of what the user has told the assistant. An important rule is that the user is the one to decide how long the conversation takes;
3. Setting User Expectations – designing with breadth known as discoverability. The users should know when they can speak and what they can say. Another important element in expectations is the task, if a task can be accomplished there should be the corresponding task that goes with it;
4. Using Design Tools:
    a. Sample dialogs – a key way to design an entire conversation.
    b. Visual mock-ups (wireframes) – visualize the user experience.
    c. Flow – diagrams that illustrate the paths than the VUI system can take.
    d. Prototyping tools – such as: Tincan.ai, Api.ai (owned by Google).
    e. Confirmations – confirming input, making sure the user is understood;
5. Conversational Markers – let the user know that they're in a conversation and that they are understood. The markers include: timelines (first, finally), acknowledgement (thanks, got it), positive feedback (good job, nice to hear that);
6. Error Handling – examples: no speech detected, speech detected but not recognized, something was recognized correctly but the system does the wrong thing with it, something was recognized incorrectly;
7. Don't blame the user – error messages affect the user's perception of the system and the user's performance;
8. Novice and Expert Users – it is important to include different design strategies: shorten explanatory prompts, adapt the prompts based on user interaction, adapt to users'

behaviour, take advantage of the concept of priming (exposing someone to a particular stimulus will influence their response to a later stimulus;

9. Keeping Track of Context – the use of two different terms to refer to the same thing is called coreference and is an essential part of communication;
10. Help and Other Universals – repeat, main menu, help, operator, and goodbye. It's important that users can get help when needed;
11. Latency – determine whether the system will or will not have latency or delays and ensure that the system has a way to handle it. Latency is generally caused by: poor connectivity, system processing, or database access;
12. Disambiguation – rely on any known information to determine the answer without having to ask the user. Examples include: locations, or actions;
13. Design Documentation – there are tangible things to think about:
    a. Prompts – what the system can say to the user (complete sentences or snippets). They serve multiple purposes such as a list of voice talent to record or getting sign off from the user.
    b. Grammars/Key Phrases – specify complete grammars in the dialog, such as: Yes: "yep", "yeah", "uh-huh", "sure", and so on;
14. Accessibility – the best practices for VUI design informed by accessibility:
    a. Interactions should be time efficient.
    b. They should provide context.
    c. They should prioritize personalisation over personality.

Except these principles there are as well some other thoughts on VUIs for accessibility that are of great importance. Creating systems with voice input and output can be an enabling option for users who are visually or motor impaired. The universal design principles apply to VUIs as well (Pearl, 2016, p. 68). Multimodality is recommended because some information is preferred to be heard and other information is preferred to be seen. The key is that "multimodality should ensure that all modalities are available at all times, and that designers should not make assumptions about how all users will prefer to access these systems" (Pearl,

2016, p. 68). Failing to incorporate these design principles could become barriers to inclusion for users with disabilities.

6.2. Persona Testing

The results from persona testing showed that it wasn't possible to write to Siri in iOS 10.3.1 at all times. It was only possible to change the recognized sentence by taping in it. When iOS 11 was released, Apple introduced a new Siri interface in which was possible to type to Siri at all times. Meanwhile, it is not possible to type to Siri in the MacBook Pro with touch bar with the latest operating system.

Although it is possible to type to Bixby in Galaxy S9, results indicated that it wasn't possible to do so in S Voice in Galaxy S5. As for the Alexa powered devices, it is possible to control the Echo by using the Alexa app on iPhone or Android. However, the application doesn't allow to command it by typing. On the other hand, it is neither possible to type to Alexa in the user interface of the Echo Show. As for the Google Assistant in the Pixel 2 phone, the test showed that it was possible to type to the assistant. The summary of the results can be seen in the table below.

Device	IPA	Communicate by Typing
Apple iPhone 7	Siri	No (iOS 10.3.1), Yes (iOS 11)
MacBook Pro 13 Touch Bar	Siri	No
Galaxy S5 and S9	S Voice and Bixby	No (S Voice), Yes (Bixby)
Amazon Echo	Alexa	No
Amazon Echo Show	Alexa	No
Google Pixel 2	Google Assistant	Yes

Table 6.2. Persona Testing Results

6.3. The Qualitative Data

These results are based on the answers of five blind participants. Results show that three of the blind participants are users of Siri, one is a user of Bixby (was a user of S Voice), and one is a user of Google's Assistant.

The participants started using IPAs in different times and all had been using IPAs every day. Two of the participants started using Siri since it became available in Norwegian in 2012.

One participant had been using Siri since 2016. One had been using OK Google and Google's Assistant for one year. One participant had been a user of S Voice for three years and lately had started using Bixby.

The participants used and liked using their IPAs for a number of reasons, such as: sending messages, searching for information online, liked the quality of the voice, opening apps, asking about the weather, making phone calls, reading messages and emails, finding or adding appointments in the calendar, calculating, finding restaurants or hotels, setting alarms, using it as a shortcut for enabling and disabling settings, and other purposes.

All the participants thought that it is very important that the IPA is easy to access. They preferred to use a button to activate their IPA because they wanted to be sure it is activated and found it easy to activate. Two participants hadn't considered to switch to another IPA, one was waiting to switch to Google's Assistant when it will be available in Norwegian, and two participants had other primary reasons for switching phones such as the fingerprint reader. All the participants stated that they can't use IPAs in noisy environments. All the participants stated that they are not in a good level of knowledge or familiarity with their IPA, and they didn't know what their virtual assistant can and can't do.

The categories were created by taking words, phrases and sentences of the same meaning. Furthermore, the categories were named using in vivo coding, and to each category was assigned the corresponding data.

7. Analysis and Interpretation

7.1. Systematic Literature Review

The results of the systematic literature review are very important because they represent various themes that are interconnected with themes of the results of the interviews and observations. Two topics of the systematic literature review indicate the need of the development of new algorithms to improve and enhance IPAs. Based on topic 6.1.1.2, the new algorithms should extend the dialog history, among other improvements, and algorithms for processing and understanding free-forms of conversations. Similarly, topic 6.1.1.3 points out the need of development of new machine learning algorithms that would increase the intelligence of IPAs with the goal of establishing a communication between devices and smart objects. The algorithms fall under the second principle by Pearl in 6.1.1.11, conversational design.

Topic 6.1.1.4 focuses on the usage barriers of non-frequent users which make seventy percent of all users of IPAs. Some of the issues, among several of them, found in this study are problems around integration with third party applications, issues of trust in Siri performing tasks, and problems with regards to speech recognition. These issues are the same ones found in the interviews of this thesis presented in section 7.3. If these barriers were to be disabled it would mean leading to high adoption. Topic 6.1.1.5 focuses on situationally-induced disabilities in noisy environments. The problem at hand of this topic complies with the findings of the interviews. Based on this study, a possible solution for blind users were using voice user interfaces to complement touch screen interaction when impaired by the situation and the users found them to be beneficial. This means that, voice user interfaces are essential in IPAs and they help the blind.

The findings of the study of 6.1.1.7 about the users of smart speakers comply with this thesis interview findings and resulted in almost identical experience of users. The users felt that they should be understood by their speaker, trust their speaker, and minimize misunderstandings. In order to be understood and minimize misunderstandings, the barrier of speech recognition when dictating found in the interviews of this thesis has to be disabled. A very important study is also the study by the two Microsoft researchers in 6.1.1.9. The users of IPAs in this study had poor mental models and saw the interactions with the IPA as secondary

tasks. Additionally, the users didn't trust the system to do complex tasks because the system might not get the task done correctly and saw IPAs as an interface that had to be learned. Some of these findings are in compliance with this thesis findings especially the trust in IPA to correctly complete tasks. The study suggested that there is the need to rethink the system's feedback and design goals. Lastly, a pattern that occurred in several topics and also in the findings of this thesis is that the users of IPAs couldn't trust their IPA to complete tasks and couldn't rely on the IPA to get the job done correctly.

7.2. Persona Testing

The results from persona testing show clear signs of improvements in the design and development of IPAs that happened this year and the past year. Furthermore, they indicate that Apple was among the first to offer the possibility to type to their assistant at all times together with Google's Assistant, and followed by Samsung's Bixby. Apple is a leader in inclusion because their assistant Siri speaks twenty languages as presented in table 4.5.1 and they are reaching five times as many people as its competitors and offers the largest number of methods of activating Siri. As mentioned in section 4.6, Apple's SVP of Marketing stated that IPAs are better with a screen and with their new interface of Siri in iOS 11 they prove to be leaders in making the assistant's screen usable, this automatically includes the non-verbal as potential users. Also, as mentioned earlier in section 6.1.1.11, the first principle points out the importance of the visual capabilities in which case the IPA and the visual are essential components and it is vital to design together. On the other hand, Siri in MacBook Pro couldn't be commanded by typing.

Similarly, Bixby showed the same nature in improvement as Siri comparing to how their S Voice was and how it is today in the aspect of typing to Bixby. However, as seen in table 4.5.1 Bixby and Alexa are very limited in the languages they support and are in a disadvantage. They fail to include as many people as possible and the language barrier is eminent. Amazon's Alexa in the Echo and Echo Show speakers, shows that even though it is possible to control them through an application or browser they cannot be controlled by typing. Leaving the non-verbal aside. As well, has the language barrier and fails to include as many people as possible. As for Google's Assistant, it was possible to type to the IPA in Google's Pixel 2 phone and that makes possible the interaction of the non-verbal people with their IPA.

7.3. Qualitative Data

The analysis and results of the semi-structured interviews and observations represent the barriers encountered by the blind participants when interacting with their IPAs. The barriers are: (1) voice recognition when dictating, (2) language, (3) acquire information online, (4) IPAs communication with the 3rd party applications, (5) issues or problems of the system, and (6) trust. These barriers are extracted from the categories. There isn't a category that is more important than the other categories and neither is there a hierarchy among categories. Each category is important in what it represents. Also, the categories are linked with one another in such a way that they remain vague on their own. It's important that the categories are looked as a whole and as separated with the same level of noteworthiness. The categories are listed below.

7.3.1. Category 1 – Voice Recognition when Dictating

The keywords of this category are: dictating, recognition, issues, errors, and correctly. Also, two codes into one ned code are: written correctly, issues with the keyboard, type errors, and punctuation marks. The sentences of this category, phrases and the statements particularly stating importance are:

1. "My blind friends post on Facebook with type errors";
2. "Siri is not so good to understand what I'm saying";
3. "Dictating can be improved";
4. "Better voice recognition when dictating";
5. "It is difficult to know if Siri has written correctly the message";
6. "Dictating function could be better";
7. "Dragon Dictate has better recognition";
8. "When dictating messages, I wish the recognition of the voice could be better and be less fault";
9. "It's very important that the message is written correctly including the punctuation marks".

7.3.2. Category 2 – Language

The keywords are: Norwegian, English, pronounce, and understand. The new code is Norwegian availability. The sentences of this category, phrases and the statements particularly stating importance are:

1. “Waiting for Google Assistant to be available in Norwegian”;
2. “Alexa is not available in Norwegian”;
3. “When Siri was only available in English, I had to pronounce Norwegian names differently so that Siri would understand”;
4. “I think Siri is better in English than Norwegian”;
5. “Maybe if Google Assistant is available in Norwegian I would buy a smart speaker to use at home”.

7.3.3. Category 3 – Search for Information Online

The keywords are: search, internet, things, information, and answer. The new codes are: search information, and answer things. The sentences of this category, phrases and the statements particularly stating importance are:

1. “When Siri is available in Norwegian, it is better for information that has to do with Norwegian artists for example”;
2. “If I ask who the prime minister in Norway is, it wouldn’t answer the question directly but say: this is what I found in the internet about prime ministers in Norway”;
3. “It doesn’t give you a straight answer, makes you surf on the internet”.

Other data part of this category is that one participant on the third task of the observation asked a particular question about sport and Siri listed the sports Siri can answer. Then, the participant asked particularly about Tennis and Siri said “I can’t answer that question”. After that, the participant asked particularly about football in Norway and Siri again listed the sports Siri can answer questions about. Finally, the participant asked about the rank of the football club Liverpool in Premier League and Siri gave a correct answer.

7.3.4. Category 4 – 3rd Party Applications

The keywords of this category are: functions, available, communication. The sentences of this category, phrases and the statements particularly stating importance are:

1. “I couldn’t interact with the phone and couldn’t tell Google Now to start an app because of Samsung limitations”;
2. “Not all Google functions are available to Samsung because Samsung wants to promote their own systems so they may be ‘dispromoting’ Google things”;
3. “You don’t need to use S Voice on Samsung but you can use Google”;
4. “Use Siri in other apps, for example if I could ask Siri about ‘ruter’ or ‘buy ticket’”;
5. “I wish I could use Siri for example to ask for phone numbers in applications that offer numbers instead of opening the apps and finding it”.

7.3.5. Category 5 – Issues/Problems

The keyword of this category is: problems. The new code is: type errors. The sentences of this category, phrases and the statements particularly stating importance are:

1. “One thing I have problems with is the calendar. It doesn’t understand the 24h time format. If I say 08:30 in the calendar it will set the meeting into 20:30 in the evening”;
2. “The phone can misunderstand because of other voices in the room”;
3. “Issues with the keyboard”;
4. “Siri only understands me if it is quiet”;
5. “When I say “punktum” (Norwegian for full stop), the assistant actually writes the word punktum instead of adding a full stop”;
6. “The assistant can’t understand me if I am in a noisy place”;
7. “Sometimes Siri reads old notifications”;
8. “The phone can misunderstand because of other voices in the room”.

Other data part of this category is that one of the participants in the second task of the observation asked Siri about the calendar today and Siri said that there was nothing there meanwhile there was something.

7.3.6. Category 6 – Trust

The new code of this category is: difficult to know. The sentences of this category, phrases and the statements particularly stating importance are:

1. “I would have used it much more if it was good enough”;

2. “It is really important that I can trust what’s on the screen, I don’t want to lose all control”;
3. “I prefer to read it on braille display”;
4. “When I ask Siri to do something, like calendars or alarms, I always double check the times set by Siri”;
5. “It is difficult to know if Siri has written the message correctly, and it affects me by deleting the message and typing by dictating”.

7.3.7. Category 7 – IPAs and their User Interface

The sentences of this category, and the statements particularly stating importance are:

1. “I am a little bit unfamiliar with the voice assistant. The problem is that I have to access it with a screen reader. Talk back can read to content in the layout. I am a little bit worried that the assistant would get the commands from talk back”.
2. “I don’t know what Siri can do or can’t. I don’t see the screen. I am not sure if voice over and Siri talk to each other. I don’t know, I use voice over all the time”.

Based on the first category, IPAs have clearly a hard time understanding their users when they dictate, be it a message, email, or a Facebook post. One participant explicitly stated that Dragon Dictate had a better recognition, and the other had noticed that his friends on Facebook post with type errors and was sure about what was going on because this issue had been discussed among them. For the participants, voice recognition was very important and unfortunately had concerning issues. The participants said that it was difficult to know whether the IPA has written the message correctly. This links this category with the trust category. This means that the participants cannot rely on their virtual assistant to get this kind of job done.

According to the second category, it brings up the language barrier. The language barrier was eighty percent discovered by asking them if they would want to switch to another IPA. The participants said that Google Assistant and Alexa are not available in Norwegian. This means that the participants didn’t have such a nice experience with their current assistant and were not satisfied by it and were willing to try something new. One participant mentioned that Siri was better in English than Norwegian. This means that IPAs are good based on the language that the user has set to use and are better at finding information that has to do with that

language, e.g.: if English is used, it is more effective for finding American artists. This fact links this category with the third category about searching or acquiring information online. One participant was willing to buy a smart speaker if Google Assistant was available in Norwegian. The language barrier leads to mass non-adoption of IPAs.

The third category discovers the barrier of acquiring information. As mentioned earlier in section 5.1.2.2.1, the idea of the third task was to see if the participants actually get what they are searching for. On the other hand, the idea of including the observations in the first place was to listen to the perceptions (Crabtree & Miller, 1999). While observing a participant doing the third task, we both saw that the IPA wasn't able to answer a particular question about sport and ended up doing several attempts to get that information. Moreover, one participant said that when asking about the prime minister of Norway the assistant wouldn't answer the question directly but show what it had found on the internet. Also, another participant said that the assistant doesn't give a straight answer and makes them surf the web. This means that the assistant cannot distinguish opinions and facts and is unsure about the information it feeds and prefers to make the user see for themselves and to believe what they see. It is important that the assistant knows when the user has really space to decide whether something is true or not and needs to do research of their own in order to make their own standing about a phenomenon. But, when it's the case of facts, such as with the prime minister, the participants feel that they should be provided with a straight answer. The users of IPAs want a straight answer about a non-changeable or non-arguable fact. This category as well is linked with the trust category and according to the participants there should be mutual trust.

The fourth category describes the barrier of the communication between the IPA with the third-party applications. Except that in one case there was the limitation of the phone itself and the issue wasn't only prone to the IPA. The participants expressed the need of having an IPA that would get the information from a third-party application and offer to the user. One participant explicitly stated that this communication is very important with the example of the 'ruter' application. Also, another participant gave the example of the same nature of getting phone numbers from the application that has this information, instead of opening the application, finding it, and then closing it.

The fifth category represents general issues or problems encountered by the participants while using their IPAs. These issues are of different contexts not related with the categories, except one issue with the keyboard and the assistant not recognizing the word “punktum” in Norwegian which means a full stop mentioned by only one participant which could somehow be linked with voice recognition. One participant encountered problems with the calendar by saying the IPA didn’t understand the 24-hour format and ended up creating events in the evening instead of morning. This means that the participant didn’t know that it is required to specifically say “in the morning” or “in the evening” or “PM” or “AM” when creating events using the IPA. Otherwise, by only commanding the IPA to create an event at eight thirty the IPA would randomly create the event in the morning or in the evening. Another participant explicitly stated an issue that Siri was reading old notifications. At first, this sounded strange and left me wondering about its meaning. Then, I found out a news page explaining that Siri was reading the notifications hidden from the lock screen notifications section and was going to be fixed in the next update (Morris, 2018). On the other hand, another issue concerning the calendar was raised when one participant in the second task of the observations asked Siri to show the events of that day and it said there wasn’t any but the participant was sure there were some events. At this point, I didn’t ask further questions and didn’t show interest because that would require to check the participants phone and that would be unethical. It’s a challenge to know the real issue and its meaning in this case. Most importantly, this category brings up the situational barrier of the usefulness of IPAs in noisy situations. Based to their opinions, their IPA couldn’t understand them and was of no use when it came to noisy rooms, or noisy situations.

The sixth category represents the barrier of trust. The participants expressed concerns whether the message was written correctly, whether the time of the event or alarm was set correctly and ended up double checking. One participant preferred to use braille display instead of using the IPA because it allowed the participant to have more control on the display. According to one participant, IPAs didn’t deserve to be used as much as they are using them because they are not good enough. The problem with trust is that once its broken it’s very difficult to go along well in the future. That’s why the participant had chosen to use braille most of the time and sometimes the IPA. Meanwhile the other participants, when in doubt preferred

to double check. Most probably, the participants double checked only the most important events, because nobody would want to double check everything when it's better to save the effort and use voice over instead if the trust is completely broken.

Based on the seventh category, two participants didn't have a fruitful experience with the user interface of their IPAs. Regarding the other participants, unfortunately I wasn't able to get much information about this. However, based on only two responses the participants didn't know the capabilities of their assistants because they were unable to see the suggestions of what the assistant can do. Moreover, they were concerned that their IPA would get the commands by voice over. The participant preferred to ask other people about the possibilities that the IPA offers.

8. Discussion

As interpreted throughout the seventh section, the results of this thesis comply and are in agreement with the previous research, at least partly. The gap in literature is filled since this thesis, unlike the previous research, views the problem from another perspective, that being the concept of universal design of ICT, legislations, regulations, and standards. The major difference between this thesis and previous research is that it serves as a combination of improving the technology of IPAs to be attracting, accessible, and useful, at the same time in accordance with the law and ethics while accommodating to the broadest possible scope of users.

The strong aspects of this thesis are that the findings are based on the actual thoughts and experiences of the blind people. However, it's a challenge to generalise these findings to the whole blind population because they are based on a relatively small number of participants. Moreover, technology is rapidly and continuously changing and improving/updating and the findings might get out of date very fast because assistive technology, and IPAs particularly, might soon have new interactions and introduce new capabilities. Anyways, as of today, the results are significant and bear importance because they represent the latest barriers of IPAs.

The research questions posed are answered. The findings suggest that IPAs don't completely ensure accessibility and usability to the technology they are present for the blind and non-verbal people because they suffer from various problems and issues and complications with other devices or applications. Regarding the sub-questions, this thesis presents the barriers encountered by both the blind and non-verbal as well as the usefulness in noisy environments and further discusses creative solutions. Based on the literature, the findings, media reports, and design principles, the interfaces of IPAs facilitate accessibility and usability by considering the visual an essential component.

Another weak aspect of this thesis is that, it cannot represent all the IPAs available and attempts to cover only the most used ones. Also, it is both a challenge and a downside that the interviews only cover a lot of data about Siri and leave undiscovered in depth the experiences and barriers of the users of Alexa, Bixby, Google's Assistant, and Cortana.

Some of the findings extend to new practices such as the suggested improvements of IPAs in the next sub section based on principles of universal design. It is possible to do a thorough evaluation by putting each IPA against the principles of VUI design and compare a set of questions and responses or even a whole conversation of each IPA to assess accessibility and usability in more details, and this thesis could be criticised for not doing this. Also, there could be added more IPAs of other devices in the persona testing such as Huawei's P20 Pro smartphone which has a different IPA than the ones mentioned until now.

The findings also revealed new gaps in literature. There are limited resources that promote the universal design of intelligent personal assistants. There exists the need of the creation of precise guidelines and standards by authorities like the World Wide Web Consortium in order to set the foundations of the UD of IPAs.

8.1. Possible Solutions

When thinking about possible solutions there are a few things to have in mind. It's important to consider a combination of findings in order to come up with the best solutions. These combinations of findings include: (1) the themes of the topics of the systematic literature review and how they relate with the design principles by Pearl (2016), (2) the findings of persona testing and how they relate with the literature review and systematic literature review, (3) the findings of the interviews and observations and how they relate with the design principles by Pearl (2016), and (4) the concept of universal design with its components.

Hence, the following suggestions for solutions or improvements can be drawn: (1) IPAs should support as many languages as possible – such as the opportunity to consider integrating the power of Fluent.ai mentioned in 4.5 (2) speech recognition when dictating using IPAs should be improved to minimize misunderstandings and errors, (3) enhance IPAs by developing new algorithms or by using the sample design to distinguish facts while having in mind the accessibility principle fourteen by Pearl (2016) without making the user surf, by being time efficient and requiring less effort, (4) establish a communication between devices where IPAs are implemented, especially devices operable only by voice such as smart speakers, i.e.: possibility of using Siri with AirPlay to communicate with other devices like the HomePod or Apple TV, (5) improve integration of 3rd party applications – especially the developers of 3rd

party applications should allow IPA integration, (6) improve IPA user interface to complement the interaction when the user is impaired by the situation, (7) allow typing to IPA to facilitate the non-verbal, (8) fix general IPA problems to build trust, such as: coordination of Siri with the time set in the system (phone settings) to fix the calendar time format issue, and fix bugs like reading old notifications, (9) build on the trust between users and IPAs by using conversational markers to improve reliability in IPAs, i.e.: timeliness, acknowledgements, and positive feedback. The findings of this thesis can help and support companies and developers come up with solutions and improvements regarding the universal design of intelligent personal assistants.

9. Conclusion

As technology is evolving continuously, there are emerging many opportunities to advance the development of search systems. In the next decade or beyond, there will be new ways of interactions with the technology systems (Pearl, 2016). These interactions will be new, richer and more natural interaction capabilities. This also applies to knowledge, and the opportunity to access knowledge mined from both human resources and data with the goal of answering searchers' questions.

Intelligent Personal Assistants are not only for smartphones anymore. In time of this writing, they are expanding to the wrist, car, and even the refrigerator and toilets (Bonnington, 2018). As far as it is possible, people with disabilities should have the same experience using IPAs as other people. The blind shouldn't feel as unequal with the sighted (Hansen, 2017). When everybody is included, it benefits the individual and it benefits all society (NDA, 2018). By providing more options and improvements for the fully blind and non-verbal users whom life relies on assistive technology it will help the usability of IPAs by making it a better experience for all users. It is very important to not lose sight of the goal: making them easy to use and enjoyable.

Apple (2018b) states that "the most powerful technology in the world is technology that everyone, including people with disabilities, can use". This thesis, is going to support and help future researches explore and learn about new topics and domains and facilitate people with disabilities.

10. Appendices

10.1. Appendix A – Informed Consent

10.1.1. Purpose of the Research

I, Regont Kurtishi a master's student in Universal Design of ICT in OsloMet University (HiOA), am doing this research study for my thesis to find out what barriers do fully blind people experience when using intelligent personal assistants. With a better understanding of what prevents access to this technology, I can come up with ways to improve the user experience. I hope that the results of this study will have beneficial effects to improve intelligent personal assistants and make the life of people with disabilities easier.

10.1.2. Procedures for Participants

You are asked to perform a set of four tasks using the personal assistant in your smartphone for approximately 7 minutes prior to the interview. After performing the tasks, you are asked to be interviewed for approximately 20-30 minutes regarding your own experiences with your personal assistant.

10.1.3. Confidentiality

Participation in this study is voluntary. All information will remain strictly confidential. Although the description and findings may be published, your name or any other identification will not be used. You are at liberty to withdraw your consent to the study and discontinue participation at any time without prejudice.

I have read and understood the information on this form and had all my questions answered.

_____

Place and Date

_____

Signature

10.2. Appendix B – NSD Notification

Anthony Giannoumis
Postboks 4 St. Olavs plass
0130 OSLO

Vår dato: 26.01.2018 Vår ref: 57937 / 3 / EPA Deres dato: Deres ref:

Tilrådning fra NSD Personvernombudet for forskning § 7-27

Personvernombudet for forskning viser til meldeskjema mottatt 20.12.2017 for prosjektet:

57937	*Universal Design of Voice Assistants*
Behandlingsansvarlig	*Høgskolen i Oslo og Akershus, ved institusjonens øverste leder*
Daglig ansvarlig	*Anthony Giannoumis*
Student	*Regont Kurtishi*

Vurdering
Etter gjennomgang av opplysningene i meldeskjemaet og øvrig dokumentasjon finner vi at prosjektet er unntatt konsesjonsplikt og at personopplysningene som blir samlet inn i dette prosjektet er regulert av § 7-27 i personopplysningsforskriften. På den neste siden er vår vurdering av prosjektopplegget slik det er meldt til oss. Du kan nå gå i gang med å behandle personopplysninger.

Vilkår for vår anbefaling
Vår anbefaling forutsetter at du gjennomfører prosjektet i tråd med:
• opplysningene gitt i meldeskjemaet og øvrig dokumentasjon
• vår prosjektvurdering, se side 2
• eventuell korrespondanse med oss

Meld fra hvis du gjør vesentlige endringer i prosjektet
Dersom prosjektet endrer seg, kan det være nødvendig å sende inn endringsmelding. På våre nettsider finner du svar på hvilke endringer du må melde, samt endringsskjema.

Opplysninger om prosjektet blir lagt ut på våre nettsider og i Meldingsarkivet
Vi har lagt ut opplysninger om prosjektet på nettsidene våre. Alle våre institusjoner har også tilgang til egne prosjekter i Meldingsarkivet.

Vi tar kontakt om status for behandling av personopplysninger ved prosjektslutt
Ved prosjektslutt 30.04.2018 vil vi ta kontakt for å avklare status for behandlingen av personopplysninger.

Se våre nettsider eller ta kontakt dersom du har spørsmål. Vi ønsker lykke til med prosjektet!

Dokumentet er elektronisk produsert og godkjent ved NSDs rutiner for elektronisk godkjenning.

NSD – Norsk senter for forskningsdata AS Harald Hårfagres gate 29 Tel: +47-55 58 21 17 nsd@nsd.no Org.nr. 985 321 884
NSD – Norwegian Centre for Research Data NO-5007 Bergen, NORWAY Faks: +47-55 58 96 50 www.nsd.no

Figure 11.2 The NSD notification receipt for conducting the study.

10.3. Appendix C – Interview Guide

10.3.1. The Opening

The opening of the interview guide is all about getting to know the participant a little bit. It consists of friendly questions and answers, such as: how was your day? Any difficulties coming here? The weather today, and so on. But most importantly, making sure that the participant is comfortable and in a good mood.

10.3.2. The Body

The body of the guide includes the interview schedule with 8 questions. It is important not to rush the questions and keep a nice climate of trust.

10.3.3. The closing

The closing was done by thanking the participants and handing out their reward. Wishing them a wonderful day.

11. References

Abdolrahmani, A., Kuber, R., & Hurst, A. (2016). *An empirical investigation of the situationally-induced impairments experienced by blind mobile device users.* Paper presented at the Proceedings of the 13th Web for All Conference.

Amazon. (2017). Why Alexa? Retrieved from https://developer.amazon.com/alexa

Apple. (2017). "Hey Siri, wake me up at 7 AM tomorrow". Retrieved from https://www.apple.com/ios/siri/

Apple. (2018a). iOS Feature Availability. Retrieved from https://www.apple.com/ios/feature-availability/

Apple. (2018b). We believe that technology should be accessible to everyone. Retrieved from https://www.apple.com/lae/accessibility/

Berg, B. L., & Lune, H. (2004). *Qualitative research methods for the social sciences* (Vol. 5): Pearson Boston, MA.

Bergen, M., & Moritz, S. (2017). Google's Big Marketing Push Pays Off for its Pixel Phone Over Holiday. Retrieved from https://www.bloomberg.com/news/articles/2017-01-19/google-s-big-marketing-push-pays-off-for-its-pixel-phone-over-holiday

Bonnington, C. (2018). Alexa, Lift the Toilet Seat. *Slate*.

Bourne, R. R., Jonas, J. B., Flaxman, S. R., Keeffe, J., Leasher, J., Naidoo, K., . . . White, R. A. (2014). Prevalence and causes of vision loss in high-income countries and in Eastern and Central Europe: 1990–2010. *British Journal of Ophthalmology*.

Britt, A. (2018). 7 Key Predictions For The Future Of Voice Assistants And AI. Retrieved from https://clearbridgemobile.com/7-key-predictions-for-the-future-of-voice-assistants-and-ai/

Burzagli, L., & Emiliani, P. L. (2013). *Universal access: a concept to be adapted to technological development and societal change.* Paper presented at the International Conference on Universal Access in Human-Computer Interaction.

Cain, A. (2017). The Life and Death of Microsoft Clippy, the Paper Clip the World Loved to Hate. Retrieved from Visual Culture website: https://www.artsy.net/article/artsy-editorial-life-death-microsoft-clippy-paper-clip-loved-hate

Chen, L. (2017). Alibaba Challenges Google, Amazon With New Echo-Like Device. Retrieved from Bloomberg Technology website: https://www.bloomberg.com/news/articles/2017-07-05/alibaba-challenges-google-amazon-with-new-echo-like-device

Chris, L. (2014). Why Cortana Assistant Can Help Microsoft in the Smartphone Market. The Street.

Colorado State University. (2006). Conceptual Analysis. Retrieved from https://writing.colostate.edu/guides/page.cfm?pageid=1309&guideid=61

Cowan, B. R., Pantidi, N., Coyle, D., Morrissey, K., Clarke, P., Al-Shehri, S., . . . Bandeira, N. (2017). *What can i help you with?: infrequent users' experiences of intelligent personal assistants.* Paper presented at the Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services.

Crabtree, B. F., & Miller, W. L. (1999). *Doing qualitative research*: sage publications.

Davis, K., Biddulph, R., & Balashek, S. (1952). Automatic recognition of spoken digits. *The Journal of the Acoustical Society of America, 24*(6), 637-642.
Denes, P. (1959). The design and operation of the mechanical speech recognizer at University College London. *Journal of the British Institution of Radio Engineers, 19*(4), 219-229.
DiCicco-Bloom, B., & Crabtree, B. F. (2006). The qualitative research interview. *Medical education, 40*(4), 314-321.
Dolcourt, J. (2017). Samsung's Bixby assistant needs to grow up. Fast (hands-on). Retrieved from https://www.cnet.com/news/samsung-bixby-assistant-galaxy-s8-hands-on/
Doody, O., & Noonan, M. (2013). Preparing and conducting interviews to collect data. *Nurse researcher, 20*(5), 28-32.
Dudley, H., Riesz, R., & Watkins, S. (1939). A synthetic speaker. *Journal of the Franklin Institute, 227*(6), 739-764.
Easwara Moorthy, A., & Vu, K.-P. L. (2015). Privacy concerns for use of voice activated personal assistant in the public space. *International Journal of Human-Computer Interaction, 31*(4), 307-335.
Disability Act 2005, 14 C.F.R. (2005).
Ferguson, J. (1980). Hidden Markov analysis: an introduction. *Hidden Markov Models for Speech, 14*.
Fletcher, H. (1922). The nature of speech and its interpretation. *Bell Labs Technical Journal, 1*(1), 129-144.
Fluent.ai. (2018a). About Us. Retrieved from https://www.fluent.ai/about-us/
Fluent.ai. (2018b). Fluent.ai Brings Universal Speech Recognition to MWC. Retrieved from https://www.fluent.ai/2018/03/09/fluent-ai-brings-universal-speech-recognition-to-mwc/
Fontana, A., & Frey, J. H. (2000). The interview: From structured questions to negotiated text. *Handbook of qualitative research, 2*(6), 645-672.
Forni, A. A. (2016). Gartner Reveals Top Predictions for IT Organizations and Users in 2017 and Beyond [Press release]. Retrieved from http://www.gartner.com/newsroom/id/3482117
Frick, K. D., & Foster, A. (2003). The magnitude and cost of global blindness: an increasing problem that can be alleviated. *American journal of ophthalmology, 135*(4), 471-476.
Fuglerud, S. (2014). *Inclusive design of ICT: The challenge of diversity.* PhD dissertation. Oslo: Department of Media and Communication, University of Oslo. Available full-text DOI: 10.13140/2.1. 4471.5844.
FutureLearn. (2016). Digital Accessibility: Enabling Participation in the Information Society. Retrieved from https://www.futurelearn.com/courses/digital-accessibility/
Glasser, A., Kushalnagar, K., & Kushalnagar, R. (2017). *Deaf, Hard of Hearing, and Hearing Perspectives on Using Automatic Speech Recognition in Conversation.* Paper presented at the Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility.
Google. (2017). Meet the Google App. Retrieved from http://www.google.com/landing/now/

Google. (2018). Google Assistant. Retrieved from https://assistant.google.com/#?modal_active=none
Gurman, M., & Stone, B. (2018). Amazon Has a Top-Secret Plan to Build Home Robots. *Bloomberg Technology*.
Hansen, M. (2017). On Siri, Voice Control, and How the Blind Use iPhones. Retrieved from https://www.applevis.com/blog/ios-apps-opinion/siri-voice-control-and-how-blind-use-iphones
Haslam, O. (2017a). Amazon Echo Show Announced With 7-Inch Display, Video Calling, More. Retrieved from Redmond Pie website: http://www.redmondpie.com/amazon-echo-show-announced-with-7-inch-display-video-calling-more/
Haslam, O. (2017b). Apple's Phil Schiller On Echo, Home: Voice Assistants Are Better With A Screen. Retrieved from Redmond Pie website: http://www.redmondpie.com/apples-phil-schiller-on-echo-home-voice-assistants-are-better-with-a-screen/
Haslam, O. (2018). You Can Now Create Your Own Alexa Skills With Amazon's New Skill Blueprints. Retrieved from http://www.redmondpie.com/you-can-now-create-your-own-alexa-skills-with-amazons-new-skill-blueprints/
Henry, S. L. (2006). Understanding Web Accessibility. Retrieved from http://uiaccess.com/understanding.html
Henry, S. L. (2007). *Just ask: integrating accessibility throughout design*: Lulu. com.
High, R. (2012). The era of cognitive systems: An inside look at IBM Watson and how it works. *IBM Corporation, Redbooks*.
Holloway, I., & Galvin, K. (2016). *Qualitative research in nursing and healthcare*: John Wiley & Sons.
Holsti, O. R. (1968). Content analysis. *The handbook of social psychology, 2*, 596-692.
Huffman, S. (2014). OMG! Mobile voice survey reveals teens love to talk. Retrieved from https://googleblog.blogspot.no/2014/10/omg-mobile-voice-survey-reveals-teens.html
International Business Machines. (2012). IBM Shoebox. Retrieved from http://www-03.ibm.com/ibm/history/exhibits/specialprod1/specialprod1_7.html
International Organization for Standardization. (1998). Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability (Vol. 9241-11).
International organization for Standardization. (2011). 9999 Assistive products for persons with disability -- Classification and terminology: Bruxelles: ISO.
Jacob, S. A., & Furgerson, S. P. (2012). Writing interview protocols and conducting interviews: Tips for students new to the field of qualitative research. *The Qualitative Report, 17*(42), 1-10.
Jöst, M., Häußler, J., Merdes, M., & Malaka, R. (2005). *Multimodal interaction for pedestrians: an evaluation study.* Paper presented at the Proceedings of the 10th international conference on Intelligent user interfaces.
Juang, B.-H., & Rabiner, L. R. (2005). Automatic speech recognition–a brief history of the technology development. *Georgia Institute of Technology. Atlanta Rutgers University and the University of California. Santa Barbara, 1*, 67.
Jurafsky, D., & James, H. (2000). Speech and language processing an introduction to natural language processing, computational linguistics, and speech.

Kastrenakes, J. (2016). Siri's creators will unveil their new AI bot on Monday. Retrieved from https://www.theverge.com/2016/5/4/11593564/viv-labs-unveiling-monday-new-ai-from-siri-creators
Kratzenstein, C. G. (1782). Sur la naissance de la formation des voyelles. *Journal de Physique, 21*, 358-380.
Kulatunga, U., Amaratunga, R., & Haigh, R. (2007). Structuring the unstructured data: the use of content analysis.
Kuniavsky, M. (2003). *Observing the user experience: a practitioner's guide to user research*: Morgan kaufmann.
Lazar, J., Feng, J. H., & Hochheiser, H. (2017). *Research methods in human-computer interaction*: Morgan Kaufmann.
Lee, D. J., Gómez-Marín, O., Lam, B. L., Zheng, D. D., & Caban, A. (2005). Visual impairment and morbidity in community-residing adults: the national health interview survey 1986–1996. *Ophthalmic epidemiology, 12*(1), 13-17.
Lee, K.-F. (1988). On large-vocabulary speaker-independent continuous speech recognition. *Speech communication, 7*(4), 375-379.
Lewin, M. R., & Lowitz, J. N. (2016). *Amazon Echo - What We Know Now*. Retrieved from http://files.constantcontact.com/150f9af2201/70c07fdd-a197-4505-9476-e83aa726f025.pdf
Lewis, J. R. (2006). Sample sizes for usability tests: mostly math, not magic. *interactions, 13*(6), 29-33.
Lofland, J. (1995). Analytic ethnography: Features, failings, and futures. *Journal of contemporary ethnography, 24*(1), 30-67.
Luger, E., & Sellen, A. (2016). *Like having a really bad PA: the gulf between user expectation and experience of conversational agents.* Paper presented at the Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.
Marchick, A. (2017). The 2017 Voice Report by VoiceLabs. Retrieved from http://voicelabs.co/2017/01/15/the-2017-voice-report/
Marks, D. (1997). Models of disability. *Disability and rehabilitation, 19*(3), 85-91.
Mauldin, M. L. (1994). *Chatterbots, tinymuds, and the turing test: Entering the loebner prize competition.* Paper presented at the AAAI.
McCarty, C. A., Nanjan, M. B., & Taylor, H. R. (2001). Vision impairment predicts 5 year mortality. *British Journal of Ophthalmology, 85*(3), 322-326.
Metrock, B. (2017). The Voice-First Workplace is a Big Development for 2018. Retrieved from https://www.voicebot.ai/2018/01/07/voice-first-workplace-big-development-2018/
Michael, S., Bryman, A., & Liao, T. (2004). The sage encyclopedia of social science research methods. *University of Michigan. The United States of America*.
Microsoft. (2017). Cortana is your truly personal digital assistant. Retrieved from https://www.microsoft.com/en-us/windows/cortana
Milhorat, P., Schlogl, S., & Chollet, G. (2014). *BUILDING THE NEXT GENERATION OF PERSONAL DIGITAL ASSISTANTS*. Paper presented at the 1st International Conference on Advanced Technologies
for Signal and Image Processing - ATSIP'2014, Sousse, Tunisia. http://ieeexplore.ieee.org.ezproxy.hioa.no/stamp/stamp.jsp?arnumber=6834655

Miller, C. (2017). What feature is a ‘must have’ for you to use Apple’s Siri Speaker? [Poll]. *9to5mac*.
Mohri, M. (1997). Finite-state transducers in language and speech processing. *Computational linguistics, 23*(2), 269-311.
Moorthy, A. E., & Vu, K.-P. L. (2014). *Voice activated personal assistant: Acceptability of use in the public space.* Paper presented at the International Conference on Human Interface and the Management of Information.
Morris, P. (2018). Siri Reading Hidden Notifications On iOS Lock Screen Bug To Be Fixed in Upcoming Update, Apple Confirms. Retrieved from http://www.redmondpie.com/ios-11.2.7-11.3-final-to-fix-siri-reading-hidden-notifications-on-lock-screen-bug/
Mutchler, A. (2017). Voice Assistant Timeline: A Short History of the Voice Revolution. Retrieved from https://www.voicebot.ai/2017/07/14/timeline-voice-assistants-short-history-voice-revolution/
Nagata, S. F., Oostendorp, H. v., & Neerincx, M. A. (2004). *Interaction design concepts for a mobile personal assistant*. Paper presented at the Proceedings of the conference on Dutch directions in HCI, Amsterdam, The Netherlands.
National Disability Authority. (2009). Ethical Guidance for Research with People with Disabilities.
National Disability Authority. (2018). Benefits and Drivers.
National Institutes of Health. (2016). *Ethical Challenges of Research*. Retrieved from https://obssr.od.nih.gov/wp-content/uploads/2016/05/Ethical-Challenges.pdf.
National Telecommunications & Information Administration. (1995). FALLING THROUGH THE NET:
A SURVEY OF THE "HAVE NOTS" IN RURAL AND URBAN AMERICA. Retrieved from https://www.ntia.doc.gov/ntiahome/fallingthru.html
Neuman, L. W. (2002). Social research methods: Qualitative and quantitative approaches.
Nielsen, J. (2012). Usability 101: Introduction to Usability. *Nielsen Norman Group*.
Norges Blindeforbund. (2016). Fakta og statistikk om synshemninger. Retrieved from https://www.blindeforbundet.no/oyehelse-og-synshemninger/fakta-og-statistikk-om-synshemninger
Norges Blindeforbund. (2018). Om Blindeforbundet. Retrieved from https://www.blindeforbundet.no/om-blindeforbundet
Norges Offentlige Utredninger. (1999). *Fra bruker til borger*. Oslo Retrieved from https://www.regjeringen.no/contentassets/1e18b045dd9346849813392b34c9cdc1/no/pdfa/nou200120010022000dddpdfa.pdf.
Norges Offentlige Utredninger. (2003). *St.meld. nr. 40 (2002-2003)*. Retrieved from https://www.regjeringen.no/no/dokumenter/stmeld-nr-40-2002-2003-/id197129/sec1.
Norges Offentlige Utredninger. (2005). *Likeverd og tilgjengelighet*. Retrieved from https://www.regjeringen.no/contentassets/e2361e34e995496589470336829751cc/no/pdfs/nou200520050008000dddpdfs.pdf.
Notification form for research- and student projects processing personal data, (2017).
Nwana, H. S. (1996). Software agents: An overview. *The knowledge engineering review, 11*(03), 205-244.

Olson, C. (2016). Just say it: The future of search is voice and personal digital assistants. *Campaign*.
Organisation for Economic Cooperation and Development. (2001). *Understanding the Digital Divide*. Retrieved from Paris: https://www.oecd.org/sti/1888451.pdf
Organization, W. H. (1968). *International statistical classification of diseases, injuries and causes of death*: Government Printer.
Osborne, J. (2016). Why 100 million monthly Cortana users on Windows 10 is a big deal. Retrieved from http://www.techradar.com/news/software/operating-systems/why-100-million-monthly-cortana-users-could-be-a-bigger-deal-than-350-million-windows-10-installs-1325146
Overmyer, K. (2016). What the Voice Recognition Explosion Will Mean for Content Marketing. Retrieved from https://www.skyword.com/contentstandard/marketing/what-the-voice-recognition-explosion-will-mean-for-content-marketing/
Palmquist, M. E., Carley, K. M., & Dale, T. A. (1997). Two applications of automated text analysis: Analyzing literary and non-literary texts. *Text analysis for the social sciences: Methods for drawing statistical inferences from texts and transcripts*, 171-189.
Pearl, C. (2016). *Designing Voice User Interfaces: Principles of Conversational Experiences*: " O'Reilly Media, Inc.".
Pilon, A. (2016). AI Personal Assistants Survey: Few Plan to Increase Usage. Retrieved from https://aytm.com/blog/daily-survey-results/ai-personal-assistants-survey/#sthash.qXXxsGXI.dpbs
Pinola, M. (2011). Speech Recognition Through the Decades: How We Ended Up With Siri. *PCWorld*.
Qu, S. Q., & Dumay, J. (2011). The qualitative research interview. *Qualitative research in accounting & management, 8*(3), 238-264.
Resnikoff, S., Pascolini, D., Etya'Ale, D., Kocur, I., Pararajasegaram, R., Pokharel, G. P., & Mariotti, S. P. (2004). Global data on visual impairment in the year 2002. *Bulletin of the world health organization, 82*(11), 844-851.
Ritchie, J., Lewis, J., Nicholls, C. M., & Ormston, R. (2013). *Qualitative research practice: A guide for social science students and researchers*: Sage.
Robson, C. (2002). Real world research. 2nd. *Edition. Blackwell Publishing. Malden*.
Ryan, G. W., & Bernard, H. R. (2003). Techniques to identify themes. *Field methods, 15*(1), 85-109.
Samsung. (2017). A new way to interact with your phone. Retrieved from http://www.samsung.com/global/galaxy/apps/bixby/
Santos, J., Rodrigues, J. J., Casal, J., Saleem, K., & Denisov, V. (2016). Intelligent Personal Assistants Based on Internet of Things Approaches. *IEEE Systems Journal*.
Schwartz, R., Barry, C., Chow, Y.-L., Derr, A., Feng, M.-W., Kimball, O., . . . Vandegrift, J. (1989). *The BBN BYBLOS continuous speech recognition system.* Paper presented at the Proceedings of the workshop on Speech and Natural Language.
Sentance, R. (2016). What does Meeker's Internet Trends report tell us about voice search? Retrieved from https://searchenginewatch.com/2016/06/03/what-does-meekers-internet-trends-report-tell-us-about-voice-search/

Shakespeare, T. (2006). The social model of disability. *The disability studies reader, 2*, 197-204.
Simonite, T. (2017). Facebook's Perfect, Impossible Chatbot. *MIT Technology Review*.
Smith, B. (2017). *OK, GOOGLE: DESIGNING INFORMATION ARCHITECTURE FOR SMART SPEAKERS.* (Master).
Smith, J. (2017). *THE VOICE ASSISTANT LANDSCAPE REPORT: How artificially intelligent voice assistants are changing the relationship between consumers and computers*. Retrieved from http://www.businessinsider.com/voice-assistant-report-2017-3?r=US&IR=T&IR=T
Statped. (2018). Om Statped. Retrieved from http://www.statped.no/om-statped/
Stemler, S. (2001). An overview of content analysis. *Practical assessment, research & evaluation, 7*(17), 137-146.
Sterling, G. (2016). Google says 20 percent of mobile queries are voice searches. Retrieved from http://searchengineland.com/google-reveals-20-percent-queries-voice-queries-249917
Story, M. F., Mueller, J. L., & Mace, R. L. (1998). *The Universal Design File: Designing for People of All Ages and Abilities. Revised Edition.*: NC State University.
Suzuki, J. (1961). Recognition of Japanese vowels. *Journal of the Radio Research Laboratory, 8*(37), ????
Taylor, S. J., Bogdan, R., & DeVault, M. (2015). *Introduction to qualitative research methods: A guidebook and resource*: John Wiley & Sons.
Theodoridis, S. (2003). Konstantinos Koutroumbas Pattern recognition Elsevier: Academic Press.
Thompson, T. (2012). Tech Tips: Are You Talking To Your Computer Again? Retrieved from http://www.washington.edu/doit/tech-tips-are-you-talking-your-computer-again
Thylefors, B., Negrel, A., Pararajasegaram, R., & Dadzie, K. (1995). Global data on blindness. *Bulletin of the world health organization, 73*(1), 115.
Tollefsen, M. (2011). Universell utforming som prosess, virkemiddel og mål i utvikling av ny programvare. Retrieved from http://www.medialt.no/universell-utforming-som-prosess-virkemiddel-og-maal-i-utvikling-av-ny-programvare/1016.aspx
Tundrea, D. (2017). *Artificial Intelligence is Getting Personal.* (Master), Upsala Universitetet.
UDLL. (2016). A Best Practice Guideline. *Universal Design for Learning in Higher Education – License to Learn*.
Van Der Meulen, R. (2016). Gartner Says by 2019, 20 Percent of User Interactions With Smartphones Will Take Place via VPAs [Press release]. Retrieved from https://www.gartner.com/newsroom/id/3551217
Waibel, A., & Lee, K.-F. (1990). *Readings in speech recognition*: Morgan Kaufmann.
Web Accessibility Initiative. (2005). Introduction to Web Accessibility. Retrieved from https://www.w3.org/WAI/intro/accessibility
Weber, R. P. (1990). *Basic content analysis*: Sage.
Wheatstone, S. C. (1879). *The scientific papers*: Taylor & Francis, for the Physical Society of London.
World Health Organization. (2017). *Blindness and visual impairment*. Retrieved from http://www.who.int/en/news-room/fact-sheets/detail/blindness-and-visual-impairment

Young, K. (2016). 25% of 16-24s use voice search on mobile. Retrieved from http://blog.globalwebindex.net/chart-of-the-day/25-of-16-24s-use-voice-search-on-mobile/

Young, S., Evermann, G., Gales, M., Hain, T., Kershaw, D., Liu, X., . . . Povey, D. (2002). The HTK book. *Cambridge university engineering department, 3*, 175.

Yurieff, K. (2017). Voice-powered search is replacing typing — and other internet trends. Retrieved from http://money.cnn.com/2017/05/31/technology/mary-meeker-internet-trends/index.html